Year : 2012

COMMUNITY POLICING IN SWITZERLAND'S MAJOR URBAN AREAS

KREIS CHRISTIAN

UNIL | Université de Lausanne

Faculté de droit et des sciences criminelles
Institut de criminologie et de droit pénal

Faculté des géosciences et de l'environnement
Institut de géomatique et d'analyse du risque

Community Policing in Switzerland's Major Urban Areas

An Observational Study of the Implementation and Impact Using Geospatial Data Mining

THÈSE DE DOCTORAT

présentée à la Faculté de droit et des sciences criminelles
de l'Université de Lausanne par

Christian Kreis

Licentiatus Philosophiae

Université de Zurich, Suisse

Jury:	
Président du colloque:	Prof. Marcelo Aebi
Co-directeur de thèse:	Prof. André Kuhn
Co-directeur de thèse:	Prof. Mikhail Kanevski
Expert interne:	Prof. Martin Killias
Expert externe:	Prof. Martin Schuler
Expert externe:	Prof. Henk Elffers

Lausanne, 2012

UNIL | Université de Lausanne
Ecole des sciences criminelles
bâtiment Batochime
CH-1015 Lausanne

IMPRIMATUR

A l'issue de la soutenance de thèse, le Jury autorise l'impression de la thèse de M. Christian Kreis, candidat au doctorat en criminologie, intitulée

« Community Policing in Switzerland's Major Urban Areas »

Le Président du Jury

Professeur Marcelo Aebi

Lausanne, le 15 mars 2012

Under a streetlamp there stands a drunk who searches and searches. A policeman comes along, asks him what he is looking for, and the man answers, "My keys." Now they both search. After a while the policeman wants to know whether the man is sure that he lost his key here, and the latter answers, "No, not here, back there – but there it is much too dark."

Paul Watzlawick,
The Situation Is Hopeless, But Not Serious.

Community Policing in Switzerland's Major Urban Areas

Résumé

La présente étude est à la fois une évaluation du processus de la mise en œuvre et des impacts de la police de proximité dans les cinq plus grandes zones urbaines de Suisse – Bâle, Berne, Genève, Lausanne et Zurich. La police de proximité (community policing) est à la fois une philosophie et une stratégie organisationnelle qui favorise un partenariat renouvelé entre la police et les communautés locales dans le but de résoudre les problèmes relatifs à la sécurité et à l'ordre public.

L'évaluation de processus a analysé des données relatives aux réformes internes de la police qui ont été obtenues par l'intermédiaire d'entretiens semi-structurés avec des administrateurs clés des cinq départements de police, ainsi que dans des documents écrits de la police et d'autres sources publiques. L'évaluation des impacts, quant à elle, s'est basée sur des variables contextuelles telles que des statistiques policières et des données de recensement, ainsi que sur des indicateurs d'impacts construit à partir des données du *Swiss Crime Survey* (SCS) relatives au sentiment d'insécurité, à la perception du désordre public et à la satisfaction de la population à l'égard de la police. Le SCS est un sondage régulier qui a permis d'interroger des habitants des cinq grandes zones urbaines à plusieurs reprises depuis le milieu des années 1980.

L'évaluation de processus a abouti à un "Calendrier des activités" visant à créer des données de panel permettant de mesurer les progrès réalisés dans la mise en œuvre de la police de proximité à l'aide d'une grille d'évaluation à six dimensions à des intervalles de cinq ans entre 1990 et 2010. L'évaluation des impacts, effectuée *ex post facto*, a utilisé un concept de recherche non-expérimental (*observational design*) dans le but d'analyser les impacts de différents modèles de police de proximité dans des zones comparables à travers les cinq villes étudiées. Les quartiers urbains, délimités par zone de code postal, ont ainsi été regroupés par l'intermédiaire d'une typologie réalisée à l'aide d'algorithmes d'apprentissage automatique (*machine learning*). Des algorithmes supervisés et non supervisés ont été utilisés sur les données à haute dimensionnalité relatives à la criminalité, à la structure socio-économique et démographique et au cadre bâti dans le but de regrouper les quartiers urbains les plus similaires dans des clusters. D'abord, les cartes auto-organisatrices (*self-organizing maps*) ont été utilisées dans le but de réduire la variance intra-cluster des variables contextuelles et de maximiser simultanément la variance inter-cluster des réponses au sondage. Ensuite, l'algorithme des forêts d'arbres décisionnels (random forests) a permis à la fois d'évaluer la pertinence de la typologie de quartier élaborée et de sélectionner les variables contextuelles clés afin de construire un modèle parcimonieux faisant un minimum d'erreurs de classification. Enfin, pour l'analyse des impacts, la méthode des appariements des coefficients de propension (*propensity score matching*) a été utilisée pour équilibrer les échantillons prétest-posttest en termes d'âge, de sexe et de niveau d'éducation des répondants au sein de chaque type de quartier ainsi identifié dans chacune des villes, avant d'effectuer un test statistique de la différence observée dans les indicateurs d'impacts. De plus, tous les résultats statistiquement significatifs ont été soumis à une analyse de sensibilité (*sensitivity analysis*) afin d'évaluer leur robustesse face à un biais potentiel dû à des covariables non observées.

L'étude relève qu'au cours des quinze dernières années, les cinq services de police ont entamé des réformes majeures de leur organisation ainsi que de leurs stratégies opérationnelles et qu'ils ont noué des partenariats stratégiques afin de mettre en œuvre la police de proximité. La typologie de quartier développée a abouti à une réduction de la variance intra-cluster des variables contextuelles et permet d'expliquer une partie significative de la variance inter-cluster des indicateurs d'impacts avant la mise en œuvre du traitement. Ceci semble suggérer que les méthodes de géocomputation aident à équilibrer les covariables observées et donc à réduire les menaces relatives à la validité interne d'un concept de recherche non-expérimental. Enfin, l'analyse des impacts a révélé que le sentiment d'insécurité a diminué de manière significative pendant la période 2000-2005 dans les quartiers se trouvant à l'intérieur et autour des centres-villes de Berne et de Zurich. Ces améliorations sont assez robustes face à des biais dus à des covariables inobservées et covarient

dans le temps et l'espace avec la mise en œuvre de la police de proximité. L'hypothèse alternative envisageant que les diminutions observées dans le sentiment d'insécurité soient, partiellement, un résultat des interventions policières de proximité semble donc être aussi plausible que l'hypothèse nulle considérant l'absence absolue d'effet. Ceci, même si le concept de recherche non-expérimental mis en œuvre ne peut pas complètement exclure la sélection et la régression à la moyenne comme explications alternatives.

Community Policing in Switzerland's Major Urban Areas

Summary

The current research project is both a process and impact evaluation of community policing in Switzerland's five major urban areas – Basel, Bern, Geneva, Lausanne, and Zurich. Community policing is both a philosophy and an organizational strategy that promotes a renewed partnership between the police and the community to solve problems of crime and disorder.

The process evaluation data on police internal reforms were obtained through semi-structured interviews with key administrators from the five police departments as well as from police internal documents and additional public sources. The impact evaluation uses official crime records and census statistics as contextual variables as well as Swiss Crime Survey (SCS) data on fear of crime, perceptions of disorder, and public attitudes towards the police as outcome measures. The SCS is a standing survey instrument that has polled residents of the five urban areas repeatedly since the mid-1980s.

The process evaluation produced a "Calendar of Action" to create panel data to measure community policing implementation progress over six evaluative dimensions in intervals of five years between 1990 and 2010. The impact evaluation, carried out *ex post facto*, uses an observational design that analyzes the impact of the different community policing models between matched comparison areas across the five cities. Using ZIP code districts as proxies for urban neighborhoods, geospatial data mining algorithms serve to develop a neighborhood typology in order to match the comparison areas. To this end, both unsupervised and supervised algorithms are used to analyze high-dimensional data on crime, the socio-economic and demographic structure, and the built environment in order to classify urban neighborhoods into clusters of similar type. In a first step, self-organizing maps serve as tools to develop a clustering algorithm that reduces the within-cluster variance in the contextual variables and simultaneously maximizes the between-cluster variance in survey responses. The random forests algorithm then serves to assess the appropriateness of the resulting neighborhood typology and to select the key contextual variables in order to build a parsimonious model that makes a minimum of classification errors. Finally, for the impact analysis, propensity score matching methods are used to match the survey respondents of the pretest and posttest samples on age, gender, and their level of education for each neighborhood type identified within each city, before conducting a statistical test of the observed difference in the outcome measures. Moreover, all significant results were subjected to a sensitivity analysis to assess the robustness of these findings in the face of potential bias due to some unobserved covariates.

The study finds that over the last fifteen years, all five police departments have undertaken major reforms of their internal organization and operating strategies and forged strategic partnerships in order to implement community policing. The resulting neighborhood typology reduced the within-cluster variance of the contextual variables and accounted for a significant share of the between-cluster variance in the outcome measures prior to treatment, suggesting that geocomputational methods help to balance the observed covariates and hence to reduce threats to the internal validity of an observational design. Finally, the impact analysis revealed that fear of crime dropped significantly over the 2000-2005 period in the neighborhoods in and around the urban centers of Bern and Zurich. These improvements are fairly robust in the face of bias due to some unobserved covariate and covary temporally and spatially with the implementation of community policing. The alternative hypothesis that the observed reductions in fear of crime were at least in part a result of community policing interventions thus appears at least as plausible as the null hypothesis of absolutely no effect, even if the observational design cannot completely rule out selection and regression to the mean as alternative explanations.

Acknowledgements

First of all, I would like to thank my two thesis advisors, Prof. André Kuhn and Prof. Mikhail Kanevski, for their steadfast support and encouragement. André gave of his time whenever necessary, patiently listening to my every question and offering his experience and advice. A formidable debater who will point out the slightest inconsistency in an argument, his incisive comments have been a valuable influence on my work. Mikhail has a wonderful knack of motivating his students and bringing out the best in them. I am most indebted to him for making the fascinating worlds of spatial statistics and data mining accessible to me.

I would also like to express my gratitude to Prof. Martin Killias who gently nudged me in the right direction at the very beginning of the project and who has been supportive of my research ever since. I am most grateful to him for letting me work with his extraordinary data and for all his support with the SNSF grant applications. I would also like to thank the two external experts on my examination board, Prof. Henk Elffers and Prof. Martin Schuler, for the time and energy they invested to evaluate my work and their highly useful comments, and Prof. Marcelo Aebi for presiding over the board. I am most indebted to Henk for all his efforts to make possible the postdoc fellowship at the Netherlands Study Center for Crime and Law Enforcement in Amsterdam.

In my research, I have also benefited greatly from the wealth of experience of several more scientific experts. I am heavily indebted to Dr. Jean-Philippe Antonietti for countless discussions of statistical methodology and for his encouragement to try out the statistical computing software R. Dr. André Berchtold deserves credit for teaching me the essence of complex survey analysis, and Dr. Arlette Mottaz-Baran repeatedly offered invaluable advice for the handling and coding of qualitative data. Last but most definitely not least, I would like to thank Dr. Alexandre Hirzel for sharing his passion for GIS software and offering his advice and technical support whenever necessary.

The current project was first made possible thanks to the ready cooperation of the five cities' police departments. Moving from west to east, I am particularly grateful to Didier Froidevaux and to Captain Claude Pahud of the Geneva Cantonal Police Department for their professionalism and candid discussion of the history of community policing in Geneva. In Lausanne, I owe special thanks to the Chief of Police Gérald Hagenlocher (retired) as well as to Captain Christian Pannatier and Captain Frédéric Pilloud for their candid discussions of community policing at the Lausanne Municipal Police Department. I also thank Christiane Roh of the Vaud Cantonal Statistics Office for providing me with her data on neighborhood level crime rates in the Lausanne urban area. In Bern, a special thanks goes to Stefan Krebs for the various discussions of the Bern community policing model. I am especially indebted to Thomas Suter who went the extra mile to run a special query of the Bern police crime records for the current study. In Basel, I would like to thank Corporal Max Morgenthaler and, especially, Corporal Jean-Pierre Roubaty for their warm welcome and unmatched efforts to provide me with accurate information on community policing at the Basel Cantonal Police Department. In Zurich, my thanks go first of all to Captain Jürg Müller for his candidness and support as well as to Captain Jörg Balzer and Rolf Vieli for the in-depth insights into the "SIP Züri" and "Langstrasse Plus" projects, and to Adjutant Erwin Brühlmann. My thanks also go to Peter Sonder for providing me with the police crime statistics on Zurich's administrative districts.

I also would like to express my deep appreciation to my colleagues at the ICDP for the congenial atmosphere and joyful working days they have provided me. I am particularly grateful to Anne-Laure, Aude, Benjamin, Camille, Catherine, Daphné, Hervé, Jessica, Joëlle, Julian, Loïc, Manon, Marie, Maryam, Natalia, Nathalie, Véronique, and Yann as well as to Daisy, Huguette, Isabelle, Magali, Marie-Louise, and Anne Marville. During the second half of the project, I regularly worked at IGAR where I greatly enjoyed the company of Carmen, Devis, Delphine, Giona, Jean, Loris, Marj, Michele, Nathan, Sylvain, and Carole. I am particularly grateful to Loris for the many methodological discussions and his useful comments for the postdoc grant application and to Sylvain for giving me a head start on the random forests algorithm.

Special thanks also go to Yolanda Gonzalez for polishing the original community policing implementation reports on police reform in Geneva and Lausanne, which I first drafted in French.

Finally, an especially warm vote of thanks goes to my family for their encouragement and support, to my parents for their guidance and inspiration, and for their patience and understanding that I spent the past three years in Lausanne.

* * *

This research was supported by the two Swiss National Science Foundation research grants "Community Policing in Switzerland: A Comparative Evaluation Study of the Community Policing Strategies of the Police in Major Swiss Urban Areas" (No 100015-122463/1) and "Community Policing in Switzerland: A Comparative Evaluation of the Impact of Community Policing on Neighborhood Development in Major Swiss Urban Areas" (No 100017-132675/1). The opinions, findings, and conclusions or recommendations expressed herein are those of the author and do not necessarily reflect those of the SNSF.

I want to express my appreciation to the SNSF for supporting the two research projects, which financed my Ph.D. student position for a total of 3 years over 3.5 years and thus made it possible to conduct this research in quasi ideal conditions. I would also like to thank the anonymous reviewers of the two research proposals for their helpful comments.

Contents

List of Tables

List of Figures

Acronyms

2-D	two dimensional
a.k.a.	also known as
BCS	British Crime Survey
BCU	Basic Command Unit
BMU	best matching unit
CAPS	Chicago Alternative Policing Strategy
CART	classification and regression trees
CATI	computer-assisted telephone interviews
CCTV	closed circuit television
CDRP	Crime and Disorder Reduction Partnership
cf.	see
CoA	Calendar of Action
CPD	Chicago Police Department
CPTED	Crime Prevention through Environmental Design
CRIPOL	official crime statistics of the Vaud Cantonal Police Department
DAC	District Advisory Committee
DBI	Davies-Bouldin Index
e.g.	for example
f.	following (page)
GIS	Geographic Information Systems
i.e.	that is
ibid.	cited previously
ICVS	International Crime Victimization Survey
ISP	Swiss Institute of Police
KRISTA	official crime statistics of the Zurich Cantonal Police Department
MAUP	modifiable areal unit problem
MDS	multi-dimensional scaling
NGO	non-governmental organization
NLO	neighborhood liaison officer
NPP	Neighbourhood Policing Programme
NRC	National Research Council
NRPP	National Reassurance Policing Programme
NYPD	New York City Police Department
OFS	Swiss Federal Office of Statistics
OOB	out-of-bag
p.	page
PCA	principal component analysis
PERF	Police Executive Research Forum
pp.	pages
SARA	Scanning, Analysis, Response, and Assessment
SCRIS	Vaud Cantonal Statistics Office

SCS	Swiss Crime Survey
SES	socio-economic status
SNSF	Swiss National Science Foundation
SOM	self-organizing map
SPI	Swiss Institute of Police
VI	variable importance
ZIP	Zone Improvement Program (U.S. mail delivery system)

1

Introduction

"In 1992, the *gendarmerie* scored success in the fight against petty crime. ... In our mission to maintain order and safeguard 'International Geneva', we put in more hours with fewer staff. Yet, these results are deceptive. ... While quantitatively and globally, [our] performance keeps improving, qualitatively and in terms of human relations, it keeps deteriorating, imperceptibly but surely."[1] Thus the analysis of the commander of the *gendarmerie* (patrol division) of the Geneva cantonal police, Guy Baer, in his department's in-house publication in 1993. More ominously, Baer observed that in spite of the apparent successes in combating crime, fear of crime among the general population was increasing. Determined to stop this slide, Baer announced the most ambitious reform agenda of his police organization in a generation (Baer 1993, 3f.).

Under the plan "Pégase" (Pegasus), officially launched in 1994, the Geneva police's *gendarmerie* embarked on a major internal reorganization to introduce a *police de proximité* (community policing) that should foster regular contacts between ordinary citizens and the police. The plan aimed to boost the feeling of safety of the general population by increasing police visibility through the renewed deployment of police officers on foot patrol and the creation of a new bike patrol. Moreover, the plan "Pégase" created a new position at each police station scattered across the canton's territory that was to be filled by an experienced police officer. These *îlotiers de quartier* (neighborhood liaison officers, NLO) enjoyed a high degree of autonomy within the station's hierarchy and were tasked exclusively with community outreach and problem-solving activities (Baer 1993, 6). The plan "Pégase" was hence the first formal effort to introduce the concept of community policing in a Swiss police force.

As community policing was gaining ground internationally during the 1990s, a few visionary chiefs of police in Switzerland were anxious not to miss out on this new idea whose time had apparently come. A number of cantonal and municipal police departments began to devise their own reform agendas to adopt the new policing paradigm. Over the decade between 1994 and 2004, the police departments of Lausanne, Basel, Bern, and Zurich all followed suit, each undertaking major reforms of their internal organization and operating strategies to implement community policing (Bolle & Knoepfler 2000a; Bolle & Knoepfler 2000b; Sauter *et al.* 2005).

Two years after the launch of the plan "Pégase", the Lausanne Municipal Police Department adopted its own community policing model, which was substantially expanded in 2001. In the German part of Switzerland, the cantonal police of the City of Basel and the Bern Municipal Police Department each adopted their own reform agendas in 1998 and 1999, respectively, and began implementing far-reaching organizational reforms in view of a department-wide transition to community policing. In both cities, the police for the first time divided up the city territory into

[1] "En 1992, la gendarmerie a enregistré des résultats positifs dans le domaine de la lutte contre la petite criminalité. ... En matière de maintien de l'ordre et de la sécurité de la 'Genève Ville internationale', nous avons consacré plus d'heures d'engagement avec moins d'effectif. Cependant, ces résultats sont trompeurs. ... En fait, quantitativement et globalement, la performance ne cesse de s'améliorer, mais qualitativement et sur le plan des relations humaines, elle ne cesse de se détériorer imperceptiblement, mais sûrement." (Baer 1993, 3).

separate police precincts and stationed their staff from all divisions equally over the corresponding precinct police stations. In Zurich, the municipal police department in 1999 began revamping its network of neighborhood police stations in order to boost its emergency response capabilities and to foster police-citizen contacts but stopped short of fully embracing the community policing agenda. Five years later, in 2004, the Zurich Municipal Police Department implemented another major reform of its organization and operating procedures in view of a full-blown transition to community policing.

Broadly speaking, the objectives of these organizational and strategic reforms at the five police departments were to decentralize the provision of basic police services and regular patrols, to introduce flatter command structures, to improve internal cooperation between different divisions, and to foster regular contacts between the police agencies, local interest groups, and ordinary citizens. In accord with the problem-solving doctrine of community policing, the five police departments have also forged strategic partnerships with other branches of the local government and civil society groups. Police departments have duly reformed the basic and, to a lesser extent, continued training of the patrol and neighborhood liaison officers to hone their interpersonal and conflict management skills. Finally, all five police departments have adopted strategic performance targets to reduce crime, allay fear of crime among the population, check disorder in public places, and improve or maintain good police-community relations.

The advent of community policing in Switzerland was hardly revolutionary. It is arguably more accurate to speak of a long march of community policing that began in the mid-1990s on the shores of Lake Geneva and culminated in 2004 with the introduction of a new federal examination that police recruits nowadays take upon completion of their basic training at their local or regional police academy. This federal examination, in which community policing theory and practice are mandatory subjects, has firmly entrenched community policing in the curriculum of the basic training at the local police academies. In a rare instance of national cooperation, the major cantonal and municipal police departments in Switzerland jointly drafted a community policing teaching manual to prepare police recruits for the community policing part of the national examination (Spaar *et al.* 2007).

The Swiss police forces did not draw up their community policing agendas from scratch. All of them looked abroad, seeking to emulate community policing models that were already being practiced elsewhere. In Geneva and Lausanne, due to the physical proximity and cultural affinity, the police administrators looked towards neighboring France and Quebec for inspiration. The police administrators in Basel and Zurich had sought guidance from a publication of the Federal Criminal Police Office (Bundeskriminalamt) of Wiesbaden, Germany, which had adapted the American experiences with community policing to the German context (Bässmann & Vogt 1997). (The police in Bern had largely looked towards Basel for theoretical guidance.) Moreover, police administrators from all three Swiss German cities took study trips to Amsterdam in the mid-1990s to get a glimpse of the community policing model practiced there. In order to understand the origins and the constituting elements of community policing, it is thus necessary for a moment to look beyond the Alpine meadows of Switzerland and to take a quick look at recent policing history in the United States of America, the birthplace of this newly ascendant paradigm in Swiss policing.

1.1 The Roots of Community Policing

The Community Policing Consortium, a forum created by the U.S. Department of Justice's Bureau of Justice Assistance, describes community policing as "a collaboration between the police and the community that identifies and solves community problems." (Bureau of Justice Assistance 1994, vii). Community policing is based on the recognition that the police cannot solve the problems of crime and disorder that undermine community welfare alone and therefore must involve all elements of the community in the search for solutions to these problems (Fridell 2004, 3).

Community policing was not the brainchild of a single author but the confluence of several ideas that emerged gradually over the course of the final two decades of the 20th century. As a consequence, there has never been a universally accepted definition of community policing, and the definitions offered by the leading scholars have evolved slightly over time. Trojanowicz & Bucqueroux (1994, 2) defined community policing essentially as "a philosophy and an organizational strategy that promotes a new partnership between people and their police. It is based on the premise that both the police and the community must work together to identify, prioritize, and solve contemporary problems such as crime, drugs, fear of crime, social and physical disorder, and overall neighborhood decay, with the goal of improving the overall quality of life in the area." A decade later, Skogan & Roth (2004, xvii) described community policing somewhat less ambitiously as a "new model of policing, which supplements traditional crime fighting with a problem-solving and problem-oriented approach that emphasizes the role of the public in helping set police priorities."

Skogan & Roth (2004, xvii) called community policing "the most important development in [American] policing in the past quarter century." Community policing grew out of a theoretical and empirical critique of the dominant paradigm in American policing during the 20th century (p. xviii-xix). This "reform era" or "professional" model of policing had narrowed the police mission to combating crime and reduced the tactical arsenal of the police almost exclusively to "motorized patrol, rapid response to calls for service, and retrospective investigation of crimes" (Moore *et al.* 1988, 1). In an effort to root out police corruption, which had been endemic in American policing, especially during and after Prohibition, police reformers deliberately sought to isolate the police from all political and community interference. Police managers purposely "assigned line officers to rotating shifts and frequently moved them [between patrol districts]." They tightened supervision to ensure compliance with standardized operating procedures that were meant to limit officer discretion to a minimum and to further an aura of professionalism and impartiality (Bureau of Justice Assistance 1994, 5; Moore & Kelling 1988, 4-6). The traditional function of the police constable throughout the late 19th and early 20th century as neighborhood ombudsman who "maintain[ed] public order, regulate[d] economic activity, and provide[d] emergency services" all but disappeared (Moore & Kelling 1983, 49).

The community policing paradigm emerged haphazardly – "[t]here was no master plan"– drawing on the insights from a series of policing experiments and academic writings that sought to improve on the professional model of policing but in the process undermined virtually every tenet of it (Skogan & Roth 2004, xviii-xix). First, during the 1960s and 1970s crime began to rise, throwing doubts on the crime fighting prowess of the police, until the Kansas City experiments delivered scientific proof that the two primary tactics of the professional model – preventive patrol and rapid response to calls for service – had no effect on crime (Kansas City Police Department 1977; Kelling *et al.* 1974). Then, Goldstein (1979) criticized the incident-driven approach of the professional model, arguing that efforts of the police to improve internal management were vain unless the police adopted a problem-oriented approach; instead of uniformly responding to individual calls for service, the police should try to identify the root causes and chronic hot spots that generated the calls in the first place. Finally, research in the late 1970s and 1980s showed that fear of crime in urban environments was more closely correlated with disorder than with objective risk of victimization (Moore & Kelling 1988, 8; Skogan & Maxfield 1981; Skogan 1986), prompting police departments in several cities to revive a forgotten police tactic, foot patrol. These foot patrol experiments famously inspired Wilson & Kelling (1982) to their seminal article "Broken windows", which argued that in neighborhoods where both people and buildings appear disorderly, local residents become wary about their personal safety and tend to retreat to their fortressed homes, a process which seriously weakens existing mechanisms of informal social control and opens the door to even more disorder, vandalism, and more serious crime. Based on their dynamic model of a spiral of urban decay, Wilson & Kelling (1982) called for a broadening of the police mandate beyond mere crime control, as many disorderly activities are non-criminal and fall outside the realm of law enforcement.

In the early 1990s, community policing advocates began pulling together the lessons learned from these writings and policing experiments, while adding a few new elements (Skogan & Roth 2004, xxiii). The community policing model that thus emerged involves the community as a co-producer of security, broadens the police mission beyond crime fighting to maintaining order, reducing fear of crime, and improving quality of life, and emphasizes a proactive problem-solving approach to control and prevent crime (Fridell 2004, 3f.).[2]

1.2 The Core Elements of Community Policing

It is important to note that community policing is an organizational strategy rather than a fixed program. It has three core elements: community engagement, problem solving, and organizational reform of the police. Community policing is a process rather than a final product that leaves the fleshing out of these principles largely to local communities and the police who serve them. In practice, however, these three core elements are inextricably intertwined and police departments that fail to act on any one of them will not mount a very successful program (Fridell 2004; Skogan 2006c, 27f.).

Since local communities are involved in the process of identifying and setting priorities, the implementation of community policing varies from one place to another. As a matter of fact, how community policing is implemented should vary from place to place, reflecting the unique local circumstances and solutions that the communities and their police adopt to solve local problems (Skogan 2006c, 28). This point has caused much confusion in the literature where critical assertions abound that community policing is only vaguely defined and that its manifestations vary across time and space (e.g. Weisburd & Eck 2004, 52).

1.2.1 Community Engagement

Whereas under the professional model the police were deliberately isolated from the community and committed to a standardized model of law enforcement, under the community policing model police departments regularly consult with neighborhood associations and community groups to let ordinary citizens have a bigger say in the definition of police priorities and the allocation of resources. The police no longer jealously guard their "monopolistic responsibility" for crime control but support community-based crime prevention efforts and interact with private security agencies (Moore & Kelling 1988, 12). Community policing thus revives the famous dictum by Sir Robert Peele, founder of the London Metropolitan Police, about "the historic tradition that the police are the public and the public are the police" (quoted in: Kelling & Coles 1996, vii).

A main driving force behind community engagement has been the recognition that the police cannot solve security and disorder problems alone (Fridell 2004, 4). In order to bring about a collaborative partnership to tackle crime and disorder problems, the police need to engage with the public and provide opportunities for the public to come into contact with them. A second argument for community engagement is that the public's primary concerns are not normally those of the police. People usually fret about "threatening and fear-provoking *conditions* rather than discrete and legally defined *incidents*." The police therefore need to engage with the public when they set their priorities in order to learn about the public's views (Skogan 2006c, 28f.; emphasis in original).

[2]This briefest of accounts can only touch on the milestones of the process of the emergence of community policing. More elaborate reports of the patchy history of community policing are offered by Moore & Kelling (1988), Skogan & Roth (2004), and Trojanowicz *et al.* (2002).

The expected benefits of community engagement are that it not only makes the police more legitimate in the eyes of the public but also helps the police become more effective in combating crime. Community engagement also carries risks, however. For a start, it can be difficult to achieve, especially "in areas that need it the most" or that have a history of fraught police-community relations. It may lead to inequitable outcomes, as some groups will be harder to get involved than others. There is also a risk of opposition from within the ranks of the police against any public interference with their priorities or operating strategies (Skogan 2006c, 30-34).

1.2.2 Problem Solving

As its second core element, community policing incorporates a problem-oriented or problem-solving strategy. Problem-oriented policing discards the "fragmented approach" of traditional policing that treated each crime incident as an "isolated event with neither history nor future." Under a problem-solving approach, the police seek to determine if calls for service or crime incidents are the symptoms of a larger, underlying problem that are likely to recur if the police response does not address the root causes (Moore & Kelling 1988, 10). Problem-solving requires extensive training of police officers in identifying problem situations, analyzing their root causes, and devising strategies to deal with these underlying factors. Nowadays, problem-solving is helped by the ready availability of computer-based analyses of operational data to identify hotspots that generate the bulk of crime incidents and calls for service (Skogan 2006c, 34).

A problem-orientation also implies that the solution to a crime or disorder problem is not necessarily police-based. The police must therefore forge strategic partnerships with other government agencies, which are difficult to pull off. Problem-oriented policing also brings about a renewed focus on crime prevention, which during the reform era used to be entirely deterrence based. Under a problem-solving orientation, by contrast, criminological theories such as routine activities or situational crime prevention guide crime prevention activities. Problem-oriented policing can operate without community involvement, though. Indeed, a key difference between community-oriented policing and problem-oriented policing is that the former engages the community or its representatives in the process of identifying and prioritizing problems, whereas the latter is largely driven by the computerized analysis of police operational data (Skogan 2006c, 34-36). [3]

1.2.3 Organizational Reform

The third core element of community policing is the organizational reform of the police. The most important structural reform of the police organization that is required to implement community policing is decentralization. Decentralization can be achieved at two levels within the police organization: at the middle management level, precinct commanders usually assume more responsibility for tactical decisions to deal with crime and disorder problems in their districts. At the bottom of the ladder, individual patrol officers and their sergeants enjoy greater autonomy in identifying and tackling neighborhood problems (Skogan 2006c, 36f.).

The police also need to reorganize in order to consult with the community. This normally means instituting new channels such as local advisory boards or roundtables to consult with civil society groups (Skogan 2006c, 29, 38). It also means that patrol officers are assigned to the same beat

[3] A note on terminology is in order here. The current study consistently uses the term "community policing" to mean a policing style that involves the community in consultation and in the co-production of security. Community policing is always understood also to include problem-oriented policing. Problem-oriented policing in the sense of the early writings by Goldstein (1979; 1990) is an inseparable part of community policing. In those sections where it is necessary to distinguish community policing from problem-oriented policing, the term community-oriented policing is used to express the idea that the police involve the community in the policing process. This use of the term community policing is also consistent with its use in Swiss policing circles. The police in the German part use the term "community policing", or "CP" in the jargon, to mean both community-oriented and problem-oriented policing. In the French speaking part, the term *"police de proximité"* conveys the same idea.

for longer periods on end and encouraged to have informal encounters with local residents (Moore & Kelling 1988, 12). One aspect of police reform required to implement community policing that often gets short-changed is training. Officers need to be trained in how to engage with members of the community and they also require extensive training in the problem-solving strategy (Skogan 2006c, 35f.).

1.3 Evaluating Community Policing

Evaluation is a central demand of the community policing paradigm. As the previous section explained, the community- and problem-oriented policing styles emerged in the wake of the mother of all policing experiments, the Kansas City Preventive Patrol Experiment (Kelling *et al.* 1974). The upshot of the Kansas studies, namely that the two dominant policing tactics of the reform area had no discernable effect on crime, "spurred an era of experimentation and evaluation" of novel policing strategies that eventually gave birth to the community- and problem-oriented policing paradigms (Cordner 2010, 16).

The early community policing scholars were adamant that performance evaluation under the new paradigm needed to be more comprehensive than the usual, police-internal assessment of law enforcement data such as calls for service, response times, arrests, and clearance rates. While these traditional yardsticks remain important indicators of success, they do not provide sufficient information for a full assessment of police performance. As police departments engage the community in strategic partnerships and shift their focus from fighting crime to solving problems, alternative indicators of success become more important. In the era of community policing, police performance evaluation has to complement traditional law enforcement statistics with measures of fear of crime, disorder, citizen satisfaction with police services, and the overall quality of life in neighborhoods (Kennedy & Moore 1995, 281f.; Moore & Kelling 1988, 13; Trojanowicz & Bucqueroux 1994; Kreis 1999).

In Switzerland, however, efforts to evaluate community policing only go so far. The need to monitor performance and to evaluate policing strategies using non-traditional indicators has eventually been recognized. All five police departments being studied here have contracted local polling institutes to conduct regular public surveys to assess local levels of victimization, fear of crime, disorder, and public satisfaction with the police. By contrast, the evaluation of crime prevention activities relies heavily on the feedback from the strategic partners of the police, and such efforts are rarely monitored systematically. Since 2004 the Geneva cantonal police have conducted a regional security assessment with the help of external consultants at regular intervals of three to four years (Wisler *et al.* 2004; Wisler *et al.* 2008b; Wisler *et al.* 2011). The latest 2011 edition made an effort at evaluating the impact of a hotspots policing campaign carried out in the summer of 2010 in the inner city of Geneva. In 2007/08, the Zurich Municipal Police Department commissioned the Institute of Criminology of Zurich University to evaluate a burglary prevention project (Manzoni & Thalmann 2008), and the same institute has also conducted an in-depth study of the Zurich municipal administration's inter-departmental neighborhood regeneration project "Langstrasse PLUS" (Schwarzenegger *et al.* 2009), but scientific evaluations of prevention strategies remain rare.

In Switzerland, community policing has also generated little academic interest in its own right. The current study is one of the first attempts at a comparative assessment and evaluation of the new policing paradigm. In this context, it is noteworthy that the two preceding efforts to take stock of community policing practices in Switzerland were both carried out on a contractual basis. In one case, the police departments of Basel, Bern, and Zurich tasked a group of researchers from the University of St. Gallen with a review of community-based crime prevention in general and of their implementation of community policing in particular (Sauter *et al.* 2005). In the other case, the Paris-based *Institut National des Hautes Etudes de la Sécurité* contracted two researchers from the University of Neuchâtel to draft a compendium on the state of community policing in Switzerland as part of a cross-country comparison of different community policing styles (Bolle & Knoepfler 2000a).

1.4 Research Objectives and Challenges

The current research project is both a process and impact evaluation of community policing in Switzerland's five major urban areas with a population of more than 100,000 inhabitants. It aims to establish a detailed overview of the implementation of community policing by the police departments in Basel, Bern, Geneva, Lausanne, and Zurich and to assess the impact on the well-being of the residents of the country's five largest urban areas. In a first phase, the study has tracked the organizational and strategic reforms the five police departments have undertaken to implement community policing. This first part of the study aims to answer the following research question: *When and in which parts of their city did police departments introduce which elements of the community policing strategy?* The objective has been to establish a record of community policing implementation for each of the five police departments for the period from the early 1990s to 2010.

In a second phase, the current study aims to assess the impact of the community policing efforts the five police departments have undertaken in recent years. The research question for this second part of the project is thus: *How effective have community policing efforts been in reducing levels of fear of crime and neighborhood disorder as well as in improving citizen satisfaction with police services?* The objective is to analyze the data from surveys of the residents of the five major Swiss urban areas in order to assess whether community policing has made the police forces more effective and equitable in reducing fear of crime, tackling neighborhood disorder, and improving police-community relations. Answering this research question will require a separate analysis of the indicators of three theoretical constructs of community policing impact: fear of crime, neighborhood disorder, and public attitudes towards the police.

The intelligence on police organizational and strategic reforms has been obtained directly from the police. Semi-structured interviews were conducted with key police administrators from each of the five police departments to this end. This information was complemented with additional intelligence gleaned from internal documents provided by the five police departments as well as additional public sources. The data on fear of crime, neighborhood disorder, and popular satisfaction with the police were obtained from standing survey instruments, namely the Swiss Crime Survey (SCS; Killias *et al.* 2007). The SCS has polled residents of the five major urban areas repeatedly since the mid-1980s and tracked the three theoretical constructs of community policing impact.

The choice of an appropriate research design for the present impact evaluation of community policing faces two major challenges. First of all, the present evaluation is carried out *ex post facto*. When work for it began, community policing efforts had already been underway for several years in all five urban areas. As the process evaluation has unearthed, none of the five police departments has taken provisions to prepare for a scientific evaluation, neither of the transition to community policing nor of its impact. This means that experimental designs are obviously not possible and that baseline data for any pretest-posttest comparisons will have to come from existing data sources. Luckily, the Swiss Crime Survey polled a sufficiently large number of residents of all but one of the five cities prior to or during the early implementation of community policing, providing a rich source of baseline data. After a time lag of five years, the SCS polled another random sample of the residents of the five cities in this study, allowing valid pretest-posttest comparisons at the infra-city level.

Secondly, the process evaluation has also revealed that the five police departments introduced community policing more or less rapidly across their entire jurisdiction, leaving no area within the city proper completely unaffected. After an initial, usually rather short trial run at one or two police stations, the police command in each city swiftly thereafter decided to introduce community policing to the entire department. Such a department-wide transition was in line with the recommendations of the literature (e.g. Trojanowicz & Bucqueroux 1994, 2), but it complicates the task of an evaluator. Without appropriate control areas within a given city, even the basic quasi-experimental research design is not a viable option.

Obligated to use an observational design since the aspect of evaluation had been largely overlooked, the current study tried to make a virtue of necessity. The study has been driven by the goal to enhance the validity of observational studies of complex interventions such as community policing that are challenging to evaluate through controlled experimentation even under the best of circumstances (cf. Kennedy & Moore 1995). In a nutshell, the basic approach of this observational study is to compare the implementation of community policing across the five cities and to analyze the impact of different community policing strategies thus identified across matched comparison areas across the five urban areas. Short of the randomization of treatment, it is generally considered that “the identification of well-matched comparison groups provides a high level of internal validity for a study.” (Weisburd *et al.* 2010, 144, attributing the point to Farrington *et al.* 2002).

In his seminal paper, Cochran (1965) famously advised the authors of observational studies always to ask: “How would the study be conducted if it were possible to do it by controlled experimentation?” (quoted in: Rosenbaum 2010, 4). The current study was guided by the notion of the clear superiority of the randomized trial and attempted to follow Cochran’s advice to emulate the controlled experiment design as much as possible with an observational study.

In the present context, the research design of a randomization controlled experiment would look as follows: first match the study areas pairwise on contextual characteristics, then flip a coin to determine for each matched pair which of the two study areas receives the treatment, interview a sufficient number of residents in both the treatment and control area before and after implementation to obtain panel data, and finally conduct a pairwise statistical test to assess whether treatment was effective. In a randomized study, the areas are normally matched pairwise on a limited number of covariates before the treatment assignment is decided by a random process. Randomization then ensures that both observed and *un*observed covariates are sufficiently balanced across treatment and control group, and the only systematic variation between the two is the treatment intervention (Rosenbaum 2010, Chapter 2).

As controlled experimentation is not a viable option for the current study for the reasons stated above, a programmatic outline for an observational study that aims to heed Cochran’s (1965) advice looks as follows: first, conduct a process evaluation in order to identify the areas that exhibit the biggest gaps with regard to program implementation and chose a research design with a “treatment” and a “control” area so as to compare neighborhoods that received sharply different “doses” of treatment. This is being done since a statistical test is more likely to pick up an effect – especially a small to moderate effect – if the comparison is between areas that received comparatively a lot of or little treatment (Rosenbaum 2010, 125). Second, match the neighborhoods of the entire study area based on contextual variables to achieve balance of the observed neighborhood-level covariates. However, in an observational study – contrary to a randomized experiment – there is absolutely no basis to assume that the matching of study units based on observed covariates also balances the *un*observed covariates (Rosenbaum 2010, 73-76). Therefore this matching procedure should strive to match neighborhood areas on a maximum of covariates, ideally on all the theoretically relevant dimensions. In a third step, within each neighborhood area, match the individual survey respondents from the pretest and posttest samples to achieve balance on the observed covariates at the individual level such as a respondent’s age, gender, and education level, which are known to impact their responses to the outcome variables. In a research design that balances the pretest and posttest samples on the observed covariates at the neighborhood and individual levels, these will be distributed approximately equally between the two groups and thus be independent of treatment. Then, conduct a test to assess the statistical significance of the difference in the outcome variables between the pretest and posttest samples. Finally, if this test is significant, conduct a sensitivity analysis in order to determine the degree to which inferences about program impact may be undermined by bias introduced by some unmeasured covariates. In an observational design this can never be ruled out since matching can only ever balance the observed covariates. If the response to treatment, i.e. the observed difference between the pretest and posttest samples, is robust in the face of potential bias due to some unmeasured covariate, compare this result with the same pretest-posttest difference of similar neighborhood areas in other cities. If the areas that show significant improvements on the indicators of treatment outcome compared to similar areas, are also the areas that have received comparatively higher "doses" of community policing implementation or "treatment", this will count as further evidence that the observed differences are due to shifts in policing strategy and not to some other extraneous factors.

In implementing an observational research design that emulates the template of a randomized experiment, the current study took several steps at different stages in an effort to enhance internal validity. First of all, the process evaluation produced a "Calendar of Action" (Crawley & Hope 2003) to create panel data to measure community policing implementation progress over six evaluative dimensions over time. This was done in order to obtain a clear idea of the elements of the treatment and the time when it was begun, "so there is a clear distinction between covariates measured prior to treatment, and outcomes measured after treatment." (Rosenbaum 2010, 5). In a second step, all the indicators of the four theoretical constructs of community policing impact – fear of crime, disorder, and public attitudes towards the police plus official crime rates – were subjected to an exploratory analysis in order to detect the spatio-temporal patterns in the data. This was done primarily to see if there are any macro-level trends that manifest themselves across the five cities, which would point to possible sources of bias if any such trend were not accounted for in the final analysis. If there are such macro trends in the data, it is very likely that these trends are affected by macro-level forces rather than by changes in micro-level policing strategies.

Thirdly, in order to match similar neighborhood areas across the five urban areas, the current study uses geospatial data mining algorithms to develop a typology of urban neighborhoods. Such a neighborhood classification system serves primarily to enhance the internal validity of the observational research design. The basic idea is to group neighborhoods of a similar type in order to minimize the within-cluster variance in the contextual variables that may affect the outcome variables and thus risk confounding inferences about program effectiveness in a non-experimental impact evaluation. One objective of the clustering algorithm is thus to minimize the between-cluster similarity of the contextual variables that may rival treatment as the plausible explanation for the observed variance in the outcome indicators. In other words, neighborhoods that are similar in terms of their demographic and socio-economic structure as well as the built environment must be grouped into clusters of similar type. Simultaneously, the resulting neighborhood typology should account for a maximum of the between-cluster variance in the outcome measures prior to the onset of the treatment, i.e. the survey response patterns should be similar for residents of the same neighborhood types across urban areas. In a research design based on a neighborhood typology that satisfies both these objectives, community policing will be evaluated across urban districts that not only resemble each other with regard to the ecological context, but where local residents also collectively reported similar levels of fear of crime, disorder, and satisfaction with the police at the onset of the treatment. This set up allows an evaluator to dismiss a series of threats to the internal validity that otherwise are present in an observational design.

Finally, even if the contextual variables or covariates and the outcome measures can be balanced perfectly between neighborhoods of the same type across urban areas (a big if), the current research design will still be liable to threats to the internal validity. In particular, the current study assesses the impact of treatment based on repeated random samples of the study population. However, any difference in response pattern between the two is only valid as an indicator of a treatment effect to the extent that the pretest and posttest samples are comparable (Cook & Campbell 1979, 117). Therefore, within each neighborhood type identified, propensity score matching methods are used to match pretest and posttest samples on individual-level covariates such a respondent's age, gender, and education level before conducting the statistical test of the observed difference. Moreover, since in an observational study it can never be ruled out that pretest and posttest samples differ due to some unobserved covariate, a sensitivity analysis is carried out whenever the statistical test is significant in order to assess how robust this finding is in the face of potential bias due to some unobserved covariates.

1.5 Outline of the Book

The remainder of this book is organized as follows: Chapter 2 places the current research in the wider context, first by providing a quick overview over the recent history of policing evaluation research and by looking at some of the pitfalls that make evaluating community policing a particularly challenging endeavor. The review then discusses some more practical proposals how to navigate around these obstacles for both the process and impact evaluation, particularly recent developments in the area of geocomputational profiling of urban neighborhoods and their application for police performance assessment. The remainder of this section then reviews the criminological research on the implementation and the impact of community policing in an effort to assess the available evidence, starting out with a brief discussion of the two most comprehensive such efforts, the evaluation of the Chicago Alternative Policing Strategy and the National Reassurance Policing Programme in England. This section ends with a list of the working hypotheses guiding the current research.

Chapter 3 outlines the *methodology* of the present study. It starts out with a presentation of the process evaluation, discussing the procedures followed to take stock of community policing implementation. It describes the interviewing and data gathering methods as well as the six analytical or evaluative dimensions that guided the process evaluation. It also explains how the qualitative information was coded for the subsequent quantitative analysis. The part dedicated to the impact evaluation first details the operationalization of the four theoretical constructs of community policing impact – crime, fear of crime, neighborhood disorder, and public attitudes towards the police – as well as the data used to describe the neighborhood ecological context. It then discusses in turn the different methodologies used for the exploratory spatio-temporal data analysis, the unsupervised and supervised data mining algorithm employed to develop the neighborhood classification system, as well as the propensity score matching, statistical tests, and sensitivity analysis used for the impact evaluation. This section explains these methodologies in sufficient detail, including the necessary mathematical formulas and key lines of code in the programming language R, and was written assuming the reader has a basic understanding of data mining algorithms and methodology. In the following chapters, which discuss the results from these analyses, the introductory text briefly explains these methods by drawing analogies to analytical tools of classic inferential statistics that ought to be familiar to any empirical social scientist. The section concludes with a discussion of the issues of construct, internal, statistical conclusion, and external validity pertaining to this study and defines the criteria that serve as benchmark to measure the success of community policing implementation.

Chapter 4 presents the results of the *process evaluation*, providing an overview over the community policing models in the five cities included in this comparative study. The outline of this part of the book corresponds to the six evaluative dimensions of the process evaluation, highlighting the most salient features of the implementation process in each city. This section summarizes the more detailed reports on the community policing implementation process in each of the five cities, which had to be drafted either in German or in French during the early stages of this research project. It ends with the results of the quantitative analysis of the coded implementation data, which serve to reveal noticeable parallels in the five implementation chronologies and to spot the evaluative dimensions with the biggest gaps in community policing implementation across the five urban areas.

Chapter 5 presents the empirical findings of the *exploratory spatio-temporal data analysis* of the indicators of the four theoretical constructs of community policing impact, which was carried out to spot any macro-level trends in the outcome data that could potentially confound inferences about program impact in a non-experimental research design. This part of the book analyzes both the long-term trends and the spatio-temporal patterns of the outcome variables across the five urban areas.

Chapter 6 describes the procedure followed to develop a *neighborhood typology* in order to match neighborhood contexts on observed covariates across the five urban areas for the subsequent impact analysis. This section presents the unsupervised and supervised data mining algorithms – self-organizing maps and random forests – used for the dimensionality reduction and clustering of the high-dimensional data. It explains in detail the iterative procedure followed to select the key variables in order to build a parsimonious model that makes a minimum of classification errors. It then discusses the diagnostic plots and tools used to assess the appropriateness of the current neighborhood typology in light of the optimization criteria set by the exigencies of the subsequent impact analysis. The section closes with a description of the most salient characteristics of each of the six neighborhood types determined as most appropriate by the clustering procedure.

Chapter 7 presents the results of the analyses of the *impact of community policing on neighborhood residents*. This section first briefly introduces the propensity score matching methods, statistical tests, and sensitivity analysis used to assess the impact of community policing. It then presents the results of these analyses for the indicators of the three theoretical constructs of community policing impact measured using survey items – fear of crime, neighborhood disorder, and public attitudes towards the police.

Chapter 8 first discusses both merits and shortcomings of the different methodologies used in the analytical sections of this evaluation, before revisiting and reassessing the research hypotheses formulated at the end of the literature review in light of the evidence produced by this study. It ends with a discussion of the limitations of the current research, notably due to the fact that this evaluation was envisaged only long after community policing implementation had begun in all of the five urban areas under study and thus had to rely on extant data sources that were not originally conceived for the current analyses.

Finally, Chapter 9 summarizes the results of the process evaluation and the individual elements of the impact evaluation and discusses the implications flowing from the study for both policy and research. Most importantly, this section gathers all the evidence produced by this study and tries to come up with a concise answer to the all important question of any evaluation in this era of "evidence-based" policy: Does it work?

2

Community Policing Evaluation

2.1 Theoretical Approaches to Community Policing Evaluation

2.1.1 A Short Review of Community Policing Evaluation

Evaluations of community policing are as old as the new policing paradigm itself. Indeed, as the previous chapter argued, the roots of community policing lay in the evaluations of the reform era strategies and policing experiments of the 1970s and early 1980s (cf. Chapter 1.1). This convergence of academic research and policing practice is perhaps less surprising, considering that a number of scholars who were influential in shaping the nascent community policing movement had been directly involved in these scientific evaluations (e.g. Kelling *et al.* 1974; Trojanowicz 1983).

Many of the early experiments in community policing and crime prevention were accompanied by evaluation efforts to assess the impact of these alternative crime control strategies. As the number of these evaluations was piling up during the 1980s and early 1990s, a number of narrative reviews and edited books compiled the accumulating evidence (Rosenbaum 1986, 1988, 1994; Skogan 1994). Since these early days, though, there were also some critical voices who bemoaned the poor methodological quality of many anti-crime project evaluations (Greene & Mastrofski 1988; Lurigio & Rosenbaum 1986).

In the mid-1990s, a group of researchers around Lawrence Sherman of the University of Maryland conducted a systematic review of the scientific evidence on the effectiveness of crime prevention strategies on behalf of the United States Congress (Sherman *et al.* 1997). Five years after the first review, the group of criminologists published an updated version of the report, *Evidence-Based Crime Prevention* (Sherman *et al.* 2002a), which retained the original chapter structure that arranged the available evidence on crime prevention effectiveness by seven different institutional settings: families, schools, communities, labor markets, places, policing, and courts and correctional institutions. The sheer size and breadth of these reports (in all over 500 studies were reviewed) as well as their programmatic titles meant that the reviews had a lasting impact; the quest for "what works" in crime prevention became the leitmotif of evaluation research in criminal justice policy.

The Sherman report brought about a relentless focus on the quality of anti-crime program evaluations. For their systematic reviews, Sherman and colleagues had developed the Maryland Scientific Methods Scale (SMS), which grades the methodological rigor of different evaluation studies on a scale from 1 to 5, depending mainly on the internal validity of its research design (Farrington *et al.* 2002). The randomized experiment was defined as the highest methodological standard of evaluation research ("Level 5") because it best allowed discarding the many threats to the internal validity of a program evaluation and thus made possible valid inferences about a causal effect. Levels 1 to 4 distinguished between research designs without randomization of treatment, depending on the severity of the threats to internal validity a given research design failed to control. The SMS served both as a yardstick to assess the methodological quality of evaluation studies, as well as a tool to facilitate the communication of the results of a systematic review to policy makers, politicians, and law enforcement practitioners. The hierarchy of the scale of methodological quality was essentially nothing novel. The case for randomized trial research had been made decades earlier for program evaluation in general (Campbell & Stanley 1963; Boruch 1975) and criminal justice research in particular (Maltz 1972). Nevertheless, in criminology the SMS became the authoritative yardstick to measure the quality of an evaluation design.

The second comprehensive review on the status of the evidence on police research was a report by the National Research Council's Committee to Review Research on Police Policy and Practices, which under the chairmanship of Wesley Skogan and Kathleen Frydl brought together a number of eminent policing scholars (Skogan & Frydl 2004; Lum *et al.* 2011). The report's section on the "effectiveness of policing in reducing crime, disorder, and fear" was later published as a separate article. In it, Weisburd & Eck (2004, 60) concluded that the evidence suggested that community policing did reduce fear of crime but unless it is paired with problem-oriented policing, it had no effect on crime and disorder (cf. Chapter 2.2). However, the authors noted that community policing comes in so many shapes and guises that evaluators studying it in different places or at different times were not necessarily looking at the same thing. Moreover, Weisburd & Eck (2004, 52) argued that "[b]ecause community policing involves so many different tactics, its effect as a general strategy cannot be evaluated." The authors bemoaned a gap between the importance that community policing had taken on in American policing and the dearth of methodologically rigorous studies, decrying their inability to locate a single randomization controlled evaluation (p. 59).

A third set of systematic reviews of the effectiveness of crime prevention and criminal justice policies has been sponsored by the Norwegian-based Campbell Collaboration Crime and Justice Group, which was founded in 2000 (Farrington & Petrosino 2001). The Campbell Collaboration supports authors who conduct systematic reviews and meta-analyses, focusing exclusively on high-quality experimental and quasi-experimental studies. Since its foundation, the Cambpell Collaboration Crime and Justice Group has sponsored, among other studies, systematic reviews of the impact on crime of hot spots policing (Braga 2007), problem-oriented policing (Weisburd *et al.* 2008), neighborhood watch (Bennett *et al.* 2008), and improved street lighting (Welsh & Farrington 2008).

The most recent systematic review of the scientific evidence on policing strategies was produced by Lum *et al.* (2011) who analyzed almost 100 studies of moderate to very rigorous quality. (To be included in the systematic review, a study had to score as Level 3 or higher on the SMS, i.e. a quasi-experimental design with at least one matched comparison group). Lum *et al.* (2011) observed a fairly systematic bias in the focus of these evaluations on place-based, focused, and proactive police interventions, especially those employing more rigorous research designs. The authors noted that this "bias within the more rigorous evaluation literature in policing is not coincidental, nor does it reflect the reality of police practice, which we know is remarkably individual-based, reactive, and general in nature. Rather, these overall tendencies in the research reflect the innovations of scholars and police practitioners who have tried to push the field forward through these evaluations." (p. 17f.). In other words, scholars tend to evaluate policing strategies that they think police should adopt and are suited for powerful research designs, not what police are actually doing. Furthermore, the authors also found corroborating evidence of Rossi's (1987) "iron law of evaluation", which states that "the better designed impact evaluations are, the less effective the interventions or programs seem to be." (quoted in: van der Knaap *et al.* 2008, 48f.). Lum *et al.* (2011, 18) concluded that this "provides specific and updated support from the policing literature for Weisburd *et al.*'s (2001) finding that, as studies increase in methodological rigor, they are less likely to find positive results."

2.1.2 The Challenge of Community Policing Evaluation

This seeming contradiction of a broad-based shift in criminology towards methodologically more rigorous evaluations – Levels 3 to 5 on the SMS – and the small number of sound evaluations of community policing as a community-wide, comprehensive strategy calls for an explanation. In a programmatic article, Rosenbaum (2002, 192) noted a number of obstacles that make multi-agency anti-crime partnerships much more complex to evaluate than other programs and offered advice on how not to succumb to them. Since such partnerships are an essential part of community policing, these obstacles apply quasi without amendment to its evaluation. Rosenbaum (2002, 192) identified the following obstacles in the literature:

- The complexity – both vertical and horizontal – of the interventions. Comprehensive anti-crime initiatives are multi-stakeholder initiatives that involve different organizations and intervene at different levels (individuals, families, schools, neighborhoods, and communities) (ibid.). This complicates the task of an evaluator to track who does what to implement which part of a program enormously.

- The complexity of contextual variables. Anti-crime partnerships do not exist in a vacuum but are embedded in a specific historic, political, economic, demographic, and geographic context (ibid.).

- The dynamic, changing nature of the intervention. Anti-crime partnerships are dynamic in nature and evolve over time, "making it difficult for evaluators to 'hit a moving target' " (ibid.). Reports that community policing implementation is rarely faithful to the original planning and often gets delayed or not implemented in full are legend.

- The diversity of intervention processes and outcomes. Anti-crime partnerships, almost by definition, are unique and complex and their choice of inputs, procedures, and outcomes will vary according to context (ibid.). In fact, this complexity is even encouraged by the literature that stresses that community policing is an organizational strategy rather than a fixed program that leaves implementation largely to local stakeholders and circumstances (cf. Skogan 2006c).

- The lack of optimal conditions for traditional experimental research. Comprehensive, community-wide anti-crime partnerships seriously restrict the possibility to randomize the treatment or to find equivalent comparison groups (Rosenbaum 2002, 192). As if that was not enough, community policing is meant to be a department-wide strategy; experience shows that after an initial, geographically-confined trial run, community policing is being scaled up to a police departments' entire jurisdiction, precluding experimental research designs beyond the micro and meso levels for want of suitable comparison groups/areas (cf. Chapter 2.2.1).

One further obstacle to evaluation that is peculiar to community policing is the slow pace of police organizational reform. Community policing takes a long time to implement, making it inherently unsuited for the experimental paradigm. All these factors conspire to make the evaluation of multi-agency anti-crime partnerships (like community policing) a daunting task (ibid.). The NRC's panel to review research on policing even concluded that "community policing was simply too amorphous a concept to submit to empirical evaluation and recommended that researchers evaluate it by breaking it down into more specific components." (Mastrofski 2006, 44f.; Skogan & Frydl 2004).

2.1.3 Process Evaluation

In *Evidence-Based Crime Prevention*, Sherman & Eck (2002, 301) observed that many evaluations of policing strategies did not measure police activities at all or did so by using unreliable methods. Indeed, that the threat of low construct validity is acute for community policing had long been recognized in the literature (e.g. Kennedy & Moore 1995, 285). However, if the process of implementation went un(der)reported in many policing evaluations, it was not necessarily for lack of good will on the part of the evaluators. McElroy (1998) already in the early 1980s discovered that the challenge of documenting police operations was "a good deal more difficult than it sounds." (p. 81).

Greene (1998, 151) argued that in order to establish effective systems to monitor community policing implementation, researchers needed to specify the level of intervention and the expected effects of the organizational reforms. Greene recommended evaluators distinguish four different levels of analysis within a police department undergoing a transition to community policing (Greene 2000, 321-323; Greene 1998, 152-155):

- At the *community* or *environmental* level, community policing seeks to engage the police and the community in a public-safety co-production partnership. The police must gradually open up to outside organizations and seek to establish strategic partnerships with other government agencies and civil society groups. Community engagement means the police must focus on community capacity building but also give their partners some say over police tactical and strategic decision making (Greene 2000, 321; Greene 1998, 152).

- At the *organizational* level, community policing affects police department's organizational structure, division of labor, and use of technology. The reform era model – centralized, hierarchical, and bureaucratic – will hinder the transition to community policing; departments must adopt flatter hierarchies, become more decentralized and less specialized, and create internal support structures for community policing initiatives. A police department's commitment to community policing must also become manifest in its management of human resources (hiring, training, rewarding, and promoting) and its performance appraisal systems as well as its culture and value systems (Greene 2000, 322).

- At the *precinct* or *group* level, community policing calls for the fixed geographical assignment of patrol officers. Individual working groups must set and communicate performance targets and improve interpersonal communication. Better information sharing is needed across a department's divisions, especially for problem solving (Greene 2000, 322). A crucial aspect is also the training of precinct commanders and mid-level managers to act as community policing facilitators, as the persuasive powers of the chief of police may be insufficient to reform entire departments (Mastrofski 2006, 52).

- At the *individual* level, community policing broadens the job definition of line officers, giving them greater autonomy to engage in problem solving. This implies training police officers for a greater range of tasks but also giving them more feedback regarding their community- and problem-oriented activities. This shift should make police officers not only more effective, but also increase their job identification and satisfaction (Greene 2000, 323).

Besides the organizational reform of the police, the process evaluation must also document the changes to the *operational strategies* to implement community policing. On this score, a key challenge to the process evaluation of anti-crime partnership projects is that things rarely go according to plan (Hope *et al.* 2004, 6). It is thus necessary to track actual implementation activities rather than just the original plans. Measuring the intensity and duration of treatment is a prerequisite for any causal inferences about a treatment effect (Rosenbaum 2002, 207).

In the context of the evaluation of anti-burglary strategic development projects in England, Crawley & Hope (2003, 11) were frustrated by the standardized forms generated by the Home Office's Policing and Crime Reduction Unit for project accountability and cost effectiveness, which they found inadequate for the task. They realized they needed to come up with a new instrument capable of capturing both the regular shifts in focus and "the 'non-sequential' nature of implementation" of these projects. Crawley & Hope (2003) thus developed a new tool to measure implementation activity – a "Calendar of Action" (CoA) – which allowed them to measure the "intensity of action" or dosage across space and time. The CoA was conceived as an instrument that provides the basis to construct a continuous measure of the level of activity and simultaneously helps to unearth difficulties encountered during implementation.

In practice, the CoA is simply a data matrix that records program activity over the lifetime of an anti-crime development project. The matrix tracks progress (or the lack of it) of individual project interventions in periodic intervals. It contains both quantitative and qualitative data, recording hard numbers on actual "achieved outputs" as well as contextual information on why certain project interventions got stuck or shifted in focus during implementation (Crawley & Hope 2003, 11).

2.1.4 Impact Evaluation

Since the very beginning of the community policing movement, its advocates have stressed that the new policing paradigm cannot be assessed on the basis of a police-internal appraisal of traditional law-enforcement data alone (Moore & Kelling 1988; Trojanowicz & Bucqueroux 1994). Even though the police keep a strong focus on law enforcement activities and calls for service, arrests, and response times thus remain valid indicators of success, these numbers tell only part of the story. As police officers seek to engage the community in problem-solving partnerships, and police departments focus on outcomes rather than on procedures, alternative indicators are required to evaluate their performance (Kennedy & Moore 1995; Moore & Kelling 1988; Trojanowicz & Bucqueroux 1994; Kreis 1999).

How well the police perform in their emergency function and in interventions calling for the use of force is well covered by traditional performance measures such as response times as well as arrest and clearance rates. These indicators, however, have been subjected to various criticisms: firstly, it has been contended that crime reporting rates may be affected by community policing efforts. Secondly, it has been argued that these traditional measures are inadequate in view of the fact that police officers spend the bulk of their duty time in interventions not involving force. If performance benchmarks "are to reflect what police officers actually do", they must take into account order maintenance and crime prevention activities (Fielding & Innes 2006, 128f., 138).

Moreover, as Kennedy & Moore (1995, 281f.) argued, "[c]ommunity policing has expanded the notion of effective policing to include, at the very least, combating disorder and fear of crime as key elements of the police function." In order to assess their performance on these scores, police departments have more or less uniformly resorted to public surveys to gauge local levels of fear of crime, disorder, community (dis-)organization, and public satisfaction with the police. These survey instruments typically rely on a structured questionnaire and a random sample of the resident population in order to obtain a representative picture of the prevailing attitudes (Fielding & Innes 2006, 133).

More recently, this "methodological orthodoxy" of the public survey based on a random sample of the resident population has been criticized as problematic in gauging the effectiveness of alternative policing strategies (ibid.). In the context of Britain's reassurance policing strategy, Fielding & Innes (2006, 130f.) remarked that community policing interventions often occur on a small-scale local level and that a randomized survey that uses a spatial unit of analysis that is significantly larger than the target area is unlikely to reach a sufficient number of respondents to show any effect. A second source of criticism, raised by geographers, is that much contemporary policing and crime reduction research has been too focused on the question of "what works" (e.g. Sherman *et al.* 1997), while failing to take the local, geographical context into account (Ashby 2005, 415; Williamson *et al.* 2006, 199f.). The one area that stands out from this trend in policing research has been crime mapping. Police crime analysts use Geographical Information Systems (GIS) and crime mapping techniques predominantly as a tool to identify hotspots and guide tactical reactive policing strategies. GIS applications, by contrast, are seldom employed for planning purposes or strategic performance review, which is in marked contrast to their widespread use by other branches of local government (Ashby 2005, 415).

The alternative methods of performance evaluation these critics propose vary according to their background but commonly consist of either a combination of qualitative and quantitative research methods or complementing public survey data with "surrogate" measures of the impact of policing services. Rather than polling a random sample of the resident population, evaluators could rely on judgments of "key informants" who are well informed about the area in question and are in a position to assess the quality of police performance (Fielding & Innes 2006, 135). Alternatively, the concept of surrogate measures seeks to capture a series of socio-demographic or socio-economic indicators to gauge the local impact of police interventions. A medium- to long-term impact of successful community policing may well be an increase in local house prices, reduced residential instability, or an upsurge in the number of business start-ups. Since not all of these observed changes may be the result of shifts in local policing, indicators are required that capture the variety of shifts in the social context and may be linked to indicators of police activities. Such "soft" measures that are sensitive to community policing activities could then be used as indicators to measure different aspects of police performance (Fielding & Innes 2006, 137).

One approach that takes into account environmental and socio-economic measures is the geodemographic profiling of neighborhoods. Geodemographic profiling techniques and applications were first developed in the late 1970s for deprivation studies and found widespread use during the 1980s and 1990s primarily in the retailing and financial sectors. The application of geodemographic profiling in policing and criminal justice research is based on the premise that distinctive neighborhood types can be identified successfully and that these neighborhood types "differ predictably in their crime profile and policing environment." It is grounded on the theory that neighborhoods differ both in terms of the level and the quality of social disadvantage, that they follow different trajectories, and thus "are often suited to quite different types of priority area programme." (Ashby 2005, 422f.).

Ashby (2005, 427-32) proposed three kinds of geodemographic analyses that may inform policing strategy and hence be used to evaluate and improve performance: area profiling, operational data profiling, and crime survey profiling. Firstly, the basic profiling of areas such as patrol beats or precincts to distinguish between different types of neighborhoods may inform police strategic decision making and performance evaluation. Mapping the distribution of the neighborhood types thus identified using GIS software is a straightforward procedure and adds a spatial dimension to this source of intelligence. Secondly, the profiling of crime incidents and other operational data allows police analysts to calculate the probability or expected frequency of specific crime incidents for different neighborhood types based on prior observation. Such maps can then be cross-checked with the results of hot spot analyses in order to detect areas with unexpected levels of victimization or, conversely, identify areas where policing interventions have proven particularly successful. Finally, the geodemographic profiling of survey data may help discern likely variations in attitudes to disorder, fear of crime, and the police across different neighborhood types. If the ZIP code of each respondent to a crime survey is known, these survey data can be aggregated by neighborhood type to calculate national or regional averages, which can then be extrapolated for analysis at the local level.

2.1.5 Neighborhood Profiling and Policing Evaluation

Classifying neighborhoods for reassurance policing

The rationale for the use of neighborhood profiling in criminal justice research rests on the observation that, even though poverty and crime are correlated, not all neighborhoods with similar levels of deprivation are equally crime-ridden. Geodemographic profiling in this sense is reminiscent of the social area analysis and social disorganization theories provided by the Chicago School during the first half of the 20th century (Shaw & McKay 1942/1972; Williamson *et al.* 2006, 191). This type of research has experienced a revival over the last two decades, especially regarding the crime-prevention benefits of "social capital" and "collective efficacy" (e.g. Sampson & Raudenbush

1999; Sampson *et al.* 1997). Geodemographic analyses build on these insights by incorporating in the data analysis not only neighborhood ecological characteristics such as socio-economic status or relative poverty levels but data on lifestyles and attitudes as well, and thus offer a more pertinent description of an area than deprivation indices alone (Williamson *et al.* 2006, 194).

The importance of attitudes to crime has been highlighted by the analysis of British Crime Survey (BCS) data in the early 2000s, which revealed that although actual levels of victimization recorded by the survey had fallen by almost 40 percent from their peak in 1995, two-thirds of respondents were under the impression that crime rates had risen over the previous two years. This apparent disparity between falling levels of actual victimization and the widespread notion of a rise in crime has been dubbed the "reassurance gap" and spurred much police activity under the banner of "reassurance policing", which aims both to rectify the public's perception and ultimately to provide safer neighborhoods (cf. Chapter 2.2.1; Williamson *et al.* 2006, 192). However, the apparent drop in overall crime rates across Britain does by no means imply that all areas have benefited in equal measure, and it is quite possible that falling aggregate-level crime figures actually conceal a "growing polarization between high- and low-crime areas at the local level" (ibid.).

Linking British census and BCS data, Williamson *et al.* (2006) were able to estimate local levels of social capital for different neighborhood types and to compare these values to their experience of crime. They thus showed that neighborhood types that enjoyed higher levels of social capital suffered comparatively lower levels of victimization. Unsurprisingly, the authors also found that among the areas deemed high risk, the areas with higher levels of deprivation were more prevalent, but maintained that a multivariate geodemographic typology nevertheless offered extra insights over less refined deprivation indices (199f.). Indeed, the authors found that not only did people's perceptions of crime differ according to the composition of their neighborhoods, but so did actual victimization rates, crime reporting, fear of crime, and attitudes towards the police (p. 201, 206). The authors thus concluded that geodemographic profiling offers a useful tool for reassurance policing, which, they argued, is more likely to have the effect of reassuring the public, if the strategies are developed with the particular needs of different neighborhood types in mind (p. 213f.).

Neighborhood profiling using geospatial data mining

The geodemographic profiling of neighborhoods or larger spatial administrative units for policing purposes has been developed and applied primarily in the United Kingdom in an effort to increase police accountability and improve performance. In the mid-1990s, Her Majesty's Inspectorate of Constabulary developed the first such spatial typology when it created the "most similar force" group (MSG) to assign all 43 separate police forces across England and Wales into groupings of similar type. Methodologically, the MSG typology was based on a dissimilarity matrix of five exogenous variables that described the demographic, socio-economic, and geographic context in which the 43 police forces operated. The rationale behind the development of the MSG grouping was that individual police forces faced very different "policing environments" and that for meaningful cross-sectional performance evaluations, similar forces needed to be compared. Since 1997, the Home Office has thus assessed each police force by comparing its performance on five different categories of outcome indicators to the mean values of its MSG peers (Ashby & Longley 2005, 56f.).

In a similar vein, the British Home Office in the early 2000s developed and published groupings of similar areas at a smaller spatial scale than police force areas. First, the Home Office made an effort to list and classify all of the 376 Crime and Disorder Reduction Partnership (CDRP) across England and Wales and shortly thereafter applied the same clustering methodology to group the 318 Basic Command Units (BCU) into "families" of similar type (Sheldon *et al.* 2002). BCUs are "the principal organizational units of the police below force level." The CDRPs were created in the wake of the 1998 Crime and Disorder Act and regulate the mechanisms of cooperation between local authorities – police forces, health authorities, and probation committees, which the new law obligated to collaborate to tackle local crime and disorder problems. BCUs and CDRPs are similar in size but their borders are not always congruent (Harper *et al.* 2002, 2f.).

The methodology used to cluster BCU and CDRP areas across England and Wales into families is explained in Harper *et al.* (2002). The Home Office researchers used primarily data from the British census to come up with these spatial typologies. Harper *et al.* (2002) pre-selected 20 variables capturing the demographic, socio-economic, and built environment characteristics of these areas, based on their correlation with neighborhood levels of crime and disorder. Harper *et al.* (2002) then used two different clustering procedures to develop the neighborhood typology, k-means and self-organizing maps (SOM). Their optimization criterion for the clustering procedure was to create a typology of BCUs and CDRPs so as to reduce the variance in crime rates inside a given family. The rationale was that the families should describe the various policing environments that different BCUs and CDRPs represented. In all, Harper *et al.* (2002) identified 14 types of BCU families and 13 CDRP families for the entire territory of England and Wales.

The methodology of developing fixed general-purpose clusters for classifying BCUs and CDRPs for comparative performance assessment was later criticized. Ashby & Longley (2005, 59) cautioned that for comparisons between BCUs or CDRPs of the same family to be meaningful, the correct assignment of each spatial unit to the appropriate cluster was absolutely critical. However, this may not always be the case, especially for BCU or CDRP areas that fall inside the empty attribute space or grey zone between clusters and thus "share important characteristics with units in adjacent families."

Outside the field of criminology and policing evaluation, a few authors have recently employed artificial neural networks or data mining procedures to develop typologies of spatial units of analysis. Li & Shanmuganathan (2007) used the SOM algorithm for a social area analysis to classify 163 census tracts in Beppu, a city of some 120,000 thousand inhabitants in western Japan. Their dataset contained 90 demographic and socio-economic variables from the local census data, and the authors subsequently used GIS software to represent the results of the clustering procedure as a map. Spielman & Thill (2008) employed the SOM algorithm for a geodemographic classification of 2217 census tracts in New York City. Using a dataset with 79 attributes from the U.S. Census to describe the ecology of these census tracts, they used GIS software to visualize the results of the clustering procedure, which revealed that the spatial units that were most similar in terms of their social attributes were not necessarily closest in geographic space. For both of these studies, the choice of the variables to be included in the dataset was informed by expert opinion, and neither of the two did apply procedures to select key variables from among the datasets. Spielman & Thill (2008, 120) acknowledged that the SOM algorithm was an exploratory technique that is useful during the preliminary phase of data analysis but does not lend itself to variable selection or for conforming theory.

Neighborhood profiling in Switzerland

The profiling of territorial areas such as counties and urban neighborhoods is something Swiss geographers are quite familiar with. The Swiss Federal Office of Statistics (OFS) has developed a county typology which regroups Switzerland's 2,500-odd counties according to a classification system of 9 and 22 different county categories, respectively. The OFS typology aligns the counties on a center-periphery continuum based on 1990 and 2000 census data. In the context of the National Research Program "Sustainable Development of the Built Environment" (National Research Programme 54), a group of researchers led by Martin Schuler of the Swiss Federal Institute of Technology in Lausanne developed a typology of urban neighborhoods in 2008. This classification system expanded the OFS county typology to urban neighborhoods (the OFS had originally only classified counties and entire cities but not urban neighborhoods). The neighborhood typology classifies the statistical neighborhoods (census tracts) of the country's 17 biggest cities according to 13 categories (Dessemontet *et al.* 2008). However, considering exclusively the five cities being studied here, the typology counts nine different neighborhood clusters.

In similar fashion, the geography departments of some Swiss universities have created their own neighborhood/county typology to profile the local area. In one such study, Heye & Leuthold (2006) used principal component analysis of 13 indicators from the 1990 and 2000 census counts to develop a typology to profile urban neighborhoods and surrounding suburban counties in the

Zurich metropolitan area. One particular finding of this area profiling study was that the disparities between urban neighborhoods are greater than between the surrounding suburban communities.

Most recently, Swiss geographers have also taken up gecomputational data mining to develop such area typologies. Tuia *et al.* (2009) used the SOM algorithm paired with hierarchical agglomerative clustering to classify the 427 municipalities of the Vaud and Geneva cantons in western Switzerland based on 75 socio-economic variables. They discovered that five clusters were enough to represent the socio-economic structure of this region and visualized the resulting typology as a map using GIS.

In addition to these general-purpose neighborhood or county typologies, a number of authors have conducted social area analyses of urban areas in Switzerland. One finding from these studies that is relevant for the present study is that whereas during the 1970s and 1980s the neighborhoods that hosted more than a fair share of the socially disadvantaged and non-assimilated populations tended to cluster in the city centers and adjacent areas, since the turn of the century the city outskirts and surrounding suburban communities are more affected. Among the cities included in this study, this shift has been particularly pronounced for the 1990-2000 period in Basel and Zurich (Arend 2007; Arend *et al.* 2005).

Among Swiss criminologists, attempts at neighborhood profiling remain excessively rare, however. Manzoni & Thalmann (2008, 166) used cluster analysis to find suitable neighborhoods for their evaluation of a burglary prevention project in the City of Zurich. Two decades earlier, Riva (1988) studied the link between neighborhood ecological characteristics and property crime levels for the City of Lausanne. This study concluded that the two variables measuring the percentage of buildings with mixed or non-residential use and the share of the active working population relative to the resident population correlated most strongly with the rates of different property crimes across Lausanne neighborhoods (p. 190).

2.2 Assessing the Evidence on Community Policing

Following this brief overview over the challenges of community policing evaluation and some of the practical suggestions how both the process and impact evaluations might be improved, the following section of the literature review looks at actual evaluations of policing strategies in an effort to assess the available evidence. Sherman *et al.* (2002, 9) noted that evaluations of crime-prevention efforts that include measures of both implementation and impact provide the most insights but such studies are rare. This section thus first looks to Chicago and to England, which have been the backdrop of the two most comprehensive efforts to evaluate community policing that looked in depth at both the process of implementation and the impact. After this short discussion of the experiences in Chicago and England, the remainder of this review looks at what is known about the process of implementation and the impact of community policing in general. The evidence on implementation is structured around the three core elements of community policing – community engagement, problem solving, and police organizational reform; the evidence on its impact is organized around the four key constructs used to evaluate it – crime, fear of crime, neighborhood disorder, and public attitudes towards the police. This second part draws on a wider array of criminological studies and is not limited to evaluations of community policing.

2.2.1 Comprehensive Evaluations of Community Policing

Chicago Alternative Policing Strategy

The longest running and arguably most comprehensive evaluation of community policing has been conducted in Chicago, where a team of evaluators around Wesley Skogan has studied the Chicago Alternative Policing Strategy (CAPS) for more than a decade (Skogan 2006b; Skogan *et al.* 2004; Skogan & Hartnett 1997). The Chicago Police Department (CPD) launched its community policing strategy in 1993 in five police districts and expanded it to the entire city a couple of years later

(Skogan *et al.* 2004, 153). During the early stages of CAPS, Skogan & Hartnett (1997) were able to use a quasi-experimental research design, but this later collapsed as the program was expanded to encompass the entire city and all of its communities. Despite this apparent loss of design strength on the SMS, the analyses of CAPS remain one of the most comprehensive and methodologically sound evaluations of community policing. Over the years, the team of evaluators observed hundreds of beat meetings and conducted thousands of interviews with CAPS activists, police personnel, and ordinary residents (Skogan *et al.* 2004, 153).

In Chicago, community involvement has been taken further than elsewhere. The best known feature of CAPS, the CPD's monthly beat meetings where patrol officers meet with ordinary residents to discuss local security issues, were introduced in 1995 and have been institutionalized across the entire city. A second key element are the district advisory committees (DAC), which represent the community and whose members meet regularly with police district leaders to discuss local priorities (Skogan *et al.* 2004, i-iii). However, the evaluation in 2004 concluded that these institutions do not serve their purpose of shaping police priorities very effectively. The high turnover in the police officers attending beat meetings makes it impossible for local residents to get to know them, and the issues raised there were not well recorded in police memoranda, all the while too many DACs lacked in initiative and leadership (Skogan *et al.* 2004, 154).

By contrast, the CAPS evaluation in 2004 awarded the Chicago police the highest marks for its partnerships with other city government agencies. The CPD's CAPS Implementation Office coordinates inter-agency efforts to tackle crime and disorder problems, which in the past scored some notable successes in dealing with graffiti or bad buildings. In Chicago, every relevant agency has been involved in coordinated problem-solving efforts, making CAPS increasingly "the city's program" rather than "the Police Department's program" (Skogan *et al.* 2004, 154f.).

However, even though every staff of the CPD has been instructed in a five-step problem-solving procedure and beat officers can rely on a burgeoning web of supportive city services, the Chicago police earned only a passing grade for their problem-solving activities (Skogan *et al.* 2004, i). The team of evaluators analyzed hundreds of beat-level problem solving plans and found many of them wanting and locally identified problems to persist for years. Simultaneously, the importance of beat meetings to identify problems and set the police agenda has gradually slipped over the years, squandering the opportunity to engage local communities in a security partnership (Skogan *et al.* 2004, 155).

The CPD reorganized successfully to create CAPS support structures, though. Line officers were grouped together in newly created "beat teams", allowing them to focus on their assigned neighborhoods, with a sergeant supervising their problem-solving activities. At the precinct level, a district lieutenant took on responsibility as a "CAPS management team leader" for all aspects of management. However, at the department level, the CPD in 2000 instituted a new management accountability process not unlike New York's Compstat that, according to the evaluators, "has shifted the focus of headquarters to day-to-day crime fighting." This shift in emphasis notwithstanding, police headquarters organize regular review panels in order to keep up the pressure on district commanders to respond to local concerns and to coordinate police activities with other city agencies through the CAPS Implementation Office. Moreover, police internal inspectors regularly examine CAPS operations inside the CPD (Skogan *et al.* 2004, 155).

In terms of the impact of CAPS, the evaluation noted that in Chicago, like elsewhere in the United States, both violent and property crime steadily declined after a peak in the early 1990s. The CAPS evaluation also tracked fear of crime, which fell sharpest for those most prone to it – women, the elderly, and African-Americans. Regarding physical and social disorder, white Chicagoans reported fewer concerns during the early 1990s, yet still observed improvements over the following years of CAPS implementation. African-Americans were confronted with many more problems at the outset, but also noticed much bigger improvements over subsequent years. Finally, in terms of public attitudes towards the police, the survey results indicated steady improvements since the beginning of the CAPS program until approval ratings leveled off at an unprecedented level around

2000. However, the city's large and growing Latino population appeared left behind: Latinos – especially Spanish- rather than English-speakers – over the same period witnessed only marginal reductions in fear of crime, observed many more problems of both physical and social disorder without noticing much improvement, and were substantially less satisfied with the police (Skogan *et al.* 2004, iv-vi).

National Reassurance/Neighbourhood Policing Programmes

Ironically, at the time that the NRC's Committee to Review Research on Police Policy and Practices recommended dropping efforts to evaluate community policing as a comprehensive strategy (Skogan & Frydl 2004; cf. Chapter 2.1.1), arguably the methodologically soundest evaluation of community policing as a neighborhood-based strategy was underway across the Atlantic in England (Tuffin *et al.* 2006). At any rate, the evaluation of the National Reassurance Policing Programme (NRPP) was the only one among several place-based evaluations that used wards (neighborhoods) rather than smaller-scale places or hot spots as the basic unit of analysis that was included in the Campbell Collaboration systematic review of problem-oriented policing (Weisburd *et al.* 2008; Weisburd *et al.* 2010).

The "reassurance policing" strategy was the result of a collaboration between the University of Surrey and the Surrey Police that aimed to tackle the gap between the widespread perception that crime was on the rise even though actual crime rates were falling (Quinton & Tuffin 2007, 150). The strategy drew on the "signal crimes" perspective proposed by Innes (2004), which holds that some crimes and disorders are more obtrusive to the public than others and should be targeted by the police in order to reduce perceptions of risk and boost feelings of security. Crucially, according to this argument, ordinary citizens do not firmly discern crimes from disorder, and those events that act as "signals" as to their personal safety are not normally those that top the agenda of the police (Quinton & Tuffin 2007, 153).

The NRPP was a national pilot of the reassurance policing strategy across sixteen trial sites in eight police forces in England. A national program team developed program activity, building on the signal crimes perspective (Innes 2004) and insights from previous models of community policing such as the Chicago Alternative Policing Strategy. The NRPP rested on two strategic pillars: (a) the visible presence of police officers and Police Community Support Officers, which was meant to facilitate the involvement of the community in the process of identifying and prioritizing local crime and disorder issues, and (b) targeted policing activity and problem solving via processes of co-production between the police, partner agencies, and the public to tackle these crimes and disorders (Quinton & Tuffin 2007, 150f.).

Of the sixteen trial sites, six were matched to suitable control areas. The trial sites were selected from among police forces that volunteered to participate and matched on the basis of population density, ethnic composition, and the percentage of the resident population in managerial employment. The evaluators also strived to achieve balance between rural and urban as well as affluent and deprived areas, and consulted with police in order to select areas with similar crime levels (Quinton & Tuffin 2007, 152).

The NRPP was accompanied by a close process and impact evaluation (Tuffin *et al.* 2006). The process evaluation gathered police force and program documentation and conducted semi-structured interviews with police representatives in order to gather detailed information on the three main NRPP activities – increased police visibility, community engagement, and problem solving. The impact evaluation relied on police crime statistics as well as data from telephone surveys of 300 residents in each of the trial and control sites before and after the end of the 12-month program (Quinton & Tuffin 2007, 152f.).

The impact evaluation revealed that two of the six trial sites saw significant reductions in crime rates as a result of the program and, at a program level, there was a significant positive effect on five of the eight categories of anti-social behavior (physical or social disorder). Across the trial sites, the program had effectively boosted feelings of safety after dark and managed to narrow the "reassurance gap." The program had a positive effect both on police visibility and on the familiarity with the police, as well as on the public's perception of community engagement and police responsiveness across a range of measures. The program also had a measurable impact on how the public rated the effectiveness of the police. The evaluation concluded that the approach combining foot patrol, community engagement, and problem solving was crucial to achieving the desired outcomes and that involving the community not only in the identification of local problems and the setting of police priorities but also in the co-production of solutions, can have a positive impact on victimization, perceptions of anti-social behavior, and other perception measures (Quinton & Tuffin 2007, 154-159).

The NRPP trial was run as a pilot program at the ward (neighborhood) level between autumn 2003 and spring 2005. In the wake of the successful NRPP pilot, the Neighbourhood Policing Programme (NPP) was officially launched in the spring of 2005, which was supposed to introduce neighborhood policing to all police forces across England and Wales over a three-year period. Like the NRPP, the NPP was meant to promote the three key neighborhood policing activities of increased police visibility, community involvement in the definition of local priorities, and collaborative problem-solving efforts with strategic partners and the public to attend to those concerns (Quinton & Morris 2008, iv).

Like its small-scale predecessor, the NPP has been accompanied by continuing process and impact evaluations. The first of these evaluations was the NRPP follow-up study, which focused on the four most successful of the original six NRPP trial areas. The purpose of this follow-up evaluation was to test whether the improvements achieved over the first year could be sustained over the longer run. For the follow-up study, the trial wards remained matched to their comparison sites, thus maintaining the original quasi-experimental design. A third wave of telephone interviews was carried out after the second year of implementation. The data form this third round of interviews suggested that the achievements of the first year could by and large be sustained. The trial wards in London and Greater Manchester even saw significant reductions in victimization rates not seen over the previous twelve months, suggesting there were some lagged benefits. However, crime perceptions deteriorated slightly during the second year and respondents rated police visibility significantly lower. The authors thus concluded that community involvement and problem solving were more important in sustaining improvements in the long run than police visibility, and that "a concerted effort to implement neighborhood policing at a local level can deliver improved outcomes, and sustain them over time." (Quinton & Morris 2008, 9-13).

The second evaluation was carried out after the first year of the NPP implementation. This evaluation used the much larger basic command units (BCU) as the unit of analysis rather than wards. The sites included in the BCU evaluation contained between 140,000 and 360,000 residents and were thus several times bigger than the study units of the NRPP evaluations. The BCU evaluation tried to maintain a quasi-experimental research design by focusing on five experimental BCUs. These "Pathfinder" BCUs were selected by the research team primarily on the basis that they stood the best chance of fully implementing neighborhood policing over the first year of the program. A secondary concern was to include "a broad range of policing environments" by considering urban and rural areas, different regions, and the ethnic composition (Quinton & Morris 2008, 16f.).

The impact evaluation compared outcomes in the Pathfinder BCUs with controls areas that belonged to the same BCU family (cf. Chapter 2.1.5) but lagged in the implementation process and were thus "uncontaminated" by efforts to introduce neighborhood policing. However, contrary to the NRPP evaluations, where the control sites were committed to the status quo, under the NPP all police forces in England and Wales were supposed to adopt the new policy. Unsurprisingly, the process evaluation observed that implementation progress had been uneven across study sites and sometimes varied even within individual BCUs. Moreover, the Pathfinder BCUs showed no "consistent pattern" compared to their controls and few differences were statistically significant (Quinton & Morris 2008, 17-19).

The third evaluation of the NPP was the national evaluation, which studied neighborhood policing at BCU and force level across all 43 forces and 244 BCUs[4] in England and Wales. The process data for the national evaluation were obtained through repeated on-line surveys of the police forces that measured the "*self-reported* level of implementation", which the research team used to create implementation scores. Also the national evaluation tried to maintain a quasi-experimental design by comparing differences in outcomes between BCUs and police forces that – based on their implementation scores – were deemed to have implemented neighborhood policing with areas that had not yet already done so (Quinton & Morris 2008, 22-24; emphasis in original).

The data for the impact evaluation of the national study were official crime statistics as well as British Crime Survey (BCS) data on public confidence in the police. The national evaluation found that at the BCU level, the overall pattern of change was more encouraging than in the BCU evaluation, as all the outcome measure apart from the total crime rate and victim satisfaction items pointed "in the direction consistent with neighborhood policing having a positive impact." However, none of the observed changes was statistically significant, nor were there any positive effects observed at the level of police forces. Moreover, the national evaluation suffers from the limitation that the BCS data cannot be aggregated to the BCU level but only to the larger force level, and that even at this level the survey samples remain relatively small (Mason 2009; Quinton & Morris 2008, 26).

2.2.2 Community Policing Implementation

Community engagement

Community engagement is a hallmark of the community policing philosophy. Community policing is "defined in part by the efforts of the police to develop partnerships with both community members and the civic organizations that represent many of them collectively." (Skogan 2006c, 28). Police need to change their organizational structure to create opportunities for police-citizen contacts that foster such exchanges. In the United States this appears to be happening: in a national survey in 1999 almost all police departments that served cities of 50,000 inhabitants or more indicated they regularly held meetings with civil society groups (Skogan 2006c, 29f.).

In practice, though, community engagement is hard to pull off. Research shows that community-building is particularly challenging in underprivileged areas, which may be further plagued by a record of poor police-community relations (Wells 2009a, 9; Skogan 2006c 32). Resistance to community involvement may also come from within the police's own ranks. The police are normally disinclined towards programs conceived of by civilians and they are particularly loath to any provisions that give civilians some say in setting their operational priorities or in assessing their performance (Skogan 2006c, 33). However, there is also evidence to suggest that police officers' views of the public and of community policing itself change for the better as a result of being involved in community policing activities. Police personnel engaged in community policing projects were more satisfied with their jobs, viewed police-community relations more positively, and expected the community to get engaged in problem-solving activities (Lurigio & Rosenbaum 1994; Skogan & Hartnett 1997; Skogan 2006c, 34).

Problem solving

Problem solving is the second core pillar of the community policing edifice. Problem solving is the process of analyzing problem situations in order to identify the proximate causes that produce crime or generate calls for assistance and to develop strategies to deal with them or to alleviate their detrimental effects. In recent years, problem solving has been helped by the development of

[4] The author remains puzzled as to the correct number of BCUs in England and Wales. Whereas the number of police forces is consistently reported by multiple studies to be 43, the numbers quoted for the BCUs fluctuate considerably.

computer software to analyze operational data in order to locate crime and disorder "hot spots" that concentrate the bulk of complaints reported to the police (Skogan 2006c, 34).

Problem-oriented policing recognizes that the solution to problem situations may require the assistance of other agencies and that this response may not be police-based in nature. In order to enhance their problem-solving capabilities, police departments must establish a system of strategic partnerships with other branches of the local government. Such interdepartmental partnerships are not easy to pull off but are vital for community policing to succeed (Skogan 2006c, 34f.). In practice, police departments commonly do engage in partnerships with external organizations, but their responses to identified problems still frequently draw on traditional policing methods (Wells 2009c, 25-27).

Getting officers to engage in problem-solving activities is not easy either and requires substantial training (Skogan 2006c, 35). Research has revealed that what passes as problem solving in many police departments often still remains shallow and systematically "shortchanges" the phases of analysis and response assessment of the SARA[5] problem-solving method. In practice, problem solving also poses serious difficulties in both officer supervision and performance evaluation (Wells 2009c, 17, 33).

Problem-oriented policing has also reinvigorated police crime prevention efforts, which under the professional policing model had been primarily deterrence-based (Skogan 2006c, 36). Researchers have alternately linked problem-oriented policing to routine activity theory, the rational choice approach, and situational crime prevention (Weisburd & Eck 2004, 56). Problem solving has acquainted police with crime prevention theories, prompting them to extend their focus beyond offenders to address the routine activities of crime victims or to take advantage of the pivotal role of the managers of problem places such as landlords or shop-keepers (Skogan 2006c, 36).

Organizational reform

According to Mastrofski (2006, 47), community policing heralds two kinds of programmatic and structural changes of police departments. First, community policing broadens the police mission to encompass a much wider range of objectives compared to the narrower "professional" model. Second, police departments are expected to implement substantial reforms to their organizational structure in order to accomplish those goals.

The first element of reform, American police departments have embraced with alacrity. In 2000 more than 80 percent of large departments and more than 60 percent of small ones had amended their mission statement to reflect community policing values (Mastrofski 2006, 47). However, when it comes to adapting their organizational structure, police departments have shown less reformist zeal. The most comprehensive analysis of police organizational reform during the community policing era found mixed results. On the one hand, large municipal police departments failed to cut the height of their hierarchies and to become less functionally specialized. On the other hand, the agencies made significant progress toward the community policing ideals of decentralization, less bureaucratization, and greater civilianization (Mastrofski 2006 48f.).

Decentralization means delegating more power to make decisions to the people who actually carry them out at lower organizational levels rather than centralizing it all at the top. In community policing, decentralization is often synonymous with geographic accountability (Maguire & Gantley 2009a, 35). Decentralization is usually achieved by delegating more responsibility to the mid-level managers in charge of a city's policing precincts and granting individual patrol officers and their sergeants more leeway to identify and attend to community problems. In practice, this has most often meant assigning patrol officers to fixed beats and keeping them there for longer. This is also the level at which the police can consult and engage with residents and organize community groups to address the problems of their concern. Recent national surveys have found that virtually all

[5]SARA (Scanning – Analysis – Response – Assessment) is a four-step approach that police officers are instructed to follow, when dealing with a specific problem they are called upon to resolve.

larger American police departments assign patrol officers to specific geographic areas and regularly confer with local advisory councils representing the different communities in an area (Skogan 2006c, 36-38).

Of potentially even bigger value to evaluate a police department's commitment to community policing is information about the performance appraisal systems it has put in place to monitor overall and officer performance as well as the investments it makes to train its staff. Earlier studies of the subject were not especially heartening as they revealed that a preponderance of police departments were still keeping tabs on traditional crime and enforcement statistics rather than trying to track neighborhood quality of life indicators (Mastrofski 2006, 47). Recent research is more encouraging, though. While police departments commonly struggle to find satisfactory performance indicators because "problem solving effectiveness is hard to quantify", several of the agencies Wells (2009, 74-77) analyzed had made headway with adapting their assessment systems to community policing. Furthermore, several police departments had moved beyond traditional performance indicators such as clearance rates, arrests, and response times and had resorted to community polls to evaluate departmental performance, even though it remained unclear whether the agencies had fully exploited the survey data to improve police functions.

Recruit training plays a pivotal role as the principal management tool to inculcate department values, yet little is known about the impact of police training in general, let alone community policing training in particular. Critics bemoan that community policing training seeks to instill a set of beliefs rather than teach officers the necessary skills to act on them. The more fundamental problem about training, however, may be that police administrators often regard it as the "sole or principal mode of changing police culture", when supervisors, managers, or performance appraisal systems do not reinforce community policing principles. The extent of a potential misalignment between police values and management systems remains undocumented, but case studies of disappointing attempts at community policing implementation point in this direction (Mastrofski 2006, 51f.).

2.2.3 The Impact of Community Policing

Crime

Community policing is not "soft" on crime. The police maintain a firm focus on law enforcement activities and traditional indicators such as calls for service, response times, and arrests remain valid indicators of their performance. Under the new paradigm, they are not the only important indicators, however (Kennedy & Moore 1995, 281; Trojanowicz & Bucqueroux 1994, 6f.).

Community policing rose out of a theoretical and empirical critique of the professional era model of policing. At the theoretical level, community policing scholars argued that the narrow, reactive focus on serious crime led police to ignore the family contexts and neighborhood conditions that spawn crime (e.g. Trojanowicz *et al.* 2002, 167-174). At the empirical level, they pointed at the professional model's inability to stem the rising tide of crime since the 1960s, bolstered in their criticism by the results of the two famous policing experiments in Kansas City, which were seen as prove that the chief tactics of the standard model – randomized motorized patrol and rapid response to calls for service – had virtually no impact on crime rates (Kansas City Police Department 1977; Kelling *et al.* 1974; Weisburd & Eck 2004, 49f.).

In the wake of the Kansas studies, a gloomy view that "nothing works" prevailed among American policing scholars. However, this period of despondency spurred an "era of experimentation and evaluation" of novel policing strategies that eventually gave birth to the community policing and problem-oriented policing paradigms (Cordner 2010, 16). The policing scholars of the time were greatly influenced by the succession of several new criminological theories that appeared between the 1970s and the early 1990s: defensible space (Newman 1972), routine activities (Cohen & Felson 1979), situational crime prevention (Clarke 1980), broken windows (Wilson & Kelling 1982), environmental criminology (Brantingham & Brantingham 1991), and the rational choice perspective

(Clarke & Cornish 1986). What all these new approaches had in common was a shift of focus away from the person of the offender to the context of offending. These theories thus offered entirely new avenues for crime prevention. By contrast, under the standard model of policing, crime prevention is based more or less exclusively on the deterrent effect of the threat of arrest and punishment (Weisburd & Eck 2004, 45).

In their systematic review, Sherman & Eck (2002) identified four major hypotheses of how community policing should help bring crime rates down. Firstly, the increased surveillance of residential neighborhoods resulting from the instauration of neighborhood watch programs should deter potential offenders, because it increases their risk of being observed. Secondly, the community-based intelligence that police gather at neighborhood meetings, storefront offices, or through informal encounters between citizens and officers on foot patrol should increase the probability of arrest of suspects. This not only incapacitates past offenders but should have an additional deterrent effect. Thirdly, community policing also increases the flow of police intelligence back to citizens. If the police offer updated information on crime patterns and risks as well as advice on crime prevention or target hardening, citizens or business owners should be better able to protect themselves against crime. Finally, the enhanced legitimacy that community policing agencies enjoy in the eyes of the public is hypothesized to have a direct impact on crime. The rationale behind this fourth hypothesis differs from the previous three. The claim actually goes beyond the mere observation that the police must win the public's trust and confidence in order to ensure community cooperation in the law enforcement process; it states that if people perceive their treatment at the hands of the police as fair and equitable, they become more willing to obey the law itself (p. 299f.).

Regarding problem-oriented policing, Sherman & Eck (2002) observe that there are myriad specific hypotheses how the strategy can help prevent crime but maintain that these essentially come down to one overarching idea: "The more accurately police can identify and minimize the proximate causes of specific patterns of crime, the less crime there will be." In the past, problem-oriented approaches have taken two basic forms: attempts by the police to remove criminogenic substances such as guns, valuable goods, or moveable property lest they instigate crimes or be used to commit them, and efforts to hinder the intersection in space and time between a motivated offender and a suitable target (p. 300). For the sake of clarity, a distinction can be made between problem-oriented policing strategies that targeted the deployment of police resources on hotspots of criminal activity, those that involve local partners of the police in a process of the co-production of security, and situational crime prevention measures that block or limit opportunities to offend. In practice, however, there is often a considerable overlap between these approaches and the fault-lines are not clear-cut.

Until quite recently, the conventional wisdom among criminologists held that community policing was largely ineffective at preventing crime. In a narrative review published in 2004, Weisburd & Eck concluded that "[o]verall, the evidence does not provide strong support for the position that community policing approaches impact strongly on crime or disorder" (p. 52). In the years since then, however, a series of sound evaluations as well as several systematic reviews and meta-analyses have produced an impressive body of evidence showing that a surprising number of strategies associated with both community- and problem-oriented policing can have sizable crime reduction benefits.

Any discussion of the impact of community policing on crime is bound to begin with the subject of foot patrol. Foot patrol played a crucial role in the theoretical work and early stages of community policing implementation during the 1980s, and virtually all the evidence on the impact of foot patrol on crime levels still stems from that period. The Police Foundation's evaluation of the Newark Foot Patrol Experiment (Pate *et al.* 1981), which famously inspired Wilson & Kelling (1982) to their seminal article "Broken Windows", concluded that foot patrol did not affect crime, although it did impact levels of fear of crime and disorder. However, even though the evaluation was methodologically sound, with hindsight there are lingering doubts about the validity of its conclusions. Crucially, the Newark experiment relied on relatively large police patrol beats as units of analysis, which in light of more recent evaluation results appear to be ill-suited to detect crime reduction benefits of police patrol strategies (Sherman 2010, 603). Subsequently, a non-experimental evaluation of foot patrol in Flint, Michigan, reported a reduction in reported crime, whereas a study in Boston found no discernible reductions in crime or disorder as result of a dramatic shift from motorized patrol towards foot patrol. The main verdict from these experiments was thus that foot patrol had no impact on crime (Bowers & Hirsch 1987; Trojanowicz 1983; Weisburd & Eck 2004, 52f.). Until most recently, that was still the consensus view, even though the quality of the evidence used to be only moderately strong.[6] However, Ratcliffe *et al.* (2011) reported on a randomized controlled trial of the efforts of more than 200 patrol officers on foot patrol across 60 violent crime hotspots in Philadelphia in 2009. Using GIS software to carve out their spatial units of analysis, the authors after 12 weeks of treatment observed a net 23 percent drop in violent crime events at the target locations compared to suitable controls, leading the authors to conclude that targeted foot patrol can significantly reduce violent crime in high-level crime hotspots.

The empirical evidence on the crime reducing benefits of the second high-profile police tactic often associated with community policing – misdemeanor arrests or "zero tolerance" policing – was long mixed and hotly debated. In more recent years, however, a series of observational studies and a randomized experiment have shown that focusing police attention on disorder and minor criminal offenses has a small but significant effect on more serious crime (cf. Chapter 2.2.3).

The empirical record on the effect of *involving the community* in the policing process on crime rates remains mixed. Apart from the difficulties of mobilizing the local community into action discussed in the previous section (cf. Chapter 2.2.2), most such strategies appear to have no impact on actual crime rates. In general, strategies that merely aim to increase the *flow of information* between the police and ordinary citizens and vice-versa, although popular with the public, are ineffective at reducing crime. There is consistent and relatively strong evidence that such communication channels as police newsletters, storefront police stations, or regular meetings between police officers and neighborhood residents – as exemplified by Chicago's beat meetings – have little impact on crime (Sherman & Eck 2002, 315-317; Weisburd & Eck 2004, 52).

Dispensing advice on target hardening or actively engaging local residents in crime prevention, on the other hand, can have crime reduction benefits. Firstly, door-to-door visits by the police to advice citizens on security risks and crime prevention measures have proven an effective stratagem to lower victimization rates (Sherman & Eck 2002, 317; Weisburd & Eck 2004, 52). Secondly, and more surprisingly, recent evidence suggests that criminologists need to revise their standing views of neighborhood watch: whereas earlier systematic reviews dismissed neighborhood watch as ineffective at preventing crime (Sherman 1997; Sherman & Eck 2002, 315-17), a more recent meta-analysis of 12 studies covering a total of 18 evaluations concluded that the strategy can have sizeable crime reduction benefits of between 16 and 26 per cent (Bennett *et al.* 2006; Bennett *et al.* 2008, 2, 34).

Another promising approach for community policing to directly influence crime levels is through enhanced police legitimacy. Reviews of the subject have consistently found a robust link between the perceived legitimacy of the police and judicial authorities and people's willingness to obey the law (Sherman & Eck 2002, 318; Weisburd & Eck 2004, 59). In his work on "procedural justice", Tyler (1990; 2004) found a significant correlation between citizens' assessment of the legitimacy of legal authorities and their willingness to obey the law and concluded that the perceived fairness of how the police exercise their authority shapes the public's attitudes towards them independently of how effective they deem the police to be in fighting crime. While some authors question the extent to which community policing makes police officers engage in actions that enhance police legitimacy (e.g. Mastrofski 2006, 54f.), in a systematic review of six community policing evaluations, Skogan

[6] A more recent narrative review of thirteen studies evaluating police interventions that implemented foot patrol carried out in the United States, the United Kingdom, and Australia between 1986 and 2005 reported that none of them attempted to measure crime reduction benefits of foot patrol (Wakefield 2007, 348).

(1994, 175f.) found that people's assessment of police effectiveness and legitimacy had improved in all six treatment areas, most of which were compared to suitable controls.

There is a growing body of evidence based on strong evaluations that *problem-oriented policing* strategies are an effective way of reducing crime and disorder (Sherman & Eck 2002, 319-321; Weisburd *et al.* 2010; Weisburd & Eck 2004, 59). Whereas the earlier research was based on quasi-experimental designs, recently a number of randomized experiments have demonstrated its effectiveness. A Campbell Collaboration systematic review of ten experimental studies (four randomized trials and six quasi-experiments) concluded that problem-oriented policing approaches are an effective way of reducing crime and disorder. Calculating a standardized effect size over all the outcome measures of each study, the authors found a modest but significant effect of the treatment intervention in comparison with the control condition. Results were similar for a separate systematic review of 45 non-experimental pre-post or time-series evaluation designs conducted in parallel to the meta-analysis of the experimental studies (Weisburd *et al.* 2008; Weisburd *et al.* 2010, 153, 162).

Whereas Weisburd *et al.* (2010, 144) examined the effectiveness of problem-oriented policing as an overall strategy, the available evidence suggests that the approach works just as well as a tactic to combat specific crime and disorder problems. There is sound empirical evidence showing that targeting police efforts at high activity crime places or "hot spots" significantly reduces crime and disorder (Sherman & Eck 2002, 319; Weisburd & Eck 2004, 53f.). The first randomized trial of the effectiveness of the dosage of uniformed police patrol at such locations, the Minneapolis hot spots experiment, was also the first study that used hot spots rather than police patrol beats as units of analysis (Sherman & Weisburd 1995; Sherman 2010, 604). This study not only found that adding uniformed patrol had an impact – reducing total crime rates between 6 and 13 percent in the treatment areas – but also identified the "Koper curve", which holds that in order to achieve maximum residual deterrence, a patrol officer should remain for about 15 minutes on a high crime location (Sherman & Weisburd 1995; Sherman 2010, 599).

Although the Minneapolis experiment has never been fully replicated, since then a series of similar tests of the effect of additional uniformed police patrols on hot spots have been conducted and consistently reported benefits of reduced crime and disorder (Braga *et al.* 1999; Weisburd & Green 1995; Sherman 2010, 605). In three systematic reviews, Braga (2001; 2005; 2007) evaluated a growing number of studies that analyzed the effects of focusing police activity on crime and disorder hotspots. In the latest review, seven out of the nine selected studies (four quasi-experimental designs and five randomized experiments) reported substantial reductions in crime- and disorder-related calls for service in the treatment areas (Braga 2007, 19).

Previous research also suggests that the police stand a better chance of tackling some of the most protracted crime and disorder problems if they approach them in collaboration with other local government agencies or private actors. For instance, there is a sizeable body of research on police responses to drug dealing – whether on open-air drug markets or inside "drug houses." This body of evidence suggests that building up sustained pressure relying on civil statutes and regulations such as building codes or public health regulations is a more effective approach to control drug problems than more traditional police enforcement tactics such as surveillance, arrests, and search warrants (Mazerolle *et al.* 2000, 234f.; Sherman 2010, 612). In two recent systematic reviews that compared the relative impact of different street-level drug law enforcement strategies on different types of drug market problems, Mazerolle *et al.* (2006; 2007) concluded that "proactive interventions involving partnerships between the police and third parties and/or community entities appear to be more effective at reducing both drug and nondrug problems in drug problem places than are reactive/directed approaches." (2007, p. 138).

Furthermore, evaluations of situational crime prevention strategies provide additional support to the hypothesis of the effectiveness of problem-oriented policing. Systematic reviews of opportunity-blocking strategies at small places reported a substantial crime reduction effect following their installation, even though many of the reviewed studies employed weak research designs (Eck 2002, 281f.; Weisburd & Eck 2004, 56). Improved street lighting, in particular, has long been recognized as an effective way of reducing property crime in open public spaces (Eck 2002, 273). More recently, a Campbell Collaboration systematic review and meta-analysis of thirteen individual studies confirmed this view: it concluded that improved street lighting significantly reduced crime in both urban centers and residential areas. Moreover, the authors observed that crime decreased in about equal measure during daytime and nighttime, suggesting that the social benefits of the "community investment in the area" in terms of increased community pride and informal social control may weigh more heavily than the increased surveillance and deterrence resulting from higher luminosity during nighttime (Welsh & Farrington 2008, 2f.). Welsh & Farrington (2009) recently also updated an earlier systematic review of the impact of closed circuit television (CCTV). In their meta-analysis of 44 evaluations they concluded that surveillance cameras led on average to a modest but significant decrease in crime of 16 percent compared with control areas. The crime reduction effect was more pronounced for property offenses, especially inside car parks, whereas in more open public settings in urban areas and public housing estates the impact was much smaller and statistically insignificant. Larger effects were also observed for CCTVs installed in public transportation schemes, even though those effects too failed to muster statistical significance (Welsh & Farrington 2004, 2009). Interestingly, for both improved street lighting and surveillance cameras, the crime reducing effects turned out to be more pronounced in the United Kingdom than in the United States (Welsh & Farrington 2008, 3; Welsh & Farrington 2009).

The usefulness of problem-oriented or place-focused tactics would obviously be diminished if crime and disorder were merely displaced elsewhere (Weisburd & Eck 2004, 54). However, also on this score the existing research is encouraging. The first randomized experiment that explicitly tested for such a displacement effect, the Jersey City drug market analysis, found that calls for service dropped not just in the experimental locations but in the surrounding catchment areas as well (Weisburd & Green 1995; Sherman 2010, 605). This pattern has since been reported time and again. In Braga's (2007) systematic review, none of the five hot spots experiments that looked into the matter reported substantial spatial displacement of crime, and four studies found a possible diffusion of benefits associated with the police interventions (p. 15f.). More convincingly, Guerette & Bowers (2009) reviewed 102 evaluations of situational crime prevention projects with a total of 574 observations. Their findings suggest not only that displacement effects were equally likely as a diffusion of benefits (each occurring about 25 percent of the time) but also that if displacement did occur, it tended to be less important than the crime reductions in the target location itself, suggesting an overall net gain.

Fear of crime

Community policing is concerned with fear of crime chiefly for three reasons. First of all, surveys regularly find that fear of crime is widespread, affecting sizable parts of the general population, especially in urban areas. Secondly, fear of crime imposes a real cost on both individuals and society. If fear induces people to avoid certain areas or places, or makes them stay home altogether, it harms not only those directly afflicted by it but also society at large. If fear becomes endemic in an area, it undermines community organizational life, pushes up residential mobility, depresses the local housing market, and stunts economic activity (Skogan 1986, 207-209, 215-222, with quoted references). Thirdly and most importantly, it has been suggested that fear of crime affects people's behavior and social interactions in ways that may increase actual victimization rates (Hale 1996, 83). According to the "broken windows" hypothesis, the behavioral response to the fear induced by signs of disorder traps neighborhood residents in their fortressed homes, a process that weakens informal social control, makes an area vulnerable to criminal invasion, and may ultimately lead to neighborhood decay (Skogan 1990, 1986; Wilson & Kelling 1982). Under rival social disorganization theory, the corrosive effect of fear on cohesiveness and trust among neighbors may undermine the collective efficacy of communities and thus reduces their ability to fend off more serious crime (Sampson & Raudenbush 1999, 631).

In the United States, fear of crime began to rise in the wake of the surge in crime rates during the 1960s. Research interest in the subject has been driven by a series of "riddles" that have eroded the commonly held assumption that fear of crime was caused primarily by criminal victimization (Moore & Trojanowicz 1988, 2f.; Taylor & Hale 1986, 152). Most notable among the paradoxes unearthed by some of the earlier research were the following:

- victimization surveys consistently found that many more individuals said to be fearful than had recently been victims of crime or would be justified by objective risk of crime (Taylor & Hale 1986, 152f.); simultaneously, victims of crime were only marginally more fearful than non-victims (Garofalo 1979; Hale 1996, 112);

- some socio-demographic groups, notably women and the elderly, were consistently more fearful even though they suffered comparatively lower rates of actual victimization (Skogan & Maxfield 1981; Warr 1984);

- whereas fear levels tended to rise in parallel to increases in the level of reported crime, fear levels did not drop in line with subsequent falls in the crime rate (DuBow *et al.* 1979, quoted in: Taylor & Hale 1986, 151);

- likewise levels of fear of crime did not covary spatially with actual levels of crime, i.e. the neighborhoods with the highest levels of reported crime where not necessarily the areas were people were the most fearful (Lewis & Maxfield 1980; Skogan & Maxfield 1981; Taylor & Hale 1986, 153).

Pondering the seeming contradiction that fear is only moderately correlated with past victimization or the objective risk of crime, researchers – primarily from the United States but more recently also from other parts of the world – have offered various explanations of the causes of fear of crime. Chief among them is the notion of "vulnerability", introduced by Skogan & Maxfield (1981, 74-78) to describe the fact that their frailness left women and the elderly less able to cope with a physical attack and its aftermath, and hence more afraid of potential victimization. Skogan and Maxfield distinguished this *physical* vulnerability from *social* vulnerability, a term they invented to describe the fact that minorities and the poor were often forced to reside in areas with higher levels of crime and thus exhibited higher levels of fear of crime. A second important source of fear of crime is actual *victimization* (Skogan 1987). Although long debated in the literature, more recent research has established that victimization – particularly by serious crime or repeat victimization – makes individuals more fearful. A third explanation of fear of crime is the *indirect victimizatio*n model, i.e. when people learn about nearby crimes or victims through their social networks rather than becoming victims themselves (Taylor & Hale 1986, 156f.). Skogan & Maxfield (1981, 180) pointed out the significant impact of such vicarious experiences, which unlike actual victimization are relatively widespread. One aspect of indirect victimization is how crime reporting in the media affects local levels of fear of crime. In this respect, it appears that local outlets have a bigger impact than national media (Hale 1996, 112). Finally, some studies suggested that socio-psychological factors such as loneliness and lack of community attachment contributed to fear of crime among elderly urban dwellers, especially women (Silverman & Kennedy 1985, quoted in: Skogan 1986, 208).

Besides these individual-level factors, the literature has identified a set of sources of fear of crime linked to the characteristics of the local physical and social environment. The origins of these research strands, which Taylor & Hale (1986, 153-156) dubbed the "disorder" and the "community concern" perspective, actually preceded work on the individual-level factors. The disorder perspective links the presence of social or physical incivilities to heightened levels of fear. It was stated differently by different researchers and will be dealt with in the following section (cf. Chapter 2.2.3). The community concern perspective holds in short that neighborhood residents' fear of crime is more influenced by their perception of community dynamics than by crime trends (Taylor & Hale 1986, 155f.). Furstenberg (1971, reprinted in: Ditton & Farrall 2000) is commonly acknowledged as the first researcher to distinguish fear of victimization from broader measures of fear as *concern* about crime. Garofalo & Laub (1978) suggested that it is a more general "urban unease" that inspired fear among urban dwellers rather than a specific concern about any past or future victimization, famously asking whether "fear of crime" included more than "fear" of "crime" (quoted in: Taylor 1999, 66). Lewis & Maxfield (1980) argued that citizens' assessment of personal risk is not so much shaped by the prevalent crime rate in an area but rather by the perceived level of incivilities in their communities.

More recent research on fear of crime brought methodological advances, notably the use of hierarchical models in order to disentangle the influence of individual- and neighborhood-level factors (e.g. Robinson *et al.* 2003; Wyant 2008). Earlier studies had at times treated survey data taken across neighborhoods as normal random samples, ignoring possible nesting effects. A second development concerns the operationalization of fear. Fear is no longer conceived of as a permanent condition but rather as an emotion whose intensity varies over time. Farrall & Gadd (2004) introduced new survey items that do not merely record the prevalence of fear of crime but attempt to measure the frequency and severity of fearful episodes.[7]

As the empirical findings began to poke holes in the notion of fear as a largely "rational" reaction to crime, some authors were quick to grasp that fear of crime could be tempered in other ways than by reducing actual victimizations (e.g. Henig & Maxfield 1978). In accord with the broadening perspective on the causes of fear of crime, the proposed fear reduction strategies ranged from social control strategies over community-based responses to changes to the built environment (Hale 1996, 121-127). Goldstein (1977, quoted in: Hale 1996, 127) argued that the police would do well to adopt new strategies to allay fears, yet the notion that the police should concern themselves with fear of crime has never been without its detractors (Cordner 2010, 4f.).

Under the professional model, the police did not have any specific strategies to combat fear of crime. Fear is assumed to flow from actual victimization, and hence if the police managed to reduce the latter, the former would fall as well (Moore & Trojanowicz 1988, 4). During the 1980s, community policing scholars turned this traditional view that reducing crime leads to lower levels of fear of crime on its head. Building on research that showed that fear stemmed from many factors other than from actual victimization and objective risk of crime, Moore & Trojanowicz (1988, 3) theorized that "if fear could be rationalized and constructively channeled, not only would fear and its adverse consequences be ameliorated, but also real levels of victimization reduced", thus reversing conventional policing wisdom: instead of controlling crime to reduce fear, the police would henceforth attempt to control fear in order to reduce crime.

There are at least five causal models of how community policing is supposed to affect fear of crime (cf. Cordner 2010, 15). The oldest and most widely-held idea is that the increased police-citizen contact that community policing fosters assuages fear. Second, community policing is hypothesized to boost public confidence in the police, which in turn should reduce fear of crime. Third, if police use their regular contacts with neighborhood residents and community groups to better inform the public on actual crime and victimization rates, they may contain rumors of nearby crimes and victims and thus be able to reduce fear. A fourth avenue for the police to influence the prevailing levels of fear of crime is indicated by the broken windows hypothesis. If signs of disorder or unchecked decay make ordinary residents weary about their personal safety, steps by the police to combat disorder or neighborhood clean-up drives should improve the quality of life and reduce fear in the area. Finally, if the police reorient patrol activities to target crime hot spots or use the intelligence gathered at community meetings to attend to the nuisances that instill fear in the community, such a problem-oriented approach may successfully reduce it.

From its inception, community policing featured a varied mix of fear reduction strategies. The earliest and best-known tactic was foot patrol, which was not only meant to increase police visibility

[7] This discussion of fear of crime does not delve into the issues of operationalization and measurement, which are dealt with in Chapter 3.2.2.

but also allowed officers better to attend to neighborhood disorder problems. However, as foot patrol may not be desirable everywhere, especially in areas of low population density, police departments soon begun to experiment with other strategies to increase the number of non-confrontational encounters between citizen and their police: mobile police stations or storefront offices, community meetings, citizen contact patrol, bike patrol, etc. As part of their strategic communication efforts to reassure citizens, police departments have distributed police newsletters, made door-to-door visits, or taken advantage of community meetings to inform the public on crime rates and current trends and to advise them on crime prevention strategies (Moore & Trojanowicz 1988, 5f.).

The strategies community policing scholars have advocated to reduce fear of crime have evolved over the last two decades. A more recent addition from Britain is the "reassurance policing" strategy, which aimed to rectify the problem of a widening gap between the public's perception that crime was on the rise even though actual crime rates were falling (cf. Chapter 2.2.1). By 2010, Cordner advocated a problem-oriented approach that combines the well-known SARA process with elements of the community-oriented and professional policing models as the most effective strategy for fear reduction (pp. 15-23).

Besides these police-based fear reduction strategies, situational or community-based crime prevention efforts are also theorized to have an added benefit of reducing fear of crime (cf. Cordner 2010, 15). Target hardening, crime prevention through environmental design (CPTED), or situational strategies such as improved street lighting in dimly light areas have all been shown to have the potential to reduce fear of crime. In a similar vein, it has been suggested that actively involving the community in crime prevention efforts reduces fear of crime.

The empirical case that community policing reduces fear of crime is by now well-established and widely accepted. Several reviews of the available research (Dalgleish & Myhill 2004; Weisburd & Eck 2004; Zhao *et al.* 2002) and sound evaluations of community policing programs (Quinton & Morris 2008; Tuffin *et al.* 2006; Skogan *et al.* 2004) have accumulated a large body of evidence in support of the hypothesis that community policing reduces fear of crime. This body suggests in a nutshell that community policing programs designed to make police more visible and accessible or – in Pate *et al.*'s (1986, 35) famous words – "to increase the quantity and improve the quality of contacts between citizens and the police" successfully lower the prevailing levels of fear of crime in the community. In a systematic review of 26 studies conducted between the mid-1970s and the end of the 1990s, Zhao *et al.* (2002, 280f.) reported that a majority (59 percent) found increased police presence had a reassuring effect on the public independently of the study setting. Programs that went beyond mere targeted patrol, involving proactive, community-oriented policing strategies had an even higher rate of success (74 percent). Reviewing a series of quasi-experimental studies, Weisburd & Eck (2004, 53) found that policing strategies that increased community-police interaction and citizen involvement, such as citizen contact patrol, police community stations, and coordinated community policing, reduced individuals' levels of fear of crime and concern about crime in the neighborhood. In a similar vein, Cordner (2010, 22) recently pointed out that taking a problem-oriented approach to fear reduction has also proven to be a viable path in different contexts. In Baltimore in the 1980s, problem-oriented policing was noticeably more effective in allaying fear of crime than earlier efforts of "saturation patrol and traditional crime prevention efforts." However, Hale (1996, 130) cautioned that care must be taken in evaluating the impact of police presence on fear of crime. Police strategies to crack down on disorder may – "if handled insensitively" – have the perverse effect of undermining police-community trust and stoking fear of crime, a point recently confirmed by Hinkle & Weisburd (2008).

Regarding the impact of community-based or situational strategies to reduce fear of crime, the empirical record is patchier but on balance still positive. A relatively strong case can be made for improved street lighting, which has repeatedly been demonstrated to reduce fear of crime (Ramsay & Newton 1991, 20; Tien *et al.* 1979, 93f.). Smarter street lighting improves women's perceptions of safety at night (Atkins *et al.* 1991, 20) and generally leads to lower levels of fear and greater pedestrian use of public space in urban or residential settings after dark (Painter 1996, 200). As to community-based strategies, research has long pointed out that neighborhoods enjoying better neighborly relations and greater social cohesion also have lower levels of fear of crime (Taylor 2002, 787f.; Taylor *et al.* 1984, 324f.). It is, however, rather less clear which strategies are effective to increase social cohesion in areas that seriously lack it. What is more, not all community-based strategies have the desired effect on fear. Neighborhood watch, for instance, the best known and most popular community crime prevention scheme, appears to make those involved more aware of the crime risks lurking in their community and more fearful as a result (Cordner 2010, 18; Hale 1996, 126).

In Switzerland, a number of studies have analyzed fear of crime over the last two decades (Clerici & Killias 1999a; Schwarzenegger 1992). Most prominently, the Swiss Crime Survey has always included several items measuring fear of crime (Killias *et al.* 1989, 2007). For 2005, the SCS observed that between a fifth and a quarter of respondents indicated feeling insecure walking alone in their neighborhoods at night. This number increased slightly over the 2000 edition, whereas the perceived risk of a burglary victimization and the percentage of respondents who avoided certain areas to reduce the risk of victimization dropped over the same period. At the individual level, the Swiss studies by and large replicated the results from research in Anglo-Saxon countries, finding that women, the elderly, and those under 20 years of age are more afraid. The SCS also noted repeatedly that victims of crime, especially violent crime, were more fearful, as were respondents who had narrowly escaped a victimization experience (Killias *et al.* 2007, 74-78; Killias *et al.* 2011, 348).

At the community level, the SCS observed that respondents who spotted signs of physical or social disorder in their neighborhoods were more likely to be afraid and to avoid certain areas to avoid crime, with the most prominent neighbourhood-level sources of fear being the presence of "bothersome" or potentially threatening people as well as insufficient street lighting. Unsurprisingly, residents of bigger cities (with more than 50,000 inhabitants) on average felt less secure (Killias *et al.* 2007, 77-83).

More remarkably, in two separate studies Killias found evidence of a close correlation between the level of fear and the objective risk of victimization across neighborhoods in the City of Zurich. Based on data from representative surveys contracted by the city government, Killias observed that, aggregated at the neighborhood level, the percentage of respondents who had been the victim of an assault or threat over the previous five years correlated strongly with the percentage of respondents who took avoidance strategies to reduce the risk of victimization. Killias conducted the same analysis twice with two separate data sets from multiple years and both times reached the same conclusion (Killias 2003, 16-19; Killias *et al.* 2011, 351-353).

For the City of Lausanne, Kuhn & Viredaz (2004) in a first survey of the resident population observed that the respondents who reported feeling unsafe in their neighborhoods indicated that more *convivialité* (good neighborly relations), improved street lighting, a more conspicuous police presence, and more CCTV surveillance cameras were measures that would make them feel more secure. In two separate non-experimental pre-post panel design studies, the authors subsequently tested the impact of the first two of these proposed measures. They found that whereas better neighborly relations had a positive impact on quality of life and led to a slight reduction in fear of crime, improved street lighting did not (Kuhn & Viredaz 2007a; Kuhn & Viredaz 2007b).

Disorder

Like fear of crime, neighborhood disorder is another crucial construct of community policing theory that in the past did not figure high on police agendas. Loitering bands of teenagers, dilapidated buildings, and trash lying around in vacant lots or parks do not pose direct threats to public safety, but the new policing paradigm treats them as menaces to community welfare nevertheless.

The rationale behind the notion that the police must help maintain or restore order in neighborhoods has become generally known as the "broken windows" hypothesis (Wilson & Kelling 1982;

Kelling & Coles 1996). The main line of argument of this theory claims that, at the community level, disorder and crime usually go hand in hand, "in a kind of developmental sequence." Disorder and disorderly behavior, if they go "untended", lead to the "breakdown of community controls" and open the door to more serious vandalism, crime, and ultimately neighborhood decline. According to the theory, disorder by itself does not inevitably mean that "serious crime will flourish or violent attacks on strangers will occur." But in a disorderly environment, many residents become wary about their personal safety and "will modify their behavior accordingly." They will tend to retreat to their fortressed homes, and when out on the street, they will try not to get involved with any strangers. Families move out, "unattached adults" to whom the neighborhood is not their "home" but the "place where they live" move in. This growing atomization undermines the mechanisms of community control, making the neighborhood "vulnerable to criminal invasion." Neighborhood disorder thus acts a powerful signal of weakened informal social control that attracts more serious crime; just as the metaphor of a broken window that is left unrepaired signals "that no one cares" and that mischievous behavior costs nothing. More windows will soon be broken (Wilson & Kelling 1982, 31f.).

Although broken windows is by far the best known theoretical argument linking disorder and neighborhood development, Taylor (1999) maintained in a review article that it forms part of a bigger family of similar theories, which he subsumed as the "incivilities thesis." The incivilities thesis branched off from the fear of crime debate in the mid-1970s, seeking to explain the common finding of the first victimization surveys that many more people are fearful than are actually victimized. What binds these authors together, is a shared vision that "physical deterioration and disorderly social conduct each contribute independently to fear, neighborhood decline, and crime," and that, by implication, initiatives to check incivilities will bolster neighborhood stability and safety and allay fear of crime (Taylor 2006, 98). Several elements of the broken windows hypothesis such as the link between incivilities and fear and the crime preventing effects of informal social control had indeed been described before (Garofalo & Laub 1978; Hunter 1978 quoted in: Taylor 1999, 67f.; Jacobs 1961; Lewis & Maxfield 1980; Wilson 1977). The novelty of broken windows was that it introduced a dynamic, multi-step model of how persistent disorder and minor crime undermine a community's self-regulatory capacity and may lead to neighborhood decline. Whereas previous work was cross-sectional and focused on the individual, Wilson & Kelling (1982) shifted the unit of analysis to the community-level and widened the focus on group behavior and the ecological context (Taylor 1999, 65, 68).

In marked contrast to the tremendous impact broken windows has had on public policy and the heated debates this has provoked, the empirical foundations behind the theory remain tenuous. Most fundamentally, the core assertion of a "developmental sequence" tying together reactive, reform-era style policing, spreading neighborhood disorder, and rising crime has never been demonstrated by a single study (Cordner 2010, 20; Harcourt & Ludwig 2006). Both proponents and critics generally agree, though, that there is relatively strong empirical evidence in support of the front end of the hypothesis (Cordner 2010, 20; Sousa & Kelling 2006, 83). Several studies both from the United States and elsewhere have found that disorder begets more disorder (Keizer *et al.* 2008); disorder provokes the breaking of both informal and legal rules (Keizer *et al.* 2008); disorder stokes fear of crime (Hope & Hough 1988; Lewis & Maxfield 1980; Skogan 1990 Lewis & Salem 1986); and fear provokes withdrawal from public space and the outmigration of families (Skogan & Maxfield 1981).

The evidence on the back end of the hypothesized spiral of neighborhood decay – that disorder results in a breakdown of community control and a rise in serious crime – is still being debated controversially. Whereas the earliest empirical tests were supportive (Hope & Hough 1988; Skogan 1990), other studies have questioned the disorder-crime link (Greene & Taylor 1988; Matthews 1992). Moreover, Harcourt (1998), upon re-examining Skogan's (1990) data, maintained that the findings of his multi-site study unduly depended on the sample neighborhoods from a single city (Newark). Although Harcourts' study has in turn itself been criticized (Xu *et al.* 2005), the upshot of this debate may well be that Skogan's original sample was too sensitive to outliers (Eck and Maguire 2000, quoted in: Sousa & Kelling 2006, 84). What is more, all of these studies employed cross-sectional research designs. More in line with the dynamic nature of the theory, Taylor (2001) conducted a longitudinal study across several Baltimore neighborhoods but found that the influence of disorder on subsequent neighborhood development was minimal and far outweighed by other structural characteristics (Taylor 2006, 100f.). Other longitudinal studies did not find evidence of a strong or consistent impact of disorder on subsequent crime levels or neighborhood structure (Robinson *et al.* 2003; Brown *et al.* 2004).

Arguably the most serious challenge to the broken windows hypothesis was posed by the renaissance of social disorganization theory, first developed by the Chicago School of urban sociology during the first half of the twentieth century. Social disorganization theory had lost some of its shine when Robinson (1950) demonstrated that Shaw & McKay (1942/1972) in their seminal work had committed the ecological fallacy, but since the late 1980s the theory has experienced a revival (Bursik 1988; Sampson & Groves 1989; Robitaille & Séguin 2007, 93f.). In a much-cited study, Sampson *et al.* (1997) refined the original analysis by using multi-level modeling and by developing a survey instrument to measure the level of community cohesion dubbed "collective efficacy." Collective efficacy, i.e. the capacity of neighborhood residents to achieve common goals, namely to live in a safe and orderly environment, was defined as a combined index of the two underlying constructs "informal social control" and "social cohesion and trust" among neighbors, each of which were measured by five-item Likert-type scales and aggregated to the neighborhood level (p. 919f.). For the City of Chicago, Sampson *et al.* (1997) found collective efficacy a robust predictor of lower rates of violence across neighborhoods. Two years later, Sampson & Raudenbush (1999) introduced systematic social observation as a methodology to the study of neighborhood disorder based on expert observation rather than individual perceptions. They hypothesized that public disorder was not the cause of more predatory crime but that both "are manifestations of the same explanatory process, albeit at different ends of a 'seriousness' continuum." In other words, disorder and crime both belong to the *same* latent construct and are driven by the same antecedents, namely low levels of collective efficacy and structural disadvantages, in particular concentrated economic disadvantage, residential instability, and non-residential land-use patterns (pp. 608-610). The second study replicated the findings of the key significance of collective efficacy. The authors concluded that the association between disorder and crime was largely mediated by collective efficacy – except for the case of robbery – and hence a spurious one (pp. 627-630).

Sampson & Raudenbush (1999) have often been cited as offering the most compelling criticism of the broken windows hypothesis, but they are not without critics of their own. Sousa & Kelling (2006, 84) criticized that the systematic social observation of neighborhood disorder only extended during daytime hours and berated Sampson and Raudenbush for their "casual" dismissal of the significant link between neighborhood disorder and robbery. Other authors criticized Sampson & Raudenbush's (1999) model specification, namely that they failed to include a feedback loop between disorder and collective efficacy (Xu *et al.* 2005), or that – because according to the theory the link between disorder and serious crime is mediated by weakened informal social control – they were misguided to model disorder as a direct predictor of crime (Jang & Johnson 2001, 114). Somewhat surprisingly, Sampson & Raudenbush's (1999) critics have dwelt little on the cross-sectional nature of the study, which measured both disorder and collective efficacy simultaneously, and was hence ill-suited to prove causality, as the authors acknowledged themselves (p. 638f.).

From the practitioner's point of view, these debates about the validity and direction of any disorder-crime link could appear futile, but they cannot be dismissed as academic hairsplitting, because the policy recommendations flowing from the two models vary substantially. As Nolan *et al.*, (2004) pointed out, the recommended police intervention emanating from the broken windows perspective "was, and still is, simply to *repair the disorder*." By contrast, for adherents of social disorganization theory, community policing initiatives that focus exclusively on disorder and fail to take into account neighborhood-level collective efficacy, may actually work to its detriment (p. 100f., emphasis in original).

In the policy domain, Wilson & Kelling's (1982) article proved highly influential to the nascent community policing movement in the United States, which was still in fairly incoherent shape when the article was written. Broken windows has spawned a series of policing innovations, which under various names such as "order maintenance policing", "broken windows policing", or "quality-of-life policing" brought about a renewed police emphasis on public disorder and misdemeanor offenses. During the 1990s, New York's then mayor Rudolph Giuliani and Police Commissioner William Bratton repeatedly declared broken windows as key to their policing strategy (Sousa & Kelling 2006, 77-80, 94). Also the strategists behind Chicago's Alternative Policing Strategy (CAPS) made the broken windows view on disorder and crime a central plank in their thinking about neighborhood development (Skogan *et al.* 2004, 75-77).

In parallel, community policing has also brought about novel police activities drawing on social disorganization theory that are targeted more specifically on community mobilization and the strengthening of existing social ties. There are myriad ways for the police to get involved in such local community-building efforts but few of them are compatible with the traditional roles of police officers in the reform-era policing model (Rosenbaum 1998, 15f.). A better-known, recent example of this approach is the reincarnation of community policing in England and Wales under the new guise of "reassurance" or "neighborhood policing." The reassurance policing model gives communities a primary role in identifying and prioritizing local crime and anti-social behavior problems, which are then addressed in a collaborative effort between the police, partner agencies, and the public, which can be interpreted as a deliberate attempt to foster civic involvement (Quinton & Tuffin 2007, 150f.; cf. Chapter 2.2.1).

The empirical evidence in support of community policing strategies to tackle disorder is much stronger than the research on broken windows as a theory. In their systematic review, Weisburd & Eck (2004, 52) concluded that the available research generally supports the view that community policing activities such as community meetings, neighborhood watch, storefront offices, police newsletters, or door-to-door visits – although mostly ineffective as measures to prevent crime – nevertheless mitigate perceptions of disorder.

Regarding the impact of the aggressive enforcement of public order using citations or arrests over minor criminal conduct, most of the empirical evidence still stems from New York. In *Fixing Broken Windows*, Kelling & Coles (1996, 151-156) argued that a police crack-down on fare-beating youths had led to a permanent reduction of robberies in the subway system. Evaluating the New York City Police Department's reform activities, Kelling & Sousa (2001, quoted in: Sousa & Kelling 2006, 86) found that order maintenance policing had significantly reduced violent crime net of other drug-related and contextual variables. Likewise, Corman & Mocan (2002), also using the number of misdemeanor arrests as a proximate measure of order maintenance policing, concluded that the NYPD managed to bring down robberies and motor vehicle thefts (though not murders, assaults, or burglaries). More recently, three separate, methodologically sophisticated studies found that the NYPD's order maintenance policing strategy during the 1990s had a small but significant impact on homicide rates (Cerdá *et al.* 2009; Messner *et al.* 2007; Rosenfeld *et al.* 2007). The only dissenting voice finding no impact of the NYPD strategy was Harcourt & Ludwig's (2006) study (Braga & Bond 2008, 580f.). Nevertheless, the debate on the evidence of the specific impact of order maintenance policing on serious crime is likely to continue, not only because misdemeanor arrests alone may be too crude a measure of policing strategy but also since crime fell in all of the United States during this period. Moreover, the shift in strategic focus coincided with other substantive changes at the NYPD, notably the introduction of the CompStat management and accountability system.

The strongest evidence of the benefits of police efforts to tackle disorder does not come from New York but form a randomization controlled trial of a "policing disorder" strategy in Lowell, Massachusetts. For this study, Braga & Bond (2008) matched 34 discrete crime and disorder hotspots into 17 pairs in a randomized complete block design and used citizen emergency calls for service and systematic observation of social incivilities as outcome measures. After the treatment intervention, which lasted for one year, total citizen calls for service were down by 19.8 percent in the treatment areas compared to the controls, and also signs of social and physical disorder were significantly reduced. Distinguishing the impact of different elements of the policing strategy, the authors concluded that situational, opportunity blocking strategies produced the biggest crime-prevention gains, whereas misdemeanor arrests were less effective, and social service strategies produced no immediate gains at all. The authors concluded that "policing disorder can generate crime-prevention gains", lending support to the broken windows hypothesis (p. 598f.).

As to the impact of community policing on community mobilization and capacity building, the jury is still out. Earlier writings warned that previous experience cautions against overly optimistic expectations, as the rhetoric of citizen participation in community policing activities "has far exceeded the reality" (Rosenbaum 1998, 15f.). In a process evaluation of community policing across eight cities, Sadd & Grinc (1994) concluded that community participation in community policing had been limited. More recent research offers advocates a beacon of hope that community policing might help rebuilding the social fabric of neighborhoods, thus empowering residents to maintain order in their community (Sampson *et al.* 1997; Skogan 2006c, 31). The British NRPP evaluation study found that after program implementation, more people said they trusted many or some of the people in their area. However, no more thought theirs was a tight-knit community, nor was there any sign of greater involvement in community or voluntary activity (Tuffin *et al.* 2006, 61).

In sum, empirical proof of the broken windows theory remains a daunting task, but the evidence on broken windows as a policing strategy is more encouraging. This has prompted some authors to argue that broken windows should be judged on the effectiveness of the policing strategies it inspires, not the empirical validity of the theory. Taking this argument further, they claim that disorder is a serious quality-of-life issue and reducing it has its own merits regardless of any impact on more serious crime (Thacher 2004; Sousa & Kelling 2006, 90).

Nevertheless, for all the evidence of its effectiveness, order maintenance as zero-tolerance policing remains controversial. Wilson, Kelling, and Coles have always acknowledged that order maintenance raises complex legal questions and have addressed these issues at length (Wilson & Kelling 1982; Kelling & Coles 1996, Chapter 2). Sherman & Eck (2002, 315) expressed concern about the long-term effects of misdemeanor arrests; as an arrest record has a sizable negative effect on an individual's labor market participation, zero-tolerance may permanently lower police legitimacy among those arrested for minor offenses and thus lead to more serious crime in the long run.

Cordner (2010) thus has a point when he argues that broken windows "does not have to lead inexorably to strict enforcement, crackdowns, or zero-tolerance policing." Recalling that broken windows arose from the evaluations of early foot patrol programs, he notes that foot patrol officers – by relying on both formal and informal social control – inspired confidence among weary local residents that disorder would not spread, which does not necessarily imply an extensive use of citations or arrests. "In this sense, Broken Windows is very complementary to community policing, and represents a powerful approach to reassuring the public and making people feel safer" (p. 20f.).

Public attitudes towards the police

Research interest in public attitudes towards the police first emerged during the 1960s in the wake of the urban race riots, which exposed the strained relationship between the police and minority communities in various American cities (Schafer *et al.* 2003, 441). Since then, an impressive amount of research on public opinion of the police has been produced, both inside and outside the United States (Skogan 2006a, 101). Yet despite this sizable body of previous research, the concept is not being operationalized consistently in the literature and, as a consequence, the many factors and mechanisms that shape public attitudes towards the police are still imperfectly understood (Dalgleish & Myhill 2004, 16; Schafer *et al.* 2003, 442).

What is known, is that people on the whole have a rather positive opinion about their police. No matter how the concept is being operationalized, surveys regularly find that a majority of respondents are rather satisfied with or feel confident in their police, or have a positive image of them (Schafer *et al.* 2003, 442f.). Beneath the surface of widespread approval, though, lie some important differences, some of which have to do with *personal characteristics*. The most decisive such factor is an individual's age: study after study finds that the elderly are more favorably disposed towards their police (Bridenball & Jesilow 2008, 22). In the United States, race is an equally important predictor of attitudes towards the police, with ethnic minorities in general and African-Americans in particular being more critical (Schafer *et al.* 2003, 459f.). Other personal attributes, notably gender, education, and socio-economic status, have also been found to covary with attitudes towards the police, but the links are weaker than for age and ethnicity and do not consistently point in the same direction (Bridenball & Jesilow 2008, 172; Skogan 2006a, 101f.). Furthermore, in their review of the literature, Bridenball & Jesilow (2008, 151) observed that positive assessments of the police are more common among those more attached to the political system or the government, whereas negative attitudes are more prevalent among those who disapprove of the state's legitimacy in general.

Besides personal characteristics, everyday experiences or a *recent victimization* also bear on an individual's opinion of the police. Individuals who report higher levels of fear of crime are typically less satisfied with the police and demand more policing (Tseloni & Zarafonitou 2008, 403f.). In a similar vein, victims of crime generally are less satisfied with the police. Victims of crime typically are more "process"-oriented than they are "outcome"-oriented in their dealings with the police. Seeing somebody arrested or getting lost property back often matters less to them than how they are being treated by the authorities. Police willingness to offer advice or to update victims on the status of their case thus has a great impact on victim satisfaction (Skogan 2006a, 104f.). However, some authors point out a complicating factor regarding victims of crime, namely that at least part of the them have also been stopped or even arrested in the recent past, affecting how they judge the quality of police services (Maxfield 1988, quoted in: Skogan 1999, 51).

Another common finding in the literature is that the influence of personal attributes weakens or disappears altogether once an individual's *neighborhood context* is taken into account (Cao *et al.* 1996; Reisig & Parks 2000, 624f.; Schafer *et al.* 2003, 459). In areas with higher levels of crime, fear of crime, and more perceived community disorder, people usually hold less favorable attitudes towards the police (Bridenball & Jesilow 2008, 22; Reisig & Parks 2000, 620; Skogan 2006a, 103). Other neighborhood characteristics, namely concentrated economic disadvantage, were also found negatively to influence attitudes towards the police (Reisig & Parks 2000, 627f.; Sampson & Bartusch 1998). Cao *et al.* (1996) concluded that community context, especially perceptions of disorder, was the most important determinant of public confidence and that citizens thus hold the police "at least partially responsible for the disorder – the 'broken windows' – in their neighborhoods." (p. 12f.). More recently, researchers both in the United Kingdom and in the United States pointed out that earlier research probably overstated the influence of crime rates on public attitudes. Jackson *et al.* (2009, 108f.) suggested that public confidence in the police was less eroded by concerns about crime or the fear of crime than by perceptions of disorder, weakening social cohesion, and declining informal social control. The authors concluded that both fear of crime and confidence in the police are rooted in public assessments of non-criminal aspects of their neighborhood. In another recent article, Skogan (2009, 312) went further arguing that causality actually runs the other way. Using a structural equation model to analyze panel data, which enabled him to separate cause and effect, he found that public confidence in the police had assuaged people's concern about crime, whereas the link in the opposite direction was weaker and not statistically significant. It thus appears that the public is less holding the police accountable for neighborhood crime rates but rather being less concerned about crime if they are confident in their police in the first place. This is in accord with an earlier study by Reisig & Parks (2004), which found that residents who believe that police–community partnerships are healthy, perceive lower levels of disorder and report less fear as a result. This "reassurance model" of policing – the idea that public confidence in the police alleviates concern about crime – underpins community policing and argues for using public satisfaction with the police as an indicator of community policing impact (Skogan 2009, 302f.).

There may be other mechanisms at work accounting for the observable differences in attitudes towards the police in different neighborhood contexts, most notably the *behavior of the police* themselves. Several studies have shown that police officers' decisions to stop, search, or arrest people depend on the neighborhood context (Mastrofski *et al.* 2002; Skogan 2006a, 103; Terrill & Reisig 2003; Weitzer & Tuch 2004). Furthermore, it has been argued that local communities develop "neighborhood cultures" reflecting their collective experiences with crime, disorder, and the police. Such cultures may aggravate the impact of differing policing strategies as stories about police misconduct spread through neighborhood social networks, affecting the views of those who are exposed to them and perpetuating impaired police-community relations in certain neighborhoods (Sacco 1998, 133). Furthermore, in recent years the problem-oriented policing model has led to a surge in the number of police crack-downs on crime hot spots, yet little is known about the impact of such increased enforcement efforts upon citizen perceptions of police legitimacy in high-crime areas, as studies of the effectiveness of such tactics traditionally have not addressed this question (Braga 2005, 337).

Not in doubt on the other hand is the impact of a recent *personal encounter with the police* on an individual's assessment of the quality of police services. In this context, the literature traditionally distinguished between voluntary, i.e. citizen-initiated, and involuntary, i.e. police-initiated, contacts, with the latter being more likely to contribute to negative views of the police (Bridenball & Jesilow 2008, 159; Sacco 1998, 131; Skogan 2006, 104). More recent research suggests, however, that this distinction may be less important than meets the eye. Firstly, a number of studies found that respondents who have recently been in direct contact of any kind with the police are more critical towards them than those with no such experience (Skogan 2009, 312f.), although other studies found that a recent personal encounter was associated with both higher and lower satisfaction with the police (Bridenball & Jesilow 2008, 170). Secondly, far more consequential in terms of the impact of a personal encounter on attitudes is whether an individual felt satisfied about the treatment received by the police, regardless if the contact was solicited or not. On this score, there is mounting evidence that the impact of personal contact with the police may be highly asymmetrical, meaning that a bad experience may weigh heavily on an individual's opinion, whereas if the police deliver professional service, most people give them virtually no extra credit. This asymmetry appears to apply for both police- and citizen-initiated contacts (Schafer *et al.* 2003, 460; Skogan 2006a, 112, 118).

Public attitudes towards the police do matter. Research suggests that if the police are seen as legitimate by the public, people are more willing to obey police orders or requests, more willing to cooperate with the police by reporting crimes or working with them in their communities, and more willing to grant the police greater discretion in exercising their statutory duties (Tyler 2004, 89; Sunshine & Tyler 2003, 534f.; Tyler 1990). In personal encounters with the police, individuals are more willing to accept police officers' decisions if they feel having been treated fairly and openly by them (Tyler 2004, 91f.; Tyler & Huo 2002) It has already been noted that there is a growing body of evidence to suggest that greater legitimacy of the police and judicial authorities increases the willingness of people to obey the law. In other words, by focusing on police action that fosters police legitimacy, the police can enhance their public image and be objectively more effective in enforcing the law (Sunshine & Tyler 2003, 534f.; Skogan 2006a, 118f.). By contrast, if they lack legitimacy in the eyes of the public, the police have their work cut out. Systematic observations of police-citizen encounters have revealed that if police officers fail to show respect toward citizens, the latter become less willing to obey police requests (Mastrofski *et al.* 1996; Weisburd & Eck 2004, 53). In areas with a history of impaired police-community relations, where people challenge the police's authority and legitimacy, the police are unable to function effectively. At worst, widespread discontent among minorities with police action or the wider criminal justice system may flare up as violent protests as happened repeatedly in the United States since the urban riots of the 1960s.

Community policing from the very beginning set great store on greater engagement with ordinary citizens and on restoring or improving police-community relations. Community policing scholars have long found fault with the professional policing model's overreliance on motorized patrol as the main culprit that has unduly detached police officers form ordinary citizens and driven a wedge between the police and local communities. Buoyed by the findings from some of the earliest evaluations of foot patrol experiments, scholars have argued that increasing the rate of non-emergency contacts with ordinary citizens will improve public satisfaction with the police. Under the banner of community policing, police departments across America and beyond have since tried to capitalize on this by forcing police officers out of their patrol cars and implementing a series of strategies such as foot patrols, citizen contact patrols, or storefront or police mini-stations with the common objective of increasing the number of official yet informal contacts (Sacco 1998, 132).

Over the last three decades, scholars have accumulated a large body of evidence indicating that community policing does indeed improve public attitudes towards the police. Several systematic reviews (Dalgleish & Myhill 2004; Myhill 2006; Rix *et al.* 2009; Skogan 1994) as well as a series of sound evaluations of community policing programs (Bennett 1991; Skogan *et al.* 2004; Tuffin *et al.* 2006) all concluded that these programs do indeed have the hoped for effect of improving public attitudes towards the police, even when this was not the primary objective of the program (Bennett 1991; Sacco 1998, 132). In a narrative review of six quasi-experimental community policing evaluations, Skogan (1994, 175f.) found that people's assessment of police effectiveness and legitimacy had improved in all treatment areas across six American cities as a direct result of the community policing intervention. In the United Kingdom, the NRPP evaluation study, using the British Crime Survey measure of the percentage of people who think their local police do a good job, found an overall effect of community policing on the proportion of people who rated the effectiveness of the police in their area as either good or excellent. The effect was significant in four of the six experimental areas compared with control areas (Tuffin *et al.* 2006, 50f.).

In a recent systematic review of the impact of specific community policing strategies on public attitudes towards the police, Rix *et al.* (2009, 9-12) concluded that the most effective strategies were initiatives aimed at increasing community engagement. The authors found the strongest evidence in support of community policing programs that combined targeted foot patrol, community engagement, and effective problem solving. The authors noted that a sustained high-quality implementation was necessary to deliver the desired boost in confidence in the police, for if police efforts are sagging, this may dim the public's view. Furthermore, the authors argued that the police would be ill-advised to assume that the same strategy will work everywhere and have to tailor their response to each area or situation. The review also found that high-quality community engagement such as contact patrol, locally-based communications, or newsletters as well as efforts at restorative justice involving the police increased public confidence. With all these efforts, it appeared that it is the quality rather than the quantity of the police-citizen contacts that makes the difference. These results are in accord with the findings from an earlier systematic review, which concluded that merely increasing police visibility – while sometimes successful at allaying fears – does not have a direct positive impact on public satisfaction with the police (Zhao *et al.* 2002, 281).

In still the same review, Rix *et al.* (2009, 12-17) identified another clutch of policing strategies that look promising, even if the evidence of their impact on public confidence in the police remains inconclusive. One strategy tried out in the United Kingdom consists of mapping problems in order to target scarce problem-solving resources to the areas that need them most. This "reassurance mapping" had a positive impact but had been evaluated using a weak study design. Two more promising community policing initiatives are the use of multiple methods of public consultation (beat meetings, internet, surveys) and the training and education of members of the public in order to reach a wider range of community members and achieve better representation in community policing activities. A third related strategy is to improve the community engagement skills of police officers. Myhill (2006, 34) found that police officers often received insufficient training for their community policing roles, even for such basic tasks as organizing and running a public meeting. Finally, joint public consultation and communication efforts by different government agencies also look promising but the hoped for increase in public confidence does not always materialize. The main barriers cited by the researchers include the usual pitfalls of government inter-agency cooperation: reluctance of some agencies to participate or share information; desire to protect budgets, retain leadership, and get credit; and finally, over-reliance on informal exchanges between key individuals that collapse if they move on to other positions (Rix *et al.* 2009, 17).

Finally, Rix *et al.* (2009, 17f.) found that community policing is by no means a panacea. On the contrary, the available evidence suggests that perfunctory implementation actually risks undermining public confidence in the police. The first pitfall identified is that public consultation conducted under community policing is rarely representative of the community it serves. Skogan *et al.* (2004, iii), who looked into the issue for more than a decade, found that beat meeting attendance and membership on the district advisory committees in Chicago is highly skewed towards the elderly, the better off, and long-term residents, whereas Latinos are dramatically under-represented in most neighborhoods. Second, highlighting crime and anti-social behavior too much at such public consultation meetings may stimulate feelings of fear or threat among the listeners and in turn undermine public confidence. Finally, both police officers talking negatively about their own organization in public as well as a police culture resistant to change that does not recognize community policing as important police work, seriously risk undermining its proven capacity to improve public attitudes towards the police (Rix *et al.* 2009, 18f.).

In Switzerland, several studies of public attitudes towards the police have been conducted over the last three decades. First of all, as part of their efforts to implement community policing, all five police departments under scrutiny in the present study have contracted regular surveys of the resident population to gauge local levels of fear of crime and public attitudes towards the police (cf. Chapter 4.1.5). Secondly, at the national level, the Swiss Crime Survey has consistently featured questionnaire items probing for popular satisfaction with the police since its inception in the mid-1980s (Killias *et al.* 2007, 1989), the data of which have been used for additional follow-up studies (Clerici & Killias 1999b). In addition, there have been several other scientific studies of public attitudes towards the police in some of the cities covered in this study and other Swiss urban or cantonal police departments (Eisner & Manzoni 2000; Kuhn-Roux & Kuhn 1994).

The findings of these Swiss studies grosso modo tally with the results from the international research as far as the influence of personal and neighborhood characteristics are concerned (Killias *et al.* 2007, 2011, 375f.). The Swiss findings stand out though for the rather high approval rating of the police – rates of 80 percent or more of respondents declaring themselves satisfied or very satisfied with their local police are not unusual – and the fact that first-generation immigrants have a somewhat more favorable opinion of the police than native Swiss residents. Tellingly, the latter are also more likely to believe that the police discriminate against foreign residents than immigrants themselves. Moreover, the Swiss Crime Survey has noted differences in attitudes towards the police between the country's three main linguistic regions, a phenomenon for which no convincing explanation has been offered (Killias *et al.* 2012, 421; Killias *et al.* 2011, 375).

No previous study has attempted to assess the impact of specific policing strategies on attitudes towards the police in Switzerland. In some cities where the police have conducted regular surveys to gauge local public opinion, these polls have witnessed an improvement in the overall rating of the police since community policing implementation began. The Swiss Crime Survey also noted that the attitudes of victims of crime, especially violent crime, have grown significantly less critical of the police over the years (Clerici & Killias 1999b; Killias *et al.* 2007, 2011, 375).

2.3 Concluding Remarks

Greene's (2000, 331) remark that knowledge about how community policing is reshaping American police organizations and service delivery is quite limited to a large extent still rings true today. As Mastrofski (2006, 49) pointed out, what would be needed for a valid assessment of the success of community policing implementation is a sufficiently large, representative sample of police organizations monitored through on-site observation over time. However, the majority of existing evaluations of the implementation process are either in-depth case studies or cross-sectional comparisons of police departments, most of which were singled out for study as models of successful community policing implementation.

What emerges from the literature, though, is that the process of implementing a full-blown community policing strategy is risky and demanding, and plans can go awry (Mastrofski 2006, 49; Skogan 2006c, 41). Efforts to introduce earlier forms of community policing in England and Wales were characterized by police internal marginalization of the units concerned and resulted in implementation failure (Quinton & Morris 2008, 3). More recent process evaluations of community policing still regularly note serious problems in the implementation of key elements, particularly community involvement and problem solving mechanisms (Quinton & Morris 2008, 4; Sadd & Grinc 1996, 3).

What is known for a fact, by contrast, is that implementation of community policing takes a long time even under the best of circumstances. Historically, the pace of organizational changes in policing has been "glacial – slow and at times torturous." (Greene 2000, 331). Major changes in policing strategy have not occurred as cataclysmic shifts but have come about over extended timeframes, generally of more than 15 to 20 years (p. 301). The consequences for evaluation research are obvious: no experimental design however carefully planned and executed can be sustained for so long.

As for the evaluation of community policing impact, the complaint that there has not been enough good-quality research to tackle the effectiveness question is a common refrain in the literature (Greene 2000, 327; Mastrofski 2006, 65; Skogan 2006c, 41). In their systematic review, Weisburd & Eck (2004, 59) expressed their astonishment that while community policing had become the newly dominant paradigm in American policing and had received unprecedented levels of federal government funding for police agencies, there was no strong body of research that would allow researchers or practitioners to rate the effectiveness of community policing with strong confidence: "Given the importance of community policing, we were surprised that more systematic study was not available" (ibid.).

Then there are the numerous complaints about the weakness of the study design of many evaluations. Weisburd & Eck (2004, 59) reported that they did not find a single randomized experiment evaluating community policing. Quasi-experimental designs are more common, but the process of matching trial and control groups is seldom made on the basis of more than a handful of variables and rarely discussed at great length. Furthermore, quasi-experimental designs may collapse if a community policing program is being scaled up – after a successful trial run – to an entire city or much larger jurisdictions, as happened both in Chicago and with the British NRPP/NPP programs (Quinton & Morris 2008, 17).

Skeptics like to add that claims about community policing's effectiveness in reducing fear of crime or popular perception of disorder remain shaky as long as even the strongest research designs fail to control for the possibility that these improvements in people's perceptions were the result of a "Hawthorne effect" (Mastrofski 2006, 56). This criticism is not far-fetched, as the following comment from an evaluation of community policing across eight American cities attests (Sadd & Grinc 1996). Observing that most respondents thought community organization and involvement had increased since the start of the community policing project, the authors quote a local police administrator as saying that the various interventions taking place "made residents feel 'there [was] some interest in them.' Residents of many of the [community policing] neighborhoods to whom the police were paying attention for the first time may have felt that any intervention was better than no attention at all." (p. 16).

As this brief overview of some recent research has revealed, there remains ample scope for further research on community policing evaluation. Thirty years after the emergence of the concept, community policing enjoys widespread popularity but the scientific evidence in its support is at best moderately strong.

Various questions regarding community policing implementation and its impact on both the police and the public remain open and require further study. There is scope for more comparative evaluations of the process of community policing implementation across multiple sites over extended periods of time. There is a particular need for innovation to improve the quality of long-term impact evaluations. The time-frames required for community policing to take hold within a police department make field experiments hard to sustain and call for novel methods to enhance the validity of observational designs.

Last but not least, at a time when much of the relevant research still originates in Anglo-Saxon countries there is a need to conduct the current study in Switzerland, where the first police departments began to experiment with community policing as early as the mid-1990s but where the strategy has never been subject to an empirical evaluation.

2.4 Research Hypotheses

Based on the theoretical writings on community policing implementation and impact as well as the insights gained from previous process and impact evaluations and other empirical studies, the following research hypotheses were formulated to guide the current research:

2.4.1 Neighborhood Characteristics

- Areas with higher levels of crime have higher levels of fear of crime and lower levels of satisfaction with the police.
- Areas with higher levels of fear of crime have lower levels of satisfaction with the police.
- Areas with higher levels of disorder have higher levels of crime and of fear of crime.
- Areas with higher levels of social cohesion have lower levels of crime and of fear of crime.

2.4.2 Implementation and Impact of Community Policing

Organizational strategy and partnerships

- If community outreach and problem-solving activities are the prerogative of a few specially trained officers, there is an increased risk of their marginalization within the police organization and the abandonment of all community-oriented policing functions.
- With the adoption of community policing, police departments will increase the number of strategic partnerships with other branches of the local government over time.
- Involvement of the general public in community policing activities will not be representative of the overall population and be difficult to sustain in the long run.
- With the adoption of community policing, the police will broaden the definition of their mission to include combating fear of crime and disorder alongside combating crime.

Decentralization and deployment

- Areas with a neighborhood police station and permanently assigned community liaison officers have lower levels of fear of crime and higher levels of satisfaction with the police.
- Areas with higher levels of police visibility, notably of officers on foot patrol, have lower levels of fear of crime and higher levels of satisfaction with the police.

Broken windows approach

- Areas with campaigns to remove signs of physical disorder have lower levels of crime and of fear of crime.
- Areas with intervention teams attending to social disorder have lower levels of crime and of fear of crime.

Problem solving

- The better police can identify and attend to the proximate causes of crime patterns, the lower the levels of crime.
- The more strategic partnerships the police have concluded with other branches of the local administration, the less law enforcement-centered strategies to deal with disorder will be.
- The more strategic partnerships the police have concluded with other branches of the local administration, the more likely they are to be involved in the planning and steering process of neighborhood regeneration projects.
- With the adoption of community policing, the police will step up crime prevention activities.

Equity and legitimacy

- If the police train officers in the psychology of victimization and update victims on the status of their case, victims of crime will be more satisfied with them.
- In jurisdictions where the police train officers in intercultural awareness and have formalized procedures to handle citizen complaints of misconduct, the general public will be less likely to think the police discriminate against particular groups and be more satisfied with them.

3

Methodology

3.1 Process Evaluation

Prior to an evaluation of its impact, the process of program implementation must first be studied and documented. Analyzing the effects of a program remains of limited value if an evaluator is unable to specify the activities that brought them about (Lipsey *et al.* 2006, 278). A thorough process evaluation, on the other hand, enhances the external validity of the study. If the program proves effective, it can easily be replicated elsewhere. If the program shows no effect, lessons can still be drawn whether flawed theory or faulty implementation is to blame, the classic distinction between “program failure” and “theory failure” (Weiss 1972, 38).

A process evaluation can fail in more than one way that will undermine the validity of the impact evaluation. One is that program activities are not documented at all, or not in sufficient detail if the services delivered vary substantially from one program participant to another. A particular challenge poses the evaluation of programs that are being implemented across several sites, where scope and activities may vary considerably from one site to another. A global impact evaluation will be largely restricted to documenting average effects, but with a thoroughgoing process evaluation, detailing local program variants, and information on separate effects, more differentiated inferences about program impact become possible (Lipsey *et al.* 2006, 279).

In a systematic review, Sherman & Eck (2002, 301) bemoaned the poor quality of process evaluations in police research, which often leaves police action unmeasured. The studies the authors reviewed largely failed to track police activities and those that did often did so using methods that did not allow valid inferences about actual implementation. Studies of the police varied less by the methods of measurement than by the research designs used for drawing inferences. Yet however strong the research design of the impact evaluation, “the measurement of police activity remains the Achilles heel of police research.” (ibid.).

The threat of weak process evaluations is particularly acute in the area of community policing, where actual implementation rarely goes according to the original plans. Studies looking into the process of implementation regularly identify an important gap between a program’s design on paper and its execution in the field (McElroy 1998, 81). In the early stages of implementation, community policing programs almost invariably face the teething problems of overcoming patrol officer resistance and stimulating community involvement (Sadd & Grinc 1996, 3). At later stages, there is evidence to suggest that programs get delayed or not implemented in full, particularly as far as problem-solving activities are concerned (e.g. Hope *et al.* 2004, 6; Quinton & Morris 2008, 18f.). In other words, a process evaluation that carefully tracks what has *happened* rather than what had merely been *planned* is crucial to the validity of an impact evaluation.

3.1.1 Data Collection

Large-scale surveys of police departments across the United States found that in the year 2000 the overwhelming majority of police departments in large American cities claimed to have “incorporated community policing values in their mission statements" and to be practicing some variant of it (Mastrofski 2006, 47; Skogan 2006c, 27). However, parallel studies of police organizational structure and operational strategies observed little sign of a corresponding shift of police priorities from crime fighting to order maintenance and service provision (Zhao *et al.* 2001, 373; Zhao *et al.* 2003, 715f.) and virtually no sign of the far-reaching organizational changes within police departments (Maguire 1997, 572), as the community policing literature would have it. These apparently contradictory findings underline the importance of the data collecting method and highlight the need for valid and reliable indicators of community policing implementation.

Methodologically, process evaluations of community policing typically fall into one of two categories: case studies that look at the implementation of one or two individual police departments (e.g. Skogan *et al.* 2004; Wycoff & Skogan 1994). Case studies mostly include on-site inspections by the researchers often combined with in-person interviews with police administrators. By contrast, cross-sectional studies of community policing implementation across a wider array of police departments commonly rely on written or on-line surveys of police administrators (e.g. Maguire 1997; Quinton & Morris 2008; Zhao *et al.* 2001, 2003). More recently, there have been a number of process evaluations that combined on-site inspections or in-depth interviews with a cross-sectional study design (e.g. Maguire & Wells 2009; Sadd & Grinc 1996; Tuffin *et al.* 2006).

Another important aspect of a process evaluation of community policing is the length of the study period. Apart from some notable exceptions (e.g. Skogan *et al.* 2004; Zhao *et al.* 2003), community policing evaluations seldom study a period of more than one or two or (rarely) three years. This is in striking contrast to the community policing literature that stresses the long-term nature of the concept. For instance, as one of the early reformers, Sparrow (1988, 2) harbored no illusions about the order of the challenge and the time required to alter an organizational philosophy. Later studies of the subject only confirmed this assumption. Greene (2000, 309) noted that shifts in policing had historically occurred over long timeframes of "generally more than 15 to 20 years", and Maguire & Gantley (2009, 55) mused whether it would "take an entirely new generation of police officers, trained for community policing, to overcome some of the enduring values of the current culture."

From the beginning, it was thus decided that for the current study the information on community policing implementation would have to be obtained through visits of the five police departments under study. On-site observation is commonly considered a more valid approach to assess the process of community policing implementation (Mastrofski 2006, 49; Rosenbaum 2002, 197f.; Sherman & Eck 2002, 301). In view of the cross-sectional study design and a study period extending over several years, it was important to document each police organization's implementation of community policing in its unique local setting. With very little prior information to go by, this would have been difficult to accomplish by means of a mail, online, or telephone survey administered from afar. Capturing this local variation required a qualitative approach (cf. Maguire & Wells 2009, xvii).

The process evaluation aimed to trace the origins of each department's community policing efforts and to track the history of implementation over the subsequent years. The objective was to gather factual information on reforms of the police organization and operating strategies in order to trace the actual level of implementation over time (cf. Quinton & Morris 2008, 22).

Interviews

The author visited all five municipal or cantonal police forces and held interviews with one or more department representatives well-informed about the local community policing strategy. As a matter of fact, all of the author's interlocutors counted several years of experience in their current function and for the most part had been closely involved in their department's strategic planning process to prepare the community policing transition process. The following police representatives were interviewed:

- Basel cantonal police – Head of the Community Policing Unit in Kleinbasel and a Community Liaison Officer from Grossbasel
- Bern cantonal police – Head of the Community Policing Unit for the region of Bern
- Geneva cantonal police – Head of the Unit for Strategic Analyses
- Lausanne municipal police – Chief of Police
- Zurich municipal police – Head of the Crime Prevention and Community Policing Unit, plus the project managers of two community policing projects

Semi-structured interviews were conducted with all of these police representatives using an interview guide based on the six analytical dimensions of the process evaluation (elaborated in the next section). These interviews lasted between three to five hours. The guidelines for the interviews were sent to the author's interlocutors ahead of the actual interview. In addition, all interview partners received a copy of the original research proposal detailing the objectives and the methodology of the study translated into French or German. The first round of interviews was conducted between February 2009 and March 2010.

Written documents

The objective of the interviewing process was to obtain "factual" information about a police department's reforms rather than an interviewee's "subjective assessments" of implementation (cf. Quinton & Morris 2008, 22f.). The information obtained from the interviews with the police administrators was therefore complemented with additional information gleaned from written documents available for each city. Special care was taken to buttress interview information if possible with written documents dating from the time of reform. This was deemed necessary since the period of interest stretches back to the mid-1980s and human memories are inevitably fallible over such long time spans. The written documents reviewed included police strategy papers and internal memoranda as well as annual reports or books published by the police departments themselves. The study also relied on police external documents such as the minutes from contemporary debates in local parliament or government bodies, or academic texts such as master's theses or graduation papers, when these were available.

Follow-up interviews

All the intelligence gathered through the interviewing process and document reviews was assembled in a separate community policing implementation report for each city. (In the case of Zurich a separate special report was dedicated to each of two innovative community policing projects). The chapter structure of these community policing implementation reports mirrored the six evaluative dimensions. All five reports plus the two supplementary reports were written in the local language, i.e. either in French or in German. The original plan to conduct the interviews in the local language and subsequently to draft the reports in English had to be abandoned to accommodate the preference of some interview partners to correspond in their native tongue.

These draft reports were sent to the author's erstwhile interlocutors ahead of a second interview. These follow-up interviews lasted between two to four hours. This second round of interviews served to clarify outstanding issues from the first interview and to verify the accuracy of the chronology of events presented in the draft reports. The second round of interviews took place between March and April 2010. The complete community policing implementation reports for all five cities plus the two separate project reports were included as appendices to the final research report submitted to the SNSF in May 2010 (Kreis *et al.* 2010).

3.1.2 Evaluative Dimensions

To its critics, community policing has long "suffered from conceptual confusion in both research and practice." (Gianakis & Davis 1998, 486). According to Mastrofski (2006, 44), the lack of a commonly accepted operational definition has resulted in pointless debates over "what community policing 'really' means", when, in fact, it was this "all-things-to-all-people character" that explained some of the concept's enduring appeal with public leaders across the political spectrum. At the same time, these authors caution that the variability of community policing is so great as to defy efforts of evaluating its effect as a global strategy (Mastrofski 2006, 44f.; Weisburd & Eck 2004, 52).

Advocates commonly retort that community policing is not a "set of specific programs" but rather an organizational strategy that gives wide discretion to local residents and their police to set priorities and to decide the means to achieve them (Skogan 2006c, 27f.). That community policing has been implemented in a wide variety of ways is therefore neither cause for concern nor can it disguise the fact that there has been growing agreement as to what the concept stands for. According to these authors, community policing has three core elements: community engagement, problem solving, and organizational reform of the police. While Rosenbaum (1998, 7) spotted these elements as part of an emerging consensus in the literature over a decade ago, more recently a number of authors have explicitly defined them as the constituting elements of community policing (Bureau of Justice Assistance 1994; Fridell 2004; Maguire & Wells 2009; Skogan 2006c).

In practice, these three dimensions are inextricably intertwined and police departments must act on all three of them in order to mount an effective community policing strategy (Skogan 2006c, 28). This complexity, however, makes it difficult to define the analytical dimensions of a process evaluation and to select indicators without at least some overlap between them. This in part explains why no common specification of the evaluative dimensions of community policing has emerged, even though some authors have proposed templates (e.g. Greene 1998, 151-157; cf. Chapter 2.1.3).

The analytical dimensions of the current process evaluation of community policing in the five Swiss cities were drawn up along the outline of the original SNSF research proposal from March 2008. These were based on the different perspectives on community policing that the Police Executive Research Forum (1996) identified in the literature in the introduction to a compendium of different case studies of community policing in the United States and Canada. Key indicators were defined for each of the following six evaluative dimensions: "Organizational Strategy and Partnerships", "Decentralization and Deployment", "Broken Windows Approach", "Problem-solving", "Performance Appraisal Systems", and "Equity and Legitimacy".

The little information on community policing in Switzerland available at that time suggested that implementation had remained uneven across the five cities and that the operating community policing programs varied considerably from one place to another. It was therefore decided not to impose a strict working definition of community policing as a benchmark, against which to measure each police department's progress (cf. Maguire & Wells 2009, xviii). Instead, the goal was to study each city's community policing model and to document the organizational and strategic reforms the five police departments had implemented, using as a reference a list of indicators of the "few guiding precepts" of community policing (Kennedy & Moore 1995, 279) most commonly found in the literature. The process evaluation also would not disqualify programs or practices on the grounds that a department came short of the reformers' ideal of implementing all three core elements of community policing simultaneously. Instead, the study simply recorded all implemented reforms that qualify as community policing under any one of the evaluative dimensions (cf. Maguire & Wells 2009, xviii).

Organizational strategy and partnerships

At the outset, the process evaluation tracked the history of the adoption of community policing in each city, from the early pilot projects to the major organizational reforms the police departments undertook to implement the new operational strategy. The goal was to understand both the timing and impetus for reform, and to document changes in a police department's internal organization, work flows, or chain of command.

The process evaluation also looked at the investments each department made to train new and standing staff in the community policing strategy. This includes both changes to the curriculum of new police recruits as well as any continued training seminars for standing police officers. The study also tracked whether future community liaison officers or problem-solving specialists as well as future police cadres receive any specific training to prepare them for their functions in a police department formally dedicated to community policing.

Another key element of the community policing strategy are strategic partnerships, which are vital to both the information gathering and problem solving processes of the police. The study aimed to record the nature of the partnerships the five police departments had established with other branches of the local administration, business groups, neighborhood associations, and the general public. The first such strategic partnership was typically established between a city's police, social welfare, and public health departments to implement the federal drug policy during the 1990s (Ernst *et al.* 1999a; Ernst *et al.* 1999b). The process evaluation tracked how the police department had since expanded their network of strategic partnerships also to include the local park services, garbage collection, and the planning and urban development departments, before embracing neighborhood cultural centers and a city's schools. The process evaluation also looked at whether the five police departments had forged partnerships with private partners such as neighborhood associations or business groups, and how they were using traditional venues such as festivals or trade fairs as well as old and new media channels to reach out to the general public.

In the context of community policing implementation, it was of interest whether the police department had drafted a strategy paper to guide the reorganization process and, if so whether this had been distributed to the department's rank and file. All such strategy papers were analyzed for a department's stated strategic and operational objectives with regard to community policing (e.g. reduce crime, combat fear and disorder, forge partnerships, etc.) and to see if they made any references to criminological theory.

Decentralization and deployment

Decentralization is another pillar of the organizational adaptations required to make community policing work, and the current process evaluation thus tracked any major shifts in the spatial organization of the five police departments. Possible changes concern the layout of police districts, the opening or closing of neighborhood police stations, the permanency of the beat assignments of patrol officers (to allow them to become familiar with a beat's residents and their preoccupations), or the stationing of a department's generalist and specialist divisions across its jurisdiction.

Community engagement is a quintessential activity of any police department committed to community policing and the study aims to assess whether the five police departments practiced a generalist or specialist model of community outreach, or any "hybrid" of the two. Under a generalist model, all sworn police officers are supposed to engage in community outreach and problem-solving activities, whereas under the specialist model all but the most common such tasks are the prerogative of a few specially trained officers (Maguire & Gantley 2009b).

Given the eminent role of the traditional police officer on foot patrol during the nascent community policing movement, the study also tracked any sign of a renaissance of this police tactic in Swiss cities and recorded other novel forms of police patrol.

Broken windows approach

The process evaluation tracked whether a city's authorities consider signs of physical disorder such as littering and graffiti as serious nuisances to neighborhood livability and have adopted measures to clean-up streets, parks, and public places. As such schemes typically are multi-stakeholder initiatives, the process evaluation aimed to document their organizational structure, degree of institutionalization, and the role and involvement of the police. A further aspect was whether the sphere of action was limited to public buildings and places or encompassed private property as well. Besides measures to tackle physical disorder, the study also noted whether city authorities had taken steps to alleviate the effects of social disorder.

Problem solving

The process evaluation aimed to establish who was being instructed in the problem-solving method SARA and who would apply it in day-to-day operations. The study also tracked if police departments have established a central analysis unit that monitors and analyzes operational data to identify crime hotspots and to provide intelligence to formulate policing strategies. In this context it was of interest how these units were making use of Geographic Information Systems (GIS) to visualize the patterns of criminal activity and study longer-term trends.

The study also tracked how police departments had fostered cooperation between different branches of the local administration to boost their joint ability to respond to security threats. Modes of cooperation were analyzed with regard to the handling of the local drug markets, commercial sex industry, and youth violence, as well as geographically confined neighborhood regeneration projects.

The process evaluation also noted renewed efforts in the area of crime prevention. Such investments could come in the form of new infrastructure such as mobile police stations, as strategic partnerships with business owners to coordinate the fight against shoplifting, or as ad hoc campaigns to offer crime prevention advice to the general public or vulnerable groups.

Performance appraisal systems

Regular evaluation is critical to the viability of community policing and the process evaluation thus tracked how the five police departments monitor performance and evaluate the impact of policing strategies. The study distinguishes between monitoring the overall performance and the evaluation of specific crime prevention projects. The study also noticed if the community is being heard as part of the police performance assessment through regular surveys, polling the resident population on local levels of victimization, fear of crime, neighborhood disorder, and public satisfaction with the police.

Equity and legitimacy

Finally, the process evaluation focused on measures adopted by the police departments to grant victims access to adequate care and minimize suffering, noticing measures that go beyond the stipulations of the 1993 Federal Victims Assistance Act.

Police recruits nowadays are instructed in ethics and the human rights aspects of policing as part of the curriculum preparing them for the federal examination as police officers. The process evaluation noted the existence of internal or external review bodies and how they handled complaints of alleged police misconduct. Moreover, it established whether the police departments had a charter or ethics code encoding their principles of professional conduct and what mechanisms had been put in place to uphold it.

Finally, the process evaluation examined whether police officers receive specialized training to raise their awareness of cultural differences in their interactions with members of minority groups. The study also tracked (nascent) efforts to establish links with ethnic or religious communities and inquired about affirmative action programs or hiring quotas police departments may have adopted to obtain a more diverse police force.

3.1.3 Data Analysis

The qualitative information contained in the community policing implementation reports on each city was coded for subsequent analysis. To this end, a list of indicators has been specified for each of the six evaluative dimensions. A scorecard is used for each city to track the process of implementation over the period from 1985 to 2010 in intervals of five years. The idea of the community policing scorecard as a data matrix with police activity on one dimension and time on the other draws heavily on the concept of the "Calendar of Action" that Crawley & Hope (2003, 11) developed as a tool to monitor implementation progress for their evaluation of crime prevention projects in the United Kingdom.

The scorecard simply records for each point in time whether a given community policing element had been implemented. If a set of policing activities was confined to a specific area of a city, the corresponding ZIP code number was recorded instead. The main rationale for this rather straightforward coding procedure was to record the implementation data at a level of detail that can reliably be measured *ex post facto* and is appropriate for analysis together with the data available to measure treatment outcome. The community policing scorecards for each of the five cities are added as appendices to the current study (cf. Appendix B).

Between November and December 2011, the author met his erstwhile interview partners from the five police departments a third time to discuss the scorecard in order to ensure a high degree of inter-city coherence in the coding of the implementation data.[8] After this process had been completed, the finalized scorecard constituted the data base for further analyses of the implementation process. In a first phase, descriptive, cross-sectional and longitudinal analyses served to unearth notable parallels and differences in the history of community policing implementation between the five cities. In light of this analytical narrative, it was possible to validate or reject the working hypotheses on community policing implementation for the case of Switzerland's major urban areas.

In a second phase, the scorecard data was aggregated to form index scales measuring the level of community policing implementation. Each cell of the scorecard matrix was coded as 1 if the policy element in question had been implemented at that moment in time and 0 otherwise. This coding procedure is comparable to the approach taken by Quinton & Morris (2008, 25) who in the context of their evaluation of neighborhood policing in Britain scored police forces and basic command units as "either zero or one for each sweep depending on whether they had implemented the minimum requirement." (cf. Chapter 2.2.1). For the current evaluation, the scorecard points were added up for each of the six analytical dimensions, and this sum was normalized against the maximum value, i.e. the number of indicators per evaluative dimension. This procedure reduced the number of data points substantially, and the normalized index figures were subsequently displayed as a parallel plot in order to visualize a maximum amount of information in a single chart. Using separate colors for different years, the parallel plot provides a detailed overview of the process of community policing implementation over all six evaluative dimensions over time.

The summing up, normalization, and visualization of the implementation data was performed using the R language for statistical computing (`http://cran.r-project.org/`; R Development Core Team 2011). The parallel plot displays a panel for each city, with colored lines indicating the status of community policing implementation at the six points on the time scale (cf. Fig. 4.1). The `parallel` plotting function in the 'lattice' package (Sarkar 2008) performs both the normalization and plotting of the index scales automatically.

```
> parallel(~temp[,9:1]|area,data=temp,groups=temp$year,...)
```

The purpose of this quantitative analysis of the process evaluation data is not only to gain a better understanding of an otherwise bewildering array of data. The results, foremost the parallel

[8]The erstwhile interview partners from the Zurich and Lausanne municipal police departments having retired in the meantime, the scorecard was discussed with the Deputy Head of the Crime Prevention Unit in Zurich and the Heads of the Patrol Division and of the Community Policing Unit in Lausanne.

plot, also serve to check the plausibility that significant shifts in the outcome variables over time can be linked to progress made in the implementation of community policing. The basic idea is that the parallel plot helps an evaluator to spot the biggest gaps in the "dosage" of community policing over time and across different cities, and thus to identify the most interesting constellations for the impact evaluation. It is generally good scientific practice in evaluation research to "compare two treatments that are as different as possible" to prove any effect, regardless of whether the design is a controlled experiment or an observational study (Rosenbaum 2010, 125).

Earlier plans for more sophisticated quantitative analyses of the developed index scales, especially the idea to combine the indices of the six evaluative dimensions into an overall index of community policing implementation, were dropped. This decision appeared justified both on theoretical and practical grounds: one the one hand, since the author of this study did all the coding himself, inter-rater reliability cannot adequately be assessed. On the other hand, community policing theory is not sufficiently developed as to the impact mechanisms to guide the construction of such an overall index.

3.2 Impact Evaluation

3.2.1 Data

Survey data

Citizen surveys have been widely used in community policing evaluation research because they bypass the two major pitfalls that undermine the validity of official crime statistics and enable researchers better to evaluate claims about the effect of policing on crime and fear of crime (Skogan 1999, 37).

For the same reasons, the Swiss Crime Survey (SCS) offers a good opportunity for community policing impact evaluation.[9] The SCS has polled residents of the five major urban areas under study here over a period of two decades and tracked several key indicators of community policing impact. The surveys provide data on levels of crime, fear of crime, neighborhood disorder, social cohesion, and popular satisfaction with the police. The beginning of the SCS predates the introduction of community policing efforts by the majority of Swiss municipal and cantonal police departments, allowing a longitudinal impact evaluation. Furthermore, the SCS data sets have always retained the ZIP code of a respondent's address, which can be used as a geo-reference for neighborhood-level analysis. The questionnaire also inquires basic demographic data, so that disparities among respondents of different age, gender, level of education, or place of birth can easily be detected.

In 1984 the University of Lausanne conducted the first national victimization survey in Switzerland, for which some 3,000 persons in the French-speaking region were randomly selected and interviewed by telephone. In 1987, the university repeated the survey with a sample of 3,500 individuals in the German- and Italian-speaking regions. In 1989, 1996, 2000, and 2005, the University of Lausanne administered four additional victimization polls as part of the International Crime Victimization Surveys (ICVS) with samples of 1,000; 1,000; 4,200; and 3,900 respondents, respectively. In 1998, the University of Lausanne autonomously conducted one more national victimization poll with a sample of some 3,000 respondents (Killias *et al.* 2007, 13f.).[10]

A salient feature of the administration of the SCS in the most recent years has been that a number of cantonal and municipal police departments financed extra samples of interviews in their jurisdiction. In 2005, these were the police forces of the cantons of Fribourg, Vaud, Bern, and St. Gallen as well as the cities of Lausanne and Zurich. In 2000, in addition to these, the police of

[9]This subsection is based on the author's unpublished graduation paper for the Master of Advanced Studies MAS *Urbanisme durable* (Kreis 2009).

[10]The data from the most recent edition of the Swiss Crime Survey, polling for which took place in early summer of 2011, became available too late to be included in the current study.

the cantons of Ticino, Basel-City, and Geneva as well as the City of Bern provided extra funds to augment the number of interviews taken on their territory (Killias *et al.* 2007, 17).

These cash infusions, which boosted the samples in the cities of interest here to between 200 and 500 interviews, first made possible the neighborhood-level analyses undertaken by the present study. However, none of the five cantonal or municipal police departments contributed consistently to the survey administration and refrained on at least one occasion from financing extra interviews. For those years in which a city's police department abstained, its sample size was usually too small for any meaningful neighborhood-level analyses, which limits the number of valid cross-sectional and longitudinal comparisons considerably.

The selection of data samples available for individual cities is reduced further because the data base of the first victimization survey of the Romandie in 1984 has been lost. Also, the comparatively small samples of 1,000 interviews for the whole of Switzerland of the first two ICVS surveys in 1989 and 1996 preclude any analyses of a sub-sample of individual cities. This leaves the data sets from 1987 for the German- and Italian-speaking regions as well as the national data sets for 1998, 2000, and 2005. Table 3.1 indicates the total sample size for each of the five cities under study by survey year.

Table 3.1 Total SCS sample size by city and survey year

	1987	**1998**	**2000**	**2005**
Basel city/canton	148/163	40/42	415/466	42/50
Bern	123	43	194	81
Geneva city/canton	n/a	329/567	217/522	37/99
Lausanne	n/a	20	200	244
Zurich	283	507	90	482
Switzerland Total	3509	3041	4234	3898

In order to allow for valid longitudinal analyses, a victimization survey must be based on a representative sample of the overall population and apply the same crime definitions and questionnaire formats consistently over time. Beginning in 1989, the authors of the Swiss victimization surveys adopted the ICVS methodology and have applied it consistently in subsequent years, apart from minor modifications of individual items of the survey questionnaire. The overall structure of the questionnaire has remained unaltered since then, polling interviewees first for any victimization experiences over the previous five years and – in case of an affirmative answer – probing for further details on the circumstances of the incident. Also, the computer-assisted telephone interview (CATI) method has been used throughout the different sweeps of the Swiss Crime Survey (Killias *et al.* 2007, 15f.).

Police crime statistics

Desirable though it would be to use Swiss Crime Survey data on victimization experiences as indicators of the prevailing crime rates, the present study has to rely on official police crime records to gauge the level of criminal activity across the five cities. The SCS was not originally conceived for analysis of urban neighborhoods, and given the low base rates of many crimes, the samples for the five cities of interest here are too small to serve as reliable indicators of the crime rate at the level of neighborhoods or ZIP code districts.

For all five cities, the official crime records available for analysis are time-series of the total number of incidents per year by type of criminal infraction, covering a period of at least eight years. In every case except for Geneva, the beginning of the time-series precedes or coincides with the respective police department's first major organizational reform to implement community policing. For all the cities except for Lausanne, the data sets contain information on violations of both the Federal Penal Code and the Federal Law on Narcotic Substances. The unit of analysis varies from one city to another, however. Table 3.2 summarizes the availability of official police crime records for the present study.

Table 3.2 Official police crime records by city

	Basel	**Bern**	**Geneva**	**Lausanne**	**Zurich**
Infractions	Penal Code, narcotics crime	Penal Code, narcotics crime	Penal Code, narcotics crime	Penal Code	Penal Code, narcotics crime
Time period	1997-2006	2000-2010	1998-2005	2001-2008	1996-2008
Unit of analysis	Canton (n=1)	ZIP code districts (n=13)	OFS census tracts (n=60)	OFS census tracts (n=17)	Administrative districts (n=12)

For the Canton of Basel-City, official crime statistics are only available for the entire canton. The Cantonal Prosecutor's Office, which in Basel gathers the official crime records, has provided annual summary statistics only at the cantonal level for the period between 1997 and 2006. However, due to changes in the counting method starting in 2005, only the data for the 1997-2004 period are used by this study.

Like in Basel, the police in Bern do not yet routinely produce crime statistics at the neighborhood level. However, for the purposes of this study, the crime analysis unit of the Bern Cantonal Police Department in December 2010 ran a special query of its crime record data base to come up with annual summary statistics for the city's thirteen geographic ZIP code districts (by using a data base entry's date and ZIP code number as sorting variables). These neighborhood-level data account for criminal activity during the period from 2000 to 2010, albeit imperfectly. Alas, the validity of the data from the earlier years is compromised because infractions are automatically deleted from the police data base after their statute of limitations has expired, and less serious offenses are deleted more rapidly. In order to keep this systematic bias towards more serious crime in the early years of the data base in check, the ZIP code-level data have been compared with the officially published annual crime data for the entire city of Bern. These city-level crime records were published each year and hence do not suffer from the same distortion. They have been obtained from the Bern City Statistics Office for the years from 1997 to 2009.

The Geneva cantonal police provided the most detailed official crime statistics of all five cities for the present study. The data cover both infractions of the Penal Code and the Federal Law on Narcotic Substances for the entire territory of the canton for the period from 1998-2005. The data were aggregated by calendar year at the level of OFS census tracts to make the Geneva data comparable with the crime data available for the other cities in the current study.

In October 2010, the Vaud Cantonal Statistics Office (SCRIS) published a report on the distribution of the criminal activity across Lausanne's neighborhoods and the surrounding urban area (Roh 2010). For this report, the SCRIS' authors had analyzed police crime records using the same data categories as the Vaud Cantonal Police Department had employed in its official crime statistics CRIPOL until 2008. This special analysis was conducted on the basis of Lausanne's seventeen OFS census tracts and covers the period from 2001 to 2008. The data include infractions of the Penal Code but no instances of narcotics crime. The SCRIS has made its data base available for the present study.

Finally, the Zurich Cantonal Police Department publishes crime data for the City of Zurich at the level of administrative districts (Stadtkreise) in its yearly official crime records KRISTA. These data are available for analysis for the 1996-2008 period. The City of Zurich counts twelve administrative districts, which are a slightly bigger unit of analysis than OFS census tracts or ZIP code districts (Each administrative district encompasses between two and four OFS census tracts).

Census statistics

The data on the composition of the neighborhood population and housing structure have been gathered from the OFS population census counts from 1990 and 2000 and the housing census count from 1990, respectively, which were available from the University of Lausanne's GIS database in raster format for analysis in GIS software. The raster data were aggregated to the level of ZIP code districts or administrative units in order to obtain summary statistics for the geographic units of analysis of the subsequent multivariate analyses.

However, in the OFS population census data in raster format, any absolute data value of 0, 1, or 2 is automatically substituted by a value of 3 for reasons of data protection. This measure, alas, can grossly distort summary statistics of the data, especially if it is highly disaggregated such as the population count by age group, which is subdivided into no fewer than 18 5-year classes. Upon request, the Federal Office of Statistics thus conducted a special query of the 1990 and 2000 population census in order to provide accurate summary data on the socio-demographic composition and socio-economic status of the population at the level of census tracts of the five cities.

3.2.2 Indicators

Crime

Most evaluations of policing strategies use either official crime reports or victimization surveys of the public to measure crime (Sherman & Eck 2002, 301). Victimization surveys avoid two major pitfalls that seriously undermine the validity of official crime records: citizen reporting and police recording practices. Victimization surveys repeatedly shed light on the fact that a majority of crime incidents are never reported to the police and that citizen reporting rates differ markedly by type of crime and by population group: serious property offenses tend to be reported more often, whereas the rates for assaults or robberies are much lower. Men, the working class, and the young are less likely to notify the police if they are victimized (Skogan 1999, 37f.). This problem of a systematic bias in official crime records resulting from citizen non-reporting is exacerbated further by police recording practices. The number of reported and founded incidents are never identical as the police must decide which reported incidents actually get recorded. Mostly they act in good faith but at times police officers have been found to respond to pressure to "keep the crime count down", shifting reported incidents between categories, "mostly to downgrade them", or dismissing them as unfounded without a proper investigation (Skogan 1999, 38).

Victimization surveys circumvent the problems of citizen reporting and police recording practices but are not without their own shortcomings. Apart from the problems of sampling bias and non-response bias, which are inherent to survey research in general, victimization surveys tend to underestimate certain types of offences such as assaults and domestic violence and are virtually useless for measuring drug offenses and non-residential vandalism. Moreover, surveys are ill-suited accurately to measure levels of repeat victimization, which have a disproportionate effect on the overall crime rate, especially in high-crime areas. Finally, most victimization surveys just sample households in the study area, leaving out schools and business or commercial establishments that experience even higher victimization rates than residential homes (Skogan 1994, 178; Skogan 1999, 39f.).

An effective strategy to countervail the various hazards of measuring victimization rates is data triangulation, which relies on multiple indicators of crime. If several disparate measures point in the same direction, the conclusion that crime has risen or dropped appears more robust, even if the individual indicators are flawed (Skogan 1999, 41). Alas, data triangulation is not a possibility for the present study. As previously remarked, the small sample size of the Swiss Crime Survey paired with the low base rates of many crimes precludes analysis of victimization data at the neighborhood level.

At the national level, however, the authors of the SCS previously compared survey victimization rates with official police records for different types of offenses over the period from 1984 to 2005 (Haymoz *et al.* 2009, 132-137; Killias *et al.* 2007, 103-114). Regarding property offenses, the trends by and large followed the same pattern, even though absolute levels of recorded crimes differed markedly, notably concerning theft of motorbikes and bicycles. Things looked somewhat differently for offenses against persons. With regard to robbery, survey and police data indicated comparable trends, but the actual number of incidents registered by the police was extremely low compared to survey rates. Regarding assaults and threats, police figures were not only equally low when compared to survey figures, the trends also only remotely resembled each other, both indicating a global increase between 1984 and 2005. Accounting for these discrepancies, the authors explain that, according to police sources, in the context of the open-air drug markets that had appeared in Swiss cities during the first half of the 1990s, police departments applied stricter criteria in recording offenses against persons, which for this period may thus be artificially low (Haymoz *et al.* 2009, 137).

Killias *et al.* (2007, 114-116) also looked at changes in the crime reporting rate (i.e. the percentage of victimization experiences actually being reported to the police) for these offenses over the same period. They found that crime reporting rates plunged (from 88.6% to 31.3%) for robbery, whereas they dropped only slightly for property offenses (from around 90%) and remained more or less stable for assaults and threats (hovering around 30%).

Apart from the vagaries of citizen reporting and police recording, using official crime records as a gauge of neighborhood-level crime rates in Switzerland is fraught with still more difficulties. Until most recently, police crime records were not directly comparable across cantons or cities because recording practices varied considerably from one canton to another (Killias *et al.* 2007, 9). For example, in cases of multiple infractions by the same perpetrator, some police departments counted every infraction as a separate incident, whereas others only recorded the most serious offense. Equally problematic, whereas some cantons compiled their statistics by the number of perpetrators, others counted incidents by the number of victims of crime.

Such differences in police recording practices undermine the validity of cross-sectional analyses of the original data. In order to circumvent this problem and make the data comparable, several steps were taken to standardize police statistics across cities. First, the local counting methods for ten different criminal offenses were harmonized as best possible. Switzerland has a Federal Penal Code, so the definitions of criminal offenses do not vary per se. However, when compiling their crime statistics, police departments summarize the number of violations of individual articles under separate category headings, the definition of which vary slightly from one city to another. The list of these ten offenses plus their definition in the local context are included in Table 3.3.

In a second step, standardized crime rates were calculated for the cross-sectional comparisons undertaken by the current study. The standardization method varied for the city-level and the neighborhood-level analyses. For the long-term trend analysis at the city level, the crime rate for each type of offense was calculated per 1000 inhabitants. These yearly crime rates were normalized against the crime rate of the first year for which data were available for each city according to Equation (3.1), where C_t stands for the total number of recorded incidents in year t. The crime rate was set at 100 for the first year for which data was available in order to visualize the trend of the eight criminal offenses over subsequent years.[11]

$$cr_t = \frac{\frac{C_t}{Pop} * 1000}{\frac{C_{t_0}}{Pop} * 1000} * 100 = \frac{C_t}{C_{t_0}} * 100 \qquad t = 1, 2, \ldots, T \tag{3.1}$$

[11] Initially, the definition according to the local police statistics of *ten* types of criminal offenses have been harmonized across the five cities. However, sexual assaults and narcotic crime were dropped from further analysis (cf. Chapter 6.1.2).

Table 3.3 Types of criminal offenses by city

	Basel	**Bern**	**Geneva**	**Lausanne**	**Zurich**
Homicide	Tötungsdelikt	Tötungsdelikt	Meurtre, assassinat, meurtre passionnel	Meurtre, assassinat	Vorsätzliche Tötung, Mord, Totschlag
Assault	Körperverletzung, Tätlichkeit	Tätlichkeiten, Gefährdung des Lebens	Lesions corporelles graves/simples, voies de fait	Lesions corporelles graves/simples, voies de fait	Schwere/Einfache Körperverletzung, Tätlichkeiten
Burglary	Einschleichediebstahl, Einbruchdiebstahl	Einschleichediebstahl, Einbruchdiebstahl	Vol par effraction, vol par introduction furtive	Cambriolage habitation	Einbruchdiebstahl
Car theft	Wegnahme von Lastwagen, Personenwagen, Motorrad	Fahrzeugdiebstahl, Entwendung zum Gebrauch, Motorraddiebstahl	Vol d'un véhicule	Vol camion, moto, voiture	Fahrzeugdiebstahl, Entwendung zum Gebrauch
Robbery	Raub, Entreissdiebstahl	Raub, Entreissediebstahl, einfacher Diebstahl	Vol à l'arraché, brigandage	Vol à l'arraché, brigandage	Raub, Entreissdiebstahl
Vandalism	Sachbeschädigungen	Sachbeschädigung	Dommages à la propriété	Dommages à la propriété	Sachbeschädigung
Extortion	Erpressung	Erpressung	Chantage, extorsion	Extorsion, chantage, racket	Erpressung
Threat	Delikte gegen die Freiheit	Drohung, Nötigung	Menace, contrainte	Menace, contrainte	Drohung, Nötigung
Sexual assault	Vergewaltigung	Sexuelle Nötigung, Vergewaltigung	Contrainte sexuelle, viol	Contrainte sexuelle, viol	Sexuelle Nötigung, Vergewaltigung
Narcotic	Widerhandlung gegen das Betäubungsmittelgesetz	Widerhandlung gegen das Betäubungsmittelgesetz	Infraction loi fédérale sur les stupéfiants	n/a	Widerhandlung gegen das Betäubungsmittelgesetz

The methodology employed to calculate standardized neighborhood-level crime rates for the four cities for which such data were available is borrowed from the field of epidemiology, where rate standardization is a common mechanism to make disease incidence rates comparable across different study populations. Under so-called "indirect rate standardization", public health analysts estimate the expected number of incidents under the assumption that members of the study population contract a disease at the same rate as the standard population (Waller & Gotway 2004, 11-17).

In close analogy, the *expected number of crime incidents* $\mathrm{E}(C_{in})$ is the number of infractions of type i one would observe for neighborhood n if crime was equally distributed across a given city. In a city with N neighborhoods, let C_{in} denote the number of recorded incidents of type i in neighborhood n and Pop_n the resident population of neighborhood n. Then $\mathrm{E}(C_{in})$ is calculated as the city's average crime rate – defined as the total number of recorded incidents divided by the total city population – multiplied by the neighborhood population:

$$\mathrm{E}(C_{in}) = \frac{\sum_{n=1}^{N} C_{in}}{\sum_{n=1}^{N} Pop_n} * Pop_n \qquad n = 1, \ldots, N \qquad i = 1, \ldots, I \tag{3.2}$$

for each crime category $i = 1, \ldots, I$. Since crime incidents tend to cluster spatially, the observed number of incidents $\mathrm{O}(C_{in})$ will differ from the expected number across a city's neighborhoods. The number of observed incidents is divided by the number of expected incidents in order to calculate the standardized neighborhood crime rate (Equation 3.3):

$$cr_{in} = \frac{\mathrm{O}(C_{in})}{\mathrm{E}(C_{in})} \qquad n = 1, \ldots, N \qquad i = 1, \ldots, I \tag{3.3}$$

This standardized crime rate varies between zero and infinity. It is a unit-less index value that no longer depends on the different counting methods of the cantonal police crime records. The standardized crime rate expresses the neighborhood crime rate as a multiple of the city average, i.e. a value of, say, 1.53, indicates that the number of incidents per resident is 1.53 times higher than the city average; a standardized crime rate of 0.37 suggests that a neighborhood suffered only 0.37 times as many criminal infractions per head as the city overall. A standardized crime rate of 1 would imply that a neighborhood crime rate is exactly the same as for the city as a whole.[12]

Since the differences in police recording practices preclude direct comparisons of the original data, no valid information is lost in this process of rate standardization. The standardized crime rates still reveal, which neighborhood is relatively more affected by a given category of crime and, more pertinently, allow meaningful comparisons of the spatio-temporal patterns of criminal activity across the four cities under study.

As a first preliminary analysis, the standardized neighborhood-level crime rates were plotted against a time line. Interestingly, the emerging patterns were virtually identical for all four cities. Basically, the neighborhood-level trend lines fluctuated considerably for low-frequency offences such as homicide or extortion. By contrast, for offenses committed more frequently such as burglaries, assaults, or acts of vandalism, the trend line of the standardized crime rates was virtually flat. In each city, there are between one and three neighborhoods that show much higher victimization rates than the city average. Over time, though, there was little sign of major increases or decreases of individual

[12] The expected number of crime incidents was calculated by multiplying a city's average crime rate with the neighborhood resident population. Using the resident population as the denominator to compute standardized rates may not always be ideal, though. For instance, crime rates in business or commercial districts with relatively few inhabitants may be artificially inflated. Moreover, for certain types of crimes, there may be a more appropriate denominator (e.g. number of households for burglaries). Still, finding the best denominator is not always straightforward (cf. Killias *et al.* 2011, 83-86). More importantly, if the goal is to compare rates for different types of offenses for a given area, some meaningful common denominator has to be found, which is why criminologists most commonly use the number of inhabitants.

neighborhood-level crime rates. Plots of the standardized, neighborhood-level crime rates of the eight different offenses are added as appendices to the current study (cf. Appendix C).[13]

Since the long-term trends proved rather stable, it was decided to aggregate the number of incidents for multiple years and to compute standardized crime rates in order to smooth the erratic fluctuations of the less frequent offenses. The data were thus aggregated for three separate periods: 1998-2001, 2002-2005, and 2006-2008. Given the availability of the crime data and the need to obtain reliable crime rates for analysis together with the survey data, these time periods appeared a logical choice.

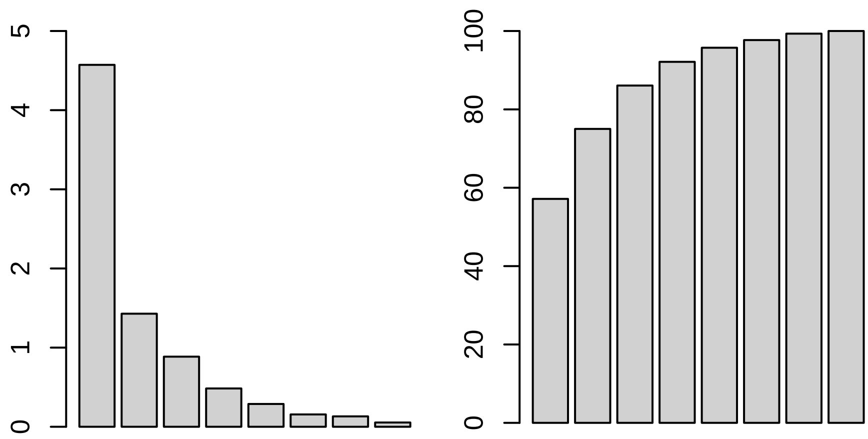


Figure 3.1 Principal components analysis of the police crime statistics. Eigenvalues (left) and percentage of cumulative variance explained (right).

As a second step, a correlation matrix for the neighborhood-level crime data aggregated for the 1998-2001 period was calculated. This analysis revealed that the standardized crime rates of several types of offenses had very high correlation coefficients of 0.9 or higher. The crime data were thus subjected to a principal components analysis, which produced only two components with eigenvalues greater than one, which jointly accounted for over 75 % of total variance (Fig. 3.1). It was therefore decided to substitute the standardized crime rates of the eight criminal offenses of individual neighborhoods with their component scores on the first two principal components.

Fear of crime

Traditionally, researchers have distinguished four different dimensions of the concept of fear of crime: *concern* about crime, *risk* of victimization, *threat* of crime, and *behavioral* changes in reaction to crime. Discriminating between these four definitions of fear is crucial because they require different operationalizations and the use of one definition versus another can substantially alter the conclusions from the research (Skogan 1999, 47).

The first dimension of fear, concern about crime, refers to the extent to which people fret about crime and disorder in general as serious problems facing their community or society. Perceived risk, by contrast, taps into more cerebral judgments about the likelihood of actually becoming a victim of crime. The third cognitive concept, threat of crime, focuses on “the potential for harm people feel crime holds for them.” Threat of crime is distinct from concern and perceived risk; people may

[13] The current study used the data from the OFS census count from 2000 to calculate the neighborhood resident population. Relying on the decennial census count misses out on the gradual changes of the resident population over subsequent years, so using annual population figures would hence have been more accurate. However, such data would first have to be compiled from local statistics bureaus (contrary to the national census statistics). Moreover, a quick glance at the standardized neighborhood-level crime rates advises against such a step. Firstly, the logarithmic trend lines are virtually flat for all high volume crimes. Secondly, even though the victimization rates in the worst afflicted downtown areas may be artificially inflated, they are several times the rate of low-crime neighborhoods, so that using a different denominator would not substantially alter the results of the analysis.

avoid activities that would expose them to a heightened risk of victimization and hence reduce the likelihood of actually being victimized, but they may still consider the potential for harm as serious if they exposed themselves to risk. The fourth, behavioral dimension of fear focuses on the changes people adopt in their daytime or nighttime activities to avoid crime (Skogan 1999, 47f.).

From its inception, the questionnaire of the Swiss Crime Survey included several items measuring fear of crime.[14] It has consistently included at least one questionnaire item covering *risk* of victimization, *threat* of crime, and *behavioral* changes in response to crime. In 1984/87 and 2005 the questionnaire also included an item probing for *concern* about crime. The wording of these questions and the answer categories have evolved over the years, most notably between the first national poll in 1984/87 and the first ICVS survey in 1989. Since then question formats have changed little and remained unaltered since 1998.

The question measuring threat of crime asked respondents: "How safe do you feel walking alone in your neighborhood after 10 pm? Do you feel very safe, quite safe, somewhat unsafe, or very unsafe?" Two additional answer categories not read by the interviewer recorded if respondents did not go out at all at night either because of worries about their safety or for some other motive not related to security. In earlier versions of the questionnaire, the wording of the question did not specify the time "after 10pm" but simply asked "in the evening". In the first poll from 1984/87, people were first asked if they do go out at night occasionally, which was followed up by a question asking if there were any places within one kilometer from their home where they are or would be afraid to be alone at night depending on their answer to the preceding question.

The wording of the question about risk of victimization since 1998 has been as follows: "What would you say are the chances that over the next twelve months someone will try to break into your home? Do you think this is very likely, quite likely, rather unlikely, or very unlikely?" In 1984/87 the risk question suggested a list of four more or less serious crimes including burglary and asked respondents to name the offense they thought held the greatest risk to them. Due to the mutually exclusive answer categories, this earlier question cannot be compared with the item from subsequent years.

The item measuring behavioral changes in response to crime asked respondents: "Walking alone in your neighborhood after 10 pm do you stay away from certain streets, areas, or people to avoid crime?" The answer is in yes/no-format with the two additional answer categories for respondents who do not go out at all at night as with the item measuring threat.

The SCS questions on the fear of crime are similar to the questionnaire items from the U.S. National Crime Survey and the General Social Survey. LaGrange & Ferraro (1989) criticized this question format on the grounds that it overestimates levels of fear of crime of women and the elderly compared to alternative instruments. More recently, Farrall & Gadd (2004) proposed a new question format in an attempt not just to measure the prevalence of fear of crime but also to assess the frequency and severity of fearful episodes. The merits of this debate are beyond the scope of this study, which due to its *ex post* nature is obligated to work with standing data sources. While the indicator of fear as the threat of crime plays a pivotal role in the development of the study design, namely the selection of the optimum number of neighborhood clusters (cf. Chapter 3.2.5), its ultimate use is as an outcome measure of the impact of community policing on neighborhood residents. From this perspective, the chief concern appears to be that the questionnaire format and answer categories have not changed between different sweeps of the SCS, which has been the case for the crucial 1998-2005 period.

Disorder

Even though the concept of disorder may be hard to define, there is surprisingly little disagreement about it in theory. Skogan (1990, 4), who coined the term, distinguished two general classes of

[14]The paragraphs describing the SCS survey items were taken from the author's unpublished graduation paper for the Master of Advanced Studies MAS *Urbanisme durable* (Kreis 2009).

disorder – social and physical – which have been widely adopted. According to this definition, social disorder is a matter of improper public behavior. This includes loitering bands of teenagers, panhandling, public drunkenness, prostitution, and open drug use. Physical disorder, by contrast, refers to "visual signs of negligence and unchecked decay": junk and trash laying around in parks or vacant lots, decaying or boarded-up buildings, abandoned cars, graffiti, and other forms of vandalism (Skogan 1999, 42).

In spite of the theoretical consensus, *measuring* disorder is by no means an easy task. As the rules on seemly conduct in public places are for the most part unwritten, the number of norms that may be flouted is difficult to circumscribe (Skogan 1990, 5). Moreover, since most disorderly activities are not unlawful, they are not recorded by the police or the authorities. As a consequence, evaluators are left to their own devices to come up with measures of their prevalence in the study area. Two common approaches to mount to the challenge are field observations by trained observers and surveys of the resident population (Skogan 1999, 42-45).

The phenomenon of disorder was first introduced in the SCS questionnaire in 1998.[15] Respondents were asked whether close to their home there were any graffiti on the walls, lots of rubbish lying around, or groups of disreputable people loitering on the streets. That year, it was simply recorded whether respondents answered in the affirmative. In 2000 and 2005, the same item featured in the questionnaire again, but the answer categories now distinguished between graffiti, litter, or groups of disorderly people, allowing respondents to answer the question more specifically. For the data sets from 2000 and 2005, the variable capturing disorder was recoded into two dummy variables indicating the presence of signs of physical or social disorder in the neighborhood, respectively. The former encompasses instances of graffiti or litter, whereas the latter captures loitering groups of disreputable people.

Until 1998, the SCS questionnaire also included two items measuring the levels of social cohesion and trust among neighbors. The wording of the survey item probing for social cohesion remained unaltered between 1987 and 1998, asking respondents whether in their neighborhood people tended to look after each other or rather minded their own business. There were three answer categories: (a) "People around here for the most part help each other out"; (b) "Around here it is every man for himself"; or (c) "Something in between". The second question measured trust among neighbors, namely if respondents asked their neighbors to keep an eye on their home during a period of absence. Between 1984/87 and 1998 the wording of this second question essentially remained the same; however, the answer categories changed substantially between the two polls, making direct comparisons fraught with difficulties and the results hard to interpret.

Both items measuring social cohesion are important from a theoretical perspective, however, because they serve as proxies for informal social control among neighbors, which plays a central role in the spiral of urban decay as theorized by the broken-windows-hypothesis. Social cohesion is also a key element of rival social disorganization theories of neighborhood development that have found the combined measure of informal social control and cohesion and trust a robust predictor of lower rates of violence across neighborhoods (Sampson *et al.* 1997; Sampson & Raudenbush 1999). Alas, 1998 was the only year for which the SCS survey questionnaire contained a measure of both disorder on the one hand and social cohesion and trust among neighbors on the other, precluding any comparative trend analyses of these concepts with measures of fear of crime or neighborhood-level rates of criminal activity.

The question which method is more appropriate to measure neighborhood disorder – systematic observation or surveys – is still the subject of debate. Observational measures of disorder must ensure acceptable levels of inter-observer agreement, which is not easy to achieve, and they are less suited to track the more transitory phenomena of social disorder that "vary enormously by the time of the day, the day of the week, and the weather." Survey instruments, by contrast, run the risk that perceptual measures of disorder reflect respondents' biases rather than its true extent. If

[15]The paragraphs describing the SCS survey items were taken from the author's unpublished graduation paper for the Master of Advanced Studies MAS *Urbanisme durable* (Kreis 2009).

disorder is largely a matter of one's disposition, then the usefulness of surveys to gather data on neighborhood problems is limited (Skogan 1999, 44f.). Sampson & Raudenbush (1999, 606) made the case for systematic observation, arguing that fearful residents tend to see more disorder, even though they report on the same conditions as their less worried neighbors. By contrast, Taylor (1999, 75-78) observed that measures from systematic assessments of incivilities cannot readily be separated from the social structure of communities and neighborhood crime rates. More recent research has revealed that people's perceptions of disorder are shaped to a greater extent than had previously been thought by the neighborhood context such as the immigrant concentration (Sampson 2009, 17-19) or their daily exposure to disorder (Gau & Pratt 2010, 762). This suggests that the issue cannot be settled once and for all and that the appropriate measure has to be evaluated in light of the research questions and context of each study.

Again, the current study is obligated to work with existing data sources and can only touch on this debate. Nevertheless, a preliminary analysis revealed that measures of disorder do not covary much with demographic characteristics. Age, gender, education, and income levels had virtually no influence on perceptions of neighborhood disorder. The only significant difference is that foreign born residents appear to be less concerned about signs of physical disorder than respondents born in Switzerland. However, regarding signs of social disorder the two groups again see eye-to-eye. By contrast, regarding the indicators measuring fear of crime, the same exploratory analysis produced the usual, highly significant links with demographic characteristics (cf. Appendix D).

Public attitudes towards the police

In community policing philosophy, it is an article of faith that police performance cannot be assessed based on law-enforcement activities alone but must be evaluated in terms of levels of fear of crime, disorder, citizen satisfaction, and neighborhood livability (Moore & Kelling 1988, 13; Trojanowicz & Bucqueroux 1994). Public accountability is writ large in community policing. The development of instruments to measure public satisfaction with police services has received comparatively little regard, however (Skogan 1999, 51).

The SCS questionnaire has consistently included an item asking respondents to rate the overall quality of their local police.[16] Whereas the wording of the question "How good do you think the police in your area are in controlling crime?" has not altered much, the answer categories have shifted considerably over time. In 1984/87, respondents were asked to rate the police performance on a scale from zero to ten. In 1998, this scale was substituted by a dichotomized variable indicating whether the police were doing "a rather good job" or "a rather bad job" to be finally replaced, in 2000, by a four-level ordinal variable rating police performance in controlling crime as "very good", "quite good", "not so good", or "not good at all".

Beginning in 1998, the SCS authors introduced additional items to the survey questionnaire to measure public assessment of police performance. The first of these targeted the impartiality of the police asking respondents on a four-step scale whether they held their police to be "always impartial", "quite impartial", "sometimes partial", or "always partial". If respondents accused the police at least of occasional bias, a follow-up question asked for the suspected reasons of this discriminatory treatment. A second set of questions inquired about the quality of any personal encounter a respondent had had with the police over the previous five years. Respondents were first asked about the reason for the encounter and then whether they approved of how the police had treated them or dealt with their request. Respondents whom the encounter had left dissatisfied were asked for the reason(s) of their disappointment.

Survey instruments to measure popular satisfaction with police services have repeatedly been criticized. These criticisms range from cynical comments that such surveys amount to little more than a "public relations exercise" that produces little valuable information for the police (Sacco

[16] The paragraphs describing the SCS survey items were taken from the author's unpublished graduation paper for the Master of Advanced Studies MAS *Urbanisme durable* (Kreis 2009).

1998, 124), to more constructive discussions of satisfaction with the police as a multidimensional phenomena that simple measures fail to operationalize sufficiently (Sacco 1998, 125-127), to warnings that such measures are heavily influenced by psychological or demographic factors and hence unreliable (Reisig & Parks 2000, 626). Again, as for fear of crime and disorder, the current study has no alternative to the standing SCS survey items as indicators of popular satisfaction with police performance. A preliminary analysis found very little covariation with demographic characteristics, though, except for age, with the young being less favorably disposed towards the police than older generations (cf. Appendix D).

Neighborhood ecology

Community policing scholars have long argued for the police to be more embedded in the neighborhood context and have called for greater involvement of the community in tackling persistent problems of crime and disorder (Wilson & Kelling 1982; Greene 2000, 308). Decentralization of the police organization, paired with the assignment of individual officers to fixed geographical areas, it is argued, enable the police to address local problems that matter to specific communities (Skogan 2006c, 36-38).

In geography, geocomputational profiling or "geodemographics" refer to analytical procedures aptly to describe complex neighborhood contexts and to classify small urban areas into a manageable number of distinctive categories of neighborhood types. Geocomputational profiling typically involves clustering analyses of a series of environmental, socio-economic, and demographic indicators in order to detect similar patterns of urban structure or neighborhood composition. The successful application of geodemographic profiling techniques to the areas of criminal justice and policing research rests on the assumption that it is possible to classify urban areas into a number of distinctive neighborhood types that "differ predictably in their crime profile and policing environment." (Ashby & Longley 2005, 422f.).

The selection of indicators included in the clustering analysis to develop a neighborhood typology for the current study was informed by standing criminological theory, notably the broken windows hypothesis (Wilson & Kelling 1982) and rival social disorganization theories (Bursik 1988; Sampson & Groves 1989; Sampson *et al.* 1997). The list includes measures of the demographic composition of the resident population, socio-economic status, residential stability, and population heterogeneity, covering, albeit imperfectly, the constructs of concentrated disadvantage, immigrant concentration, and residential stability. A second set of indicators is used to characterize the built environment, including measures of the number of floors or the height of buildings, their functional use (business/commercial, residential, or mixed), as well as the period of construction of an area's housing park. These serve as measures, however flawed, of population density and land use patterns, the latter of which according to Sampson & Raudenbush (1999, 622) is a "robust but understudied" factor associated with crime and disorder.

Still, criminological theory guided the selection of indicators, but the actual clustering procedure to develop the neighborhood typology was nonetheless foremost a data-driven rather than a theory-driven process (as Chapter 3.2.5 will make amply clear). Clustering and feature selection algorithms were used iteratively to identify the indicators with the highest explanatory power in an effort to build a good and parsimonious model to classify urban neighborhoods into categories of different types.

3.2.3 Evaluation Design

Theoretical considerations

For a discussion of their suitably for evaluation research, it is helpful to distinguish three broad categories among the various research designs: randomized experiments, quasi-experiments, and observational designs. In a randomized experiment, the study objects targeted by the program being evaluated are randomly assigned to the treatment or control condition. In a quasi-experimental design, by contrast, the treatment group is typically compared to a control group that has been selected by the researchers on the basis of some criteria of similarity with the treatment group. A variant of the quasi-experimental design is the interrupted time-series design, which compares a series of observations of the treatment outcome variable(s) made prior and after program implementation. Time-series data are generally preferable over simple pretest-posttest measures, which is why this design qualifies as a quasi-experiment even in the absence of a control group. The third type of research design, observational designs, comprise the bulk of studies that attempt to model variations in outcomes as a function of varying exposure to the treatment intervention with other potentially confounding factors statistically controlled (Lipsey *et al.* 2006, 282-4).

The methodological standards of evaluation research have been cogently established already during the 1960s and the 1970s and have been reaffirmed more recently with particular reference to crime prevention programs (Campbell & Stanley 1963; Cook & Campbell 1979; Farrington 2003; Farrington *et al.* 2002; Shadish *et al.* 2002). This body of knowledge posits a clear hierarchy of the methodological quality of different research designs, with the randomized controlled trial held as the "gold standard" of scientific evaluation. This study design rules out most threats to internal validity and thus offers the highest degree of confidence in any inference about program impact (Farrington 2003, 59). Quasi-experimental designs by comparison, are subject to a series of threats to the internal validity that can lead to bias in the estimates of the treatment effects. The more sophisticated quasi-experimental designs try to control for these extraneous influences statistically, provided these sources of bias are known and the relevant data are available. By the same token, quasi-experimental designs are still preferable over observational studies, for which the internal validity risks being seriously undermined by a failure adequately to model the relation between the treatment program and the variation of outcomes (Lipsey *et al.* 2006, 283f.).

More recently, some authors have questioned the merits of this hierarchy of methods with the randomized trial at its top for the evaluation of criminal justice policies (Eck 2002, 284; Lipsey *et al.* 2006, 284f.). The counter argument ran less on theoretical grounds and more on practical considerations of actual implementation. For a start, the stringent requirements of randomized designs may result in the selection of study sites "that are not representative of the full scope of the program being evaluated." In the fields of crime prevention and policing, moreover, randomized designs have often proven difficult to implement, because experimentation is perceived as politically risky or ethically unacceptable, or both. For area-based crime prevention programs targeted at specific places or whole jurisdictions, it can be difficult to find a sufficiently large number of matching areas in order to obtain the statistical power necessary for a randomized research design to allow significant inferences about program impact (Eck 2002, 284; Lipsey *et al.* 2006, 285). This equally applies to studies where the units of analysis are social categories such as communities or neighborhoods (Welsh & Hoshi 2002, 191). As a consequence, several authors have bemoaned the dearth of methodologically sound evaluations of policing strategies (Weisburd & Eck 2004, 56) and have called for alternative methods to enhance the feasibility and validity of program evaluations (Eck 2002, 284; Lipsey *et al.* 2006, 296f.; Welsh & Hoshi 2002, 191).

Study design

The choice of an appropriate research design for the present impact evaluation of community policing in Swiss cities faces two major challenges due to the circumstances of program implementation. First and foremost, the present evaluation is carried out *ex post facto*. When work for it began, community policing efforts had already been underway for several years in all five urban areas. As the process evaluation has unearthed, none of the five police departments has taken provisions to prepare for a scientific evaluation, neither of the transition to community policing nor of its impact. This means that experimental designs are obviously not possible and that baseline data for pretest-posttest comparisons will have to come from existing data sources.

Secondly, the process evaluation has revealed that the five police departments introduced community policing more or less rapidly across their entire jurisdiction, leaving no area within the city proper behind. After an initial, usually rather short trial run at one or two police stations, the police command in each city swiftly thereafter decided to introduce community policing to the entire department. Such a department-wide transition was in line with the recommendations of the literature (e.g. Trojanowicz & Bucqueroux 1994, 2), but it complicates the task of an evaluator. Without appropriate control areas within a given city, the basic quasi-experimental research design is not a viable option.

These obstacles notwithstanding, it is contended here that by combining an appropriate study design with novel data analytical methods, it is possible to conduct an evaluation of community policing in Swiss urban areas with reasonable confidence in the integrity of its results. First of all, the Swiss Crime Survey polled a sufficiently large number of residents of all but one of the five cities prior to or during the early implementation of community policing, providing a rich source of baseline data. After a time lag of five years, the SCS polled another random sample of residents, thus allowing valid pretest-posttest comparisons at the infra-city level.

As for the study design, as will be explained in the following section, the current study uses geospatial data mining algorithms to develop a typology of urban neighborhoods in order to evaluate the variation of the outcome variables by neighborhood cluster as a pretest-posttest design. Without proper control group, though, the strength of the evidence of an impact analysis for an individual neighborhood cluster remains rather weak; on the Maryland Scientific Methods Scale (Farrington *et al.* 2002), it would probably score merely as Level 2. However, by taking all neighborhood clusters together, it is possible to track the simultaneous development of over 60 ZIP code districts – at a minimum one dozen per city – and to compare the treatment outcome by neighborhood type both within any one city and across the five cities, as well as across two linguistic regions. Considering further that data mining algorithms allow the evaluator to develop the neighborhood typology such that it accounts for a significant amount of the variance in the outcome variable at baseline, it is no longer fanciful to claim that a rather nuanced picture of how community policing strategies impact different types of neighborhoods should emerge as a result.

A further advantage of data mining algorithms is that, contrary to many other statistical analysis techniques, they do not normally make assumptions about the underlying distribution of the data. Moreover, the possibility to process complex data sets with several dozen variables or more enhances both the internal and external validity of the current study design, since the data mining algorithms save the evaluator the unthankful task of *selecting* a series of controlling variables. As Lipsey *et al.* (2006, 285) observed, this aspect should not be underrated in a discussion of the validity of quasi-experimental or observational designs particularly in the field of criminal justice, where theory is often underdeveloped, which jeopardizes the utility of such designs for evaluation purposes.

Unit of analysis

The choice of the appropriate unit of analysis is a critical aspect of any research design, "especially for evaluations that involve selecting or matching appropriate units." An important consideration is the level at which the intervention takes place, i.e. whether a program seeks to influence the behavior of individuals, families, community organizations, neighborhoods, or formal organizations. A good theory of intervention helps address this issue (Rosenbaum 2002, 194).

The current study is faced with the not uncommon problem that the units of analysis vary for different types of the data. For the survey data, the basic units of analysis are ZIP code districts as the Swiss Crime Survey has always retained the ZIP code number of each respondent's address. By contrast, the data on neighborhood ecology come from the Federal Office of Statistics' population and housing census counts and are available both as raster data with grid cells of one hectare as well as by the "statistical neighborhoods" (census tracts) as defined by the OFS for the five urban areas. For the crime data, finally, there is no consistent unit of analysis across the five urban areas.

Official crime data are recorded by the local police departments and practices vary. In Geneva and Lausanne, the data are congruent with the OFS statistical neighborhoods. The police in Bern ran a special query of their database for the purposes of this study on the basis of the city's ZIP code districts. In Zurich, official crime data are aggregated and published as a matter of course at the level of the city's twelve administrative districts.

In general, though, it was possible to aggregate the data to the appropriate level for multivariate analyses. The study used the OFS statistical neighborhoods as the basic unit of analysis and fused these as necessary to fit the shape of the bigger postal ZIP code districts or Zurich's administrative units. Also, the census raster data was first aggregated by statistical neighborhood and then summed up for the larger units of analysis as necessary. For the City of Zurich, its 35 census tracts were merged to match the shape of the 12 administrative districts, which the fused units perfectly do. For the analysis of the SCS data, an earlier study (Kreis 2009), merged a total of 166 OFS census tracts to form 111 bigger contiguous areas that are as congruent as possible with the shape of the ZIP code districts. The merged units fit the shape of the ZIP code districts rather well in the preponderant number of cases, even though the match is not always perfect. In those few instances, where a ZIP code area straddles two OFS census tracts, the two areas could still be aligned by merging both ZIP code areas and OFS census tracts. This has not been done, however, in order not to lose any degrees of freedom in the analysis of the survey data.

3.2.4 Exploratory Spatio-Temporal Data Analysis

The indicators of the four theoretical constructs used to measure the impact of community policing – crime, fear of crime, disorder, and public attitudes towards the police – were first subjected to an exploratory spatio-temporal analysis.[17] This exploratory analysis of both the police crime statistics and the Swiss Crime Survey data consisted of two separate parts: first, the long-term trends of the indicators of community policing impact were analyzed at the level of individual cities. For these analyses, each survey item was recoded as a dichotomized variable and the corresponding percentage value computed for each sweep of the SCS. The rationale of the long-term trend analysis was to get a basic understanding of whether the indicators show rising, falling, or stable trends at the city level over time.

In a second step, these very same indicators of community policing impact were analyzed at the neighborhood level, using ZIP code districts as proxies for urban neighborhoods. The rationale of the neighborhood-level analysis was to get a better understanding of the underlying spatio-temporal patterns in the data. This exploratory spatial data analysis relied on two basic tools: colored maps of the five cities to visualize the spatial distribution of the data and quantitative analyses of Moran's I to test the apparent patterns for statistical significance.

Exploratory spatial data analysis

In order to draw colored maps of the five cities, a four-step color scale was defined for each variable to display the attribute values of the spatial units of analysis, i.e. census tracts or ZIP code districts. Such colored maps are a useful tool to visualize the spatio-temporal patterns of the attribute values, namely if high or low values of a given variable are concentrated more in the center, the periphery, or some other part of an urban area, as well as to track shifts in the spatial patterns over time.

In order to minimize the subjective element in defining the color scales yet still make the thematic maps easy to read and interpret, a two-step procedure was employed: first, the "break values" were calculated for the 0.25, 0.5, and 0.75 quantiles or quartiles of each indicator in order to classify all study units across the five cities into four equally large groups. For the survey items, these break values were then adjusted manually to reconcile the twin objectives of having four class ranges that are easy to interpret and preferably comparable in size yet at the same time

[17] Section 3.2.4 is based on the author's unpublished graduation paper for the Master of Advanced Studies MAS *Urbanisme durable* (cf. Chapter 5; Kreis 2009).

contain approximately one quarter of all urban neighborhoods each (cf. Mitchell 1999, 50-54). For example, if the exact break value was – say 23 percent – the category boundary was set at 25 percent if category breaks at 25, 50, and 75 percent were appropriate. The same value could be rounded down to 20 percent if category breaks at 20, 40, and 60 percent resulted in a more equally balanced distribution of the study units over the four classes.

As a general rule, if the question format and answer categories of a given questionnaire item remained unaltered between two sweeps of the survey, the color scale for the maps was defined on the basis of the data from the earlier survey year (usually the 1998 sample) and applied identically to the data from subsequent years. This use of identical class ranges not only points out the relative position of the ZIP code districts for a given survey year, but also highlights changes in response patterns over time (Mitchell 1999, 162).[18]

This qualitative analysis of the spatial distribution of feature[19] attributes was complemented with quantitative analysis to test the apparent patterns of the thematic maps for the presence of spatial autocorrelation. Moran's Index was calculated according to Equation (3.4) for all the indicators of the four constructs crime, fear of crime, disorder, and public attitudes towards the police. Moran's I is a measure of the presence of spatial autocorrelation of the feature attributes of a study area: a positive value indicates positive autocorrelation or clustering, meaning that features with similarly high or low attributes are found near each other; a negative value signals negative autocorrelation or dispersion, meaning that nearby features have dissimilar attribute values. A value close to zero means that a pattern exhibits no spatial autocorrelation; the distribution is said to be random.

Calculating and interpreting Moran's I

For each pair of features, the GIS software subtracts the mean value of all the features in the study area $\overline{x}$ from the values of the target feature x_i and the neighboring feature x_j, and multiplies these values to obtain the so-called cross-product. The software then multiplies this cross-product with the weight attributed to this pair of features w_{ij}, repeats this process for every pair of features in the study area, and sums the results (Mitchell 2005, 121).

$$\sum_i \sum_j w_{ij}(x_i - \overline{x})(x_j - \overline{x})$$

Next, the GIS software computes the variance of all the features in the study area and multiplies this value by the sum of the weights attributed to each pair of features (Mitchell 2005, 121).

$$\sum_i \sum_j w_{ij} \frac{\sum_i (x_i - \overline{x})^2}{n}$$

Finally, the software divides the sum of the weighted cross-products by the variance of all the features multiplied by the sum of the weights to get the ratio, which is known as Moran's I. For the simplicity of the formula, the number of features variable n is moved to the numerator (Mitchell 2005, 121).

[18] If the question format or answer categories, or both, changed from one survey year to another, the class ranges of the color scales inevitably differ, as is indicated in the corresponding keys. In this case, absolute score levels or percentage values are not directly comparable. However, since the same methodology has been applied consistently in developing each color scale, a given neighborhood's position relative to others may still be meaningfully compared.

[19] In GIS terminology, the term "feature" refers to the basic spatial unit of analysis. In the context of this study, the features are the polygon shapes that delineate the OFS census tracts or postal ZIP code districts. This use of the term feature is not to be confounded with its meaning in a data mining context, where the term is used interchangeably with the term "variable".

$$I = \frac{n \sum_i \sum_j w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_i \sum_j w_{ij} \sum_i (x_i - \overline{x})^2} \tag{3.4}$$

From Equation (3.4), it is straightforward to see how the spatial distribution of the values of different features affect the value of Moran's I. As both the denominator and the number of features variable n are positive and constant for a given study area, the value of Moran's I is determined by the sum of the weighted cross-products of the feature pairs. For example, if the mean value of a variable of all features across the study area is 6, and the target feature x_i has a value of 12 and the neighboring feature x_j a value of 10, the cross-product of this feature pair is 24 (Mitchell 2005, 122f.).

$$(12 - 6) * (10 - 6) = 6 * 4 = 24$$

If the value of both the target feature and the neighboring feature are less than the mean of the study area – say 3 and 2 – the cross-product is still positive.

$$(3 - 6) * (2 - 6) = -3 * -4 = 12$$

If, on the other hand, the target feature has a value below the mean and the neighboring feature a value above the mean, the cross-product will be negative.

$$(12 - 6) * (2 - 6) = 6 * -4 = -24$$

Hence, if the distribution of the values of a variable in a study area is clustered, meaning that similarly high or low feature values are found near each other, the sum of the cross-products will be large and positive. If the distribution is random, meaning that some neighboring features have similar values but others do not, the positive and negative values of the cross-products will cancel each other out and the sum will be close to zero. Finally, if the spatial distribution is dispersed, meaning that nearby features have dissimilar values, the sum of the cross-products will be small and negative. In other words, a value of Moran's I that is greater than 0 indicates a clustered pattern, a value close to zero signals a random distribution, and a negative value suggests a pattern is dispersed (Mitchell 2005, 123).

Alongside the value of Moran's I, the GIS software calculates the corresponding Z-score on a standardized Gaussian distribution by calculating the expected value of Moran's I and the variance assuming that the spatial distribution is random.[20] The software then subtracts the expected value of Moran's I from the observed value and divides the difference by the standard deviation of the expected random distribution (Mitchell 2005, 124). The resulting Z-score indicates the level of confidence with which the null hypothesis H_0 that an observed pattern is simply due to chance can be rejected and is reported as a normal p-value here (i.e. $Z > 1.645 \rightarrow p < 0.1$; $Z > 1.96 \rightarrow p < 0.05$; $Z > 2.576 \rightarrow p < 0.01$).

The GIS software uses the weight values to implement neighborhoods in calculating test statistics such as Moran's I. For this purpose, the software calculates a table known as the "spatial weights matrix" that contains both a row and a column for each feature. In the cells of this table, the software stores the weight attributed to each feature pair according to the definition of neighborhoods being applied. For example, in an adjacency-based neighborhood, meaning a neighborhood is made up of the target feature and all its neighboring features, the cells of the spatial weights matrix take on a value of 1 for feature pairs that do share a border, and 0 for all other pairs. Thus in calculating Moran's I for adjacency-based neighborhoods, only the cross-products of

[20] The expected value of Moran's I depends on the number of features n in a study area and is calculated as $-1/(n-1)$. This is a negative number that is close to zero. The calculation of the expected variance depends on the number of features in the study area, the number of neighboring features, and the sum of the weights for all feature pairs (Mitchell 2005, 124f.).

feature pairs that have a common border will enter the computation. The spatial weights matrix also allows for other, more complex definitions of neighborhoods such as distance-based neighborhoods, which take the physical distance between features as weights, or neighborhoods based on a threshold or cut-off distance, if a feature attribute is theoretically held no longer to exercise any influence on the target feature beyond some drop-off distance (Mitchell 2005, 136-138).

For the current exploratory analysis, the spatial weights matrix for the calculation of Moran's I was based on the Euclidean distance between the centroids of the polygons representing the ZIP code districts. Fixed distance bands of 4 kilometers were used for all five cities under study here, which ensured that every target feature had at least one neighboring feature. Spatial weights were "row-standardized", which means that the sum of all the weights for each row in the spatial weights matrix must equal to one (e.g. if a target feature has four neighboring areas, each pair gets a weight of 0.25). Row standardization is used primarily with adjacency-based neighborhoods but may equally be applied for distance-based neighborhoods to create proportional weights if features have unequal numbers of neighbors (Mitchell 2005, 144).

The fixed distance band procedure was given preference over contiguity-based neighborhoods on the grounds that the actual shapes of the postal ZIP code districts are not in every instance perfectly congruent with the census tracts as defined by the OFS, which were used as the spatial unit of analysis. In particular, in some instances nearby polygons of the postal ZIP code districts do not share a common border, whereas the corresponding OFS census tracts do and vice versa.[21] A calculation of Moran's I based on contiguity-based neighborhoods would therefore affect the number of feature pairs entering into the computation and could thus alter the value of the index. With the fixed distance band procedure, the selection criterion is physical proximity (as expressed by a set distance) rather than a shared border, which should attenuate the problem of sometimes incongruent polygon shapes.[22]

All computations were made using the R language for statistical computing (R Development Core Team 2011). The survey data were weighted on the basis of official census data to correct for sampling bias in the age and gender distribution. Weights were calculated at the city level as a stratified sample of the neighborhood population and as simple random sample at the neighborhood level (cf. Appendix A.1). Percentage values and corresponding confidence intervals at the city level as well as the weighted means or percentages for individual neighborhoods were calculated using the functions `svymean` and `svytable`, respectively, of the package 'survey' (Lumley 2010). The feature classification system used to draw the thematic maps and the Moran's I test statistics were calculated and plotted using the packages 'classInt', 'spdep', and 'maptools' (Bivand *et al.* 2008).

3.2.5 Unsupervised Learning – Self-Organizing Maps

The impetus and objectives of matching urban areas for evaluation

The exploratory spatio-temporal data analysis, which had been undertaken as a preliminary study (Kreis 2009), revealed that the spatio-temporal patterns of fear of crime, disorder, social cohesion, and public satisfaction with the police exhibited great variance over the long time-span evaluated. The three indicators measuring fear of crime resulted in different patterns, yet over the study period, fear levels generally tended to drop in the city centers and simultaneously to rise in the

[21] As mentioned previously, the Swiss Crime Survey data sets retained the ZIP code number of a respondent's address, which serve as the georeference for the current study. However, shapefiles of the postal ZIP code districts were not available for the current study. More importantly, reliable demographic data for postal ZIP code districts (vital for the weighting of the survey data) would be available only at considerable expense on a commercial basis. For the present study, OFS census tracts were thus used as the spatial unit of analysis. In the five cities, several census tracts had to be merged to create bigger contiguous areas that are as congruent as possible with the polygons of the postal ZIP code districts (cf. Chapter 3.2.3).

[22] As a test, Moran's I was also calculated using row standardized spatial weights based on first-order neighbor neighborhoods. The resulting differences both in the values of Moran's I and related Z scores were for the most part quite small. Only the results of the fixed band procedure are thus reported here.

outskirts and surrounding suburbs. This shift became most apparent for fear measured as the perceived threat of crime in the three Swiss German cities of Basel, Bern, and Zurich between 1987 and 2005. In Lausanne and Geneva, for which data was available only from 1998 onwards, the spatial dynamics proved less pronounced. Likewise, if fear was measured as the perceived risk of victimization or as the behavioral response to crime, the neighborhoods with higher levels of fear tended to be found on the five cities' outskirts towards the end of the study period. The residents of the city centers, by contrast, showed the lowest levels of fear in all five agglomerations.

The spatio-temporal patterns of disorder and social cohesion, on the other hand, revealed no such clear-cut trends. The highest levels of disorder were concentrated in the city centers especially of Bern, Geneva, and Zurich and, to a lesser extent, Basel and Lausanne. Physical and social disorder appeared foremost to be urban phenomena that seriously afflicted only a handful of neighborhoods. Social cohesion was lowest in the city centers and tended to increase towards the outskirts during the observation period. Finally, the spatio-temporal patterns concerning the police proved more amorphous still. Trends in the spatial distribution of popular satisfaction with police effectiveness in controlling crime were much harder to discern. On policing, the geostatistical tools employed showed virtually no clustering in the data at all.

In light of these rather systematic spatio-temporal patterns of some of the outcome variables, it is paramount for an impact evaluation of community policing over such a long study period to control for shifting neighborhood characteristics. Otherwise, it risks being unreliable at best and positively misleading at worst. As a consequence, the study as the next step aims to develop a system to classify Swiss urban neighborhoods based on a series of environmental, socio-economic, and socio-demographic indicators. The idea is to divide neighborhoods into groups of similar type and to study community policing by neighborhood clusters in order to enhance the validity of the planned impact evaluation. The objective is thus to develop a clustering algorithm to create a neighborhood classification system that minimizes the variance in the ecological variables and simultaneously accounts for a significant share of the variance in the outcome (survey) variables at the outset of community policing implementation. In other words, the twin optimization problem is to maximize within-cluster similarity of the potentially confounding covariates all the while minimizing between-cluster similarity of the outcome variables.

To this end, the study makes use of both unsupervised and supervised data mining algorithms. In the exploratory phase, as the next section explains, self-organizing maps are being used in combination with hierarchical agglomerative clustering of the trained Kohonen map to develop a typology of urban neighborhoods (Skupin & Agarwal 2008; Vesanto & Alhoniemi 2000). During the supervised learning phase, the random forests algorithm (Breiman 2001) serves for feature selection to weed out noisy or unimportant variables in order to develop a parsimonious model with high predictive accuracy. Random forests are also being used to assess both the quality of the clustering algorithm and the importance or predictive power of individual variables.

As a final step, the resulting neighborhood typology will be visualized as a geographic map using GIS software. This map not only indicates which areas are similar with regard to both structural characteristics and the prevailing sentiment of the resident population and thus suited for matching during the following impact evaluation. Such a map may also serve to communicate the results to both policymakers and practitioners.

Training of the self-organizing map

Proposed by Kohonen (1990; 2001), self-organizing maps (SOMs) or Kohonen maps are an unsupervised learning algorithm that simultaneously performs both vector quantization and non-linear dimensionality reduction of high-dimensional data. The Kohonen algorithm embeds in the original input space a discrete lattice of prototype vectors that are iteratively fitted onto the data points and subsequently projects these prototype vectors onto a regular grid of usually two dimensions. SOMs belong to the family of topology preserving algorithms, which means that if two prototype vectors are close to each other in original input space, they remain closely together in the projected low-dimensional output space (Lee & Verleysen 2007, 136f.).

SOMs are especially suited during the exploratory phase of data analysis because they offer superb possibilities of inspecting and visualizing high-dimensional data without requiring any a priori knowledge about the underlying structure of the data (Vesanto & Alhoniemi 2000, 586). Because of its vector quantization quality, the SOM algorithm is computationally efficient and more robust than other clustering algorithms with regard to both noise and outliers. In particular, Vesanto & Alhoniemi (2000) have shown that a two-stage clustering procedure that uses SOMs as a first level of abstraction of the original data points and then uses traditional clustering approaches to partition the prototype vectors or "protoclusters" into groups of similar type, produces results that are comparable with the results obtained from direct clustering algorithms, but offers benefits in terms of reduced computational cost and noise reduction (pp. 586-588). For these reasons, such a two-stage clustering procedure has been adopted to develop a neighborhood typology for the present study.

The Kohonen self-organizing map is perhaps the best known member of the family of dimensionality reduction algorithms that rely on a predefined lattice, i.e. a lattice that does not change in size during the training process (Lee & Verleysen 2007, 134). The neurons of the SOM neural network are arranged in a regular, usually two-dimensional (2-D) lattice or grid, the structure of which is typically either rectangular or hexagonal. Every neuron thus has four or six neighboring neurons, respectively, except for those at the margins of the grid. Each neuron of the SOM grid corresponds to a prototype or weights vector $w_i = (w_{i1}, w_{i2}, \ldots, w_{ip})$, where p corresponds to the number of dimensions of the input space, i.e. the number of variables in the training data set. The prototype or weights vectors thus represent a neuron's coordinates in the original data input space (Kanevski *et al.* 2009, 218f.).

The SOM algorithm is based on competitive learning. This means that at each training iteration, the neurons are in competition with each other to best represent the original data points that are being presented to the neural network. At each iteration, only one neuron wins. This "winning" neuron – the best matching unit (BMU) – is the neuron closest to the original data point x (usually measured in Euclidian distance):

$$d(x, w_{BMU}) = \min_i d(x, w_i) \qquad i = 1, 2, \ldots, m, \tag{3.5}$$

where m is the total number of neurons in the SOM neural network (Kanevski *et al.* 2009, 219f.).

At the outset of the training process, the weights of the prototype vectors are initialized. The initial weights were chosen as a random subset of m data points of the training set. During the following iterative training process, the data points of the training set are chosen at random one-by-one and presented to the neural network. At each iteration step t, the winning neuron is identified and the prototype vectors in the SOM grid are updated according to Equation (3.6):

$$w_i(t+1) = w_i(t) + h_i(t)[x - w_i(t)], \tag{3.6}$$

where $h_i(t)$ is the so-called neighborhood function, which is defined according to Equation (3.7):

$$h_i(t) = \begin{cases} \alpha(t), & \text{if} \quad i \in R_{BMU} \\ 0, & \text{if} \quad i \notin R_{BMU} \end{cases} \tag{3.7}$$

The SOM algorithm updates not only the prototype vector of the BMU, but also those of all the surrounding neurons that fall within the area of the SOM lattice defined by the radius R of the neighborhood function $h_i(t)$. At each iteration step, the prototype vectors are moved closer to the data point in original space, i.e. the distance between the prototype vectors w_i and the original data point x is reduced at the learning rate α, where $0 < \alpha < 1$ (Kanevski *et al.* 2009, 220f.). In other words, each time the prototype vector of the BMU is updated, the prototype vectors of the neighboring units in the SOM lattice are updated too. By moving the BMU and its topological neighbors in tandem, the SOM algorithm maintains the original grid structure throughout the training process and achieves its topology preserving quality (Lee & Verleysen 2007, 137).

Over the training epochs, both the learning rate α and the radius R of the neighborhood function steadily decrease as a function of the number of iterations. At the beginning of the training process, when both R and α are large, a sizable part of the prototype vectors are updated at each iteration step t. As the training progresses, α and R decrease linearly and the learning process becomes increasingly local until, eventually, only the prototype vector of the BMU gets updated (Kanevski *et al.* 2009, 220f.). The SOM neural network is thus a flexible grid that molds itself iteratively over the cloud of data points over the course of the training process. The training stops once the grid has converged, i.e. the prototype vectors are only updated minimally, or the process has reached a pre-defined number of iterations. The SOM differs, though, from other unsupervised data mining algorithms in that no specific optimization criterion or error function is minimized during training, so that convergence cannot be evaluated objectively (Lee & Verleysen 2007, 141).

For the current study, the number of neurons in the SOM grid was set approximately equivalent to the number of data points in the training data set. This was done in order not to impose any constraints on the mapping of data points on the SOM grid and because vector quantization was not an objective of the unsupervised learning procedure. Moreover, during the exploratory data analysis, it is generally advisable to use a number of neurons that is several times the expected number of clusters in the data (Bação *et al.* 2008; Vesanto & Alhoniemi 2000). All computations of the SOM algorithm were carried out using the `som` function of the R package 'kohonen' (Wehrens & Buydens 2007). The learning rate α decreased linearly from 0.2 to 0.01. The initial neighborhood radius R covered 2/3 of all unit-to-unit distances, which corresponds to the default value of the `som` function. The number of training iterations was set to 1000, after which the SOM showed acceptable average quantization error as is indicated by Figure 3.2 a (cf. Kanevski *et al.* 2009, 221f.).

```
> som.nbhd<-som(scale(training),grid=somgrid(xdim=7,ydim=7,"hexagonal"),
+              alpha=c(0.2,0.01),rlen=1000)
```

The results of the SOM training procedure are commonly displayed by means of two visualization tools that make use of its 2-D grid structure (Kanevski *et al.* 2009, 222f.). The "hits map" displays the number of original data points attached to each neuron of the SOM grid, i.e. the number of original data points for which each neuron is BMU after the training process is completed. In the hits map, the neurons are shaded according to a color scale indicating the number of "hits". Ineffective or empty neurons are colored in grey (Fig. 3.2 b). The second complementary graphical tool used to visualize the results of the training procedure is the "unified distance matrix" or "U-matrix" (Ultsch & Siemon 1990). The U-matrix plots the sum of the standardized distances of each prototype vector to its neighboring units. The U-matrix uses a separate color scheme, in which darker colors indicate proximity, i.e. a prototype vector is relatively close to its neighbors, whereas brighter shades point out larger distances between a prototype vector and its neighbors (Fig. 3.2 c).[23]

Clustering of the SOM

The hits map and U-matrix allow for a first inspection and qualitative assessment of the underlying structure of the high-dimensional data. However, when the number of neurons is large, quantitative analysis is normally required to partition the SOM lattice and cluster similar map units (Vesanto

[23]Sometimes a third plot of the 2-D grid structure is produced to map the "quality" of the SOM, as estimated by the average standardized distance between each prototype vector and the original data points or input vectors attached to it (Wehrens & Buydens 2007, 13). For the current study, plots of the quality of the SOM mapping were inspected but found to display purely random patterns and are not reproduced here.

& Alhoniemi 2000, 586). The clustering algorithm for this second step is not pre-determined. In practice, however, hierarchical or k-means clustering algorithms are the most widely used.

The present study used a hierarchical agglomerative algorithm to cluster the prototype vectors of the trained Kohonen map. Euclidean distance was used to compute the distance matrix. The clustering itself was based on complete linkage in order to obtain compact clusters; single linkage almost invariable resulted in the "chaining" of the prototype vectors (cf. Bartholomew *et al.* 2008, 40f.).

If the SOM prototype vectors are clustered by means of an agglomerative algorithm, the topological information ingrained in the lattice can be used to impose a spatial contiguity constraint on the construction of the dendrogram (Vesanto & Alhoniemi 2000, 591). The idea to use an agglomerative contiguity-constrained clustering method to partition the SOM into segments was first proposed by Murtagh (1995). For the present study, the hierarchical agglomerative clustering was consistently computed both with and without spatial-contiguity constraints. Contiguity-constrained clustering paired with the complete linkage criterion resulted in nice compact segments of the SOM lattice.

The case for using spatial-contiguity constraints in the partitioning of the trained lattice lies in the nature of the SOM algorithm. Because of its topology preserving quality, the prototype vectors of neighboring neurons represent data points that are close to each other in original input space. A clustering algorithm that only allows merges between neighboring prototype vectors keeps this topological information; it thus prevents the linkage of prototype vectors that are quite far apart in original space but appear close based on their Euclidian distance, which is a distinct possibility, especially if the training data set has many dimensions.

Additionally, the prototype vectors representing empty or ineffective neurons were removed from the distance matrix to perform the clustering of the SOM lattice. Again, the argument lies in the nature of the SOM algorithm. During training, some prototype vectors are inevitably dragged into the empty space between clusters and thus have few data points attached to them or may be completely empty. By using this information contained in the SOM lattice together with the spatial contiguity constraint, empty neurons generally indicate cluster borders (Vesanto & Alhoniemi 2000, 591f.).

The unconstrained clustering of the prototype vectors was performed using the `hclust` function in R.

```
> #Identify empty neurons
> pos.empty<-1:nodes
> pos.empty[som.nbhd$unit.classif]=NA

> #Hierarchical agglomerative clustering of the non-empty SOM neurons
> hc<-hclust(dist(som.nbhd$codes[is.na(pos.empty),],method="euclidian"),
+                 method="complete",members=NULL)
```

Spatially-constrained clustering can be achieved using two basic methods: (a) the spatial contiguity constraint is applied only after a conventional clustering algorithm has run its course and formed clusters that violate the constraint are split retroactively; and (b) the spatial contiguity-constraints are applied progressively during the clustering process and only units that meet them are allowed to merge. The two clustering procedures often do not yield the same number of clusters (Patil *et al.* 2006, 367-369). The current study used the second method. The contiguity-constrained clustering algorithm was programmed in the R language and proceeds as follows:

1. Calculate the distance matrix of the prototype vectors based on Euclidian distance.

2. Define a contiguity matrix which has a value of 1 if two prototype vectors are topographical neighbors in the SOM lattice and 0 otherwise.

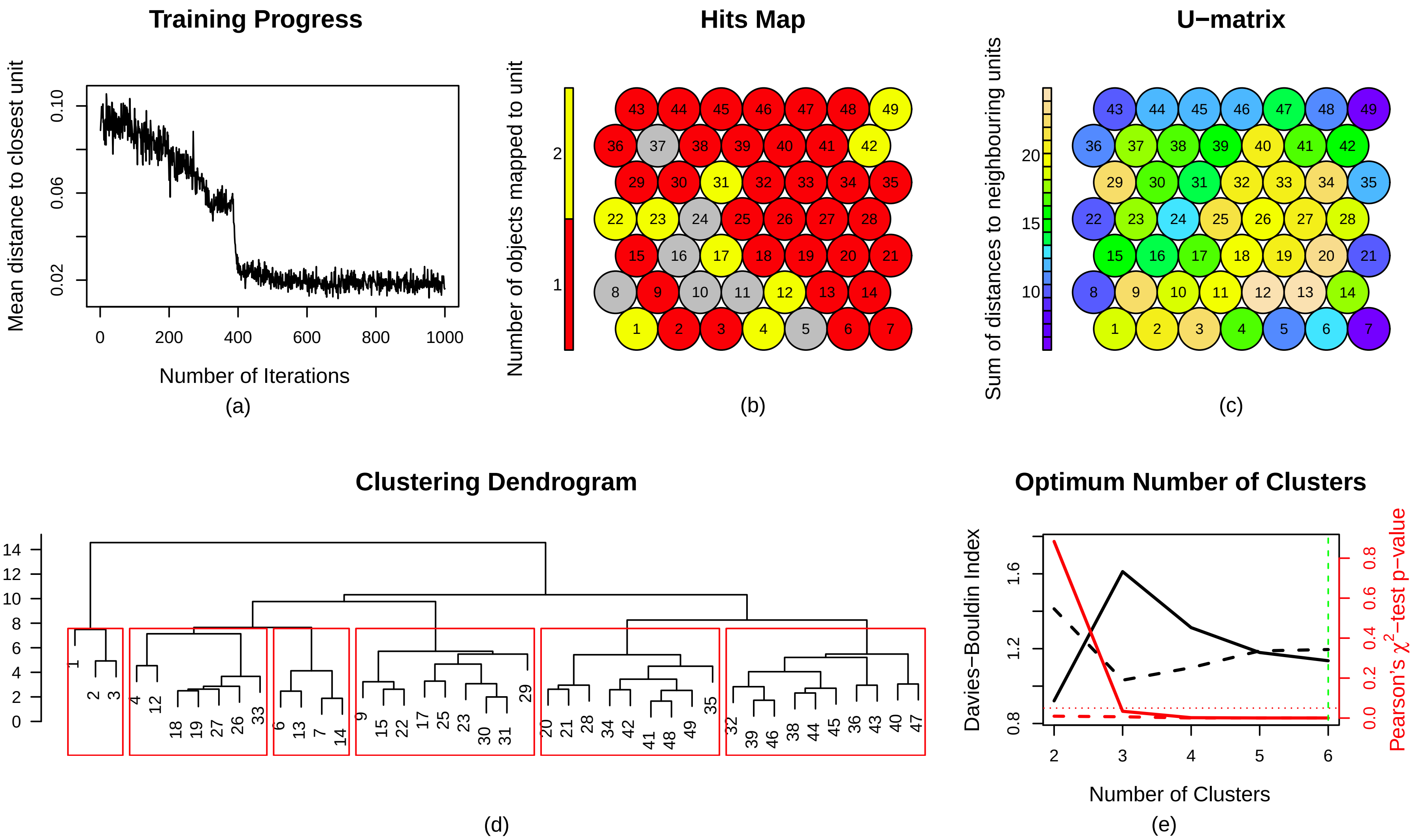


Figure 3.2 Training and clustering of the self-organizing map.

3. Merge the two singletons or clusters separated by the smallest Euclidean distance (complete linkage) subject to the spatial contiguity constraint, and update the distance and contiguity matrix accordingly.

4. Repeat step 3 until all effective neurons in the SOM neural network have been merged.

As discussed, empty prototype vectors were removed from the distance and the contiguity matrix.[24]

```
> #Attribute distance matrix
> distmat<-as.matrix(dist(som.nbhd$codes,method="euclidian",diag=T,upper=T))
> for (i in 1:nrow(distmat)) distmat[1:i,i]<-NA

> #Contiguity matrix
> contmat<-as.matrix(dist(som.nbhd$grid$pts,method="manhattan",diag=T,upper=T))
> for (i in 1:nrow(contmat)) contmat[1:i,i]<-NA
> contmat[contmat<1.4]<-1
> contmat[contmat>1]<-0

> #Contiguity-constrained clustering step
> temp<-distmat
> if (sum(contmat==1,na.rm=T)>0) temp[contmat==0]<-NA
> height[i]<-min(temp,na.rm=T)
> pos<-which(temp==height[i],arr.ind=T)
> distmat[pos[1],pos[2]]<-NA
```

Determining the optimum number of neighborhood clusters

The agglomerative clustering procedure links up all m prototype vectors of the SOM lattice but in order to separate the neighborhoods into clusters, the dendrogram tree must be cut at some level. In order to choose the best partitioning of the data points representing the urban districts with regard to the contextual variables, some kind of clustering validity index must be applied. In accord with other studies (Vesanto & Alhoniemi 2000), the current study used the Davies-Bouldin Index (DBI) to assess the quality of clustering as a function of the number of clusters used to partition the urban districts in the training data. The DBI is an index figure that relates the within-cluster variance with the between-cluster variance for different partitionings of the data (Davies & Bouldin 1979). The DBI seeks to minimize between-cluster similarity, i.e. to partition the data points into a set of compact and neatly separated clusters according to Equation (3.8):

$$\frac{1}{C}\sum_{k=1}^{C}\max_{l\neq k}\left\{\frac{S_c(Q_k)+S_c(Q_l)}{d_{ce}(Q_k,Q_l)}\right\}, \tag{3.8}$$

where C is the number of clusters, S_c is the within-cluster centroid distance and d_{ce} is the between-cluster centroid distance (Vesanto & Alhoniemi 2000, 588).

In order to calculate the optimum partitioning of the urban districts, the dendrogram resulting from the clustering of the SOM prototype vectors was cut at different levels, starting near the top at

[24]Prototype vectors that capture outliers and are completely surrounded by empty prototype vectors in the SOM lattice had no effective neighbors in the contiguity matrix. The algorithm allowed the merger of these isolated prototype vectors only after all other prototype vectors with non-empty neighbors had been clustered. Since the Euclidean distance of such isolated prototype vectors to their nearest non-empty second-order neighbors was usually rather large, the resulting dendrogram commonly retained an integral tree structure. For the time being, no second-order contiguity matrix was used in the contiguity-constrained algorithm because the trade-off between the programming effort and the limited extra insight advised against this step. Such isolated prototype vectors appeared only during the exploratory data analysis when crime data alone was being clustered.

$k = 2$ and moving down progressively until the first singleton of the tree structure was reached (at which point the DBI can no longer be calculated since there is no longer any variance within that cluster). For each number of clusters $k = 2, 3, \ldots, K$, the DBI value was calculated for both the unconstrained and the spatially-constrained clustering of the trained SOM map. All computations were made using the `index.DB` function of the R package 'clusterSim'. Figure 3.2 e plots the values of the DBI against the number of neighborhood clusters k. The solid black line represents the values for the contiguity-constrained clustering and the dashed line stand for the unconstrained clustering. In the current study, the DBI did not generally show multiple local minima.

```
> temp<-cutree(hc,k=i)
> dbi[i,2]<-index.DB(som.nbhd$codes[is.na(pos.empty),],
+           temp,centrotypes="centroids")$DB
```

A clustering procedure that is run as a matching exercise for an observational study should not only minimize between cluster similarity with regard to the contextual variables of the units of analysis (and thus reduce the influence of these potentially confounding covariates), the resulting neighborhood typology should also account for a maximum of the variance in the outcome variable prior to program implementation. In order to resolve this twin optimization problem, in a second step, the respondents of the Swiss Crime Survey were regrouped by their place of residence according to the neighborhood clusters just determined for the contextual variables. In analogy to the calculations of the DBI values, for each number of neighborhood clusters $k = 2, 3, \ldots, K$, a χ^2-independence test was taken to check whether the response patterns for the variables of community policing impact – notably fear of crime – varied significantly across the different neighborhood clusters.

As the size of the SCS survey sample was fixed, the higher the number of neighborhood clusters k, the smaller the number of survey respondents per cluster since the same overall survey sample is divided up into more groups. This inevitably increased the risk of ending up with a contingency table with too many cells with an expected frequency of less than five, which may distort the χ^2-statistic. (As a rule of thumb, 20 percent of cells with an expected frequency below five are tolerated.)

In order to circumvent this problem, the p-values of the χ^2-independence test were computed using Monte Carlo simulations. Monte Carlo simulations are implemented as a standard option in the `chisq.test` function in R. All χ^2-independence test p-values were computed on the basis of 2000 replications, which is the standard value R proposes for such simulations. In order to correct for the sampling bias inherent in a survey sample, the contingency table was weighted using the `svytable` function of the 'survey' package (Lumley 2010).

```
> des<-svydesign(id=~1,data=df,weights=df$weights)
> ct<-svytable(~var+type,des)
> test<-chisq.test(ct,simulate.p.value=T,B=2000)
```

Once the p-values of the χ^2-independence tests had been calculated, they were plotted on a chart against the number of neighborhood clusters k alongside the DBI values (Fig. 3.2 e). The solid red line again depicts the values of the contiguity-constrained clustering, whereas the dashed line displays the unconstrained clustering. The optimum number of clusters was then chosen as follows: from among all the numbers of neighborhood clusters, for which the null hypothesis of the χ^2-independence test was rejected, the optimum number was chosen so that the DBI is minimal. This optimum number is indicated as a dashed vertical green line in the chart. For the present study, the survey question regarding the threat of crime was commonly used for the χ^2-independence test to evaluate the appropriateness of the clustering with regard to the survey data.

Since the results of the SOM depend to some extent on the original data points or input vectors randomly sampled to initialize the training process, the entire SOM training procedure and determination of the optimum number of clusters was replicated 50 times. A frequency table was then calculated of the optimum number of neighborhood clusters identified over the 50 replications. Given a requirement to identify a minimum of four clusters, the number of neighborhood clusters that the algorithm suggested most frequently was six. (Less than four clusters always resulted in a pattern that pitted the urban centers against all other neighborhoods, which explained less of the variance in the outcome indicators.) From among those replications that identified $k = 6$ as the optimum number of clusters, the replication that produced the lowest value of the DBI was retained as the final clustering result.

3.2.6 Supervised Learning – Random Forests

For all its sophistication, the neighborhood clustering algorithm described so far resembles a "black box". The procedure does create a neighborhood classification system but gives only little indication how distinct the resulting clusters really are and absolutely no clue which variables are relatively more important in determining a given neighborhood's class. Furthermore, because of a phenomenon known as the "curse of dimensionality" (Lee & Verleysen 2007, 6), the clustering of high-dimensional data is fraught with difficulties. It is thus important to complement the unsupervised data mining performed so far with a supervised algorithm for feature selection.

The algorithm used for the current study is "random forests", the increasingly popular supervised learning algorithm for classification and regression proposed by Breiman (2001). The random forests algorithm belongs to the family of tree-based models. Decision trees are a mathematically simple but computationally intensive data mining algorithm for classification and regression purposes (cf. Breiman *et al.* 1984). At their most basic form, decision trees divide the training data sample into a number of sub-samples, by using a series of recursive binary splits with the objective to minimize the variance in the dependent variable y (if y is numeric) or to create sub-samples that contain only data points of a single category of y (if y is nominal or categorical). During training (a.k.a. the "growing" of the tree), the algorithm identifies at each node of the decision tree (a.k.a. "splits") the critical value in one of the explanatory variables that will produce the "best" separation of the data into two sub-samples such that the variance in y is minimized or the data points of different categories of y are neatly separated. The decision tree algorithm continues the binary splits of the resulting sub-samples until the final sub-samples at the terminal nodes (a.k.a. "leaves") contain only a handful of data points (for regression) or only data points of a single category of y (for classification). Once the training is completed, a "grown" decision tree provides a classification rule based on a series of binary splits from stem to leaf, and additional data points can be run down along the different paths of the tree according to their values of the explanatory variables for validation or prediction (Crawley 2007, Chapter 21).

Training of the random forests classifier

The random forests algorithm differs from the standard classification and regression trees (CART method) in two important ways. First, random forests belongs to the family of ensemble methods, i.e. methods that do not rely on a single model of the data but generate a large number of classifiers and then make predictions based on the aggregated results of these individual classification rules. The random forests algorithm, rather than relying on a single decision tree, grows several hundred of them, each of which is based on a bootstrap sample of the training data set, and then aggregates their results (by majority voting for classification and by taking averages for regression). Second, whereas standard decision trees always seek the best split from among *all* the variables in the data set, random forests at each node of a decision tree only considers a *randomly selected subset* of variables and seeks the best split only from among those features (Liaw & Wiener 2002, 18; emphasis added). Bootstrap sampling from the training set to grow multiple decision trees – also known as "bagging" – and random feature selection for binary splits are the hallmarks of the algorithm, hence its name "random forests" (Breiman 2001, 5f.)

In more formal terms, given a specific training data set T, random forests draws n_{tree} bootstrap training sets T_k with replacement. For each of the bootstrap samples T_k, $k = 1, 2, \ldots, n_{tree}$, the algorithm constructs an unpruned decision tree with the following modification. At each node, rather than seeking the best split from among all the explanatory variables in T, the algorithm chooses the best split from a subset of m_{try} predictors randomly chosen at that node. Finally, the thus constructed classifiers h_k, $k = 1, 2, \ldots, n_{tree}$, take a simple majority vote to form the bagged predictor (Breiman 2001, 11; Liaw & Wiener 2002, 18).

The random forests algorithm provides an estimate of the generalization error by running the data points outside the bootstrap sample down the individual classifiers, aggregating their prediction, and calculating the classification error rate over all the data points in the training data set. (As a result of the bootstrap sampling, each data point is left out – what Breiman [2001] calls "out-of-bag" [OOB] – about one-third of the time.) More formally, in order to get an estimate of the prediction error, the algorithm runs down each data point $\mathrm{y}, \mathbf{x}$ in the training data set T over all those classifiers for which T_k does not contain $\mathrm{y}, \mathbf{x}$ to get the OOB classifier. The estimate of the generalization error of the random forests classifier, better known as the "OOB estimate of error rate", is then the error rate of the OOB classifier on the training set, i.e. the number of data points that fall off the diagonal of the confusion matrix (Breiman 2001, 11; Liaw & Wiener 2002, 18).

The use of the OOB error estimate makes a separate testing data set superfluous. In his paper, Breiman (2001) developed the proof that random forests do not overfit when more trees are added but that the generalization error converges to a limiting value (p. 7; Appendix I). What is more, as a predictive tool, random forests compare favorably with many other classification algorithms and the algorithm is very "user-friendly". There are only two parameters that must be set – the number of trees to grow n_{tree} and the number of variables on which to split the data sub-samples m_{try} – and the algorithm is generally quite robust to changes of their values (Liaw & Wiener 2002, 18).

Variable importance

In addition, the random forests algorithm uses the OOB data samples to calculate a variable importance (VI) measure that is useful both for interpretation purposes and feature selection. The VI score can be used to build a good parsimonious model by retaining only the independent variables with high predictive power and getting rid of moderately correlated or noisy explanatory variables. The random forests algorithm computes the importance score of a given variable by calculating the increase in the error rate of the OOB classifier when the observed values of that variable are randomly permuted in the OOB samples, in comparison with the OOB estimate of the error rate when all variables are left unchanged (Genuer *et al.* 2010, 2226, 2229; Liaw & Wiener 2002, 18f.). The logic behind the VI score is that if a given variable has high predictive power, the OOB estimate of the error rate should increase substantially as result of this random permutation of its values, whereas in the case of a less important explanatory variable the error rate will only be moderately affected by such a random change.

The VI scores are computed during the training of the random forests classifier. After each tree h_k is grown, the values of the pth variable of the corresponding out-of-bag data points are randomly changed. These "noised up" data points are run down the classifier h_k, and the predicted class of each data point x_n is recorded. After the algorithm completed this operation for every classifier h_k, $k = 1, 2, \ldots, n_{tree}$ in the random forests, the ensemble of the out-of-bag class votes for each data point x_n is compared with its true label to give its misclassification rate. The VI score is the percent increase in the misclassification rate of the OOB classifier with the values of the pth variable randomly permuted, compared with the OOB estimate of error rate with all the variables left intact. This procedure is repeated for all the $p = 1, 2, \ldots, P$ variables in the training data set T to compute individual importance scores for each variable (Breiman 2001, 23f.)

The VI score, like the random forests algorithm itself, is highly versatile. It is applicable in standard situations where the number of data points exceeds the number of variables ($n > p$) as well as in high-dimensional regression or classification problems ($n < p$). Moreover, the relative importance of the VI scores, i.e. the ranking of the predictive power of individual variables, is robust with regard to changes in the number of data points n or the number of variables p in the training set T, the values of m_{try} and n_{trees} as well as the presence of highly correlated (i.e. collinear) explanatory variables (Genuer *et al.* 2010, 2226-2229).

Proximity matrix

The second highly useful bit of information that the random forests algorithm computes besides the VI importance score is the proximity matrix. The (i, j) element of the proximity matrix is the fraction of all the classifiers h_k, $k = 1, 2, \ldots, n_{tree}$, on which the data points x_i and x_j end up on the same terminal node. The interpretative value of the proximity matrix does make intuitive sense: similar observations should fall more often in the same terminal node than dissimilar ones, so the higher the element (i, j) the more similar are data points x_i and x_j, and the proximity matrix can thus be used to detect the underlying structure of the data (Liaw & Wiener 2002, 18f.)

The random forests proximity matrix can be visualized in a 2-D plot using multi-dimensional scaling (MDS), for a visual exploration of the quality of the random forests classification (Liaw & Wiener 2002, 21). If the preceding clustering step was done well, the individual data points of the training set will end up in compact clusters that are nicely separated from each other. The visual inspection of the random forests proximity matrix thus provides an additional check on the choice of the optimum number of neighborhood clusters.

All computations using the random forests algorithm were carried out using the R package 'randomForest' (Liaw & Wiener 2002). The R default value for the number of randomly chosen features on which to split the data is $p^{1/2}$ for classification, where p is number of variables in the training set. The recommended strategy to select the optimum number m_{try} is to try, consecutively, the default value, half the default value, and twice the default, and then to stick with the number that results in the lowest OOB estimate of the error rate (Liaw & Wiener 2002, 21). Similarly, Liaw & Wiener (2002, 21) recommend an iterative procedure to determine the adequate number of trees to grow. The trick is to compare the predictions of an entire forest with those of a subset of trees: if the prediction error remains the same, the subset number of trees will do. For the current study, however, n_{trees} was set to 1000 (i.e. twice the R default), since the random forests algorithm is used exclusively for feature selection based on the VI score and to calculate the proximity measure. This appeared more than justified since the computational burden remained fairly light and it is acknowledged that many trees are necessary to obtain stable estimates of the VI scores and the proximity measure (Liaw & Wiener 2002, 21). Moreover, as the random forests algorithm depends to some extent on the bootstrap samples drawn to train the classifier, the algorithm to compute the variable importance score was replicated 50 times and the VI scores averaged over the 50 replications in order to get more reliable estimates.

```
> classify<-data.frame(training,type=as.factor(cnbhdtype[,k]))
> rf.model<-randomForest(type~.,data=classify,
+           ntree=1000,importance=T,keep.forest=F)
```

Geo-visualization of the neighborhood typology

Once the optimum number of neighborhood clusters has been identified and the quality of the clustering algorithm assessed, the results can be visualized using different tools. Figure 3.2 d shows the dendrogram of the clustering procedure cut at the level that corresponds to the optimum number of clusters k. Since the leaves of this dendrogram represent the individual prototype vectors, the results of the clustering procedure may also be visualized using the SOM lattice. Figure 3.3 plots the SOM lattice with the neurons being colored according to their neighborhood type. In addition, each neuron displays the original data points, which it represents (i.e. for which it is BMU in the trained map). Since these original data points correspond to ZIP code or administrative districts, it is also possible to draw maps of the five cities and to display the results of the neighborhood clustering in the original geographic space (Fig. 6.6).

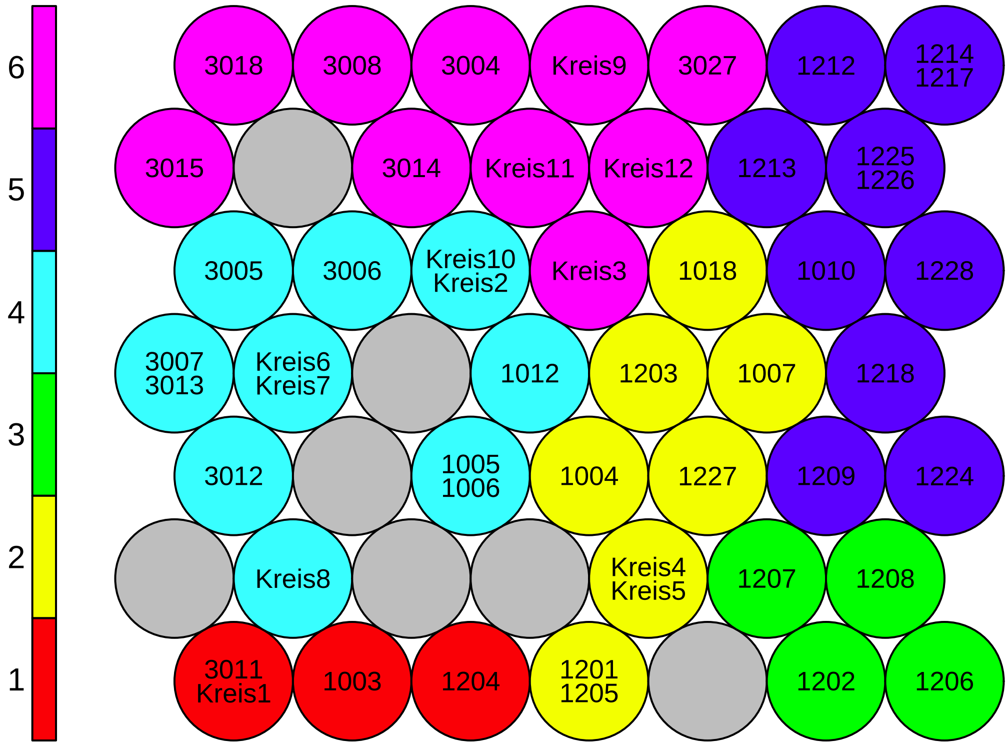


Figure 3.3 SOM best-matching units. 2-D visualization of the trained and clustered SOM lattice with each neuron displaying the original data points attached to it.

All geographic maps for the present study were produced in R using the packages 'gpclib', 'maptools', and 'spdep' (Bivand *et al.* 2008).

3.2.7 Impact Analysis

As the final step, the present study will use the resulting neighborhood typology as a tool to control for shifting neighborhood characteristics in order to boost the validity of the impact evaluation of community policing. However, even after having developed a neighborhood typology that minimizes the variance in structural variables and accounts for a significant share of the variance in the outcome variables, a number of threats to internal validity remain. For a start, if program impact is assessed based on two separate random samples of the general population, any inferences are valid only to the extent that the two samples are comparable (Cook & Campbell 1979, 117). In order to make the two samples comparable, the current study relies on propensity score matching methods to achieve balance in the demographic characteristics of the survey respondents of the pretest and posttest samples. Propensity score matching is a technique that is used in observational studies to match individuals with similar characteristics in both treatment and control group in order to reduce the risk that unaccounted for differences in observed covariates confound the observed shifts in the outcome variables.

The design of an observational study

Ideally, an observational study resembles the template of a randomized trial as closely as possible (Rosenbaum 2010, 4). In a randomized trial, the study objects are typically pre-selected based on some criteria and arranged in pairs of two, one of which is subsequently randomly assigned to the treatment condition while the other becomes the control. The effect of treatment is evaluated by comparing the difference in the outcome variable(s) between treatment and control group over all the matched pairs in the study sample. Crucially, in a randomized trial, the probability of receiving the treatment is identical for all study objects ($\pi = 1/2$). Moreover, since treatment assignment is independent of the study objects' characteristics, randomization tends to balance treatment and control groups in terms of both the observed and unobserved covariates of the study objects so that treatment is the only systematic difference between the two groups (Rosenbaum 2010, Chapter 2).

In an observational study, by contrast, the treatment and control groups are not pre-selected prior to treatment, and the probability of a study object in the general population to receive the treatment is typically unknown. In an observational study that emulates the randomized trial, treatment effectiveness is assessed by comparing the response of the treated subjects to a control group that is selected based on its similarity to the treated subjects in terms of the observed baseline characteristics (Hansen 2007, 18). Matching techniques are used to balance the distribution of the observed covariates between the treatment and control group. However, matching on the observed covariates of the study objects – contrary to the random assignment in a controlled experiment – does nothing to balance the distribution of the *un*observed covariates as well. Insofar as the unobserved covariates influence the outcome variable, an observational study is always liable to bias due to some unmeasured characteristics (Rosenbaum 2010, 73-76).

In a well-designed observational study, sensitivity analysis is used to assess just how sensitive the results are to potential bias from some unobserved covariate. In a first step, the statistical tests are carried out as in a randomized experiment, assuming that there is no bias stemming from unobserved covariates. Under this so-called "naïve model", individuals who *look* similar (i.e. have similar observed covariates) are held to *be* similar (i.e. have similar unobserved covariates as well). In a second step, a sensitivity analysis relaxes this assumption, allowing for the fact that individuals who look similar may not be similar (i.e. they may differ in the unobserved covariates), or the naïve model to be false. The purpose of the sensitivity analysis is to get an idea of how large the bias from some unmeasured covariate would have to be, "to materially alter a study's conclusions." (Rosenbaum 2010, 70-76).

Propensity score matching

In an observational study, treatment and control groups are not normally pre-selected prior to program implementation. If treatment and control groups thus differ at the outset of program implementation in significant ways, matching techniques are a tested method to make the two groups comparable (Hansen 2007, 18). Propensity score matching (Rosenbaum & Rubin 1983) is one such technique that has enjoyed increasing popularity in recent years.

The propensity score is the conditional probability for an individual to receive the treatment given their observed covariates $\mathbf{x}$. In formal terms, the propensity score $e(\mathbf{x})$ is the conditional probability of receiving the treatment ($Z = 1$) given $\mathbf{x}$, or $e(\mathbf{x}) = \Pr(Z = 1|\mathbf{x})$. In a controlled experiment, the propensity score is known and identical for all study objects because of randomization. In an observational study, on the other hand, the propensity score varies between individual study objects and normally remains unknown. However, since it is defined based on observable parameters, namely the treatment assignment Z and a study object's observed covariates $\mathbf{x}$, the propensity score can easily be estimated from those values (Rosenbaum 2010, 72).

In an observational study, the propensity score is most useful because of its tendency to balance the observed covariates between treatment and control groups (Rosenbaum & Rubin 1983; Rosenbaum 2010, 72). The balancing property of the propensity score is due to the fact that if two study objects, k and l, who have identical propensity scores $e(\mathbf{x}_k) = e(\mathbf{x}_l)$ but only one of whom received the treatment $Z_k + Z_l = 1$, are matched, the observed covariates within this pair will be similar and independent of the treatment assignment (Z_k, Z_l). If many such pairs are formed, the distribution of the observed covariates will be approximately the same for both the treatment and the control group, although the specific covariate values of the two subjects within each pair will typically not be identical ($\mathbf{x}_k \neq \mathbf{x}_l$). In other words, even though it is difficult, or neigh impossible, to match individuals on several covariates at once, it is relatively straightforward to match pairs of two subjects on one parameter, the propensity score $e(\mathbf{x})$, and doing so will tend to balance the other observed covariates $\mathbf{x}$ across treatment and control groups (Rosenbaum 2010, 72f.).

The propensity scores are estimated by means of a logistic regression of the pooled sample of the treatment and control group. The dependent variable of the model `treat` takes on the value of 1 for subjects in the treatment group ($Z = 1$) and 0 for the control group ($Z = 0$). The independent variables of the logit model are the observed covariates $\mathbf{x}$ on which the two groups are to be matched (Rosenbaum 2010, 240). For the present study, the independent variables used are a respondent's gender, age category, education, and the type of neighborhood of their area of residence. A fifth binomial variable `educmiss` was added to the model to account for individuals with missing values on education. (The other independent variables had no missing values.).[25]

The logistic regression model to calculate the propensity scores was estimated using the generalized linear model `glm` function in R. A respondent's estimated propensity score $\hat{e}(\mathbf{x})$ is the fitted value of the logit model for the pooled sample.

```
> ps<-glm(treat~gender+age.cat+educ+educmiss+ntype,family=binomial,data=temp)
```

In order to match up study objects who received the treatment with suitable controls, a distance matrix of the individual propensity scores has to be calculated and the study objects with the smallest distance set up in paired groups. The logic behind this matching exercise is that survey respondents with similar values on the independent variables will have similar estimated propensity scores. The calculation of the discrepancy matrix of the propensity scores is also the moment to impose special restrictions on the matching process such as a constraint that matching may only occur within specific subclasses of the pooled sample (Hansen 2007, 19). This is usually done by substituting the distance between two study objects with an arbitrarily high value so as to preclude the matching of unwanted pairs (Rosenbaum 2010, 168f.). As a first step, a boxplot of the propensity scores of the treatment and control groups is drawn to check whether the range of estimated values overlap sufficiently for the matching to be possible at all (Fig. 3.4; Rosenbaum 2010, 240f.).

For the matching itself, there are several strategies. In optimal pair matching, members of the treatment group are matched up pairwise with exactly one control (Rosenbaum 2010, 173-175). However, if the distance metric for matching is the propensity score, a fixed one-on-one ratio to form pairs is unduly restrictive. Subjects with large propensity scores tend to fall into the treatment group, whereas those with smaller values are more often control (since the logit model with Z as its dependent variable attempts to predict zeros and ones). The matching algorithm thus needs to be more flexible, lest some subjects will be matched poorly or not at all. With propensity scores as the distance metric, the algorithm must allow flexible ratios, producing 1:1 ratios only where $\pi(\mathbf{x})/(1 - \pi(\mathbf{x})) \approx 1$, whereas $1 : k$ matches are made where $\pi(\mathbf{x})/(1 - \pi(\mathbf{x})) \approx 1 : k$, all the while allowing l:1 ratios where the proportion is $\pi(\mathbf{x})/(1 - \pi(\mathbf{x})) \approx l \geq 2$. This so-called full matching generally produces better matched groups than pair matching if in some regions of covariate space

[25] A common way to estimate $\hat{e}(\mathbf{x})$ when some covariates have missing values is to (i) replace for each covariate all missing values with an arbitrary but fixed value, (ii) add a binomial variable that takes on the value of 1 if the original covariate value was missing and 0 otherwise, and (iii) estimate the propensity scores in the usual way by using a logit model that includes both the amended covariate plus the new binomial variable. The dummy variable indicating missing values ensures that the arbitrary values plugged in to replace a covariate's missing values do not affect $\hat{e}(\mathbf{x})$, even though they do affect the model's coefficients (Rosenbaum 2010, 193f.).

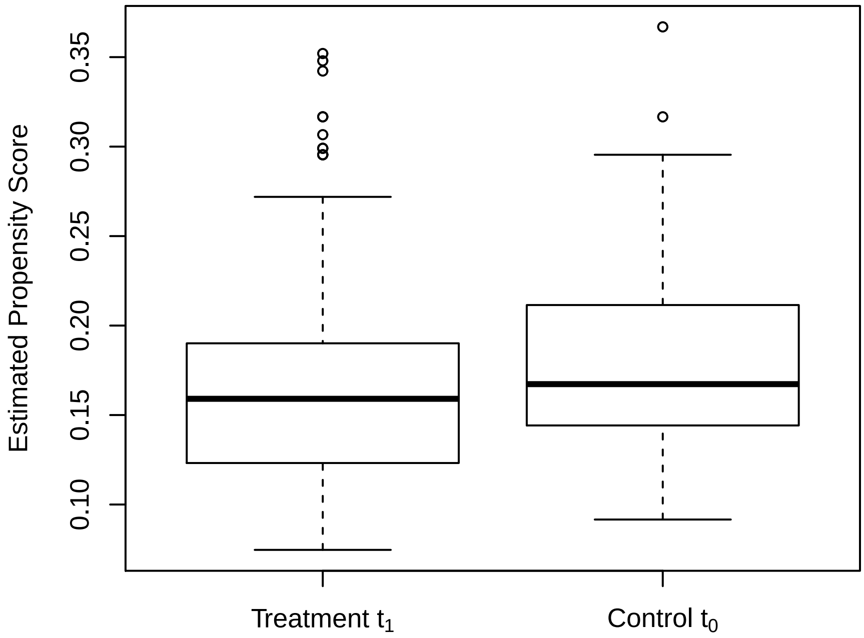


Figure 3.4 Estimated propensity scores.

treatments outnumber controls or vice-versa (Hansen 2007, 21). Full matching, however, is also slightly more complicated, because the individual responses must be weighted to reflect the varying proportions of treatments and controls within each matched group (Rosenbaum 2010, 179-183).

The matching for the present study was computed entirely in R using the 'optmatch' package (Hansen 2007). The distance matrix of the propensity scores was computed using the `pscore.dist` function, which was parameterized only to allow paired groups of respondents living within the *same* neighborhood cluster `ntype` within each city. Individual study objects with similar propensity score estimates were matched in paired groups using the `fullmatch` function with no constraints on the ratio between treatments and controls imposed. Finally, the paired groups were aggregated using the `aggregate` function, whereby, for each paired group, the median value of the outcome variable for all the treated and control subjects was retained as the response to the treatment and the control condition, respectively.

```
> dmat<-pscore.dist(ps,structure.fmla=treat~ntype)
> match<-fullmatch(dmat)

> matched<-aggregate(list(rp=temp[,1]),
+                    list(set=temp$sets,treat=temp$treat),median)
```

Statistical test – Wilcoxon's signed rank statistic

The statistical test applied to determine whether response patterns differ significantly between treatment and control groups is Wilcoxon's signed rank statistic (Wilcoxon 1945). Wilcoxon's test is based on the rank ordered difference in some (outcome) characteristic between study objects belonging to two separate groups. Unlike the more common t-test, Wilcoxon's signed rank statistic can be applied to ordinal and non-normally distributed variables alike, and it is more robust in the face of outliers than mean-based tests, which accounts for the test's enduring appeal in the field of evaluation research (Rosenbaum 2010, 36; 258; Siegel 1997, 73-80).

To calculate Wilcoxon's test statistic T, the difference in the median response between treatment and control group $Y_i = r_{Ti} - r_{Ci}$ is calculated over all the matched paired groups $i = 1, 2, \ldots, I$.

First, the absolute values of these differences $|Y_i|$ are taken, ordered from smallest to largest value, and assigned a rank number from 1 to I. Wilcoxon's signed rank statistic T then is the sum of those rank numbers for which Y_i is positive (Rosenbaum 2010, 36):

$$T = \sum_{i=1}^{I} sgn(Y_i) * q_i, \quad \text{where} \quad sgn(a) = \begin{cases} 1, & \text{if} \quad a > 0 \\ 0, & \text{if} \quad a \leq 0, \end{cases} \tag{3.9}$$

and q_i is the rank of $|Y_i|$.

The T-statistic and related p-value for the current study were computed using the `wilcox.test` function in R. All Wilcoxon tests were carried out as one-sided tests with a 5-percent significance threshold.

```
> T<-wilcox.test(matched$rp~matched$treat,alternative=c("less"))
```

Sensitivity analysis

If the T-statistic rejects the null hypothesis of no treatment effect, a sensitivity analysis is carried out to determine how sensitive this result is to bias from some unmeasured covariate. A sensitivity analysis determines how large such a bias would have to be to nullify the significant result of the test. Not unlike a power analysis of a t-test, the sensitivity analysis thus provides a measure of the confidence researchers can have in their results (Rosenbaum 2010, Chapter 14).

The sensitivity analysis calculates the confidence interval of the Wilcoxon T-statistic under the assumption that the naïve model may be false, i.e. that the odds of receiving the treatment may differ for two individuals with the same observed covariates due to some unobserved covariate. The sensitivity analysis model holds that when two subjects, k and l, with identical observed covariates $\mathbf{x}_k = \mathbf{x}_l$ are paired and one of them receives the treatment, whereas the other serves as control, then the probability that k is treated and l the control is $\pi_k/\pi_k + \pi_l$. In a randomized trial, this probability is always 1/2 for $i = 1, \ldots, I$. In an observational study, by contrast, the probability of assignment to treatment is typically unknown and varies from pair to pair but is bounded by a parameter Γ (Rosenbaum 2010, 77-79):

$$\frac{1}{1+\Gamma} \leq \frac{\pi_k}{\pi_k + \pi_l} \leq \frac{\Gamma}{1+\Gamma} \tag{3.10}$$

A sensitivity analysis shows how a study's inferences change in response to bias resulting from some unmeasured covariate u as the size of this bias increases. In a paired observational study matched for the observed covariates $\mathbf{x}$, the T-statistic and related p-value of the Wilcoxon test are correct as long as $\Gamma = 1$. However, if $\Gamma > 1$, i.e. if the odds of treatment differ for two individuals with the same estimated propensity score due to a failure of the study to control for u, the confidence interval of the T-statistic and related p-values widens. For instance, a value of $\Gamma = 2$ implies that one individual is twice as likely to receive the treatment as another with the same observed covariates due to some unobserved difference between the two. As the size of this bias increases ($\Gamma \to \infty$), the upper bound of the interval of possible p-values of the T-statistic will eventually exceed the 5 percent significance threshold, at which point the null hypothesis of no treatment effect ($H_0 : \mathbf{r}_T = \mathbf{r}_C$) can no longer be rejected. In other words, every observational study is sensitive to sufficiently large biases resulting from a failure to control for some unobserved covariate. A sensitivity analysis simply shows how large this bias would have to be to nullify the study's results (Rosenbaum 2010, 76-79).

For the current study, the sensitivity analysis was programmed in R using as the template the code proposed by Rosenbaum (2010; Appendix 3.9).

3.3 Plan of Analysis

Following this brief description of the data analytical tools employed in this study, it is now possible to give a detailed plan of analysis for the impact evaluation of community policing.[26] The individual steps of the planned analysis are as follows:

1. Develop a typology based on the ecological data to classify neighborhoods across the five cities. This first step relies on the clustering algorithm using self-organizing maps and random forests described in Chapter 3.2.5 for dimensionality reduction and feature selection of the high-dimensional data in order to develop a parsimonious model with high predictive accuracy. Make sure that the developed neighborhood typology explains a maximum of the between-cluster variance in the outcome variables prior to program implementation and, ideally, that there is no significant amount of within-cluster variance left between cities. This is done to ensure that observed differences in the outcome variables are not due to differences in neighborhood structural characteristics or important differences in popular sentiment at the outset of program implementation.

2. Pool individual survey respondents by the neighborhood clusters developed at step 1 for each city. Within each pooled group, use propensity score matching to juxtapose individual survey respondents from the samples taken at time t_0 and t_1 in order to minimize the risk that the apparent differences in response patterns over time are due to changes in the composition of the survey samples.

3. On the pooled matched survey samples, run Wilcoxon's rank sum tests to determine whether aggregate response patterns differ significantly between t_0 and t_1. If Wilcoxon's test reveals no significant differences: STOP. In this case, there is no evidence in the data to suggest that community policing had any significant impact on a particular neighborhood cluster in a given city for the time interval under investigation.

4. If the results are significant, carry out a sensitivity analysis to determine to what extent Wilcoxon's T-statistic is liable to bias introduced by some unmeasured covariate.

5. Finally, assess the plausibility that significant changes in the outcome variables are due to community policing by comparing the outcomes within a given neighborhood cluster across cities. The bigger the difference in the "dosage" of community policing between cities during the observed time interval as unearthed by the preceding process evaluation, the more plausible it appears that the observed differences are due to community policing.

This five step procedure is to be applied consistently for all comparisons, which are carried out over the 2000-2005 study period. The Swiss Crime Survey sample from 2000 serves as the baseline data t_0 for the outcome variables and to develop a neighborhood typology, and the 2005 data serve to evaluate the impact of community policing at t_1.

3.4 Validity and Reliability

Traditionally, the criteria used to assess the quality of an evaluation design were internal and external validity (Campbell & Stanley 1963). Cook & Campbell (1979) added statistical conclusion validity and construct validity. These four categories still are the benchmarks of the quality of a research design and the validity of inferences about program impact (Farrington 2003; Shadish *et al.* 2002).[27]

[26] The idea to draw up a plan of analysis was inspired by Rosenbaum (2010, Chapter 19).

[27] Section 3.4 is a thoroughly updated and expanded version of the corresponding section of the author's unpublished Master's Thesis (Kreis 1999).

3.4.1 Construct Validity

Construct validity is concerned with the quality of the match between the elements of an empirical study – the study objects, settings, treatments, and outcomes – and the constructs used to characterize and measure them (Shadish *et al.* 2002, 64f., 72). A construct valid design "allows reliable measurement of the intervention and all its functional components" and records how closely the program in practice reflects the core elements of the underlying theory. It can be clearly established whether the observed effects are the result of the intervention as a whole or just a limited number of program activities (Lurigio & Rosenbaum 1986, 30-32).

Threats to construct validity basically concern either the explication of constructs or problems related to sampling and measurement. The inadequate explication of constructs or the confounding of different constructs can lead to false inferences about the elements of the empirical investigation. The biggest pitfalls with regard to measurement are the use of only one operationalization per construct, relying on a single method of measurement, and the failure to recognize that a multi-faceted construct may only have been realized in part in a given context (Shadish *et al.* 2002, 72-76).

Evaluations of community policing have notoriously low construct validity, especially as far as the operationalization of treatment is concerned. In their systematic review, Sherman & Eck (2002, 301f.) found that most of the policing evaluations they analyzed largely failed to measure police activities, and those that did often applied doubtful methods. Most damagingly, when programs showed no impact on crime, this measurement gap left evaluators unable to say whether police activity was ineffective or insufficient. In a later review, Weisburd & Eck (2004, 52) concluded that the range of policing tactics falling under the community policing umbrella was so broad, that "its effect as a general strategy [could not] be evaluated." Furthermore, community policing strategies had evolved over time so that it was often hard to determine whether researchers studying it in different places or at different times were studying the same thing.

Advocates of community policing had long acknowledged this fact but reached diametrically opposite conclusions. Kennedy & Moore (1995, 285) argued that the procedural straightjacket required for strong evaluations may even be detrimental to most community policing interventions, which often do undergo transformations during implementation. If the police change tactics not just from one area to another but also over time, an evaluator has his work cut out. He or she may draw a "sketchy picture" of the overall process of implementation, but the evaluation can only assess the impact of community policing as a whole not of its individual operational interventions.

The current study thus took pains to document the process of community policing implementation – the "treatment" – at a level of detail that is appropriate for the following impact analysis. This was all the more important since the process evaluation has been carried out *ex post facto* and hence could only analyze organizational reforms, the forging of strategic partnerships, and strategic initiatives; it could no longer measure the *outputs* of police reforms such as patrol activities or actual problem-solving. In order to tackle the challenge of the adequate explication of constructs, six evaluative dimension of community policing have been defined, and for each of these a list of indicators has been drawn up as a benchmark against which to measure progress. This information has been coded on a time scale that draws inspiration from the "Calendar of Action" approach that allows the evaluator to study program implementation and intensity over time (Crawley & Hope 2003, 11; Hope *et al.* 2004, 7). Moreover, the current evaluation studied community policing implementation across five different cities. This not only meant that the adequacy of the evaluative dimensions was tested repeatedly, the multi-site design also allows for the analysis of the impact of different levels of implementation or "dosages" of treatment. Finally, to confront the issue of mono-method bias, the author complemented the intelligence gathered through interviews with information from written documents and met all his interview partners for a second and third time to discuss the draft reports and the final implementation scorecards, respectively.

With regard to the expected effects of the treatment, a series of research hypotheses have been formulated on the basis of previous studies of community policing implementation and impact. These hypotheses are sufficiently precise to guide the analysis of both the process of community policing implementation and the impact on neighborhood residents.

The current evaluation must confront the conceptual question, though, to what extent it is valid to operationalize urban neighborhoods as postal ZIP code districts. The problem is aggravated further since the study does not apply the same unit of analysis consistently, alternating between postal ZIP code districts and OFS census tracts. Shifting units of analysis can influence the results of area-based statistical analyses, a phenomenon well-known as the modifiable areal unit problem (Openshaw 1984). In practice, though, the question is moot because the current study is obligated to work with standing data sources. A respondent's ZIP code is the only geo-reference recorded in the SCS data base, whereas official crime records are mostly based on census tracts. Any neighborhood-level analysis of the data forcibly has to rely on these areas as unit of analysis (Kreis 2009).

3.4.2 Internal Validity

Internal validity refers to the quality of a research design to allow "inferences about whether the observed covariation" between the treatment and outcome "reflects a causal relationship." To support such an inference, a study must demonstrate precedence in time of cause to effect, covariation, and the absence of plausible alternative explanations. From this it follows almost automatically that threats to internal validity are all other extraneous factors that could have caused the observed variation in the dependent variable in the absence of treatment (Shadish *et al.* 2002, 53f.).

Arguably the most formidable threat to the internal validity of the current study is *selection.* Selection bias refers to any preexisting systematic differences between the study objects that could have produced that same observed effect. Selection bias is a threat to any evaluation without random assignment (Shadish *et al.* 2002, 55f.); and it is particularly acute in neighborhood-level research (Sampson *et al.* 2002, 466).

A set of three related threats to internal validity are history, maturation, and regression. *History* refers to all events other than the treatment that occurred between the pretest and posttest measurement that could have affected the dependent variable. *Maturation* refers to gradual changes of the study objects that occur naturally in the absence of treatment and can have a similar effect. The third threat, statistical *regression,* refers to the fact that if study objects display high or low values of the dependent variable, these values may exhibit a tendency to regress to the mean over time irrespective of actual program impact (Shadish *et al.* 2002, 56f.).

A second set of threats to the internal validity of the current evaluation design are testing and instrumentation. *Testing* refers to changes in the dependent variable that the fact of taking a test or being observed may elicit. *Instrumentation* poses a problem if changes in the measurement instrument between pretest and posttest simulate an effect in the absence of treatment (Shadish *et al.* 2002, 60). Both these threats take on a special meaning if the dependent variable is a survey instrument. Survey questionnaires that measure attitudes, perceptions, or behaviors often prompt answers that are "high in social desirability." (Lurigio & Rosenbaum 1986, 35f.). In addition, the questionnaire format and the ordering of the survey questions can influence response patterns.

As to instrumentation, a potent threat to internal validity for the current study is *sampling* bias. If pretest and posttest groups are repeated random samples of the study population, a research design allows valid inferences about a causal effect only to the extent that the two groups are comparable, for the simple reason that sample selection almost always could have produced the observed effects (Cook & Campbell 1979, 117). Even in the absence of sampling bias, the data may still be skewed due to non-response bias. If respondents randomly selected for the survey cannot be reached or refuse to participate, the sample may suffer from bias as a result.

The current evaluation has taken multiple steps to tackle these threats to internal validity. To cope with selection bias, the study first developed a typology to cluster structurally similar neighborhoods where, in addition, local residents collectively expressed similar attitudes at baseline. At the individual level, the impact analysis relied on propensity scores to match survey respondents with comparable demographic characteristics. Finally, since bias from some unmeasured covariate can never be ruled out in an observational study, all significant test statistics were followed up with a sensitivity analysis to determine just how large such a bias would have to be to nullify these results.

The threats of history, maturation, and regression are commonly countered by means of an adequate comparison group not subject to treatment. However, if a program is implemented across entire cities or jurisdictions, as is the case for the present study, a control group can only come from elsewhere, and the possibility of divergent history remained a valid threat. The current study thus relies on two standard remedies – repeated pretest measurements and multiple non-equivalent outcome variables – to blunt these threats (Shadish *et al.* 2002, 110f.). By tracking several outcome variables across more than 60 neighborhood areas, while controlling for a range of neighborhood-level variables, the risk that some local events, which affect some areas but not others, jeopardize the internal validity of the study design thus appears, if not entirely absent, at least substantially reduced.

As for the final two threats, testing cannot be ruled out categorically for the SCS data, but the threat appears rather small. Over the study period of this evaluation, the wording of the relevant questions has not changed at all, and the sequence of questions has by and large remained the same (Killias *et al.* 2007, 16). The repeated random sampling further precludes that respondents may merely have become more alert to certain treatment activities by a pretest interview, one of the few advantages of the present design over a panel study (Lurigio & Rosenbaum 1986, 36).

In order to correct for sampling bias, weights were calculated for all survey respondents based on 1990 and 2000 census data. Crucially, this weighting procedure managed to correct any gross misrepresentation of the age and gender distributions of the survey sample for all ZIP code areas (cf. Appendix A.1). Regarding non-response bias, the rate of the 2005 SCS survey was 30 percent, well within the range of comparable telephone surveys (Killias *et al.* 2007, 160f.). However, in the absence of any additional evidence on respondents who refused to participate, non-response bias cannot be ruled out completely.

3.4.3 Statistical Conclusion Validity

Statistical conclusion validity refers to the quality of statistical inferences whether the presumed cause and effect do covary and the magnitude of this covariance. A threat to statistical conclusion validity is thus any reason that leads a researcher to draw a false conclusion about this covariance (Type I or II error)[28] or to over- or underestimate its size (Shadish *et al.* 2002, 42). The two most common threats to statistical conclusion validity are the inappropriate use of statistical techniques and problems of variable measurement (Lurigio & Rosenbaum 1986, 39).

The two main threats to valid inferences from the statistical analyses are insufficient statistical power and *violations of the assumptions of the statistical tests* being employed. Most statistical tests make assumptions about the underlying distribution of the data, which may bias results if these are violated. Insufficient statistical power affects the probability of correctly rejecting the null hypothesis of no effect when it is indeed false. The problem of insufficient statistical power becomes the more serious, the smaller the samples and the smaller the (expected) effect of a program (Farrington 2003, 52).

[28]A type I error occurs if an evaluator rejects the null hypothesis of no treatment effect, when it is in fact true (i.e. the evaluator concludes that the treatment was effective, when it had no effect). A type II error is being committed if the evaluator fails to reject the null hypothesis of no treatment effect, when it is in fact false (i.e. the treatment had an effect but the evaluator fails to notice it). The probability of committing a type I error is referred to as the level of statistical significance. In the social sciences, this probability is commonly set to be no higher than 5 percent.

For the present study, the risk of violating the assumptions of the statistical test appears rather small. Wilcoxon's signed rank statistic makes no assumption about the distribution of the dependent variable and can be applied both to ordinal or non-normally distributed variables. The test is highly robust in the face of outliers and, as a non-parametric test, can be calculated correctly even for very small samples (Siegel 1997, 32f.). In addition, if there are no ties, meaning that no two of the matched pairs display exactly the same difference in response between treatment and control, the null distribution of the T-statistic is always the same. It is thus known in advance, which is why the statistic is also called "distribution-free". If there are ties between some matched pairs, which often occurs if the dependent variable is ordinal, the null distribution of the T-statistic depends not only on the number of pairs I but also on the pattern of ties, but it can still be computed correctly (Rosenbaum 2010, 39).

By contrast, the threat of *insufficient statistical power* and thus the risk of committing a type-I or type-II error appears non-negligible. At the city level, the SCS survey samples are sufficiently large, but at the neighborhood level, the number of respondents is usually rather small. The problem of insufficient power is compounded by the fact that standing evaluations of community policing do not suggest a dramatic response to treatment of the general population, at least in the short run. As Ashby (2005, 432f.) remarked, the evaluation of the impact of policing strategies at the neighborhood level is likely to be hampered by the insufficient sample size at this scale of the regular citizen victimization surveys, unless a batch of comparable neighborhoods can be lumped together for the purpose of analysis.

The current study therefore used geostatistical profiling to cluster neighborhoods by typology in order to assess the impact of specific community policing strategies at a local level, which is most useful to identify and disseminate "best practices" across different cities (Ashby 2005, 436; Williamson *et al.* 2006, 213). In other words, the preliminary efforts of this impact evaluation to develop a typology of urban neighborhoods not only serve to reduce the heterogeneity of the study objects receiving the treatment and thus to boost the internal validity of the current study. Pooling survey respondents from neighborhoods that are similar in character also helps address the problem of potentially low statistical power at the infra-city level.

Besides invalid statistical testing, program evaluations may also be undermined by *deficiencies in the measurement* of outcome or explanatory variables. The sections on individual indicators have already dealt with these issues insofar as they can be addressed in an *ex post* study. Other confounding factors such as sample selection or non-response bias, which potentially affect all survey variables, have been discussed as threats to internal validity. The challenge of measuring treatment has been dealt with as threats to construct validity.

Another threat to measurement validity is the displacement of crime. If closer police-community cooperation results in a drop in crime in the treatment area, crime may simply be displaced elsewhere. Displacement means that crime has simply been shifted in location rather than been reduced and may lead to false inferences about program impact (Lurigio & Rosenbaum 1986, 40). The current study, while analyzing crime patterns across time and space, cannot address the issue of crime displacement as a result of the implementation of community policing. However, recent evaluations of problem-oriented policing initiatives have produced a growing body of evidence that prevention initiatives or hot spot policing do not result in the displacement of crime and may even have the opposite effect, a diffusion of crime prevention benefits to neighboring areas (Braga 2005, 330f.; Braga & Bond 2008, 596f.; Eck 2002, 282f.; Weisburd & Eck 2004, 54). Crime displacement thus seems to constitute a rather diminished threat to conclusion validity.

3.4.4 External Validity

External validity refers to the generalizability of the results of an evaluation, "the extent to which a causal relationship holds over variations in persons, settings, treatments, and outcomes." (Shadish *et al.* 2002, 83). External validity refers to the robustness of the findings of an evaluation, the extent to which a specific program would lead to the same results if replicated with a different sample, at a different point in time, or in a different location (Lurigio & Rosenbaum 1986, 37f.).

External validity has never been the chief concern about validity and, with the growth in popularity of meta-analyses, the concept has slipped down further on the agenda of evaluation researchers. Campbell & Stanley (1963), while maintaining that internal validity was the "*sine qua non*" of evaluation and that the question of external validity was "never completely answerable", still argued in favor of "designs strong in both types of validity." (p. 5, emphasis in original). Four decades later, the Maryland Scientific Methods Scale (Farrington *et al.* 2002, 18) treated external validity almost as an afterthought. For Sherman *et al.* (2002, 7f.), the external validity or generalizability of an internally valid evaluation is established through replication. Most recently, however, external validity has again received increased attention, including by adherents of the experimental paradigm, as a means to facilitate the dissemination of practices that have proven their effectiveness in other contexts (Braga 2010, 175f.).

Threats to external validity are generally conceptualized as statistical interaction effects, which, if present, affect the causal relationship between treatment and outcome, leading to changes in the observed effect size. By this notion, threats to external validity are all such interaction effects of the causal relationship with the study sample, variations of treatment, outcome observations, or the study setting (Shadish *et al.* 2002, 86f.)

Program evaluations that compare multiple sites generally have higher external validity by allowing researchers to make inferences not only about the study samples at each site but also about "the more general population that these sites represent collectively." (Lipsey *et al.* 2006, 289). The external validity of the present evaluation is thus likely to be enhanced by a research design that compares treatment outcome across more than 60 neighborhood areas, which vary noticeably in demographic composition, economic well-being, and encompass suburban and downtown areas. These neighborhoods are embedded in a larger context of five urban areas, of varying size, spread across two linguistic regions, two of which are situated close to a national border. This research design should allow the evaluator not only to assess the impact of community policing on residents in a variety of neighborhood settings but, by developing a typology of neighborhoods, also to determine in what context different policing strategies have proven particularly effective and where else they may be implemented successfully (cf. Ashby 2005, 429).

Nevertheless, the scope for policy recommendations depends on the degree of construct validity of the current research design. If the operationalization of community policing across the five urban areas remains insufficiently precise, the external validity of the findings only goes so far. This underlines the inextricable link between construct and external validity of any evaluation design (Lurigio & Rosenbaum 1986, 38f.).

3.5 Defining Community Policing Success

In the current drive for "evidence-based" criminal justice programs and policies, crime prevention is assessed "not by its intentions, but by its consequences." (Sherman *et al.* 2002b, 3). The question of whether community policing "works" thus implicitly asks whether it serves to prevent crime. Crime rates hence figure prominently among the outcome measures in a community policing evaluation; they are not the only important indicator, however (Kennedy & Moore 1995, 281).[29]

Indeed, to paraphrase Kennedy & Moore (1995, 281f.), community policing has opened a normative debate on the ends of policing that expanded thinking about the police function to include, at the very least, combating fear and disorder as critical elements. Community policing has also rekindled thinking about the values of accountability, responsiveness, economy in the use of force and authority, freedom from corruption and abuse, adaptability, and the acceptability of police

[29] Section 3.5 is a thoroughly updated and expanded version of the corresponding section of the author's unpublished Master's Thesis (Kreis 1999).

behavior to communities. An assessment of community policing impact must reflect this. Therefore if community policing does not reduce violent crime but lowers levels of fear and disorder, it is not an unqualified success but worth having nonetheless. Equally, if community policing is no more effective than the reform era model at preventing crime, but does so in a manner more responsive to community concerns, less alienating to minorities, or less productive of police brutality, it realized a net gain both to society and the police.

More concretely, community policing success has to be examined in terms of police effectiveness, efficiency, and equity. These are the three dimension of a conceptual framework that allows evaluators to set benchmarks of community policing progress that can be verified empirically. These three concepts have already been applied to assess earlier models of policing, but with community policing their meaning and the indicators used to measure them have evolved (Eck & Rosenbaum 1994, 5f.).

An effective community policing strategy will reduce crime, lower fear of crime and disorder, and improve overall quality of life in an area. Since providing quality service is an important goal, citizen satisfaction becomes an important indicator of police effectiveness (Bureau of Justice Assistance 1994, 45). In the context of the present study, community policing may be deemed effective if official crime records indicate stable or falling crime rates over time, and survey respondents are less concerned about fear of crime and neighborhood disorder, and generally rate the quality of police services as high or as significantly higher than before.

Efficiency, the second criterion of success, means getting a maximum of benefits from the available resources. The crucial question facing police administrators and politicians alike, particularly during times of fiscal austerity, is "whether they can afford community policing and whether it is worth the investment." This issue is more acute because the implementation of community policing requires additional resources, at least in the short run (Rosenbaum 1998, 19f.). Since the current evaluation is not concerned with the costs or the cost-effectiveness of community policing, the question of efficiency is not elaborated any further.

Equity, the third criterion for appraising success, is arguably the most important for community policing to succeed. For community policing activities, equity has three different dimensions: "equal access to police services by all citizens, equal treatment of all individuals under the [law], and equal distribution of police services and resources among communities." Police must not give preference to one community over another; however, equity at times requires that police target interventions and dedicate a larger share of police resources to neighborhoods suffering higher rates of crime or unusual levels of fear (Bureau of Justice Assistance 1994, 49-51). From an equity perspective, community policing has thus succeeded if the indicators measuring effectiveness show comparable levels for different communities and neighborhoods, or if pre-existing gaps have narrowed substantially over time.

4

Process Evaluation

4.1 Community Policing Implementation

4.1.1 Organizational Strategy and Partnerships

In the decade between 1994 and 2004, the police departments of Basel, Bern, Geneva, Lausanne, and Zurich all undertook major reforms of their internal organization and operating strategies to implement community policing (Bolle & Knoepfler 2000a; Bolle & Knoepfler 2000b; Sauter *et al.* 2005).[30] The era of community policing in Switzerland began in 1994, when the Geneva cantonal police's gendarmerie executed the strategic plan "Pégase", which in addition to reforming the police organization also instituted the first community outreach or neighborhood liaison officers (*îlotiers de quartier*) in an inner city police station. Two years later, the Municipal Police Department of Lausanne adopted its own community policing strategy, which was substantially expanded in 2001.

In the German part of Switzerland, the cantonal police in the City of Basel and the Municipal Police Department of Bern adopted the reform agendas "4plus" and "Apollo" in 1998 and 1999, respectively, and both began implementing far-reaching organizational reforms in view of a department-wide transition to community policing. In Zurich, the municipal police department also in 1999 began revamping its network of neighborhood police stations in order to boost both its emergency response capabilities and police-citizen contacts. This reform agenda, dubbed "Phoenix Q", stopped short of fully embracing the community policing agenda, however. Five years on, in 2004, the Municipal Police Department of Zurich implemented another major reform of their organization and operating procedures under the name "Stapo 200X" in view of a full-blown transition to community policing.

Since then, the five police departments' reform efforts have continued apace. In 2005, the Geneva cantonal police department, after police internal wrangles had sapped the strength of its existing strategy, officially launched its reform agenda "Proxipol" to reinvigorate its community policing efforts. In Bern, the municipal police department was dissolved at the beginning of 2008 and merged with the cantonal police. The merged force kept intact the community policing strategy and organizational structure of the former municipal police for its newly created policing region "Bern". In Basel, the cantonal police replaced their agenda "4plus" with a new strategy "Optima" in October 2008, cutting their network of four precinct police stations down to two in response to budgetary pressures.

[30]The first part of this chapter is an extended summary translation of the separate reports on the community policing implementation process in the five cities covered by this comparative study. It gives a detailed overview of local developments regarding the elements of community policing chosen as indicators to track implementation progress by the process evaluation. The more detailed individual progress reports on the cities of Basel, Bern, and Zurich in German as well as on Geneva and Lausanne in French were drafted in 2009 and early 2010 and are included as appendices to the final research report submitted to the SNSF in May 2010 (Kreis *et al.* 2010).

Training

In the wake of the organizational reforms, the five police departments have begun to instruct police recruits in community policing. The Lausanne municipal police was the first force to introduce such instruction for all its recruits during basic training in the late 1990s. In Basel, the cantonal police introduced a course in community policing of one and half days in the basic training program for all their recruits beginning in 2001, which was extended to five days in 2003. In the City of Bern, the police in 2003 ran a pilot project using the community policing teaching module of the Swiss Institute of Police (ISP) in Neuchâtel and kept this format thereafter. In Geneva and Zurich, community policing was introduced into basic training only with the advent of a federal examination for police officers. Since 2005, all police recruits upon completion of the basic training at their local or regional police academy take a national examination administered by the ISP in order to obtain a federal diploma as police officers, in which community policing is one of the examined subjects.

With the introduction of the national examination for police officers, the basic training in community policing has been harmonized across the four police academies.[31] In an instance of national cooperation, the major cantonal and municipal police departments in Switzerland jointly drafted a community policing teaching manual for this purpose (Spaar *et al.* 2007). Police recruits nowadays generally receive two week's worth of instruction in community policing during their first year of basic training: one week of theoretical instruction in the classroom and one week of applied practice, typically in the form of mock police interventions with actors staging real-life situations or on-site visits of the premises of partner organizations of each police department.

The continued training of police officers in community policing, however, has not yet been harmonized. In Basel and Bern, the neighborhood liaison officers regularly organize training seminars in the community policing strategy for their colleagues. In Zurich, the police distributed their internally drafted community policing manual to all staff and declared this document mandatory reading but, coming at a time of budgetary cuts, the message did not always fall on fertile ground. For the time being, thus, except in Basel and Bern, continued training in community policing has not yet been institutionalized or remains a work in progress, as some of the strategy documents recount and police administrators acknowledged.

The neighborhood liaison officers, by contrast, receive specialized training to prepare them for their future community outreach and problem-solving duties in addition to the community policing instruction they have received during basic training. In Geneva, the neighborhood liaison officers have received specialized psychological training in communication, negotiation, and conflict resolution since the plan "Pégase" created this office in 1994. This course at first lasted for two days and has since been extended to three days. The Geneva NLOs also convene several times per year to discuss experiences and exchange best practices. In Basel, the NLOs took a series of training seminars at the police academy in nearby Freiburg im Breisgau in Germany. The Basel NLOs also meet monthly to exchange views and strategies and regularly hold conferences with speakers from within the police or from partner organizations. In Bern, the three NLOs and their supervisor, the community policing coordinator, meet weekly and organize a training seminar of half a day at least once a month to stay abreast with recent developments within the city and its environment and to get to know their strategic partners.

In Lausanne and Zurich, which have no specialized NLOs, community outreach activities are incumbent on the commanders of the neighborhood contact offices. These police cadres receive no specialized instruction in community policing as part of the training they undergo to prepare them for their role as police administrators, even though in Zurich, they are briefed on the strategy by the head of the specialized Crime Prevention Unit. In Lausanne, three strategically placed police cadres have been immersed in the community policing strategy. These cadres now teach the subject at the regional police academy and act as multipliers within the Lausanne municipal police.

In Basel and Bern, standing and future police cadres receive additional training in community policing. In Geneva, police cadres may take the psychological training classes dispensed to the neighborhood liaison officers on a voluntary basis, but few actually do so.

[31]The Cantonal Police Departments of Basel-City and Bern are part of a consortium of cantons which operate a joint police academy in Hitzkirch in the Canton of Lucerne.

Cooperation and partnerships

In accord with the problem-solving strategy of community policing, the five police departments have with varying degrees of success forged strategic partnerships with other branches of the local administration, neighborhood associations, local business interests, and private agencies.

First, the police departments implemented organizational reforms to improve cooperation between the different divisions within the police organization itself. In Geneva, the NLOs were freed from the hierarchical chain of command of their police station and vested with the power to call in intervention units to their patrol beats. In Lausanne, the municipal police strengthened cooperation between the intervention units, the investigative divisions, and the staff of the neighborhood contact offices. In Basel and Bern, the police spread their staff from all divisions equally over four and three precinct police stations, respectively, which were supposed to run each policing precinct autonomously.

Next, the police departments strengthened existing bonds and created new partnerships with other branches of the local administration. The strategic partnerships between the police, the social welfare, and the public health departments typically had been established to implement the four-pillar federal drug policy and for the most part predated the implementation of community policing (Ernst *et al.* 1999a; Ernst *et al.* 1999b). The police departments started to cooperate with the local park services, garbage collection, and the planning and urban development departments to run campaigns to clean up parks, streets, and public places and to remove graffiti from public buildings and private property. These partnerships often began as pilot projects and were institutionalized a couple of years later. Finally, the police departments sought to establish links with neighborhood cultural centers and a city's public schools in order to regulate the modes of cooperation and police interventions in case of legal transgressions and to allow the police youth specialists to brief middle-school aged pupils on the risks of delinquent behavior.

The five police departments have also established regular contacts with the neighborhood associations in each city. These contacts typically occur in the form of a police representative attending events or assemblies hosted by these associations or the police themselves organizing round table meetings to discuss security concerns. In all five cities, the NLOs or the commanders of the neighborhood contact offices regularly convene or participate in roundtables of the actors with a stake in public security within the different police precincts of their city.

The five police departments have striven to forge partnerships with private partners as well. In Geneva, one of the first tasks of the NLOs was to establish a registry of all businesses in their patrol beats. In Basel, the cantonal police on a voluntary basis have established a database of the owners of businesses that allow them to identify different groups and target crime alerts specifically. In Lausanne, the municipal administration signed a charter with the owners of nightclubs to facilitate cooperation between their security services and the police.

The five police departments have also invested in public relations efforts to enhance police-community relations. All five police departments regularly set up stands at trade fairs and allow public visits of police facilities to present the police organization to the general public. As the last of the five departments, the Lausanne Municipal Police Department in 2000 hired a communications specialist and created a press and communications unit to handle relations with local media and the press. More recently, the five police departments have also made increasing use of the internet to distribute information about the police to a wider audience.

Objectives of community policing

As part of the transition to community policing, all five police departments have drafted a strategy paper to guide the reorganization process (Baer 1993; Hagenlocher 1996; Kantonspolizei Basel-Stadt 1998; Stadtpolizei Bern 1999; Müller 2006). These community policing manuals have been distributed to the rank and file of the police departments, except in Lausanne where the strategy paper was distributed only among police administrators.

These community policing strategy papers explained the precepts underlying the new strategy and, in the case of Basel and Zurich, explicitly referred to criminological theories, namely the broken windows hypothesis. The theoretical underpinning of the strategy papers varied; Hagenlocher (1996) was largely inspired by Jean-Pierre Harvin, the spiritual father of community policing in neighboring France, who visited the City of Lausanne on more than one occasion. The community policing manuals in Basel and Zurich drew heavily on the community policing handbook that Bässmann & Vogt (1997) drafted on behalf of the Federal Criminal Police Office (*Bundeskriminalamt*) of Wiesbaden in Germany. For the practical know-how of community policing implementation, the police departments of Geneva and Lausanne primarily looked to France and Quebec, whereas their colleagues in Basel, Bern, and Zurich took inspiration from models practiced in the Netherlands, particularly Amsterdam.

The community policing strategy papers of the five police departments invariably declared the reduction of crime and fear of crime as objectives of the new strategy. As to their operational strategy, the five police departments all vowed to forge strategic partnerships with public and private actors in the area of public safety and to establish regular contacts and enhance relations with the community.

4.1.2 Decentralization and Deployment

Patrol districts and police stations

As part of these strategic reforms, the police in the three Swiss German cities have substantially altered their spatial organization at the end of the 1990s. In Basel, the police department in 1999 in the context of its strategic reform “4plus” shut down its extensive network of 16 neighborhood police stations and newly divided the city territory into four police precincts. The staff of all but the investigative division was stationed equally over the four precinct police stations. In Bern, where police had maintained a neighborhood police station in each of the city’s six administrative districts for decades, the police as part of their reform program “Apollo” compartmentalized the city into three police precincts and concentrated their manpower on the three corresponding precinct police stations. In both Basel and Bern, the newly created precinct police stations held far-reaching administrative and operational autonomy. In Zurich, the municipal police department through its reform agenda “Phoenix Q” gathered all its interventionist forces in five strategically located district police stations, which assumed responsibility for all motorized patrols and answered emergency calls. Contrary to Basel and Bern, however, the Zurich police did not do away with their network of existing neighborhood police stations but, after scaling them down and reducing opening hours, maintained them as neighborhood contact offices.

The layout of police districts has continued to evolve in these three cities over the past decade. In 2004, the Zurich Municipal Police Department for the first time divided the city territory into two policing precincts for the provision of all basic services. In Bern, as part of the merger between the municipal and cantonal police departments, the three existing police precincts of the city were enlarged in 2008 also to encompass the suburban communities, and were reshaped again in early 2010. In Basel, the cantonal police department in October 2008 retrenched its forces from their four precinct police stations to only two, in an effort to optimize its resource allocation across the city.

In Geneva and Lausanne, by contrast, the advent of community policing has barely altered the lay out of police districts or the deployment of police personnel. The Geneva cantonal police maintained its policy of stationing the agents of the *gendarmerie* (patrol division) at the one-dozen police stations spread all but one across the City of Geneva and the surrounding suburban communities. In Lausanne, the municipal police have also stuck to their guns and kept their interventionist staff stationed at their headquarters near Saint-Martin in the city center, detaching only a small batch of police officers to the half-a-dozen neighborhood contact offices strategically implanted near the city’s hotspots. The Lausanne police have closed two neighborhood contact offices in the city center and opened a new one instead in the Flon neighborhood to keep up with the area's rapid transformation from an industrial backyard to an uppity shopping and entertainment hub.

Foot and bike patrol

In terms of patrolling activities, all five cities have seen a renaissance of the traditional police officer on foot patrol, both as a tool to reassure a worried citizenry and as a tactical tool to control sensitive areas and combat crime hotspots. Foot patrol had been the prevalent form of patrol in Swiss cities until the late 1970s but had gradually fallen by the wayside during the 1980s (Segmüller 2005, 37; Müller 2008, 10). Testimonies from that time show that police administrators were keenly aware of the reassuring effect a police officer on foot patrol had on ordinary citizens but were struggling to use their scarce personnel resources for policing activities higher up on their agenda (Police municipale de Lausanne 1986; Vieli 2006, 20f.). By the mid-1990s, by contrast, the strategy papers the police commanders drafted to prepare for the community policing transition process put renewed emphasis on foot patrol (Baer 1993, 6; Hagenlocher 1996, 36). All five police departments have also established bike patrol units to complement traditional forms of patrol with a vehicle that combines the benefits of rapid displacement with virtually the same easy approachability of an officer on foot patrol. More recently, the cities of Geneva and Zurich have even added police officers patrolling on inline-skates to their tactical arsenals.

Neighborhood liaison officers

In accord with the fundamental precepts of community policing, the police administrators in all five cities have taken steps to foster the dialogue with neighborhood associations, local interest groups, and the general public. The organizational models adopted to this end essentially fall into two categories: in Basel, Bern, and Geneva, where the police staff is more or less evenly split between the few decentralized precinct police stations on their territories, the police departments have designated specialized neighborhood liaison officers, who are based at these precinct police stations and whose principal task is to foster a regular dialogue between the police and the institutions and organized interests in the area. In Zurich and Lausanne, which have concentrated their intervention forces at a few central district police stations but maintain a network of neighborhood contact offices spread over the city territory, the staff at these outposts typically engages in community outreach activities. In these two cities, it is incumbent on the chiefs of the neighborhood contact offices to engage in a dialogue with local institutions and organized interests. Both under the neighborhood liaison officer and the neighborhood contact office model, the police departments as a rule designate older police agents with several years of professional experience to engage in community outreach activities.

In all five cities, the NLOs and neighborhood contact office commanders not only enjoy a high degree of autonomy to go about their business, they also benefit from a position of authority to wield a substantial say on police operations in the area. Generally, it is incumbent upon them to go on foot or bike patrol for some of their duty time and to convene regular round tables with local interest groups. For the NLOs in Basel, Bern, and Geneva, community outreach and problem-solving activities are their primary responsibility, whereas in Lausanne and Zurich the contact office commanders are inevitably absorbed to some extent by their duties as police station managers.

4.1.3 Broken Windows Approach

Disorder and fear of crime

The community policing manuals or strategy papers that the administrators of the five police departments have produced, all acknowledge the special nature of fear of crime as defying simple definition. All describe the elusive nature of the concept that is only loosely related to objective risk of crime but very real to those afflicted by it. The plan “Pégase” in Geneva named rising levels of fear of crime and a deterioration of police-community relations as one of the driving forces urging for reform (Baer 1993, 3f.). In Lausanne, the 1996 strategy paper of the Chief of police also dedicated a chapter to the feeling of insecurity. Distinguishing between objective and subjective insecurity, the white paper deemed the impact of incidents of physical or social disorder as well as rumors of nearby crimes and victims as very real (Hagenlocher 1996, 4-10). In Basel and Zurich, the community policing manuals made explicit reference to the broken windows hypothesis as a valid model of urban development, recounting local experiences that signs of disorder, if they go untended, beckon more disorder. The Basel community policing manual held that many signs of disorder are not directly linked to crime but according to the broken windows hypothesis foster a climate conducive to crime, which is why police must take subjective levels of insecurity into account and step in also against minor infractions to improve the quality of life in an area. However, both the Zurich and Basel police strategy documents unequivocally rejected zero-tolerance policing as an untenable policy against disorder and petty crime (Kantonspolizei Basel-Stadt 2003, 9, 17; Müller 2006, 27f.)

Management of public spaces

The theoretical musings about the impact of environmental factors on fear of crime and public safety have been translated into plans of action. In all five cities, the authorities nowadays consider littering and graffiti as serious nuisances to neighborhood livability and have adopted measures to clean-up streets, parks, and public places. These clean-up schemes invariably are joint efforts of multiple stakeholders, involving several agencies of the municipal or cantonal administration and sometimes private actors.

The role and involvement of the police in these clean-up schemes varies from one city to another. In Zurich in 2003, the Municipal Police Department and the Department for Garbage Collection and Recycling jointly held the lead in establishing a city-wide network of local working groups involving other municipal agencies, neighborhood associations, and small business interest groups. These meet in regular intervals to discuss means and strategies to improve the cleanliness of public space and reduce subjective feelings of insecurity. In Zurich, the authorities have also set up an insurance scheme for landlords to remove graffiti from private property and run a graffiti telephone hotline. In Basel, the police and the department for garbage collection coordinate efforts to combat graffiti and littering, and the cantonal building department runs a public-private partnership to remove graffiti from walls on the city territory. In Bern, the state-run cantonal real estate insurance agency, the city’s urban planning department, and the police in 2005 launched a pilot project to remove illegal graffiti immediately after their appearance, which two years later was turned into a private association. The project was first restricted to the inner city but its domain of activity has since been extended to the entire city territory. In Geneva, the city government in 2009 launched a campaign structured as a public-private partnership to efface graffiti swiftly, beginning in the worst afflicted areas of the inner city.

Besides these elaborate and increasingly institutionalized programs to tackle physical disorder, local authorities in the five cities have also taken steps to alleviate the negative effects of social disorder. In these efforts to minimize the impact of disorderly conduct, there has been a noticeable convergence of ideas over the last decade. All five cities have established new government intervention units who patrol sensitive areas such as parks or public places daily and cooperate very closely with the local police but remain institutionally independent from them. Combining outreach social work with order maintenance duties, these street workers commonly hold no coercive power but try to coax homeless people, alcoholics, or drug addicts to call upon local counseling centers, shelters, or drop-in centers run by local government bodies or supported with public funds.

This policy began in 2000, when the City of Zurich launched a pilot project of a street worker intervention unit called “sip züri” (*Sicherheit-Intervention-Prävention* – security-intervention-prevention Zurich). These street workers patrol sensitive areas daily and step in if large gatherings of marginalized or disreputable individuals form in parks or public places, driving away other potential users. From its inception, “sip züri” cooperated closely with the municipal police, but its staff does not get involved in coercive interventions. The intervention unit is attached to the social welfare department and institutionally independent from the police.

In 2001, the Municipality of Lausanne decided to establish its own interventionist task force modeled on “sip züri” to reduce the effects of harmful behavior of drug and alcohol addicts. The Municipality’s message announcing the creation of “UNISET” (*Unité d’intervention socio-éducative de terrain* – socio-educational street intervention unit) made explicit reference to the Zurich experience (Municipalité de Lausanne 2001, 33). The City of Bern likewise introduced its street intervention task force “PINTO” (*Prävention-Intervention-Toleranz* – prevention-intervention-tolerance) in 2004. Within the municipal administration, “PINTO” forms part of the social welfare department. “PINTO” has emulated the Zurich model, the manager of “sip züri” sitting on its advisory board (Stadt Bern 2005, 7).

In Geneva, the NLOs who began operating in 1994 were initially freed form their coercive duties as police officers, but the Geneva cantonal police has since abandoned this policy. The City of Geneva and other surrounding municipalities have since established municipal security agents (*Agents de sécurité municipale*), who are nowadays close partners of the cantonal police in discharging their order maintenance duties. A 2009 cantonal law reestablished them as municipal policing agents (*Agents de la police municipale*), but they remain unarmed and have limited policing powers.

4.1.4 Problem-oriented Policing

Methods of problem-oriented policing

Since the introduction of the new federal diploma for police officers in 2005, all police recruits are instructed in the problem-solving method SARA during basic training at the police academy. SARA (Scanning – Analysis – Response – Assessment) describes a four-step approach that police officers are instructed to follow, when dealing with a specific problem they are called upon to resolve. As a rule, however, regular patrol officers only take on simple problem-solving tasks. More complex or time-consuming jobs are generally performed by the specialized neighborhood liaison officers.

All five police departments have established a central analysis unit that monitors and analyzes operational data to identify crime hotspots and to provide intelligence to formulate policing strategy. Beginning in the late 1990s, these units started using Geographic Information Systems (GIS) to visualize the distribution of criminal activity and study longer-term trends. However, the use of GIS technology remains for the most part limited to reactive hotspot policing and does not extend to profiling neighborhoods, even though some police forces (e.g. Geneva) have declared their intent to develop such analytical tools.

Interdepartmental projects

Interdepartmental cooperation between different branches of the local administration in the area of public safety goes back many years, but with the implementation of community policing the five police departments have strengthened existing and forged new strategic partnerships to tackle specific problems. In Bern and Zurich, the city governments already in the early 1990s established interdepartmental task forces both at the political and the operational level to tackle the threat to public security posed by the growth of the open air drug markets. In the cities of Basel, Bern, Geneva, and Zurich, the police cooperate with the drop-in centers for drug addicts these cities have established. More recently, the police in all five cities began cooperating closely with the street intervention units the five cities have created to combat disorder in public areas.

The same goes by and large for police activities to control the commercial sex industry, even though only the cities of Basel, Geneva, and Zurich have areas where the clustering of such establishments seriously affects the quality of life of residents in the area. In all five cities, the police brigades concerned with the commercial sex industry have regular exchanges with NGOs supporting sex workers.

As the issue of youth violence has moved up on the political agenda at both the local and national level, the political authorities in Lausanne and Zurich have established interdepartmental task forces in 2000 and 2008, respectively, which were tasked to study the phenomenon and to come up with recommendations for action. In Basel, the cantonal police in 2009 have established a new youth intervention unit that seeks to establish pro-active contacts with adolescents. In Geneva and Lausanne, the police and education departments have signed protocols governing the modes of collaboration between the police agents and public schools.

The police departments also cooperate with other branches of the local administration and private actors in the framework of specific neighborhood regeneration projects. In Zurich, the City's Police Department in 2001 established a project team to run an inter-agency project to improve the inner-city neighborhood of the Langstrasse. In Basel, the Cantonal Police Department has held the lead over an interdepartmental project to manage the area along the banks of the Rhine river notorious for littering, noise, and public drug consumption and has sat on the project team of the cantonal Construction and Urban Planning Department to improve the quality of life in underprivileged neighborhoods in the city's north. In Lausanne, the municipal police have been involved in the steering committee that oversaw the transformation of the central Flon neighborhood from derelict industrial zone to the city's entertainment hub.

Crime prevention

Like interdepartmental cooperation, crime prevention is essentially nothing new – the Zurich Municipal Police Department for instance has had a specialized crime prevention unit for over 50 years. In the process of implementing community policing, however, all five police departments have increased their efforts and resources dedicated to the prevention of unlawful acts. The police departments have reshaped the organizational structure of their crime prevention units, invested in new infrastructure for preventive campaigns and, in some instances, hired new staff to give their efforts broader reach and more heft. In Basel, Lausanne, and Zurich, the police have invested in mobile police stations to mark a more conspicuous police presence and reach out to the population in neighborhoods experiencing a surge in crime.

All five departments regularly are present at festivals or trade fairs to offer crime prevention advice to the general public, and make special efforts to reach out to vulnerable groups. In Basel and Lausanne, the police in regular intervals mail out brochures to all elderly people living in their cities informing them on the potential risks and the best precautionary measures to take. The Basel cantonal police have also forged a partnership with business owners to coordinate the fight against shoplifting and run a database of shops and businesses in the city that allow them to send out crime alerts targeted to a specific audience. In Bern, the neighborhood liaison officers are responsible for crime prevention and focus heavily on institutions. In Zurich, the municipal police have run several campaigns to prevent burglaries, alerting households in areas experiencing a temporary surge or, at one point, even slipping flyers through windows left open by absent home owners.

4.1.5 Performance Appraisal Systems

Evaluation and monitoring

The need to monitor performance and to evaluate policing strategies has eventually been recognized but is not yet implemented across the board. In Basel, the effectiveness of police strategies is being evaluated as part of the annual planning process, in which the higher echelons of the cantonal police set the strategic goals for the department for the next year. The Basel police have also established units of volunteers of all levels from the police hierarchy to monitor performance. In Bern, the head of the community policing unit has produced an annual statistic accounting for its activities since 2001. In Zurich, the municipal police in 2004 created a new internal body to analyze operational data and monitor police effectiveness. In Zurich, the police have also created new data bases to better monitor and assess performance. In Geneva, the cantonal police in 2005 recognized the need for a stricter monitoring of police performance and set out to develop a list of indicators that would allow them to do so. In Lausanne, the police complement traditional police statistics with the results of citizen surveys to assess effectiveness.

The effectiveness of crime prevention efforts is assessed based on anecdotal evidence, relying heavily on the feedback from the partners of the police, except in Zurich where they are monitored systematically.

Project evaluations

In 2004, 2007, and 2010, the police in Geneva have conducted three regional security assessments (Wisler *et al.* 2008a, 2004, 2011). The Geneva police used the latest edition to assess the impact of operation Figaro, a campaign to combat inner city hotspots that was sustained for the better part of 2010. These assessments are generally conducted with the help of external consultants. In 2007/08, the Institute of Criminology of Zurich University evaluated a burglary prevention project, and the same institute has also assessed the neighborhood regeneration project “Langstrasse PLUS” (Manzoni & Thalmann 2008; Schwarzenegger *et al.* 2009). However, scientific evaluations of community policing activities remain rare.

Citizen surveys

In all five cities, the police departments have organized regular surveys to poll the resident population on local levels of victimization, fear of crime, neighborhood disorder, and public satisfaction with the police. The Basel cantonal police have contracted the first such citizen survey as early as 1983 and have tasked the same local polling institute to repeat the survey in 1990, 1995, 2000, 2003, 2006, and 2009. All five police departments at some point contracted the University of Lausanne to conduct surveys of the resident population as part of the International Crime Victimization Surveys of 1998, 2000, and 2005. Since then, the Geneva police have independently conducted surveys in 2004, 2007, 2010 as part of their regional security assessments. In Lausanne, the city’s observatory of public safety has contracted four citizen surveys in 2003, 2006, 2007, and 2008 from the local university. In Zurich, the municipal police contracted a local a polling firm in 2004, 2006, and 2008. In Bern, the municipal statistics bureau has conducted a regular poll of the resident population on quality-of-life and security-related issues since 1999.

4.1.6 Police Equity and Legitimacy

Victims assistance

With the Federal Victims Assistance Act, which entered into force in 1993, police departments all over Switzerland were obligated to instruct their staff in the new procedures introduced by this law, which grants victims of crime access to local counseling centers. Nowadays police recruits are instructed in the appropriate treatment of victims of crime during their basic training at the police academy. These classes are usually taught by psychologists working for the police departments’ internal psychological units.

In Geneva, the then Chief of police considered it especially important that police officers be alert to the plight of victims and know how to interact with them, because their contacts with victims determine the public image of the police (Walpen 1996, 27). In Lausanne, the municipal police department set up a special room at its headquarters to offer crime victims a more comfortable welcome and to avoid that they accidentally cross their tormentors. The Lausanne police also follow a policy of keeping crime victims informed on the progress of the inquiry (Municipalité de Lausanne 2001, 20f.).

Ethics and human rights

Police recruits nowadays are instructed in ethics and the human rights aspects of policing as part of the curriculum preparing them for the federal examination as police officers. The Swiss Police Institute has recently assembled a new folder with teaching materials in French to prepare police recruits for the national examination; a German version is to follow soon.

In Geneva, efforts by the police hierarchy to ascertain the equity and fairness of police interventions go back to the early 1990s but have recently been intensified as a result of the community policing strategy "Proxipol". The Geneva police have long had an internal review body to examine instances of alleged abuse and in 2007 increased that body's staff from one to three. The Lausanne Municipal Police Department has recently gone a step further. The Lausanne police have an internal review body that examines allegations of professional misconduct, which includes an external expert in mediation. It proposes administrative sanctions against staff, which are found to have violated the code of conduct. In addition to this internal review body, the Lausanne police in a collaborative effort of all divisions have drafted a charter encoding their principles of professional conduct. The police have set up an ethics committee, comprised in equal parts of police and civil society representatives, to uphold the charter and issue guidelines on the ethical aspects of police operations.

In Basel, the cantonal justice and police department already in 1978 established a special unit to hear complaints against the police from aggrieved members of the public and to mediate between the two. The cities of Bern and Zurich have no formal police complaints board that handle allegations of police misconduct. Until recently, complaints against the police had to be filed either with the chief of police or addressed to the local political authorities or a city's ombudsperson. In both cities, the police have recently set up new units to evaluate complaints from the public.

In Geneva, the police adopted a formal ethics code as early as 1997. In Lausanne, the municipal police discharge their duty according to a binding code of conduct and a more exigent but non-binding ethics charter. In Basel, the then Chief of police already in the late 1990s drafted a manual on the applicability of the European Convention on Human Rights, detailing the implications of the convention for day-to-day police operations (Kantonspolizei Basel-Stadt 1999). In Bern and Zurich, the police have published a set of guiding principles, but these are drafted in more generic terms than in the other cities.

Minorities

In all five cities, police officers have received cultural diversity training to raise their awareness of cultural differences and to learn how to interact with members of minority communities. Such classes nowadays form part of the basic training curriculum in all four police academies preparing police recruits for the federal examination as police officers. In the country's two largest cities, special training classes to foster the intercultural understanding of police officers have a long tradition: the Geneva police introduced them already at the beginning of the 1990s (Walpen 1996, 23); in Zurich, the earliest such efforts go back even further.

Efforts by the police departments to establish links with ethnic or religious communities have also gotten under way, most notably in Basel where the cantonal police's neighborhood liaison officers meet regularly with representatives of the minority groups established in different parts of the city. In Lausanne, the municipal police have established regular contacts with the Jewish community and are reaching out to other minorities, also through their cooperation in a municipal working group on the integration of foreigners. In Bern and Zurich, similar efforts are under way, but the police have not yet established a framework of regular meetings. In Geneva, the cantonal police have made repeated attempts to install a network of “cultural mediators” between the police and immigrant communities but so far to no avail, although contacts do exist via the cantonal bureau for the integration of foreign nationals.

No police department has adopted affirmative action programs or hiring quotas to obtain a more diverse police force. All five assert, however, that their ranks are growing naturally more diverse as a result of the drafting of naturalized, second-generation immigrant police recruits. The Geneva Cantonal Police Department has publicly mused about a hiring policy to make the police force better reflect the diversity of the canton's population (Wisler *et al.* 2008a, 70). At present, foreign nationals holding a long-term residency permit may apply to the police academy in Geneva on condition that they seek naturalization before taking the oath of office, i.e. before completing their basic training. In Lausanne, Bern, and Zurich, legal provisions prohibit the employment of foreigners as police officers, although in Zurich non-nationals serve as police assistants. For the time being, the Basel cantonal police remain the only security force with non-Swiss nationals as police officers on active duty within their ranks.

4.2 Quantitative Analysis

In a second step, the qualitative information of the community policing progress reports on each city was coded for quantitative analysis. The rationale for this coding procedure was to condense the wealth of qualitative information on community policing implementation down to a manageable number of indicators. These can then be used to compare the long-term trends of community policing implementation and to unearth notable parallels or differences across the five cities.

The coding procedure was relatively straightforward: for each of the six evaluative dimensions of this process evaluation – organizational strategy and partnerships, decentralization and deployment, broken windows approach, problem-oriented policing, performance appraisal systems, and police equity and legitimacy – a list of indicators has been specified as a benchmark to measure implementation progress. A scorecard is used for each city to track the implementation process over the period from 1985 to 2010 in intervals of five years. For each five year interval, a city's scorecard simply records if a community policing element had been implemented or not at the time (e.g. if a police liaison regularly conferred with neighborhood associations or if a clean-up scheme to remove graffiti or litter from public space was in place). These scorecards for the five cities are included in the appendices of this book (cf. Appendix B).

The scorecard data was then aggregated to form index scales that measure the level of community policing implementation progress on each of the six evaluative dimensions over time (cf. Chapter 3.1.3 for details on the methodology). Figure 4.1 displays these community policing implementation scores for each of the five cities. The colored lines connect the score value on each of the six evaluative dimensions of the process evaluation for a particular year. As a police department continues to make progress in implementing community policing, its score value for a particular dimension shifts horizontally from left to right. (A “Min” value implies that a police department had met none of the implementation criteria defined for a specific evaluative dimension, whereas a “Max” value means that it fulfilled all of them.)

Figure 4.1 now serves as the basis for an analysis of the chronology of community policing implementation in each of the five cities. As can be seen from the left-most panel, in Geneva, the police took the greatest step towards community policing already between 1990 and 1995 with the implementation of the plan “Pégase”. The chart also indicates that the Geneva police back then were the undisputed leader in community policing among the five urban areas. The olive green time line for 1995 in Geneva is well to the right of the same time line for the other four police departments. The chart also highlights, however, that over the decade that followed, the Geneva police made only limited progress towards full implementation. Only with the adoption of the second community policing agenda “Proxipol” did the Geneva police reinvigorate their efforts to implement community policing reform, ending the decade on a par with their peers.

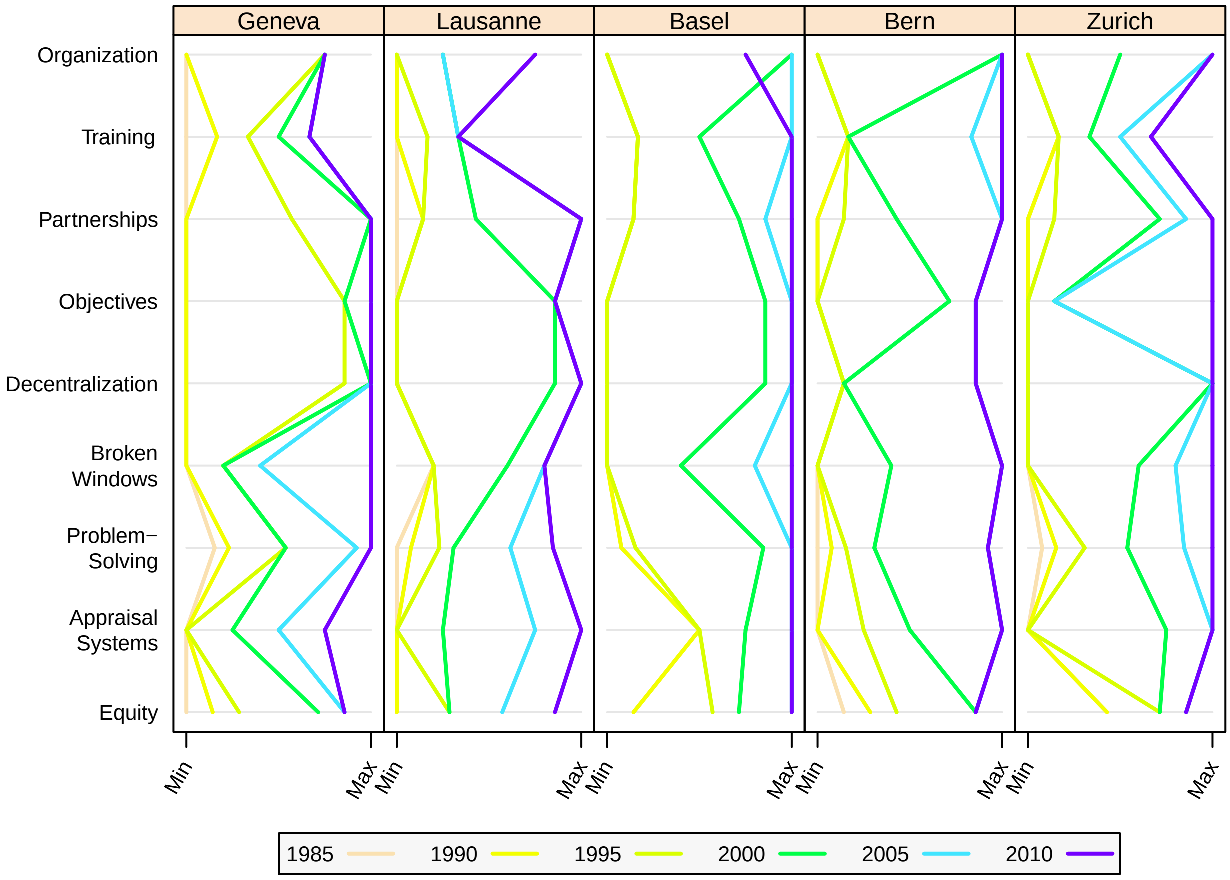


Figure 4.1 Community policing implementation scores 1985-2010. Community policing implementation progress by evaluative dimension across the five cities. (The first four items “Organization”, “Training”, “Partnerships”, and “Objectives” together make up the first evaluative dimension “Organizational Strategy and Partnerships”.) A “Min” value implies that a police department had met none of the implementation criteria defined for a specific evaluative dimension, whereas a “Max” value means that it fulfilled all of them.

The Lausanne municipal police followed a more gradualist approach to reform. The department took a first major step towards community policing between 1995 and 2000, implementing a batch of reforms. Since then, the Lausanne municipal police have made incremental but steady progress towards community policing implementation. However, the Lausanne police have spent comparatively little effort to institute training in community policing. Whereas the department was the first among the five to introduce the concept in its basic training of new recruits, nowadays it lags other departments in the training of its standing officers, police cadres, and neighborhood liaison officers.

In Basel, the cantonal police department took a quantum leap towards community policing when it launched its reform agenda “4plus” in 1998. The department has since made steady progress and today is at the forefront of community policing implementation. The one blot in an otherwise enviable record is the 2008 reform agenda “Optima”, which among its most drastic changes brought about renewed centralization by reducing the number of full-blown precinct police stations located over the city territory from four down to just two.

In Bern, the municipal police formally adopted community policing in 1999 with its reform agenda “Apollo” and took the greatest strides towards implementation over the next half decade.

Between 2005 and 2010, however, the police in Bern against the backdrop of the dissolution of the municipal police and their merger with the cantonal police have only made moderate progress towards full implementation of community policing.

The Zurich municipal police was a late convert to the community policing philosophy, adopting it only in view of the pending introduction of the federal examination of police recruits. However, the department had implemented many of community policing's core demands such as geographically fixed assignments of police officers or strategic partnerships to solve problems before its formal adoption of the new policing philosophy. So it is no surprise that the Zurich municipal police shows low scores on community policing training and objectives during the 1990s, whereas it was already ahead in terms of its approach to neighborhood disorder and problem-solving tasks.

Figure 4.1 not only helps to elucidate the chronology of community policing implementation in each urban area, it can also be used to point out a few noticeable parallels between the experiences of the five cities. Whereas most departments were quicker to institute reforms in the domain of strategic partnerships, community policing objectives, and decentralization, progress in the areas of combating disorder or broken windows policing and problem-solving policing appeared harder to achieve. Moreover, the chart also highlights that the one area, where more than one police departments has made comparatively little effort, is training.

5

Exploratory Spatio-Temporal Data Analysis

This chapter presents the results of the exploratory spatio-temporal data analysis of the variables measuring the four theoretical constructs crime, fear of crime, disorder, and public attitudes towards the police. Each indicator of these four concepts used to assess community policing impact has been analyzed both at the city and at the neighborhood level. Whereas the crime data are official police statistics, the other three constructs were measured using data from the Swiss Crime Survey.[32]

At the beginning of each section, a chart of the long-term trends of these indicators at the city level is presented. In these charts, the top row of panels displays the data for Geneva and Lausanne. Whereas for Lausanne the sample was limited to residents of the city proper, the sample for Geneva includes the survey respondents living in the suburban municipalities surrounding the city.[33] The bottom row of panels depicts the results for the cities of Basel, Bern, and Zurich. The survey samples for Bern and Zurich are confined to respondents living in the city, whereas for Basel the samples include respondents from the suburban communities of Riehen and Bettingen. Each section begins with an introductory paragraph that highlights the most remarkable aspects about the long-term trend of these indicators at the city level.

Subsequently, the results form the neighborhood-level analyses are presented as maps that illustrate the spatio-temporal patterns of the very same indicators used to measure the four theoretical constructs. The relevant test statistics of these analyses are reported in the caption beneath each set of maps. The accompanying text points out the most salient feature of each map or alerts the reader to interesting trends.

5.1 Crime

Figure 5.1 plots the standardized rates of eight different criminal offenses against a time line for each of the five cities. For this plot, crime rates were computed on the basis of the official crime statistics provided by the five cities' police departments. In order to make these police statistics comparable across cities, the absolute numbers of recorded incidents were indexed using as the baseline figure the number of incidents recorded for the first year for which data were available in each city. For that year, the rate for all eight offenses thus equals to 100. A figure of 200 during subsequent years suggests a doubling in the number of incidents according to local recording practices.

Figure 5.1 reveals a surprising degree of similarity in the long-term crime trends across the five urban areas being studied here. The most salient trend that can be observed across all five of them is a substantial increase in the number of threats and assaults recorded over the study period.

[32] The exploratory spatio-temporal analysis of the Swiss Crime Survey data is an expanded version of the author's unpublished graduation paper for the Master of Advanced Studies MAS *Urbanisme durable* (Kreis 2009).

[33] The decision which counties to qualify as suburban and hence to include in the analysis was made on the basis of the Geneva cantonal police's own categorization of the Geneva municipalities made for the purposes of their 2007 regional security assessment. The Geneva cantonal police's list of suburban counties includes the following municipalities: Carouge, Chêne-Bougeries, Chêne-Bourg, Grand-Saconnex, Lancy, Meyrin, Onex, Plan-les-Ouates, Thônex, Vernier, and Versoix (Wisler *et al.* 2008b, 3).

These increases generally appear more important in the three smaller cities of Basel, Bern, and Lausanne. In the two bigger cities, Geneva and Zurich, they are less pronounced.

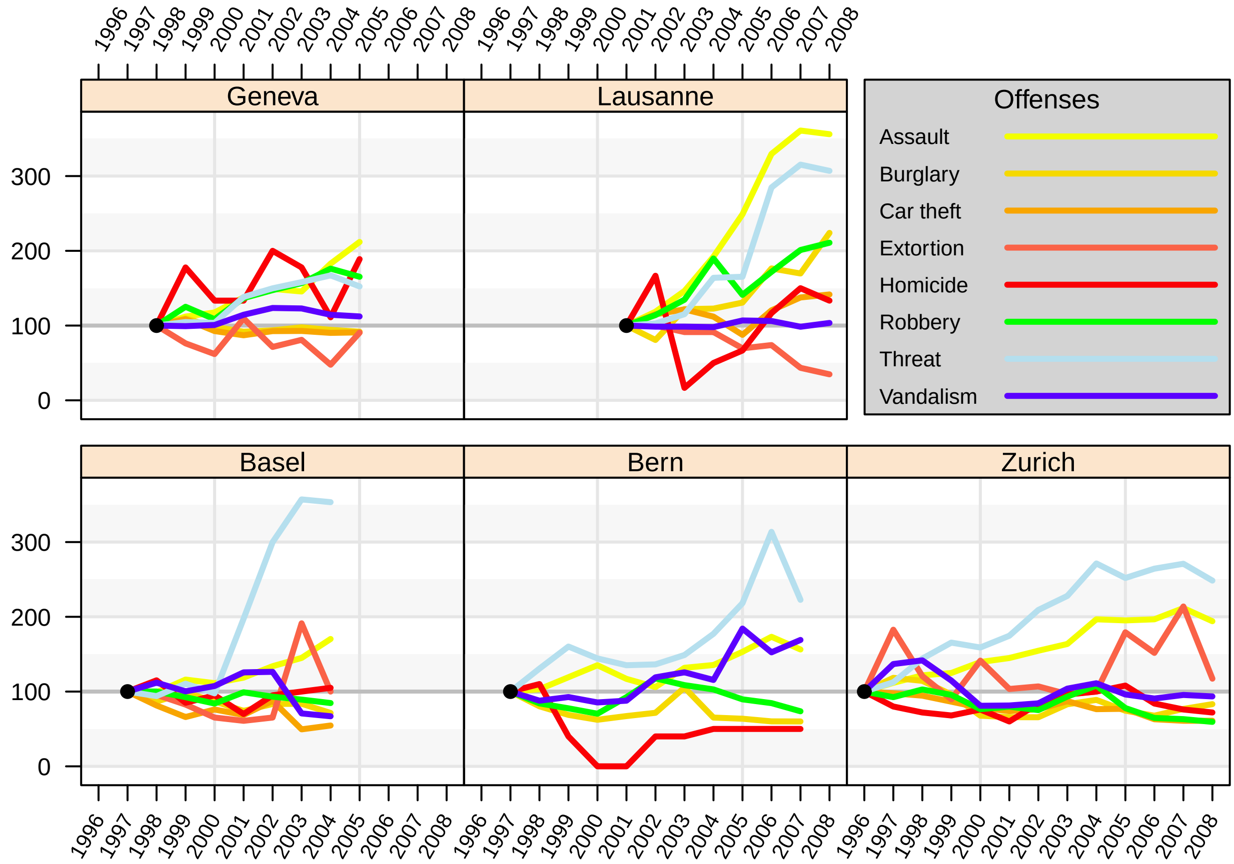


Figure 5.1 City-level crime rates 1996-2008. Long-term trends of eight different criminal offenses across the five study areas. Rates are indexed at 100 for the first year, for which data were available for each city.

As to the other types of offenses, the long-term trends are less consistent than for threats and assaults, even though they still reveal some striking parallels. For instance, burglaries were either steady or falling slightly in all cities except in Lausanne, where rates increased substantially over the study period. The same also goes for motor vehicle thefts, which fell across the board except in Lausanne. In a similar vein, robberies increased in Geneva and Lausanne but fell slightly in the three Swiss German cities. By contrast, acts of vandalism remained stable over time except in Basel and Bern, where the data suggest a slight increase. Extortions show rather erratic trend lines, but overall rates were stable or falling. Finally, homicide rates, unsurprisingly given their low frequency, showed some erratic trend lines as well, but there is no indication of a sustained increase in any of the five cities being studied here.

Figure 5.2 plots colored maps indicating the standardized neighborhood-level burglary rate across the four cities for which neighborhood-level data were available for the current study.[34] As these neighborhood-level crime figures vary considerably year-on-year, the absolute numbers of incidents were aggregated and standardized rates calculated for multiple years (cf. Chapter 3.2.2).

As with the long-term trends, the maps reveal some striking parallels as regards the spatial distribution of burglary offenses across the cities. Worst afflicted in each of the four cites are the downtown areas as well as one or two adjacent neighborhoods in close physical proximity. By

[34] For the Canton of Basel-City, official crime statistics were available only at the cantonal level (cf. Chapter 3.2.1)

contrast, the further away one moves from the city centers, the lower the burglary rate generally gets relative to the city average.

Only Geneva shows a slightly clustered pattern as indicated by the positive value of Moran's I that is statistically significant. For the other three cities, the spatial pattern is not significantly different from a random distribution. However, it is probably advisable not to overrate this finding, since the spatial units of analysis differ greatly across the four cities. For Geneva and Lausanne, the basic units of analysis are OFS census tracts, whereas in Bern the study units are ZIP code districts, and in Zurich the still larger administrative districts of the city government.

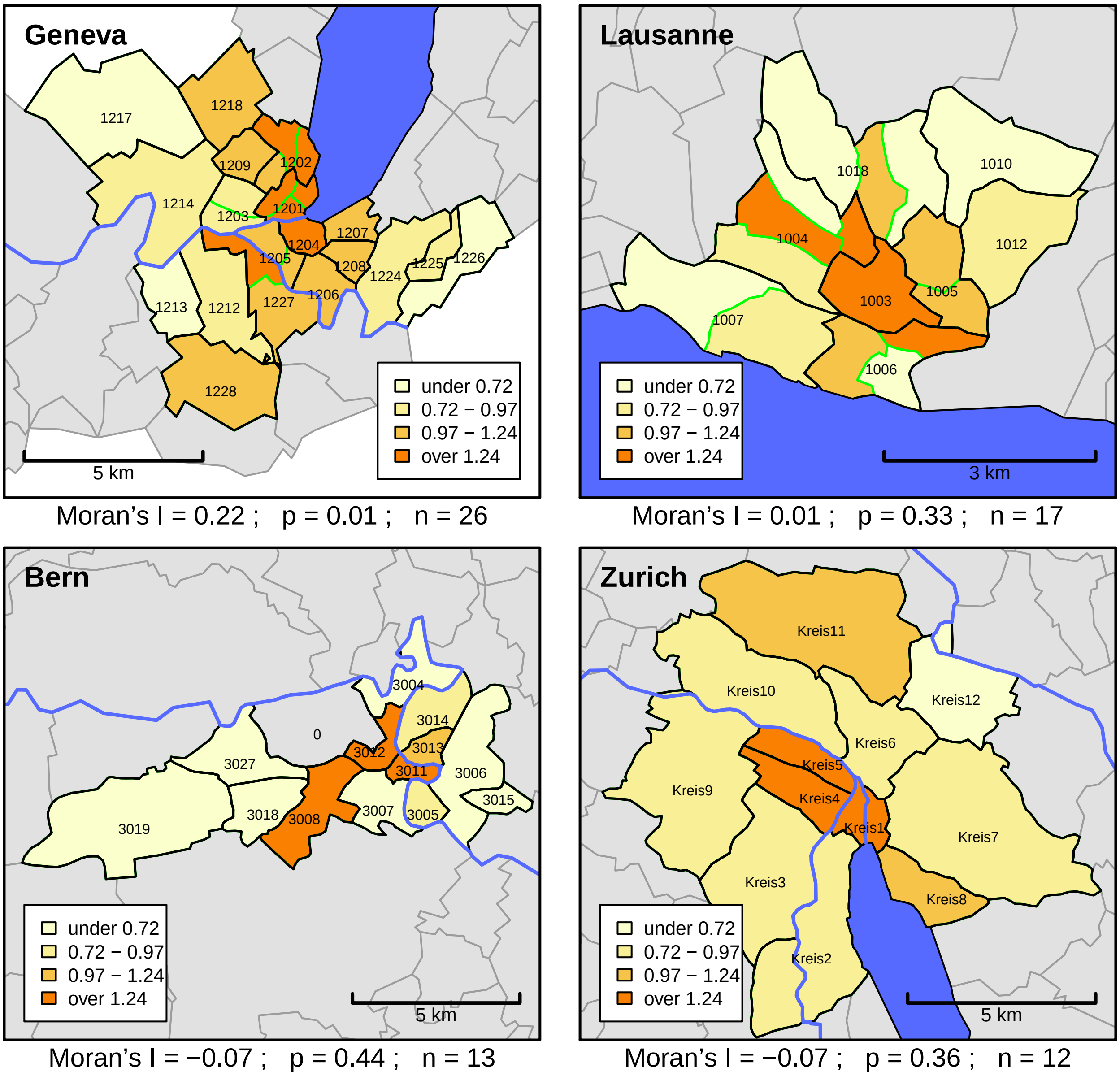


Figure 5.2 Neighborhood burglary rates 1998-2001. Standardized neighborhood-level burglary rates across four of the five cities aggregated over the 1998-2001 period (to the extent that data are available). Standardized rates indicate the neighborhood-level crime rate per resident as a multiple of the city average. For Geneva and Lausanne, crime statistics were available at the level of census tracts, the borders of which are indicated as green lines inside the ZIP code districts.

Figure 5.3 plots the standardized burglary rates aggregated over the period between 2002 and 2005. The most striking aspect about these maps is how closely they resemble the patterns observed in Figure 5.2. The neighborhood burglary rates have hardly changed at all; the rates are still notably higher in the city centers and decrease as one moves towards the fringes of each urban area.[35]

This visual interpretation is also born out by the quantitative analysis of Moran's I. Again, only Geneva shows significant clustering of the neighborhood burglary crime rates, whereas the spatial patterns for the three other cities are not significantly different from a random distribution.

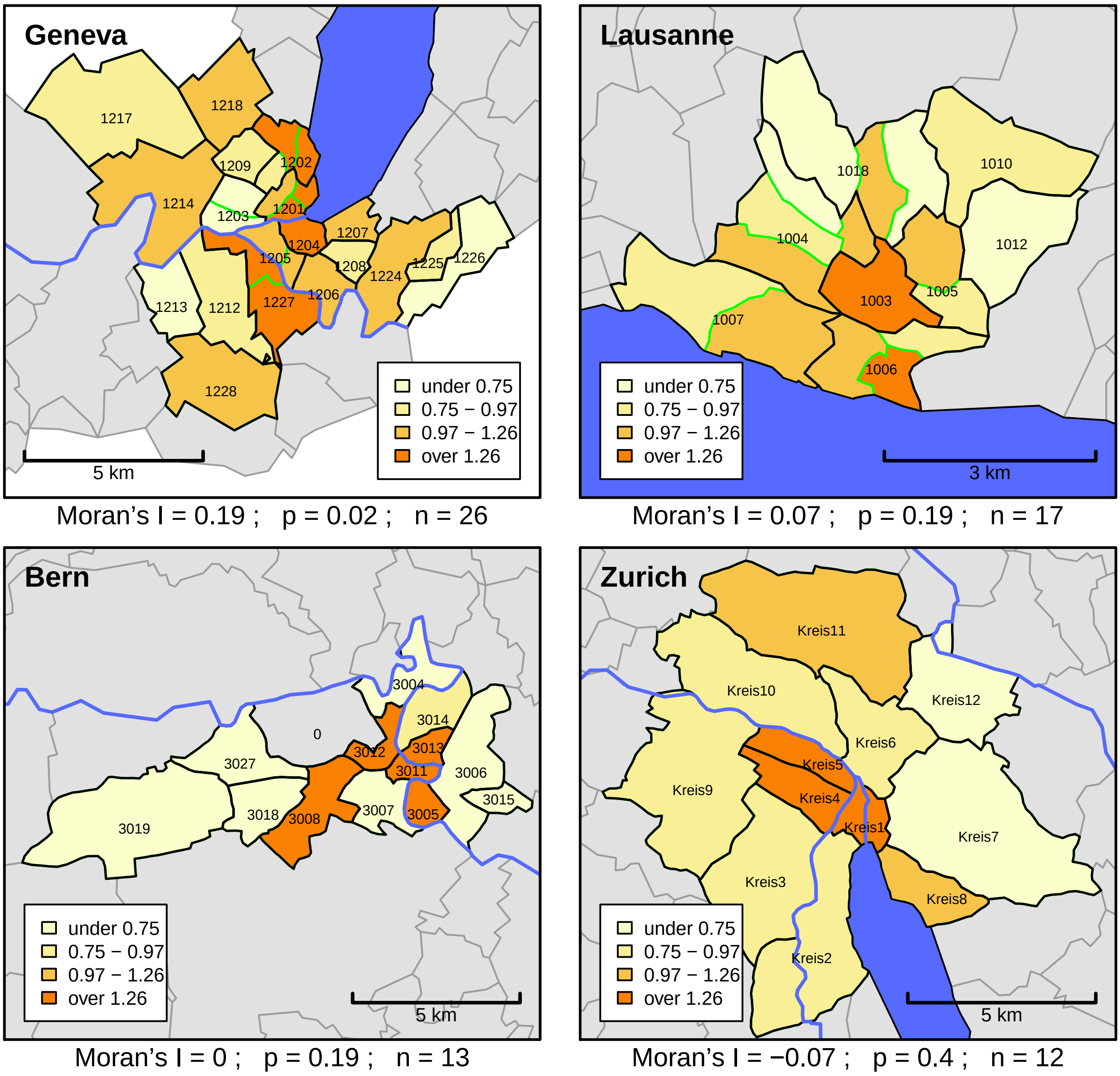


Figure 5.3 Neighborhood burglary rates 2002-2005. Standardized neighborhood-level burglary rates across four of the five cities aggregated over the 2002-2005 period. Standardized rates indicate the neighborhood-level crime rate per resident as a multiple of the city average. For Geneva and Lausanne, crime statistics were available at the level of census tracts, the borders of which are indicated as green lines inside the ZIP code districts.

Only the spatial patterns of neighborhood burglary rates are presented here. This was done mainly for two reasons: first, burglary rates are the most interesting for this analysis, because they directly relate to the SCS survey item measuring the perceived risk of victimization, thus

[35] Even more striking, perhaps, is how much the patterns of neighborhood burglary rates resemble the maps that Riva (1988) produced for the City of Lausanne based on police crime statistics from the year 1980.

allowing a comparison between the subjective and objective risk of victimization. Secondly and more importantly, the observation that neighborhood-level crime rates do not shift substantially over time also holds true by and large for the other types of criminal offenses tracked by the current study.[36] Moreover, in a preliminary analysis, the correlation between several of the eight types of criminal offenses was found to be so high that it was decided to substitute the standardized crime rates by the component scores of each neighborhood on the first two components of a principal components analysis of the crime data (cf. Chapter 3.2.2).

5.2 Fear of Crime

Since the inception of the Swiss Crime Survey, its questionnaires have contained multiple items targeting different dimensions of the construct fear of crime. This makes it possible to track the evolution of different indicators over time and to analyze the long-term dynamics of fear of crime in Swiss urban neighborhoods.[37]

5.2.1 Threat of Crime

Since its very first edition in 1984/87, the Swiss Crime Survey has included an item measuring fear as the perceived threat respondents felt crime held for them. Respondents typically were asked how safe they felt walking alone in their neighborhood in the evening.

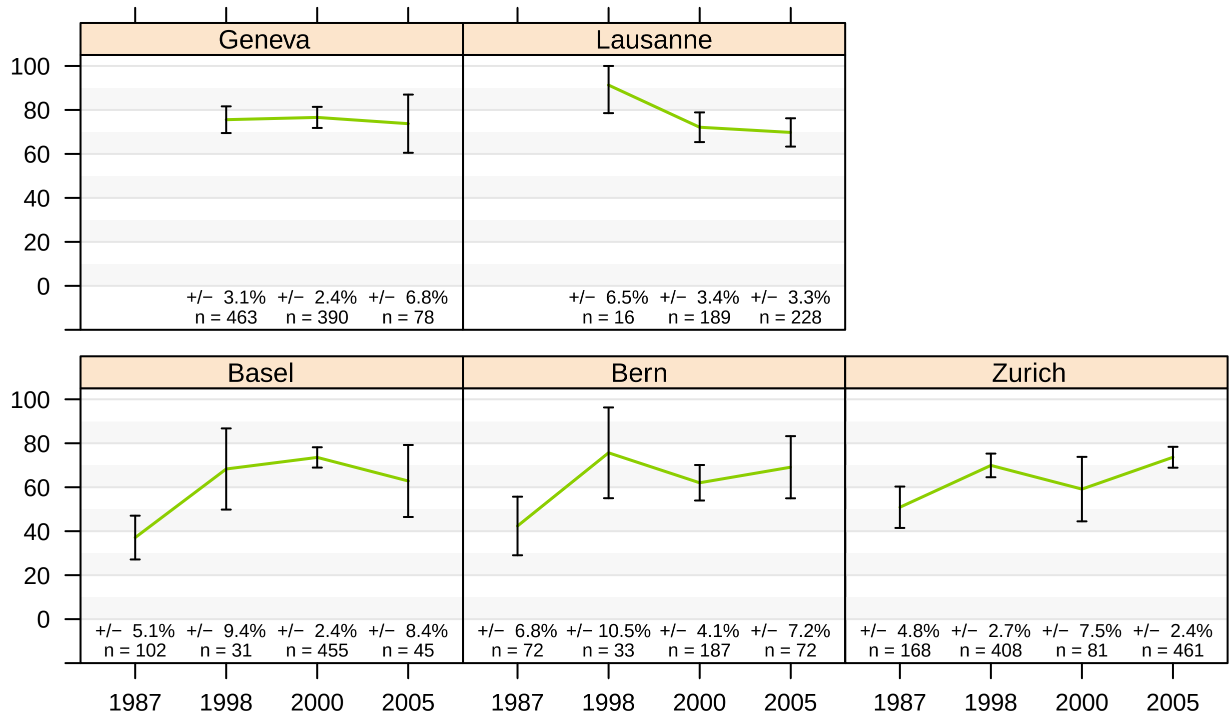


Figure 5.4 Threat of crime 1987-2005. Percentage of respondents who feel "very safe" or "quite safe" walking alone in their neighborhood at night. NOTE: The wording of the survey question and answer categories differed in 1987.

[36]The reader who is unconvinced by this assertion can inspect the plots charting the long-term trends in standardized neighborhood-level crime rates for each of the eight offenses for all four cities in Appendix C.

[37]For the analyses that follow, survey data were weighted to correct for bias in the age and gender distribution (cf. Appendix A.1).

Figure 5.4 plots the evolution of this indicator over the four sweeps of the SCS. In Geneva, the percentage of people who feel "very safe" or "quite safe" walking alone in their neighborhood after 10 pm has remained more or less stable since 1998, hovering around 75 percent, whereas in Lausanne this number dropped between 1998 and 2000. Even though the confidence intervals do not overlap – implying that the difference is statistically significant – the finding should not be overrated since the inordinatey small sample size in 1998 (n=16) makes this figure inherently unreliable. In the three Swiss German cities, the share of respondents feeling safe at night appears to have risen between 1987 and 1998, but this apparent improvement must also be taken with a grain of salt, since the wording of the survey item has changed between these two editions of the SCS. The shifts observed for these cities over the 1998 to 2005 period remain within the range of the confidence intervals so that it is not possible to determine whether the observed fluctuations in the survey responses reflect real differences or were merely due to chance.

The results from the neighborhood-level analyses are presented as colored maps over the following pages. For Geneva, the lapse in time between the two maps is a mere two years, a snapshot compared to the long study periods for the other cities. A decade ago, the territory of the City of Geneva (ZIP 1201 to 1209) appeared to be divided, with the neighborhoods on the right side of the river Rhône displaying higher average levels of fear than the ZIP code districts left or south of it (Fig. 5.5 top). Also for Lausanne, the shifts in the spatial distribution of fear over time were more moderate, perhaps not surprisingly, given that the time span between the two polls was a mere five years. The most salient feature about Lausanne is that the city was divided into two parts along an axis running from north east to south west, with the neighborhoods on the eastern side of the dividing line generally exhibiting lower levels of fear than residents in the western part. By 2005, however, fear levels had risen slightly in the eastern part and fallen in the west, narrowing the gap between the two areas (Fig. 5.5 bottom).

In Basel in 1987, the ZIP code districts with the highest levels of fear happened to be the town center (ZIP 4051) as well as three adjacent neighborhoods on either side of the river Rhine (ZIP 4052, 4057, and 4058). The neighborhoods with the lowest levels of fear, by contrast, were the suburban community Riehen in the northeast and two neighborhoods on the southern outskirts of the city (ZIP 4053 and 4059). A good decade later, the situation had been turned upside down. The city center now recorded the lowest levels of fear, and the peripheral neighborhoods reported comparatively higher levels of fear (Fig. 5.6 top). In the capital Bern, the shifts in the pattern of fear of crime over the same period followed a similar trend. As in Basel, the city center (ZIP 3011) turned into a lighter shade of green over time, whereas the suburban neighborhoods experienced a relative decline in subjective levels of security over the same period (Fig. 5.6 bottom).

In Zurich back in 1987, the average scores of the fear of crime variable revealed a city that was grosso modo divided into two parts. Neighborhoods that exhibited low levels of fear of crime clustered in the southern half. The average score on the fear scale was 1.5 or lower, which meant that at least half of the respondents in those ZIP code districts felt safe walking alone at night in their neighborhood. In the northern part of the city, by contrast, lay a batch of neighborhoods in the shape of a wedge pointing towards the city center that displayed higher levels of fear (Fig. 5.7).

By 1998, the outlook had changed substantially. While higher levels of fear still appeared to be concentrated in the northern half, the rest of the city had become more heterogeneous over the previous decade. Neighborhoods where at least half the respondents felt safe, which – due to the altered fear scale – meant a score of two or lower, were scattered over the city and no longer formed a contiguous pattern. By 2005, the pattern had altered yet again albeit less profoundly. The distribution had become even more random and the spatial autocorrelation all but disappeared. Higher levels of fear were no longer concentrated in the northern half of the city, and neighboring ZIP code districts with similar fear scores had become exceedingly rare.

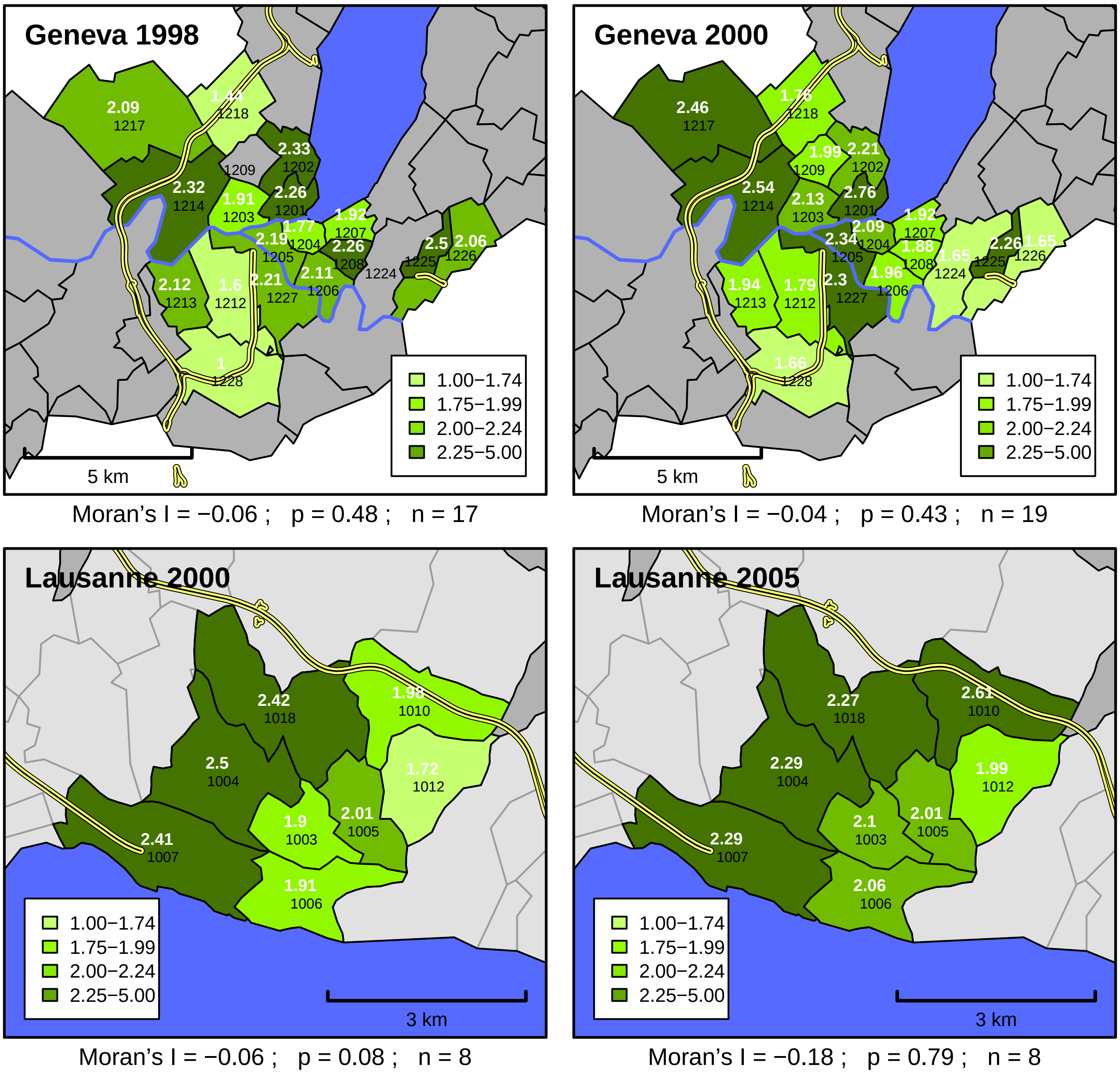


Figure 5.5 Threat of crime in Geneva and Lausanne. Average score by ZIP code district of how safe respondents felt walking around alone at night in their neighborhood on a scale from 1 to 5 (a higher score indicating higher levels of fear). For Geneva in 1998, Moran's I of -0.06 indicates a slightly dispersed spatial distribution that is not significantly different from a random pattern, though (top left). Two years on, the spatial distribution has changed only marginally; Moran's I and the corresponding p-value suggest a spatial pattern that cannot be distinguished from a random distribution (top right). For Lausanne, Moran's I of -0.06 exhibits the presence of slight negative autocorrelation in 2000 that is almost statistically significant ($p < 0.1$; bottom left). In 2005, by comparison, Moran's I of -0.18 shows a somewhat more dispersed pattern, albeit below any conventional level of significance (bottom right).

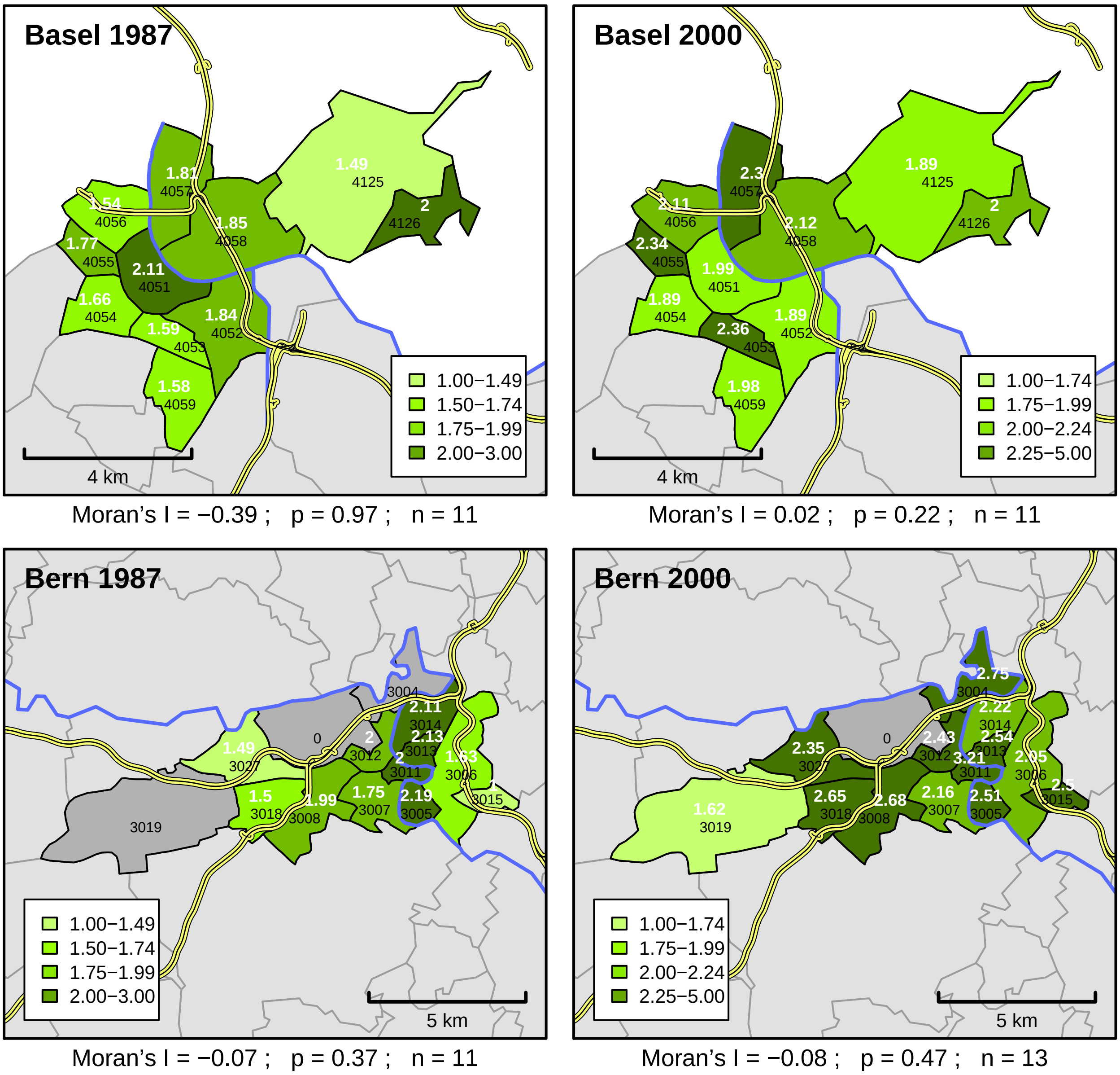


Figure 5.6 Threat of crime in Basel and Bern. Average score by ZIP code district of how safe respondents felt walking around alone at night in their neighborhood on a scale from 1 to 3 and 1 to 5, respectively (a higher score indicating higher levels of fear).[a] For Basel in 1987, Moran's I of -0.39 indicates negative spatial autocorrelation in the data that is not statistically significant (top left). In 2000, Moran's I of 0.02 indicates a random distribution that is not significant either (top right). For Bern, Moran's I of -0.07 indicates that ZIP code districts with similar values were slightly dispersed in 1987 but the pattern is not statistically significant and possibly due to chance (bottom left). For 2000, Moran's I of -0.08 again shows slight negative spatial autocorrelation in the data, but the value is again not significant (bottom right).

[a]The question format and answer categories used to measure the feeling of safety have changed between the 1987 and 2000 surveys. As a consequence, the category boundaries of the color scales are different for 1987 and 2000 as indicated in the key, reflecting the changes in the underlying scales of the ordinal variable. Absolute mean levels are thus not directly comparable. However, since the same methodology has been applied consistently in developing the color scales, a given neighborhood's position relative to others may be meaningfully compared (cf. Chapter 3.2.4 for details on the methodology of defining the color scales).

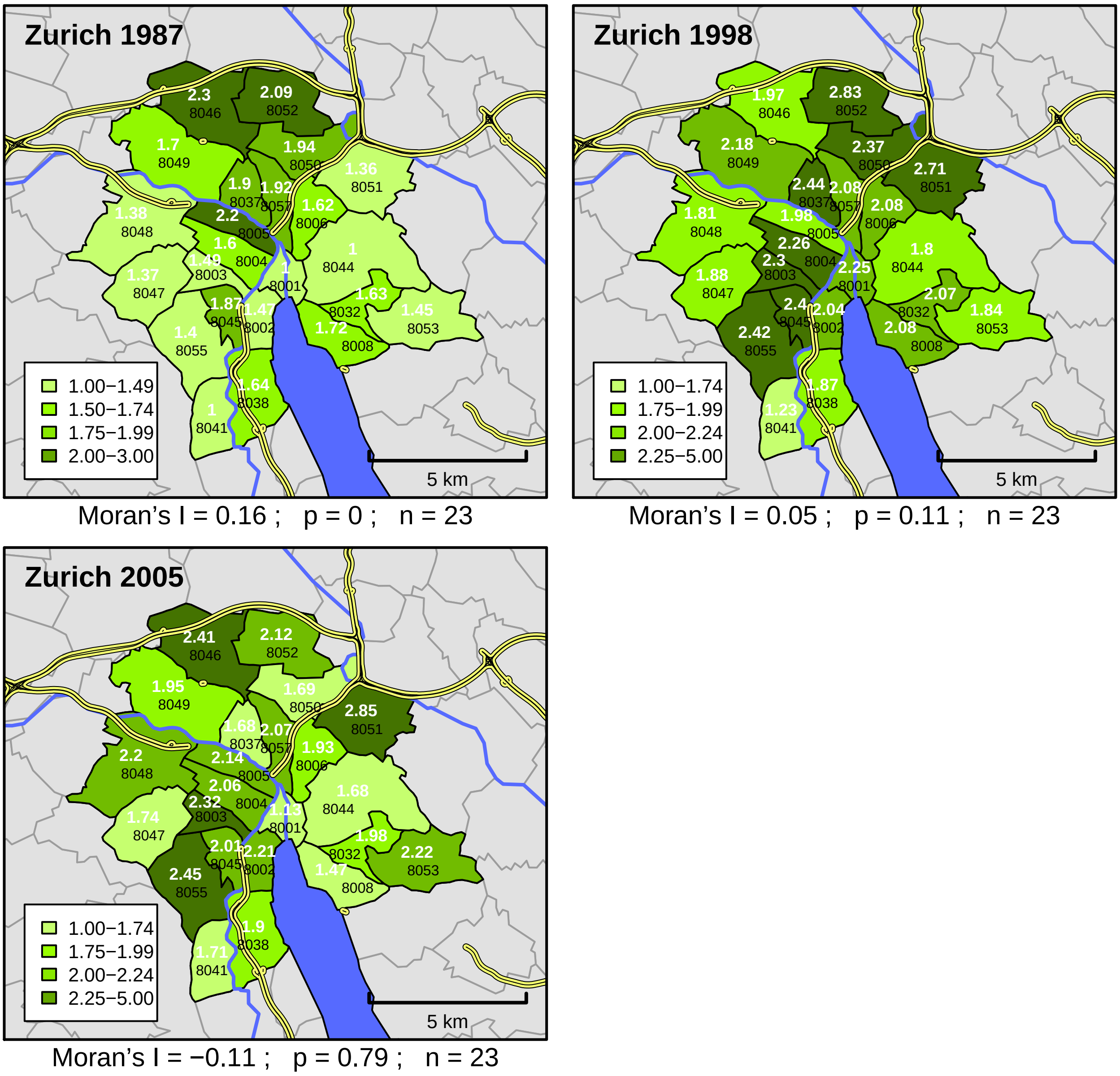


Figure 5.7 Threat of crime in Zurich. Average score by ZIP code district of how safe respondents felt walking around alone at night in their neighborhood on a scale from 1 to 3 and 1 to 5, respectively (a higher score indicating higher levels of fear).[a] A Moran's I of 0.16 that is statistically significant ($p < 0.01$) indicates positive autocorrelation or clustering in 1987, meaning that ZIP code districts with similarly high or low scores are found near each other (top left). A Moran's I of 0.05 indicates the spatial pattern had become more randomly distributed by 1998 (top right). A Moran's I of -0.11 indicates slightly negative spatial autocorrelation in 2005, but the pattern is not significantly different from a random distribution (bottom).

[a]The question format and the scales applied to measure the feeling of safety have changed between the 1987 and 1998 surveys (cf. Footnote to Fig. 5.6).

5.2.2 Risk of Victimization

Since 1998, the SCS questionnaire had included an identical item asking respondents how likely they rated the chances that someone would try to break into their home over the next twelve months. The assessment of victimization risk is the second important operationalization of fear of crime that targets more cerebral judgments than the emotive response to crime of feeling unsafe walking alone at night.

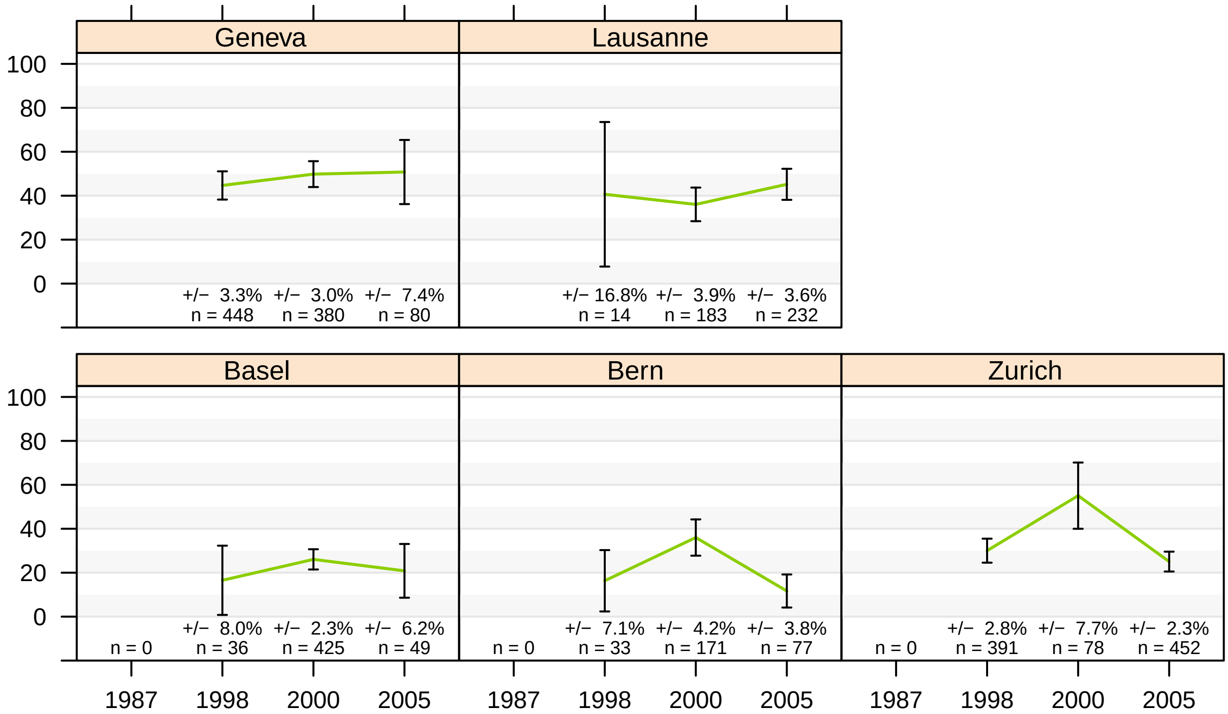


Figure 5.8 Risk of victimization 1998-2005. Percentage of respondents who think chances are “rather likely” or “very likely” that someone will try to break into their home over the course of the next twelve months.

Fear of crime measured as the expected risk of a burglary victimization over the next twelve months has remained more or less stable in all five cities under study but was more prevalent in Geneva and Lausanne. In the two Swiss French cities, more than forty percent of respondents typically deemed a burglary of their home as likely or very likely in any given survey year. The equivalent figure for Basel, Bern, and Zurich, by contrast, stayed below forty percent except for Zurich in 2000. Moreover, the figures for the Swiss German cities are for the most part significantly below forty percent and thus are due to real differences in public opinion between the two linguistic regions rather than to mere chance (Fig. 5.8).

At the neighborhood level, in Geneva, the most salient feature is how gloomy residents were about the risk of a burglary of their home. Already in 1998 their outlook was rather pessimistic and worsened still more over the following two years. In 2000, in every single ZIP code district in the City of Geneva, a third of respondents or more rated the chances of a burglary as likely. The second salient feature is that there were few disparities between neighborhoods as the value of Moran’s I close to zero attests. In general, the starkest differences were between the city proper and the suburban communities, which are even more concerned about burglary risk. Among the gloomiest assessments came from municipalities that share a physical border with neighboring France (Fig. 5.9 top).

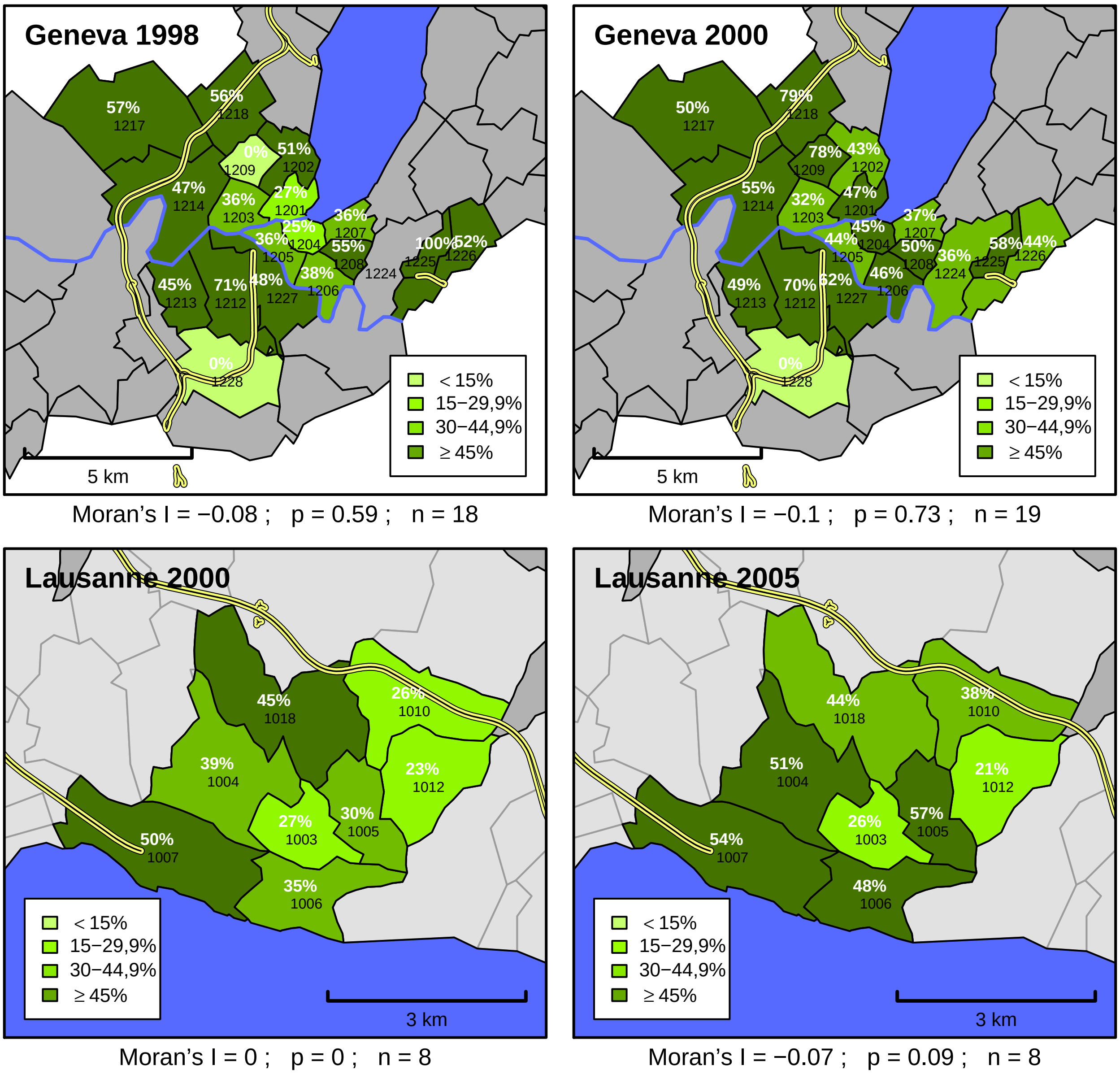


Figure 5.9 Risk of victimization in Geneva and Lausanne. Percentage of respondents by ZIP code district who think chances are "rather likely" or "very likely" that someone will try to break into their home over the course of the next twelve months. For Geneva in 1998, Moran's I of -0.08 shows a pattern that is slightly dispersed but not statistically significant (top left); in 2000, Moran's I is -0.1, i.e. virtually identical (top right). In Lausanne in 2000, Moran's I of 0 indicates a random pattern that is statistically significant (bottom left). In 2005, by contrast, Moran's I of -0.07 exhibits a slightly dispersed pattern that is not significant at the conventional 5 percent threshold ($p < 0.1$; bottom right).

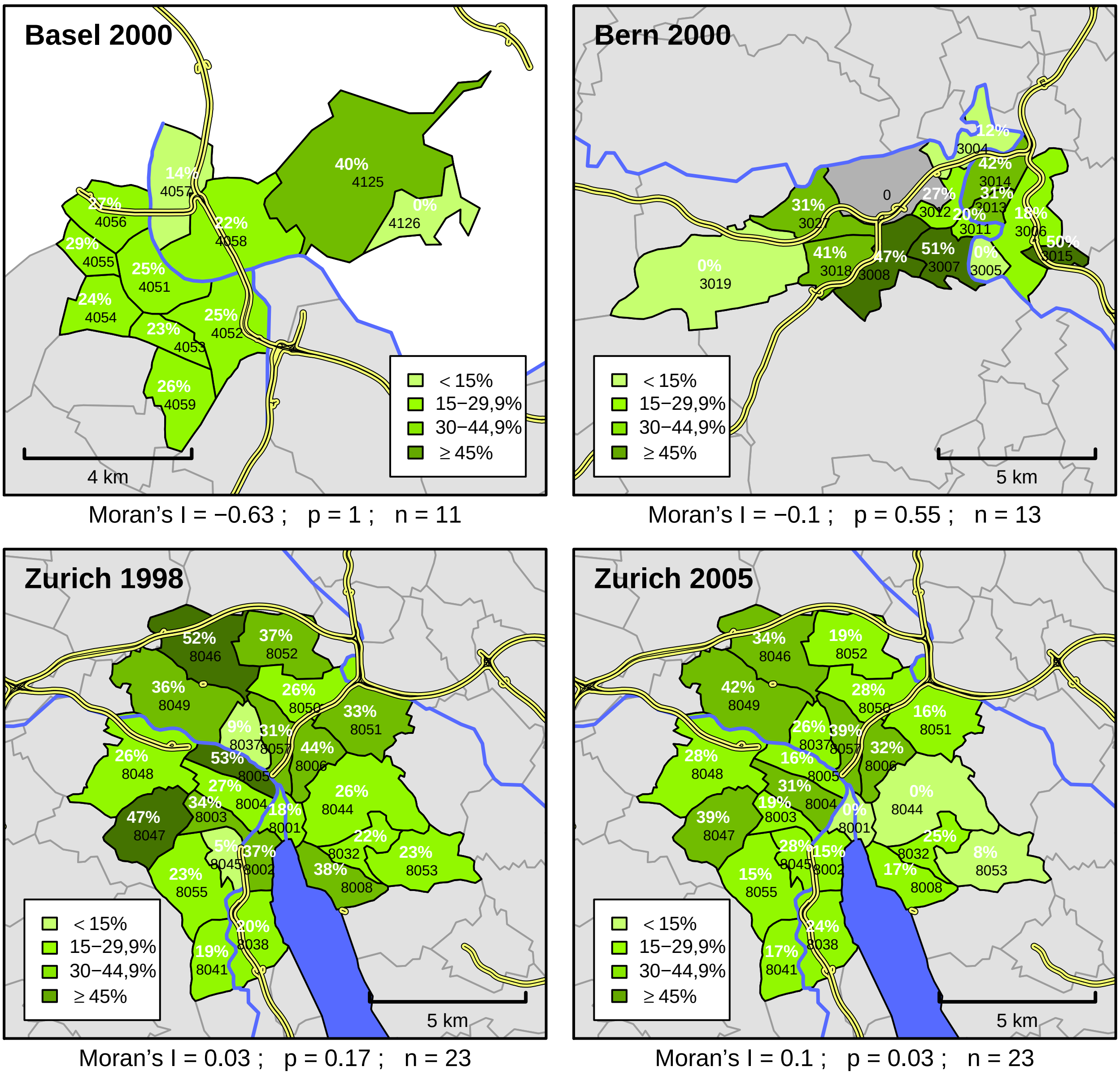


Figure 5.10 Risk of victimization in Basel, Bern, and Zurich. Percentage of respondents by ZIP code district who think chances are "rather likely" or "very likely" that someone will try to break into their home over the course of the next twelve months. For Basel, Moran's I of -0.63 shows a rather dispersed pattern that is not statistically significant, however (top left). In Bern, Moran's I of -0.1 indicates a less dispersed pattern that is not significant either (top right). In Zurich in 1998, Moran's I is 0.03, indicating a minimally clustered pattern (bottom left). In 2005, by contrast, Moran's I of 0.1 that is statistically significant ($p < 0.05$) indicates a pattern that is slightly clustered (bottom right).

In 2000 Lausanne exhibited the characteristic disparity between its eastern and western parts. Five years on, the mood had worsened substantially in two eastern neighborhoods (ZIP 1005 and 1006), resulting in a dispersed pattern. It could be interesting to try to investigate what provoked the abrupt change of opinion in these two neighborhoods, all the while the outlook in the nearby center improved over the same period (Fig. 5.9 bottom).

The spatial patterns for Basel and Bern in 2000 had more in common than meets the eye. In both cites residents in the town center (ZIP 4051 and 3011) rated the chances of their homes being burglarized as relatively low. The neighborhoods where residents were somewhat more concerned lay on the city's outskirts. There were also some differences, though: in Basel the pattern is remarkably homogenous and differences between neighborhoods are small. In Bern, the contrasts between neighborhoods are more marked (Fig. 5.10 top).

In Zurich, the patterns of 1998 and 2005 show only moderate clustering, one of which is statistically significant, but both display some interesting trends nevertheless. As with other indicators of fear, levels in the city center (ZIP 8001) fell between 1998 and 2005 and remained higher in the city outskirts. Also, the tendency of high-risk neighborhoods to cluster along the highway stretches leading into town can be observed for both survey years. A more positive development occurred in the ZIP code district 8005, where the percentage of respondents worrying about burglary crime dropped steeply from 53 percent in 1998 down to 16 percent half a decade later. It is possible that the efforts of the City of Zurich to improve safety and quality of life in the area, which had been seriously affected by the existence of the open air drug markets at the nearby Platzspitz and Letten during the early 1990s, had helped to allay some fears in this part of town (Fig. 5.10 bottom).

One hypothesis that emerges from these patterns is that for the risk assessments of respondents it matters how easily it appears for potential burglars to escape from a site. In the cities of Geneva, Lausanne, and Zurich, for which a temporal analysis of the phenomena is possible, the outlook of residents of the ZIP code districts located in close proximity to the local highway system had generally deteriorated by the time of the second survey compared to the earlier poll. In Geneva, in addition, the outlook tended to be more pessimistic in the counties with a direct border with neighboring France. If this casual observation turns out to be more than just a spurious correlation, however, the link must be more complex than one might suspect. In Basel, for instance, the neighborhoods deemed riskiest by their own residents are located neither alongside the highway nor necessarily next to the border.

5.2.3 Behavioral Response

Besides the two cognitive dimensions of fear – the feeling of safety and victimization risk – the SCS questionnaires have also included an item targeting the behavioral response to fear. The item typically asked respondents if, walking alone at night in their neighborhood, they deliberately avoided certain streets, areas, or groups of people to avoid crime.

Figure 5.11 suggests that the trends in the percentages of respondents who deliberately avoid certain areas moved in opposite directions in the Swiss French and Swiss German cities over the study period. In Geneva, the percentage of respondents taking extra precautions appears not to have shifted much; at least the observed differences are not statistically significant. In Lausanne, the figure appears to haven risen slightly, even though the small sample from 1998 (n=18) renders that figure difficult to interpret. The equivalent figure for Basel, Bern, and Zurich, by contrast, has fallen over the study period, albeit from a much higher base. This observation seems to hold true, even if one factors in that the format of the underlying questions changed between 1987 and 1998 and that the fluctuating sample sizes inevitably mean that the observed differences in Basel and Bern between 1998 and 2005 are not significant. The notable exception is Zurich, where the drop between 1998 and 2005 is statistically significant, lending further weight to this line of argument.

At the neighborhood level, in the City of Geneva in both 1998 and 2000, fewer respondents on average reported that they stayed clear of certain places in the ZIP code districts on the right

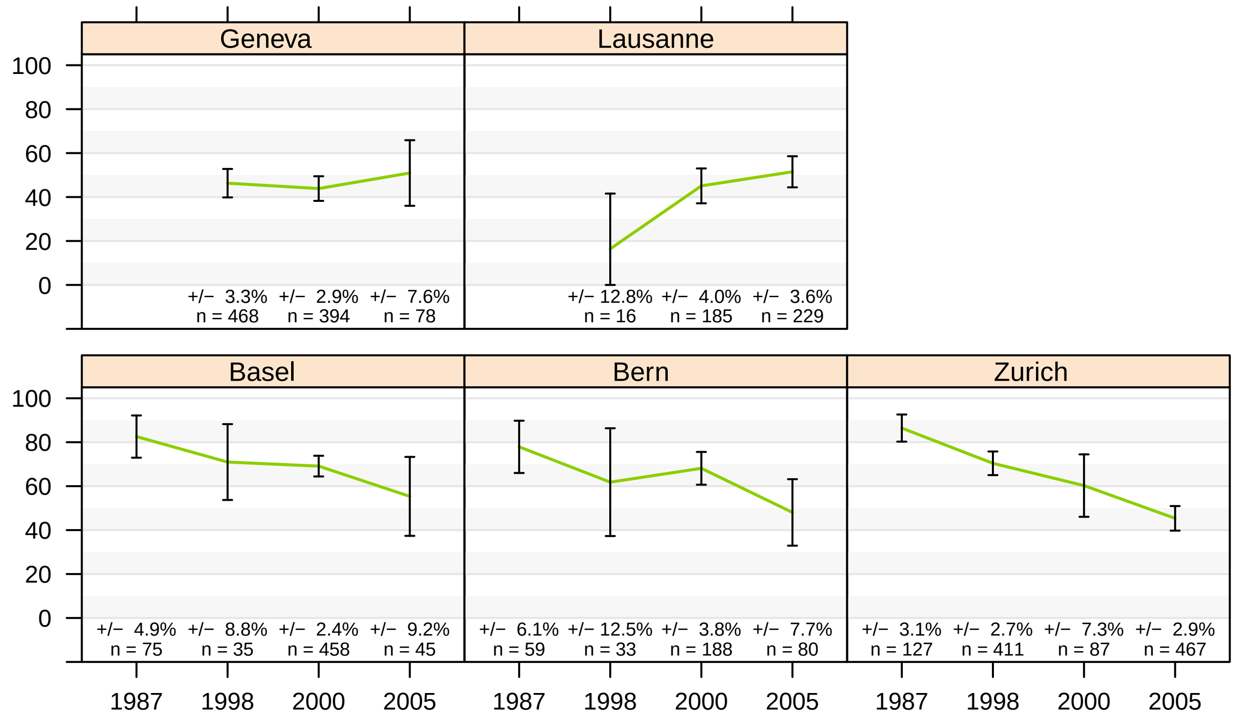


Figure 5.11 Behavioral response to crime 1987-2005. Percentage of respondents who – walking alone in their neighborhood after 10 pm – stay away from certain streets, areas, or people to avoid crime. NOTE: The wording of the survey question and answer categories differed in 1987.

or northern side of the Rhône river than on its southern bank. Given the short time interval between the two surveys, the shifts in sentiment are moderate for a majority of ZIP code districts, as evidenced by the values of Moran's Index, which were virtually identical for both years. In the suburban communities surrounding the city, however, popular sentiment had shifted over the period, mostly for the better. Between 1998 and 2000, the percentage of respondents who did not avoid certain streets had risen or remained stable in most communities, with the City of Vernier being the most prominent exception to the rule (Fig. 5.12 top).

Response patterns in Lausanne show the east-west divide that characterizes the city, albeit less markedly than for the other measures of fear. Back in 2000 in the city's western and northern ZIP code districts, which on the emotive dimension of crime generally reported higher levels of unease, more respondents also said they deliberately avoided certain places than in the eastern parts. By 2005, the three western ZIP code districts (1004, 1007, 1018) had all experienced a rise in the percentage of respondents staying away from certain places. The most positive development occurred in the city center (ZIP 1003), where the percentage of respondents not taking any steps to go out of harm's way rose from 51 percent in 2000 to 61 percent in 2005 (Fig. 5.12 bottom).

In Basel in 1987, the pattern of respondents deliberately staying away from certain areas looked similar to the pattern of how safe they felt. The highest levels of fear were found in the city center and one adjacent neighborhood. On the city outskirts, few if any respondents felt a need to stay away from some streets or areas in their neighborhood. In 2000, by contrast, fear had again shifted from the city center to the outskirts. In seven out of eleven neighborhoods, only a third of respondents or less said they did not deliberately avoid certain streets or areas. In the city center (ZIP 4051), on the other hand, these respondents formed a majority (Fig. 5.13 top).

The City of Bern followed a similar trajectory albeit from a different starting point. In 1987, some ZIP code districts in the city's central and western part including the town center had low levels of fear, whereas in some of the neighborhoods on the outskirts more people stayed away from

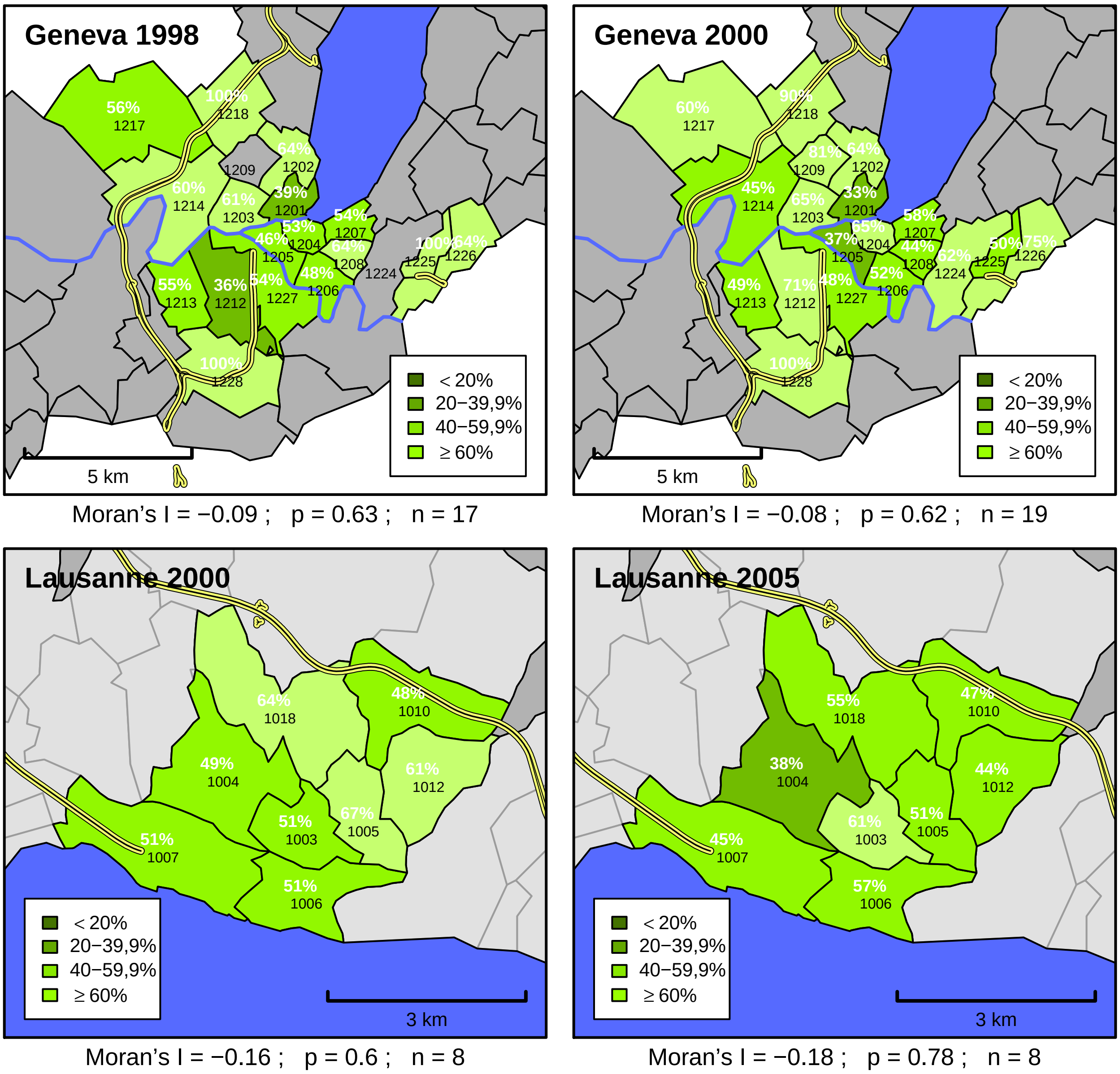


Figure 5.12 Behavioral response to crime in Geneva and Lausanne. Percentage of respondents by ZIP code district who – walking alone in their neighborhood after 10 pm – do not stay away from certain streets, areas, or people to avoid crime. In Geneva in 1998, Moran's I of -0.09 exhibits marginal negative spatial autocorrelation that is not significant (top left). In 2000, Moran's I of -0.08 shows a similar pattern that is not significant either (top right). In Lausanne in 2000, Moran's I of -0.16 suggests there is marginal negative spatial autocorrelation in the distribution of the data that is not significant, though (bottom left). In 2005, Moran's I of -0.18 indicates a pattern virtually unchanged that is not significantly different from a random distribution either (right).

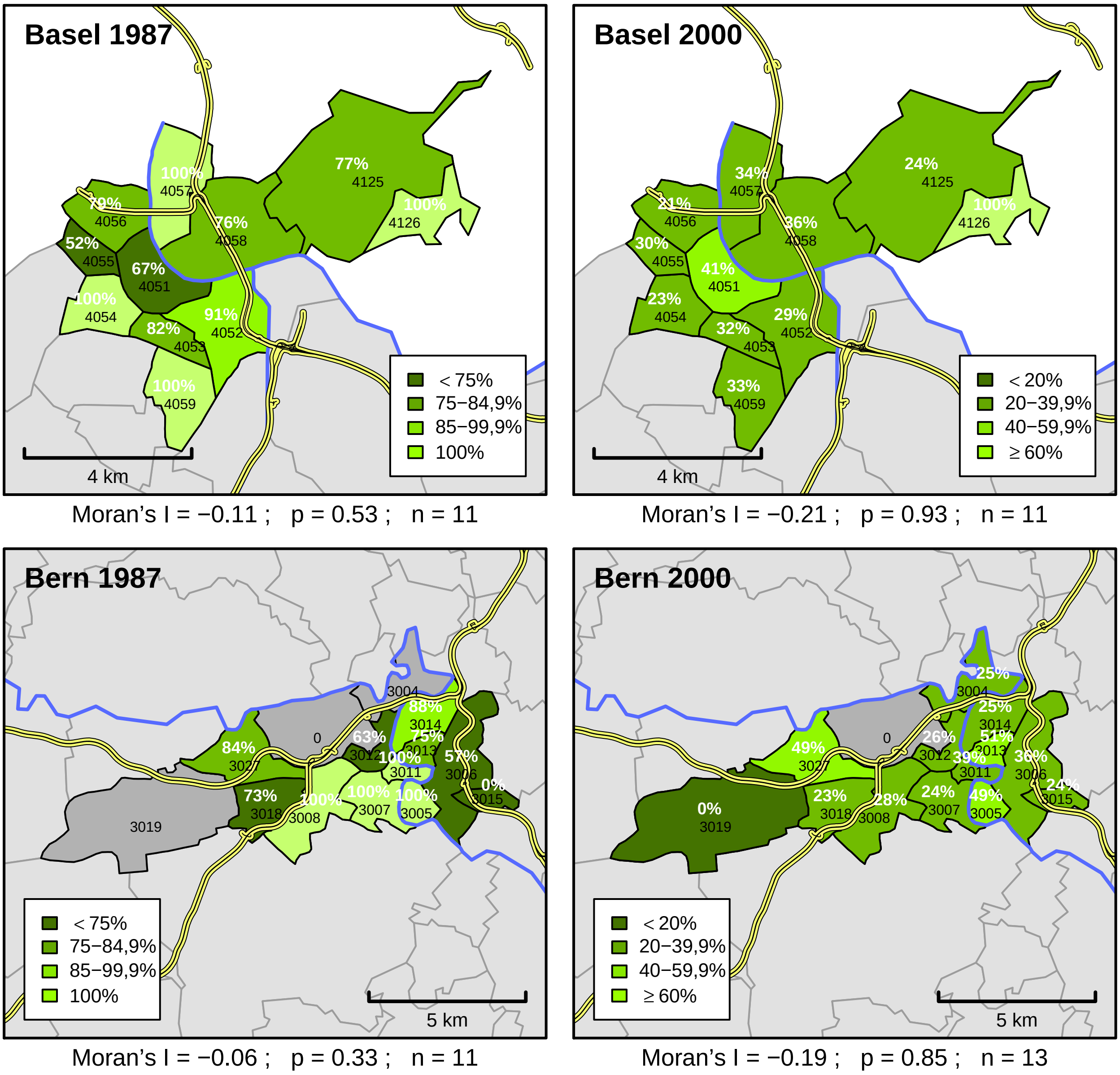


Figure 5.13 Behavioral response to crime in Basel and Bern. Percentage of respondents by ZIP code district who – walking alone in their neighborhood after 10 pm – do not stay away from certain streets, areas, or people to avoid crime.[a] In Basel in 1987, Moran's I of -0.11 shows a slightly dispersed pattern that is not significant, though (top left). In 2000, Moran's I of -0.21 indicates a slightly more dispersed distribution, but the pattern still fails a test of significance (top right). In Bern in 1987, Moran's I of -0.06 indicates a slightly dispersed pattern that is not significant (bottom left). In 2000, Moran's I of -0.19 indicates a slightly more dispersed pattern that is not significant either (bottom right).

[a]The question format and answer categories used to measure behavioral changes in response to crime have changed between the 1987 and 2000 surveys. As a consequence, the category boundaries of the color scales are different for 1987 and 2000 as indicated in the key, reflecting the changes in the underlying scales of the ordinal variable. Absolute percentage levels are thus not directly comparable. However, since the same methodology has been applied consistently in developing the color scales, a given neighborhood's position relative to others may be meaningfully compared (cf. Chapter 3.2.4 for details on the methodology of defining the color scales).

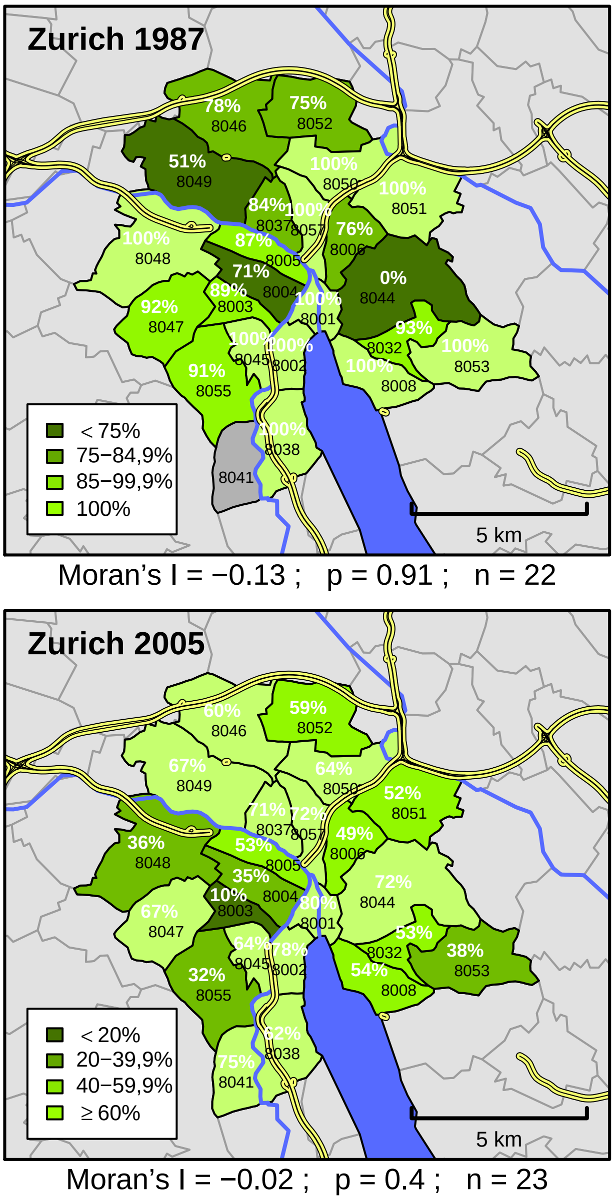


Figure 5.14 Behavioral response to crime in Zurich. Percentage of respondents by ZIP code district who – walking alone in their neighborhood after 10 pm – do not stay away from certain streets, areas, or people to avoid crime.[a] For 1987, a value of Moran's I of -0.13 indicates a slightly dispersed pattern that is not significant however (top left). For 1998, Moran's I of -0.03 exhibits a spatial distribution that is not significantly different from a random pattern (top right). In 2005, Moran's I of -0.02 again suggests a pattern that is indistinguishable from a random distribution (bottom).

[a]The question format and the scales applied to measure behavioral changes in response to crime have changed between the 1987 and 1998 surveys (cf. Footnote to Fig. 5.13).

certain areas. By 2000, the distribution of feature attributes had come more to resemble the pattern of Basel. In a majority of ZIP code districts, only a third of respondents or less didn't bother to go out of harm's way. As in Basel, the brightest spots were near the center (ZIP 3011), where 40 percent of respondents or more declared there were no streets or areas in their neighborhood they deliberately avoided (Fig. 5.13 bottom).

In 1987, Zurich like Bern showed no easily interpretable pattern. Whereas in some neighborhoods, notably in the city center, along the lake shore, and on the outskirts, no or few respondents avoided any areas, there were some pockets of high levels of fear scattered along the city's boundaries in the north, east, and west. By contrast, the chart illustrating response patterns for 1998 shows a city, where in all but one ZIP code district, a majority of respondents said they stayed clear from certain streets or areas in their neighborhood out of safety concerns. By 2005, the picture looked noticeably different again: in most neighborhoods, a clear majority of respondents indicated they did not stay away from any place in their neighborhood to avoid crime. In fact, by 2005 the pattern of fear as a behavioral response to crime had come to resemble much more the pattern of how safe respondents felt. The neighborhoods with the highest levels on both indicators of fear are found on the city's outskirts along the main traffic routes into town. Interestingly, the two ZIP code districts surrounding the Langstrasse (ZIP 8004 and 8005) that were the scene of more than a fair share of serious crime in Zurich in 1998 had the lowest percentage of respondents indicating that they avoided certain streets or areas in their neighborhood. By 2005, these figures had barely moved, but because reported fear levels dropped by so much elsewhere in town, the two ZIP code districts were now among the worst afflicted (Fig. 5.14).

5.3 Disorder and Neighborhood Development

5.3.1 Physical and Social Disorder

The phenomenon of disorder was first introduced in the SCS questionnaire in 1998. Respondents were asked whether they had noticed any graffiti or litter, or spotted groups of disorderly people loitering on the streets of their neighborhood. That year, it was simply recorded whether respondents answered in the affirmative. In 2000 and 2005, the same item featured in the questionnaire again, but the answer categories now distinguished between graffiti, litter, or groups of disorderly people, allowing respondents to answer the question more specifically.

For the purpose of analysis, the variables capturing different forms of incivilities and disorder of the 2000 and 2005 data sets were recoded into two dummy variables indicating the presence of signs of physical or social disorder in the neighborhood, respectively. The former encompasses instances of graffiti or litter, whereas the latter captures loitering groups of disreputable people (cf. Appendix A.3)

The percentage of respondents who observed physical disorder appears to have risen since 1998 in all five cities under study (Fig. 5.15). It is possible that the changes in the questionnaire format affected response patterns, but it seems implausible that these changes totally undermine the validity of the present findings. If in 1998 signs of physical and social disorder were taken together as a single survey item, whereas for the subsequent years the questionnaire distinguished between the two, the long-term trend should go down not up if only signs of physical disorder are taken into account for the analysis in 2000 and 2005 – as is the case here – and respondents had observed these phenomena in roughly equal measure.

At the neighborhood level, in Geneva, both the pattern of signs of physical and social disorder in 1998 and of physical disorder in 2000 exhibited the characteristic divide of the city between its *rive gauche* and *rive droite* parts. On the northern side of the Rhône river in 1998, the ZIP code districts 1202 and, in particular 1201 and 1209, appeared to be more affected by signs of incivilities than the rest of the city. Two years on, physical disorder seemed more concentrated in the city

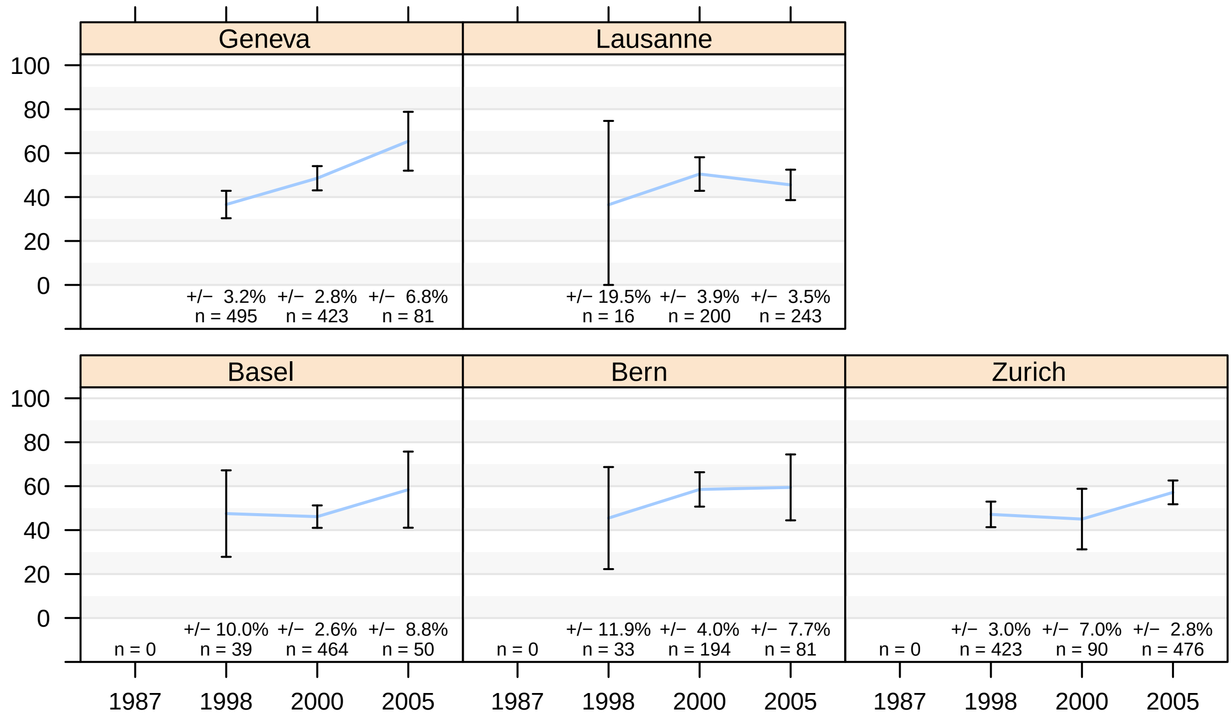


Figure 5.15 Physical disorder 1998-2005. Percentage of respondents who spotted graffiti on the walls or lots of rubbish lying around in the streets close to their home. NOTE: The wording of the survey question and answer categories differed slightly between survey years.

center and the neighborhoods around the confluence of the Rhône and Arve rivers. In contrast to the patterns of fear of crime, there appeared to be less of a gap between the city and its surrounding suburban communities in matters of disorder. Both physical and social disorder appeared to be urban phenomena that tended to cluster in some neighborhoods within the city proper (Fig. 5.16 top).

In 2000, respondents in Lausanne's northern and western ZIP code districts were on average more likely to have spotted signs of physical disorder than the residents of neighborhoods in the east. By 2005, the characteristic divide had disappeared, though. Most remarkable form a theoretical perspective, in the year 2000, those ZIP code districts with the highest levels of physical disorder corresponded for the most part to the neighborhoods that also experienced higher levels of fear of crime. By 2005, however, the link had become more tenuous (Fig. 5.16 bottom).

In Basel, the pattern of respondents who observed signs of physical disorder in their ZIP code district is remarkably homogenous. Some 50 percents on average had observed graffiti or litter in almost every neighborhood of the city proper. The percentage was slightly higher in the northernmost ZIP code district (4057), and somewhat lower in two neighborhoods on the southern outskirts (ZIP 4054 and 4059) as well as in the suburban communities of Riehen and Bettingen. In Bern, the pattern is different. The city center (ZIP 3011) appeared to be a hotbed of physical disorder and there was a belt of neighborhoods stretching from the north-eastern end of the city to its western parts, where more respondents indicated the presence of signs of physical disorder. The brightest spot close to the city center was the Kirchenfeld neighborhood (ZIP 3005), home to most of the Swiss capital's diplomatic missions (Fig. 5.17 top).

In Zurich, the spatial distribution of physical and social disorder exhibits positive autocorrelation. In 1998, disorder appeared to be more concentrated in the south western part of the city with the remainder apparently less affected, except for suburban Schwamendingen (ZIP 8051). By 2005, the pattern had become substantially more heterogeneous. Physical disorder still appeared

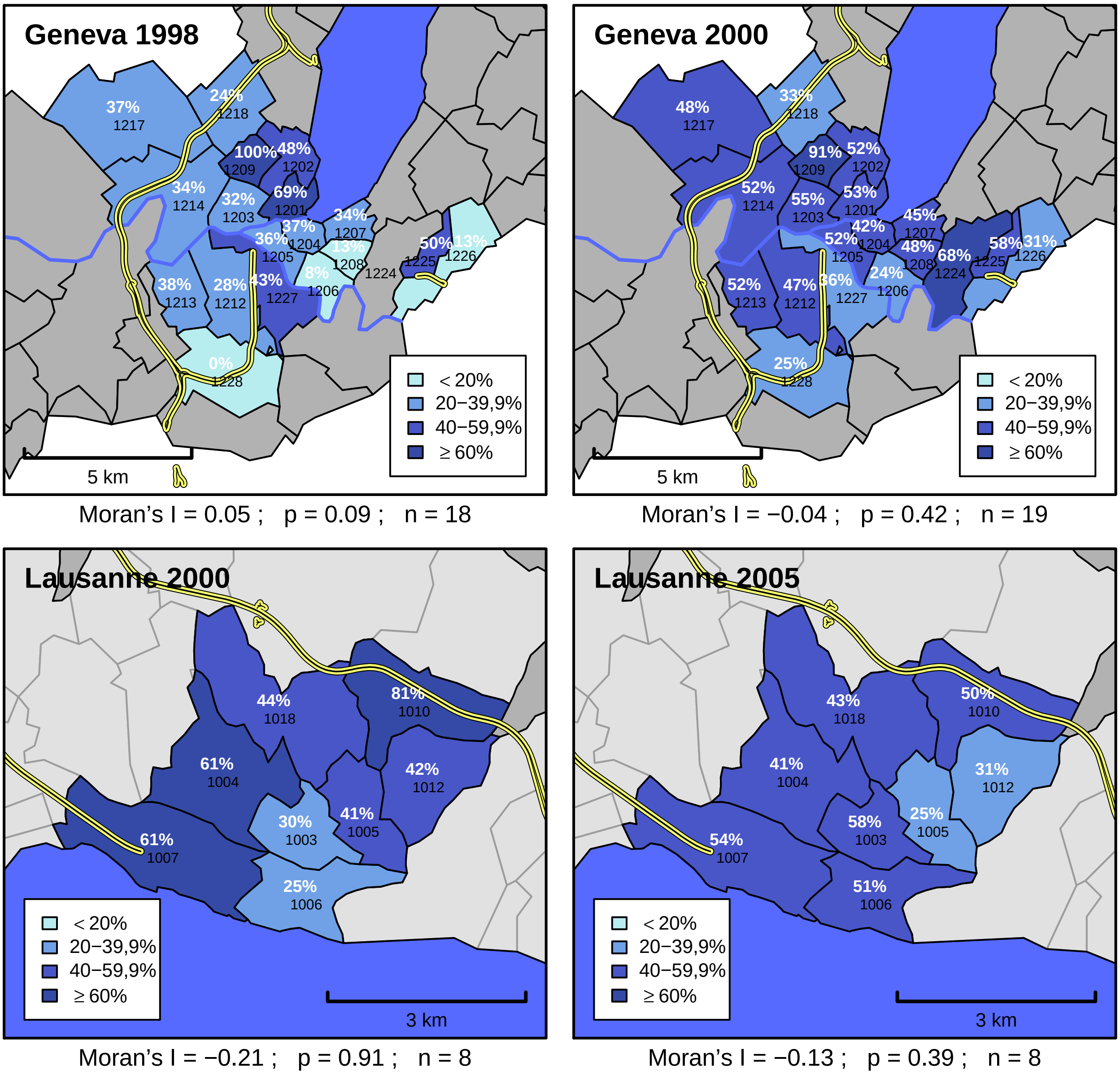


Figure 5.16 Physical disorder in Geneva and Lausanne. Percentage of respondents by ZIP code district who spotted graffiti on the walls or lots of rubbish lying around in the streets close to their home. In Geneva in 1998, Moran’s I of 0.05 shows a marginally clustered pattern that is almost statistically significant ($p < 0.1$; top left). In 2000, Moran’s I of -0.04 shows a slightly negative autocorrelation that is not significant, however (top right). In Lausanne, Moran’s I of -0.21 for the year 2000 reveals a dispersed pattern that is not significant (bottom left). For 2005, Moran’s I of -0.13 indicates a slightly dispersed pattern but the value is not significant either (bottom right).

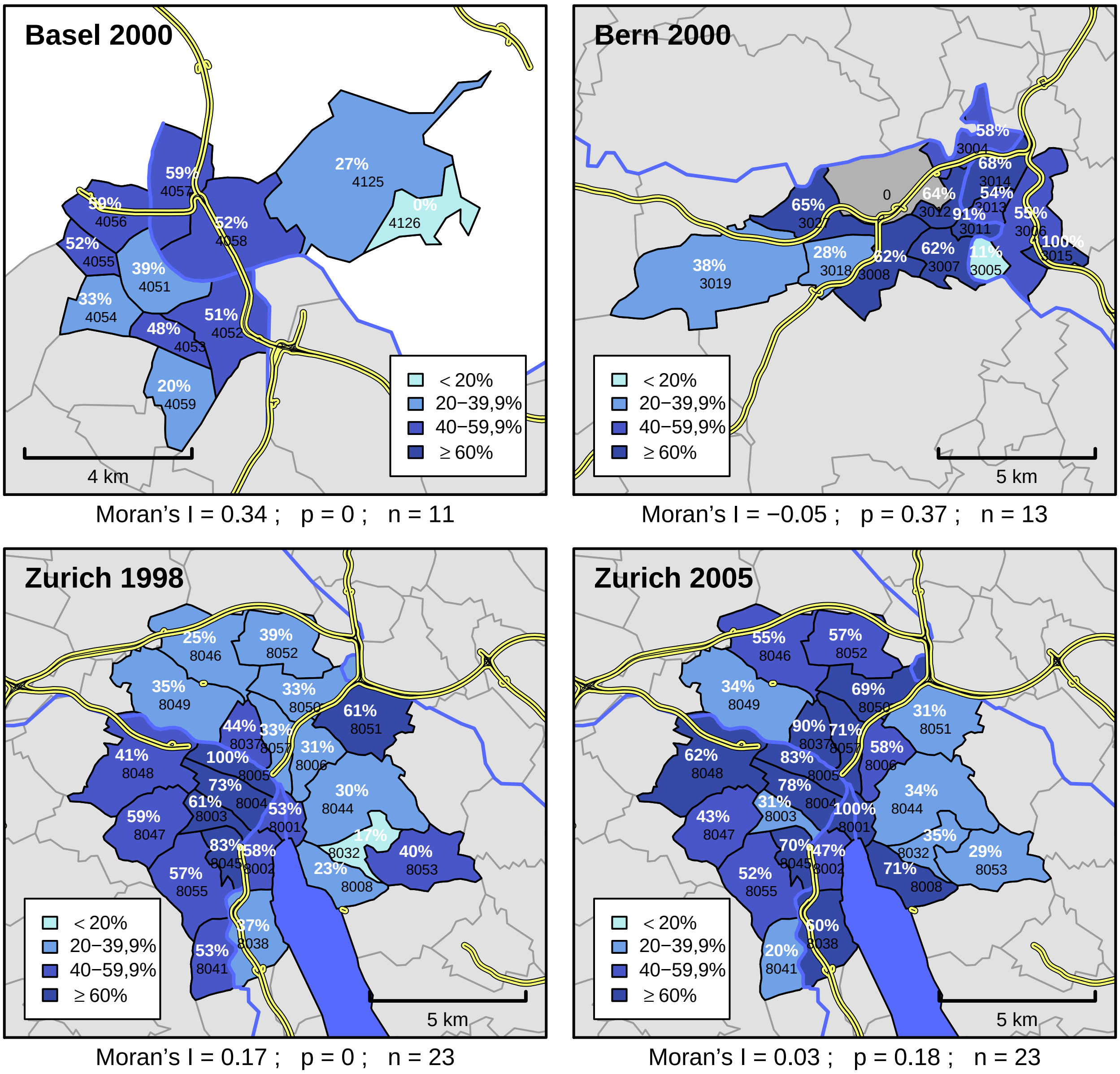


Figure 5.17 Physical disorder in Basel, Bern, and Zurich. Percentage of respondents by ZIP code district who spotted graffiti on the walls or lots of rubbish lying around in the streets close to their home. In Basel, Moran's I of 0.34 shows a clustered pattern that is highly significant ($p < 0.01$; top left). In Bern, Moran's I of -0.05 signals a slightly dispersed pattern that is not significant, though (bottom right). In Zurich, a Moran's I of 0.17 that is highly significant ($p < 0.01$) indicates a clustered pattern in 1998 (bottom left). By 2005, the percentage of respondents who spotted signs of physical disorder has risen overall, but Moran's I of 0.03 that is not statistically significant indicates the clustering has all but disappeared (bottom right).

most concentrated in the inner city neighborhoods, but it was also more frequently observed in neighborhoods on the city's outskirts. Worst affected were the inner city neighborhoods around the Langstrasse (ZIP 8004 and 8005). The impact of the open air drug markets of the early 1990s in this area of town is clearly visible. In 1998, 73 and 100 percent of respondents, respectively, reported they had witnessed signs of physical or social disorder in these areas. By 2005, the equivalent figures were a still staggering 78 and 83 percent of respondents, respectively (Fig. 5.17 bottom).

5.3.2 Social Cohesion

Until 1998, the SCS questionnaires included an item targeting local levels of social cohesion. The wording of the survey item remained unaltered between 1987 and 1998, asking respondents whether in their neighborhood people tended to look after each other or rather minded their own business.

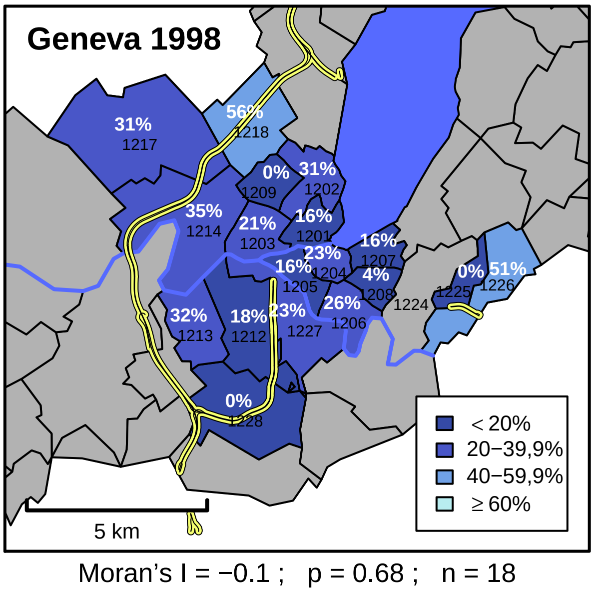


Figure 5.18 Social cohesion in Geneva. Percentage of respondents by ZIP code district who rated their neighborhood as a place where people tended to help each other out rather than mind their own business. A value of Moran's I of -0.1 suggests a dispersed pattern that fails a significance test, though.

In the City of Geneva, the characteristic split between *rive gauche* and *rive droite* also reappeared regarding social cohesion. Respondents north of the Rhône river were marginally more likely to assert that theirs was a neighborhood where residents supported each other, but the percentage differences are quite small. Unlike for the measures of fear, the differences between the city proper and the surrounding suburban communities were negligible. The contrast was much sharper with the counties located farther away from the city center, which displayed either very high or very low levels of social cohesion (Fig. 5.18).

In terms of social cohesion, Basel appeared divided back in 1987 with the Rhine river being the physical landmark of this cleavage running through the city. Residents of neighborhoods in Grossbasel, southwest of the river, were much less likely to assert that they and their neighbors were sticking it out together than the residents of the ZIP code districts north of the river in Kleinbasel and the commuter towns of Riehen and Bettingen. The neighborhood where residents were least likely to help each other out was the city center (ZIP 4051). A similar pattern of a divided city also emerges for Bern. As in Basel, respondents in the city center (ZIP 3011) rated social cohesion in their neighborhood lowest. North and to the east of the center across the Aare river lay four ZIP code districts, where residents indicated neighbors tended to support each other. In areas west and south of the center levels of social cohesion appeared to be lower (Fig. 5.19 top).

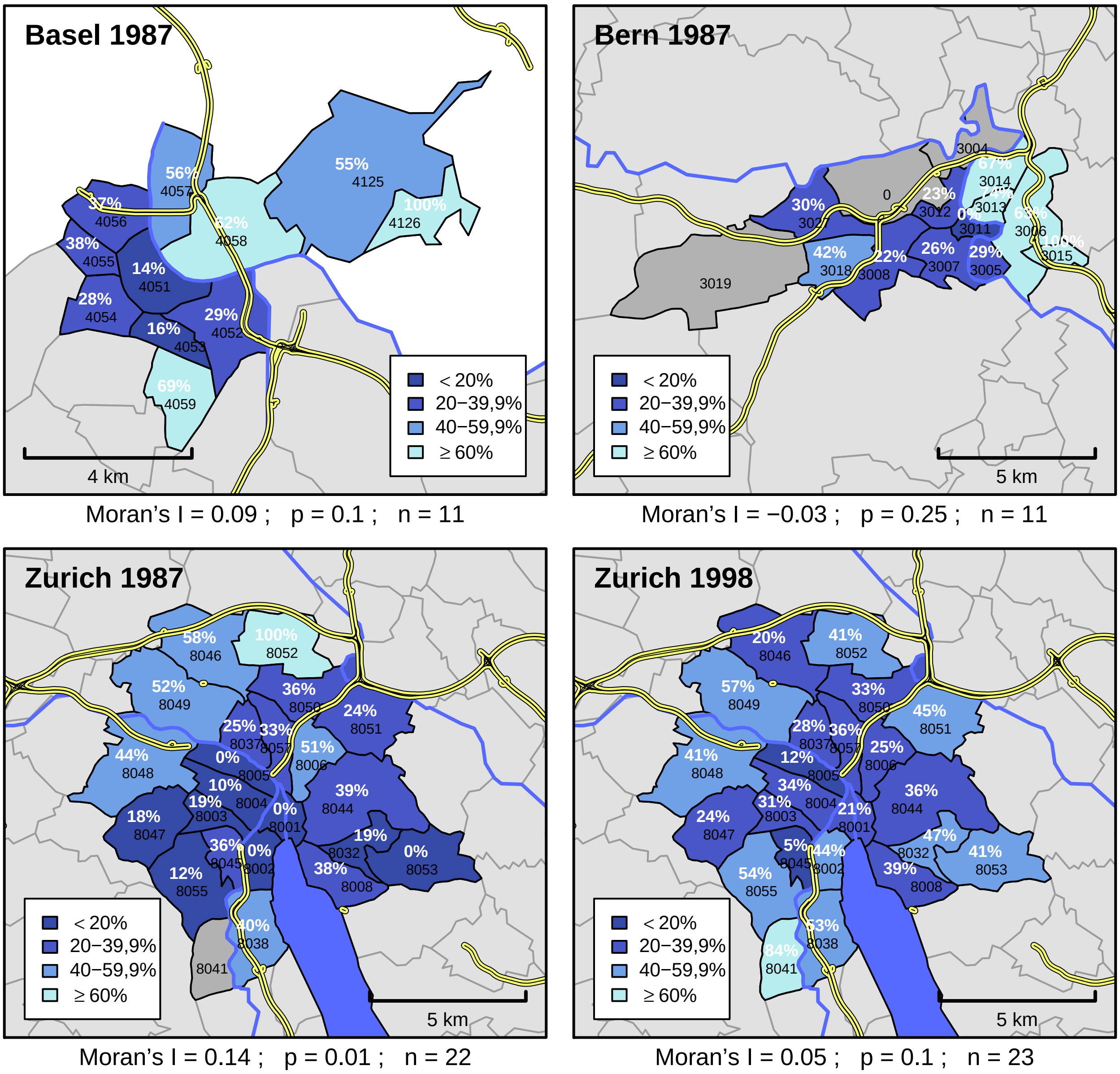


Figure 5.19 Social cohesion in Basel, Bern, and Zurich. Percentage of respondents by ZIP code district who rated their neighborhood as a place where people tended to help each other out rather than mind their own business. For Basel, Moran's I of 0.09 that is bordering on statistical significance ($p < 0.1$) indicates a slightly clustered pattern (top left). In Bern, Moran's I of -0.03 that is not significant suggests a pattern that is not different from a random distribution (top right). In Zurich in 1987, Moran's I of 0.14 reveals moderate clustering that is significant at the 1 percent level of confidence (bottom left). In 1998, Moran's I of 0.05 indicates a less clustered pattern that is also only marginally significant ($p < 0.1$; bottom right).

The pattern emerging for Zurich in 1987 is not entirely different. As in the other cities, low levels of social cohesion were concentrated in the inner city neighborhoods west of the Limmat river running from the lake through the city. The percentage of respondents who thought theirs was a cohesive neighborhood tended to rise with increasing distance from the city center. With hindsight, the particularly low levels of social cohesion in the Langstrasse district back in 1987 can be interpreted as a sign of an already weakened social fabric that boded ill for the challenges to law and public order that these neighborhoods were going to face over the following decade. By 1998, the pattern had shifted, albeit more in appearance than in substance. Low levels of social cohesion still appeared to be more concentrated in the neighborhoods of the inner city, even though the excessively low values recorded a decade earlier had improved to some extent (Fig. 5.19 bottom).

5.4 Public Attitudes towards the Police

5.4.1 Police Effectiveness

As with fear of crime, the SCS questionnaires since the beginning of the survey have included an item targeting public opinion about the effectiveness of the police in controlling crime in the area. Respondents typically were asked to rate how well the police discharged their duties in this regard. In its earliest edition, the questionnaire used a scale from 0 to 10 to gauge respondents' opinion on the effectiveness of the police, which was later replaced by a dummy variable asking respondents whether the police were doing a "rather good" or "rather bad" job. In its 2000 and 2005 editions, the survey used a four-step scale to track popular satisfaction with the police.

As the wording of the survey question and the answer categories continued to evolve over the time span of the SCS surveys, comparisons of results between different years have to be made cautiously and with the shifting answer categories in mind. No such caveats have to be made for the last two SCS editions in 2000 and 2005, which used identical question formats and answer categories.

The percentage of respondents who rate the effectiveness of the police in controlling crime in their neighborhood as good or very good has remained high in Geneva and Lausanne over the study period, hovering above 80 percent (Fig. 5.20). In the three Swiss German cities, the equivalent figure appears to have risen since 1987, albeit from a lower base. However, the changing format of the underlying question between 1987 and 1998 render direct comparisons difficult. Considering the 1998 to 2005 period only, there still appears to have been a marginal improvement in the percentage of respondents satisfied with police effectiveness. The large measurement errors make it impossible to know, though, whether these differences reflect true shifts in popular opinion or are merely due to chance, except in Zurich, where the observed difference is statistically significant. Remarkably, in 2005, over 80 percent of survey respondents were satisfied with police performance in all five cities.

At the neighborhood level, in Geneva, the usual division of popular opinion between the neighborhoods of *rive gauche* and *rive droite* did not apply to judgments about the effectiveness of the police in controlling crime. In 1998, the pattern showed a city, in which the neighborhoods on both sides of the Rhône river fell into the first and second quartile in roughly equal numbers. In 2000, most ZIP code districts now fell into the second quartile, yet the gap characterizing the city in matters of fear and disorder was still absent (Fig. 5.21).

For Lausanne the average scores may actually be compared since the underlying scales of the ordinal variables have not changed between the two polling years. In 2000, public opinion on police effectiveness was divided between the neighborhoods in the north and west of the city and their eastern counterparts. Least satisfied with police performance were residents in the city center. By 2005, the pattern had become more heterogeneous, but the pattern still exhibited the usual division along the axis from northeast to southwest (Fig. 5.22).

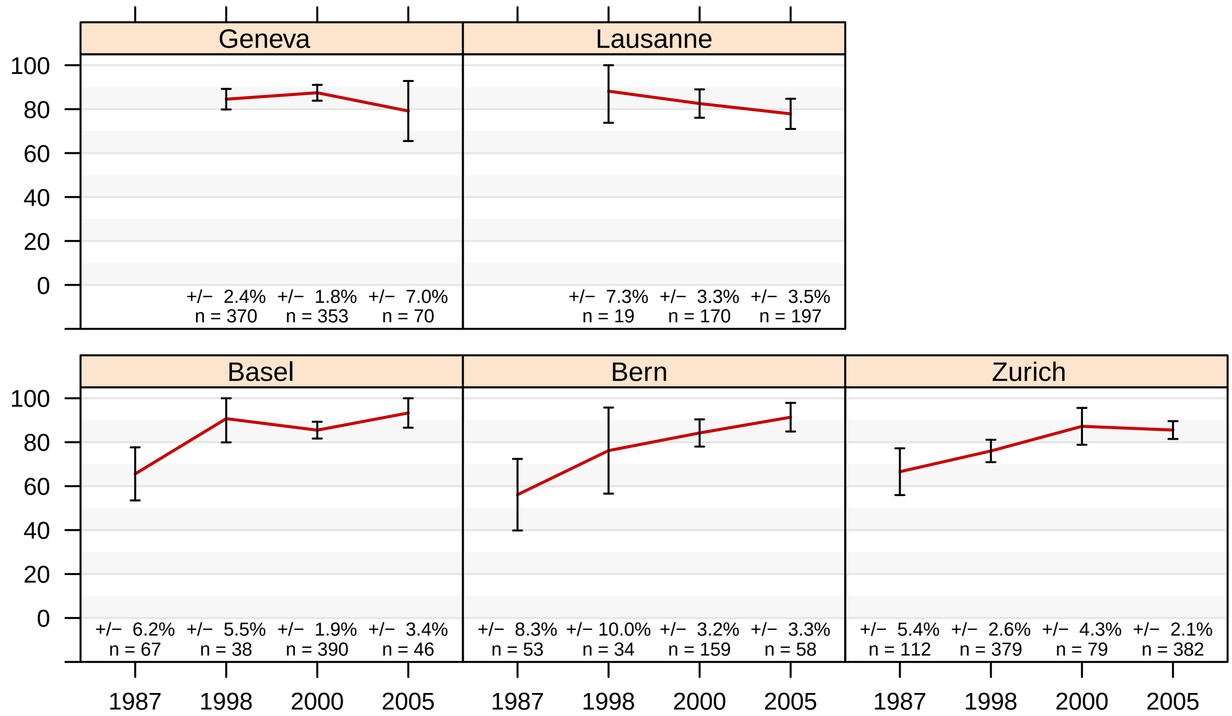
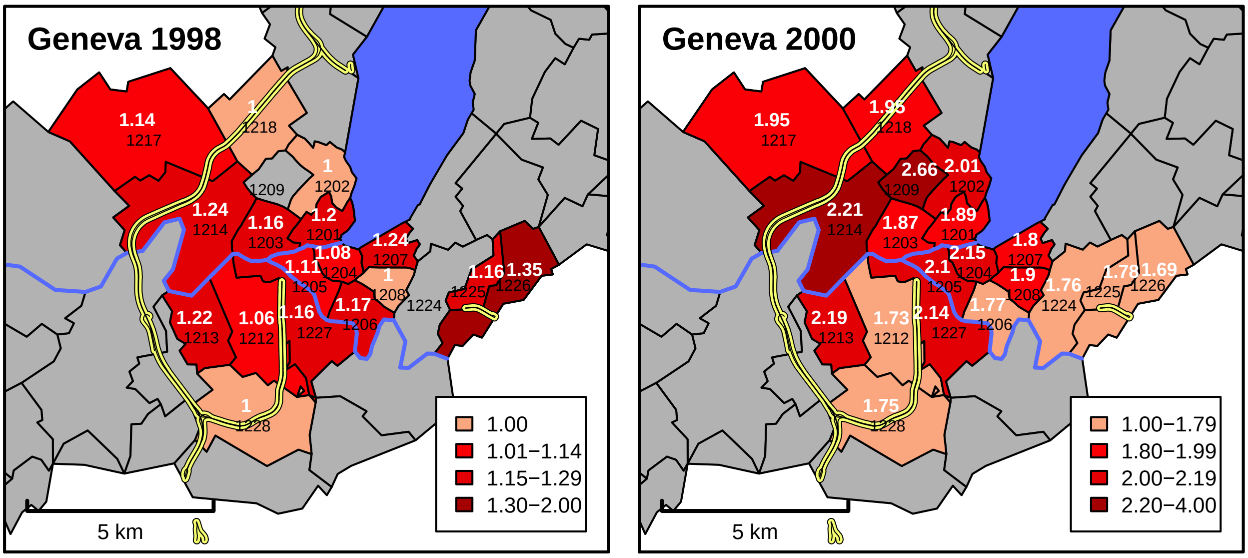


Figure 5.20 Public assessment of police effectiveness 1987-2005. Percentage of respondents who rate the effectiveness of the police in controling crime in their area as “very good” or “rather good.” NOTE: The wording of the survey question and answer categories differed between survey years.

In the City of Basel in 1987, respondents were least satisfied with the police in the city center and three adjacent ZIP code districts on the outskirts of Grossbasel. By 2000, the pattern had become much more homogenous: there were no pockets of police dissatisfaction any more, but neither were there any neighborhoods left in the top quartile rating police effectiveness unequivocally positive within the city proper. In Bern in 1987, the neighborhoods where residents were least satisfied with the police were scattered along the city’s boundaries. By 2000, the pattern like in Basel had become considerably more homogenous (Fig. 5.23).

In Zurich in 1998, popular disenchantment with the police appeared greatest in the central ZIP code districts (ZIP 8001, 8004, and 8008) as well as in suburban Schwamendingen (ZIP 8051). By 2005, by contrast, the neighborhoods, where residents had previously scored police effectiveness lowest, had all improved their relative position (Fig. 5.24).

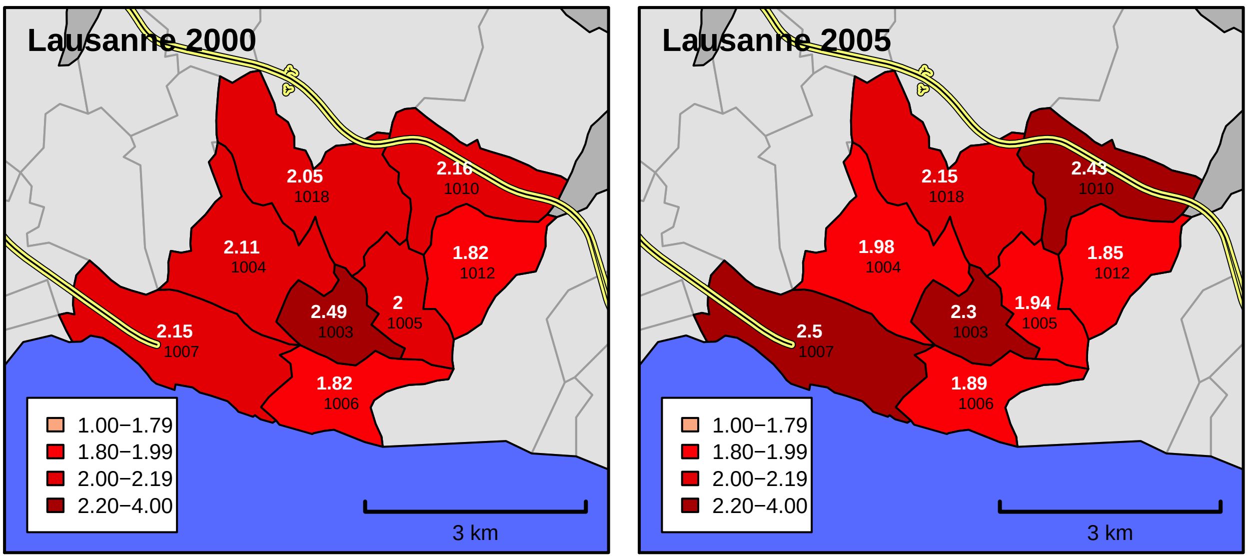
In 2005, the SCS survey contained an additional item on police performance asking respondents if they wished to see the police patrol their neighborhood more frequently. Regarding this desire for more frequent police patrols, the east-west divide that characterizes Lausanne was particularly pronounced. Almost forty percent or more of residents of the city’s northern and western ZIP code districts indicated they would like to see the police patrol their neighborhood more often, compared to barely a quarter of residents in the city center and eastern parts. Interestingly, there appeared to be an inverse relationship between how respondents in an area rated the overall effectiveness of the police and their desire to have more frequent patrols. Residents in the northern and western parts who gave the police lower marks on average for effectiveness clearly would like to see more of them, whereas residents in the eastern parts who were less critical in their judgments also appeared more satisfied with the status quo. The most prominent exception to this rule was the city center (ZIP 1003): both in 2000 and 2005 residents in this part of town were among the sternest critics of the police, yet few called for a stepping up of patrols in their neighborhood (Fig. 5.25).



Moran's I = −0.06 ; p = 0.48 ; n = 17 Moran's I = 0.13 ; p = 0.01 ; n = 19

Figure 5.21 Public assessment of police effectiveness in Geneva. Average score by ZIP code district of how respondents rated the overall effectiveness of the police in combating crime in their neighborhood on a scale from 1 to 2 and 1 to 4 respectively (a lower score indicating higher levels of approval).[a] In 1998, Moran's I of -0.06 indicates a marginally dispersed pattern that is not statistically significant, though (left). In 2000, by contrast, Moran's I of 0.13 shows a moderately clustered pattern that is statistically highly significant ($p < 0.01$; right).

[a]The scales for scoring police effectiveness have changed between the 1998 and 2000 surveys (cf. Footnote to Fig. 5.23).



Moran's I = −0.13 ; p = 0.43 ; n = 8 Moran's I = −0.18 ; p = 0.73 ; n = 8

Figure 5.22 Public assessment of police effectiveness in Lausanne. Average score by ZIP code district of how respondents rated the overall effectiveness of the police in combating crime in their neighborhood on a scale from 1 to 4 (a lower score indicating higher levels of approval). For the year 2000, Moran's I of -0.13 shows a slightly dispersed distribution that is not statistically significant, though (left). For 2005, Moran's I of -0.18 indicates the distribution is slightly more dispersed but still not statistically significant (right).

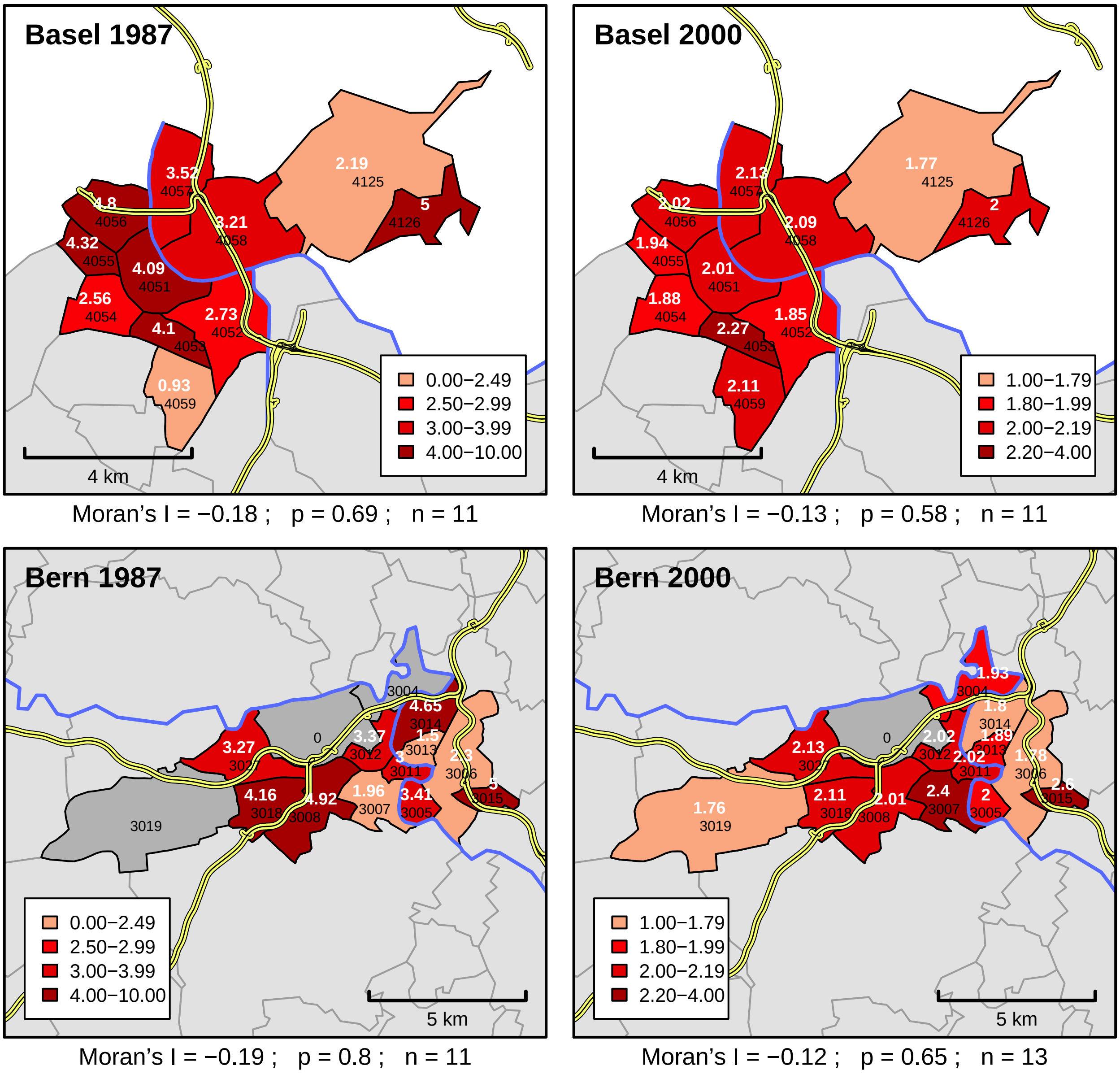


Figure 5.23 Public assessment of police effectiveness in Basel and Bern. Average score by ZIP code district of how respondents rated the overall effectiveness of the police in combating crime in their neighborhood on a scale from 1 to 10 and 1 to 4, respectively (a lower score indicating higher levels of approval).[a] In Basel in 1987, Moran's I of -0.18 indicates a slightly dispersed pattern but is not significant (top left). In 2000, Moran's I of -0.13 indicates a less dispersed pattern that is not significant either (top right). In Bern in 1987, Moran's I of -0.19 signals a dispersed pattern that is not significant (bottom left). In 2000, Moran's I of -0.12 indicates a slightly less dispersed pattern that is still not significantly different from a random distribution (bottom right).

[a]The scales for scoring police effectiveness have changed between the 1987 and 2000 surveys. As a consequence, the category boundaries of the color scales of the two maps are different, reflecting the changes in the underlying scales of the ordinal variable, and absolute score levels are thus not directly comparable. However, since the same methodology has been applied consistently in developing the color scales, a given neighborhood's position relative to others may be meaningfully compared (cf. Chapter 3.2.4 for details on the methodology of defining the color scales).

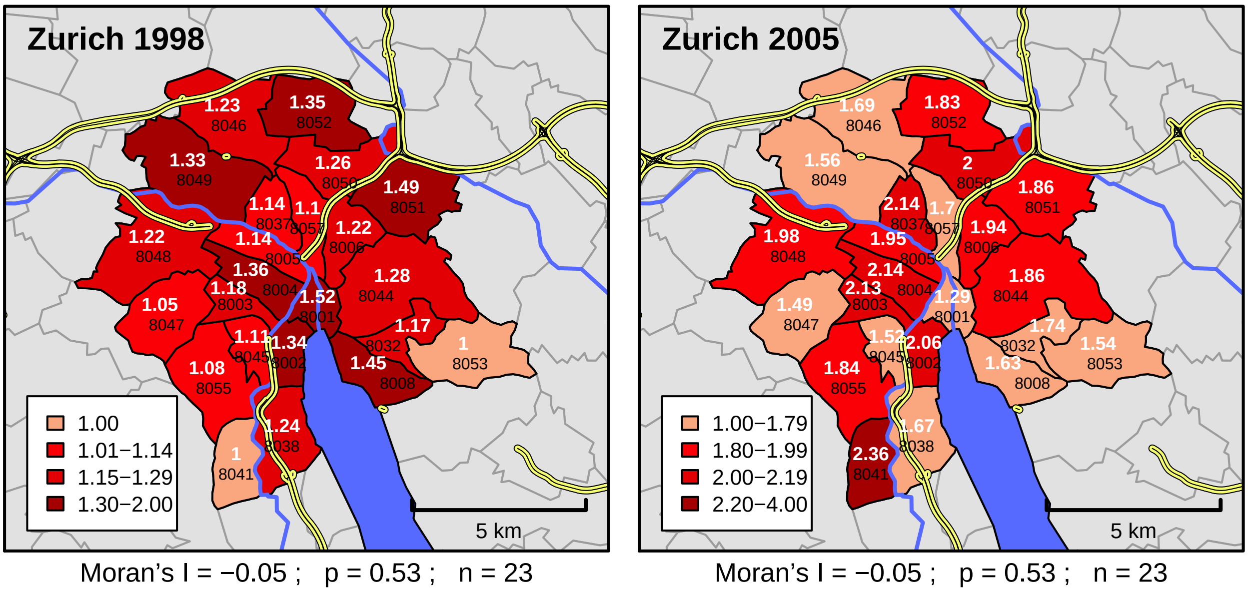


Figure 5.24 Public assessment of police effectiveness in Zurich. Respondents' average score rating by ZIP code district of the overall effectiveness of the police in combating crime in their neighborhood on a scale from 1 to 2 and 1 to 4, respectively (a lower score indicating higher levels of approval).[a] For 1998, Moran's I of -0.05 indicates a spatial distribution that is not significantly different from a random pattern (left). For 2005, both the value of Moran's I and the corresponding level of significance are virtually identical (right).

[a]The scales for scoring police effectiveness have changed between the 1998 and 2005 surveys (cf. Footnote to Fig. 5.23).

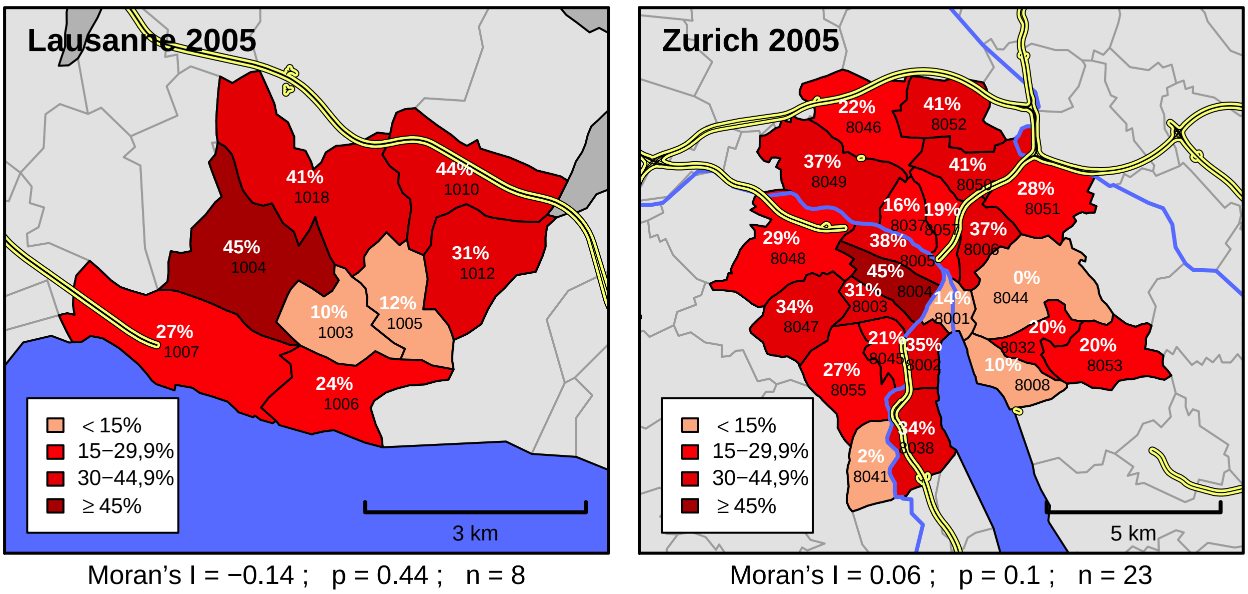


Figure 5.25 Police patrols in Lausanne and Zurich. Percentage of respondents by ZIP code district who would prefer to see more frequent police patrols in their neighborhood. For Lausanne, Moran's I of -0.14 indicates a slightly dispersed pattern that is not statistically significant, though (left). For Zurich, Moran's I of 0.06 indicates a slightly clustered distribution that is significant at the 10 percent confidence level ($p < 0.1$; right).

In Zurich, not surprisingly, the desire to have more frequent police patrols was felt most keenly in the inner city neighborhoods of the Langstrasse district (ZIP 8004 and 8005). Not far behind were the city's western, north eastern, and southern outskirts that lay around the main traffic arteries leading into town.

5.4.2 Professionalism and Impartiality

Since 1998, the SCS questionnaires have also included an item asking respondents whether they held their police to be impartial and fair or, conversely, not always immune to bias. Respondents were asked to rate the impartiality of the local police on a scale from 1 to 4, indicating whether they held the police to be "always impartial", "quite impartial", "sometimes partial", or "always partial."

For Zurich in 1998 an interesting spatial pattern emerges. The residents who rated police impartiality most critically tended to live in those ZIP code areas in or around the Langstrasse district. Apart from that, the ZIP code areas with higher average scores form an arc along the city's boundary. By 2005, the picture had brightened up a bit. The outer neighborhoods now mostly thought of the police as fairer than the ZIP code areas in the inner circle of the city (Fig. 5.26).

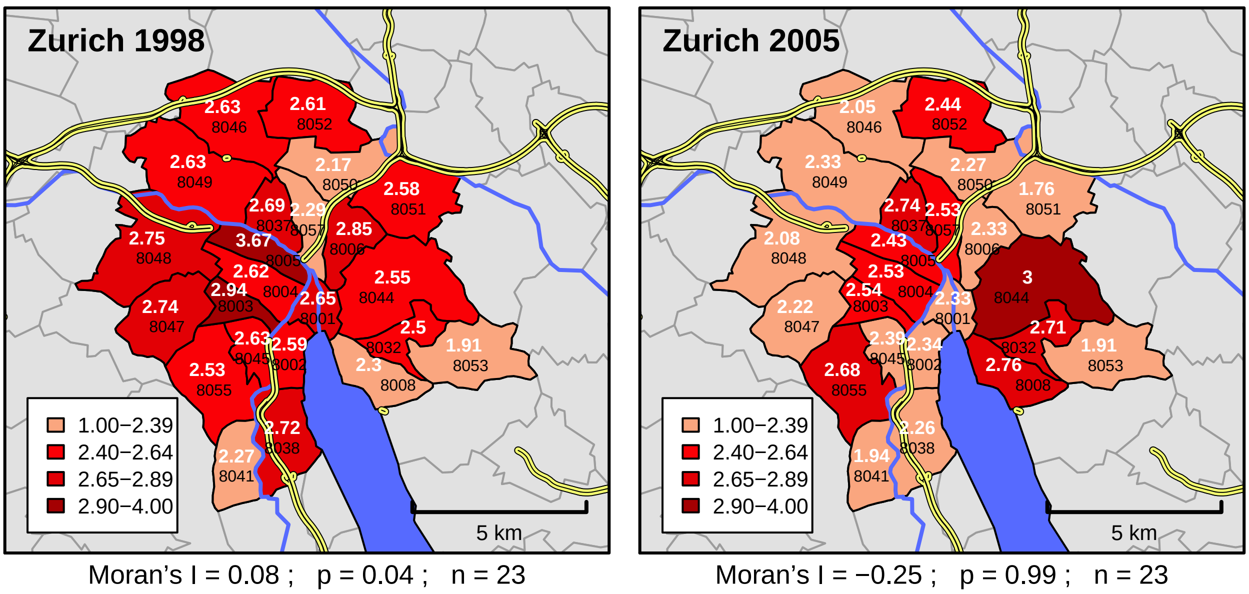


Figure 5.26 Police impartiality in Zurich. Average score by ZIP code district of how respondents rated the impartiality of the police on a scale from 1 to 4 (a higher score indicating stronger police bias). In 1998, Moran's I of 0.08 shows a slightly clustered pattern that is statistically significant at the 5 percent level of confidence (left). In 2005, by comparison, Moran's I of -0.25 indicates a slightly dispersed pattern, but the p-value suggests that the pattern could be entirely due to chance (right).

Fairness and the legitimacy of the police are critical concepts to the idea of community policing. It thus only makes sense that an evaluation of community policing that employs Geographic Information Systems to study the spatio-temporal patterns of key variables of community policing impact would also test for any "spatial inequalities", meaning any clusters of disaffection with the local police force.

The encouraging news on this score is that the pattern for Zurich in 1998 displayed above was the only example of a spatial distribution of how respondents rated the impartiality of the police that exhibited moderate clustering. For no other city or any other year, for which data was available for neighborhood-level analysis, did the value of Moran's I signal a significantly clustered or dispersed pattern at any conventional level of confidence. And this goes both for how respondents rated the impartiality of the police as well as for the indicator of how satisfied they were following a personal encounter about how the police had treated them or dealt with their demand. In other words, there appear not to have been any areas within the five cities, where popular opinion on the police was either decidedly more positive or more negative than in the city as a whole. Considering that the same cannot be said about the patterns of fear of crime or neighborhood disorder, one conclusion form this observation is that popular opinion on the professionalism and the impartiality of the police is quite evenly balanced across Switzerland's five major urban areas.

5.5 Concluding Remarks

The long-term trend analysis at the city level revealed that fear of crime diminished in the three Swiss German cities of Basel, Bern, and Zurich over the study period, whereas residents of Geneva and Lausanne on average have felt more secure all along but were more concerned with burglary risk. The percentage of residents of Basel, Bern, and Zurich who feel safe walking alone in their neighborhood after 10 pm has increased since 1987, while the percentage of respondents who deliberately avoid certain places for safety reasons at night has fallen over the study period. In Geneva and Lausanne, the percentage of respondents feeling safe on a nightly stroll has held more or less steady between 1998 and 2005, whereas the number of people avoiding some places rose slightly, albeit from a much lower base. In 2005, the percentage of respondents feeling safe according to both indicators was more or less identical in all five urban areas. By contrast, fear measured as the expected risk of a burglary of one's home has remained more or less stable in all five cities but was more prevalent in Geneva and Lausanne than in the three Swiss German cities.

The percentage of respondents who spotted signs of physical disorder such as graffiti on walls or garbage strewn in the streets near their home increased in all five cities between 1998 and 2005. There were virtually no significant differences across the five cities neither in the thrust nor in the magnitude of the problem.

Finally, the percentage of respondents who were satisfied with police effectiveness in controlling crime in their area of residence has consistently topped 80 percent in Geneva and Lausanne. In Basel, Bern, and Zurich, this figure has risen since 1987, albeit from a lower base. In 2005, over 80 percent of respondents were satisfied with police performance in all five urban areas.

Most striking, though, was the high degree of similarity of the long-term trends of these indicators at the regional level. Although there were some disparities between Geneva and Lausanne on the one hand and the three Swiss German cities on the other, the variance within a given language region was rather low.

At the neighborhood level, the study found that the three indicators used to measure fear of crime resulted in different spatio-temporal patterns. However, for all the apparent variations, there is one recurrent theme: over the study period, fear levels tended to drop in the city centers and simultaneously to rise in the outskirts and surrounding suburbs.

This shift of fear levels from the city centers to the outskirts became most apparent for fear measured as the perceived threat of crime in the three Swiss German cities of Basel, Bern, and Zurich between 1987 and 2005. In Lausanne and Geneva, for which the time span of available data was much shorter, the spatial dynamics proved less pronounced. The emerging pattern revealed a divide between the city's eastern and north western parts in Lausanne and, albeit to a lesser extent, between *rive gauche* and *rive droite* in Geneva. The study also revealed that over the period of a little over a decade, the position of a neighborhood on the fear of crime scale can shift dramatically, both up and down. Basel, Bern, and Zurich all have neighborhoods that saw their relative levels of fear of crime drop or rise precipitously over the evaluation period.

Equally, if fear was measured as the perceived risk of victimization, the neighborhoods that exhibited higher levels of fear tended to be found on the five cities' outskirts towards the end of the study period. The city centers had the lowest levels of fear in all five major agglomerations. In Lausanne and Geneva, the characteristic divide, respectively, between east and west and between *rive gauche* and *rive droite* weakened, but the two cities were on average also more concerned about burglary risk than the inhabitants of Basel, Bern, and Zurich.

Finally, if fear was measured as the behavioral response to the perceived threat of crime, the spatio-temporal patterns followed a different trajectory, but the central assertion that fear had become increasingly concentrated in the city outskirts at the end of the study period still holds true. In 1998/2000, Basel, Bern, and Zurich showed widespread, elevated levels of fear. Indeed, in the overwhelming majority of neighborhoods in these cities, two thirds or more of respondents indicated that they deliberately stayed away from certain areas or people in their neighborhoods to avoid crime. By 2005, the number of respondents going out of harm's way had plunged steeply and the pattern at least for Zurich had also become more nuanced. Geneva and Lausanne, by contrast, had not witnessed a comparable rise in city-wide fear levels. Whereas the gap between east and west all but disappeared in Lausanne, it reappeared in Geneva between *rive gauche* and *rive droite* but this time, interestingly, with inverted signs: in those parts where more respondents indicated feeling afraid walking alone at night, fewer of them on average actually stayed clear of some areas in their neighborhood.

The spatio-temporal patterns of disorder and social cohesion on the other hand revealed no such clear-cut trends. This may in part be due to the shorter observation periods for these variables that resulted from the shifting emphasis of the authors of the Swiss Crime Survey, which led them to drop some survey items and to include new ones. Nevertheless, one valid observation that can be made is that the highest levels of disorder were concentrated in the city centers, especially of Bern, Geneva, and Zurich, and to a lesser extent, Basel and Lausanne. Physical and social disorder appeared foremost to be urban phenomena that seriously afflicted only a handful of neighborhoods within any one city. A second observation is that social cohesion was lowest in the city centers and tended to increase towards the outskirts, even though this is a gross oversimplification that does scant justice to the much more complex local patterns.

The spatio-temporal patterns concerning the police proved more amorphous still. Even though the spatial distribution of popular judgments about the effectiveness of the police in controlling crime revealed some familiar patterns, notably the divide between eastern and western Lausanne or the chronic hotspot of the Langstrasse district in Zurich, trends on policing in general were much harder to discern. This result is not due to the shifting scales to measure police performance, which admittedly confounded the task of performing trend analyses of the data. The geostatistical tools employed, which proved the lack of any clustering in the data on the impartiality of and satisfaction about personal encounters with the police in Swiss cities, are not contingent on the consistency of the underlying variable scales.

6

Neighborhood Typology

The exploratory spatio-temporal analysis of the indicators of the four theoretical constructs of community policing impact – crime, fear of crime, disorder, and public attitudes towards the police – unearthed some remarkable parallels across the five cities. In particular, the spatial analyses revealed that the spatial patterns of crime rates and perceptions of disorder remained stable over the short and medium run, whereas areas with elevated levels of fear had shifted from the city centers to the peripheries between the late 1980s and 2005. Furthermore, whereas some differences between the Swiss German and Swiss French cities appeared, response patterns within a given language area proved unexpectedly homogenous. It was thus concluded that a community policing impact evaluation over a study period of several years that relies on fear of crime as an outcome measure but does not control for shifting neighborhood characteristics is unreliable at best and positively misleading at worst.

The observation that the spatio-temporal patterns of the outcome measures displayed some striking similarities between the cities under study gave rise to the idea of developing a typology of Swiss urban neighborhoods in order to get a grip on the problem of shifting ecological characteristics. This approach is based on the hypothesis that crime and survey response patterns are shaped to a non-negligible extent by the spatial dynamics of the socio-economic processes unfolding in a city, and that these processes would repeat themselves from one city to another.

The idea is thus to create a neighborhood classification system based on a series of demographic, socio-economic, and environmental indicators in order to classify the urban neighborhoods within the study area into clusters of similar type. To this end, the current study makes use of machine learning and data mining algorithms to create a classification system of urban neighborhoods, which in few dimensions aptly describes the spatio-temporal patterns observed in the high-dimensional data. This initial step reflects earlier research on the use of geodemographic profiling for resource allocation for neighborhood policing (Ashby & Longley 2005; cf. Chapter 2.1.5). In the logic of evaluation research, this process of dimensionality reduction and clustering of the high-dimensional attribute data can be thought of as analogous to finding matching pairs between treatment and control districts for area-based crime prevention programs.

The main rationale for the development of such a neighborhood typology is to enhance the validity of an observational impact evaluation of a complex intervention across multiple sites. By regrouping neighborhoods of a similar type, such a classification system minimizes the within-cluster variance in the contextual variables that affect the outcome variables and may thus confound inferences about program effectiveness in an evaluation design without proper control group(s). The clustering procedure to develop this neighborhood typology thus has a double objective: on the one hand, the clustering should minimize the between-cluster similarity of the contextual variables that may rival the program intervention as the plausible explanation for the observed variance in the indicators of treatment impact. Put differently, urban districts that resemble each other with regard to their demographic composition, socio-economic structure, and the built environment must be grouped into clusters of similar type. On the other hand, the resulting typology of urban districts should account for a maximum of the variance observed in the variables of treatment impact across the neighborhood clusters prior to program implementation. In other words, the typology should capture a maximum of the “neighborhood effects” of the different neighborhood types. In a research design that satisfies both these conditions, the impact of a given policing strategy will be evaluated across urban districts that are not only similar with regard to neighborhood context, but where local residents also collectively expressed similar subjective sentiments regarding crime or the police at the onset of the treatment. This set up allows an evaluator to dismiss a series of threats to the internal validity that otherwise beset a non-experimental research design.

6.1 Developing a Neighborhood Classification System

6.1.1 Spatio-Temporal Data Mining Algorithms

Unsupervised learning – Classifying urban neighborhoods

The current study employs geospatial data mining algorithms to develop the neighborhood typology as a means to enhance the validity of the observational research design. In geospatial data mining, unsupervised learning algorithms serve as analytical and modeling tools to identify patterns and structures in the data in order to classify study objects with (dis-)similar features in attribute or in geographic space (Kanevski *et al.* 2009). Attribute space hereby refers to the variables that describe the ecological characteristics of a given geographic area. In complex data mining models, attribute space may be “high-dimensional”, i.e. encompass several dozens of variables or more. Geographic space by contrast is usually described by means of an area’s x-, y-, and z- (in 3D-space) coordinates.

During unsupervised learning, self-organizing maps (Kohonen 1990, 2001) serve as modeling tools to identify the spatio-temporal patterns and structures in the neighborhood ecological data. The self-organizing map (SOM) or Kohonen map is arguably one of the “most widely known methods in the field of artificial neural networks.” The SOM owes this popularity to the simplicity of its algorithm, which – once understood – makes intuitive sense and performs well in the presence of both noisy data and outliers (Lee & Verleysen 2007, 135, 141).

The SOM algorithm simultaneously performs both dimensionality reduction and vector quantization. Vector quantization means the algorithm replaces a large number of input vectors or original data points with a smaller number of representative prototype vectors by “attaching” the data points to the nearest prototype vector. Dimensionality reduction means the algorithm projects the high-dimensional input space into a low dimensional (usually 2–D) output space (Kanevski *et al.* 2009).

As a dimensionality reduction algorithm, the SOM can be thought of as a non-linear but discrete Principal Components Analysis (PCA). In PCA, a hyperplane is fitted into the data cloud so as to minimize the distance between the hyperplane and the original data points, in order to replace a larger number of correlated variables with a smaller number of uncorrelated principal components. In the SOM algorithm, the hyperplane is replaced by a network or “lattice” of artificial neurons that is introduced into the input space. The segments of the SOM lattice are elastic and articulated around the neurons, which thus fit easily over curve-linear or unevenly distributed data. The process of fitting the SOM lattice is like covering the data cloud with an “elastic fishing net” (Lee & Verleysen 2007, 136).

After the training of the SOM lattice is completed, the prototype vectors are aggregated using a hierarchical agglomerative clustering algorithm (Skupin & Agarwal 2008; Vesanto & Alhoniemi 2000). The resulting dendrogram is then cut at the appropriate level so as to produce the number of neighborhood clusters that best satisfies the twin optimization problem staked out in the introductory section of this chapter. Finally, the original data points of the urban neighborhoods are assigned the label of the neighborhood cluster of the nearest prototype vector of the SOM lattice, to which they have been attached during training (Fig. 6.1).

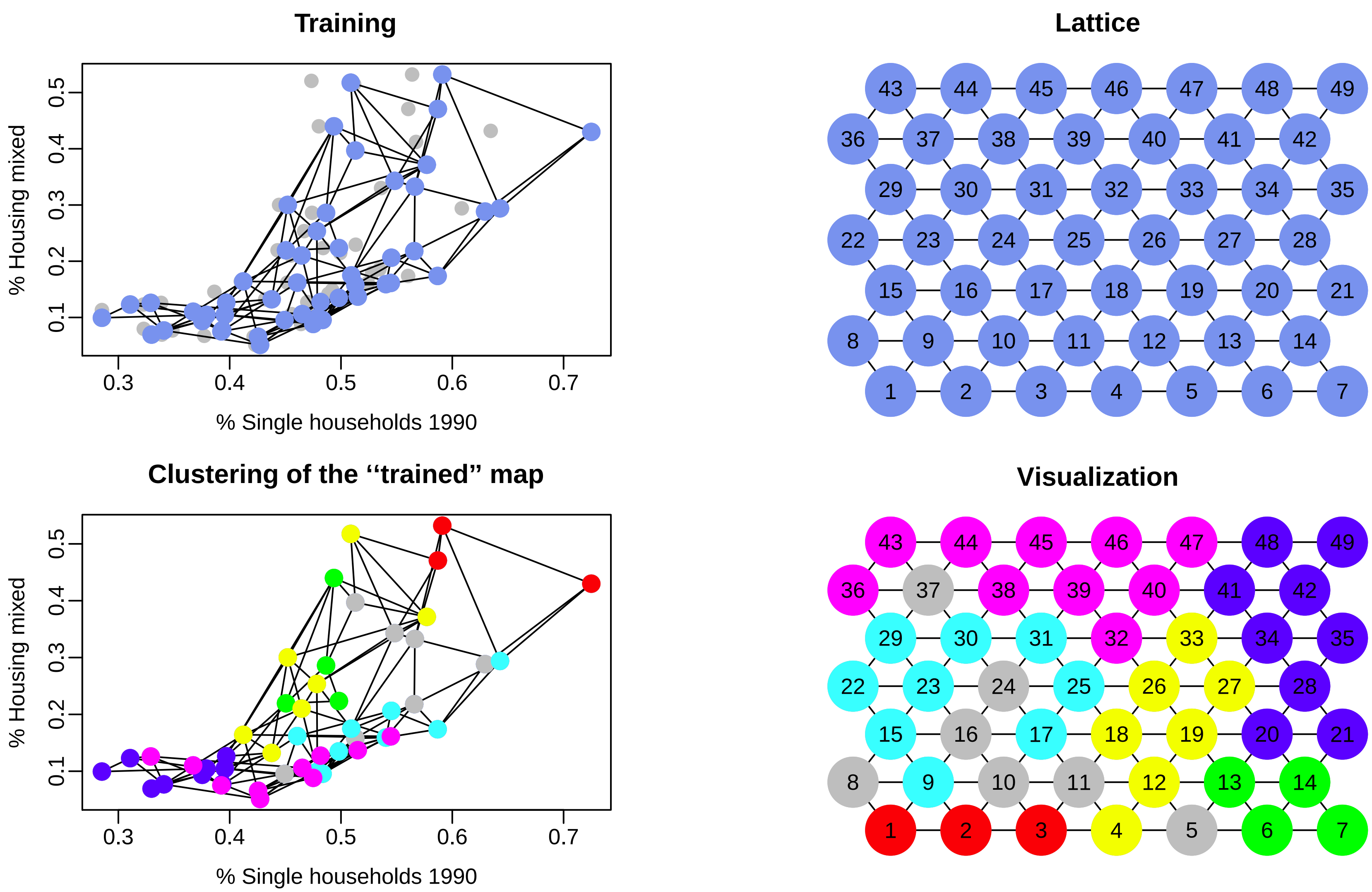


Figure 6.1 The SOM algorithm visually explained. During training, the network of prototype vectors (blue) is iteratively fitted over the original data points (top left). The structure of the SOM lattice in projected 2-D output space (top right). Following training, the prototype vectors are clustered and labeled according to their topological position inside the lattice (bottom left). Visualization of the partitioning of the lattice in output space (bottom right). NOTE: Training and clustering occur in high-dimensional input space; for simplicity, only 2 variables are plotted here.

Supervised learning – Assessing the neighborhood typology

Supervised data mining algorithms seek structures in the data while taking prior information on outcomes or classifications into account (Kanevski *et al.* 2009). During the supervised learning phase, the neighborhood classification system just developed during the unsupervised learning is imposed as a priori knowledge on the training data. During this phase, the random forests algorithm (Breiman 2001) serves to develop a classifier that predicts each neighborhood's class, on the basis of the same variables describing its ecological characteristics used during unsupervised learning.

In the current context, supervised learning can be thought of as analogous to non-linear multivariate logistic regression, in which the neighborhood classification label is the dependent variable. As a multiple regression equation is used to plug in the actual variable values in order to predict the value or class of the dependent variable, a random forests classifier develops a decision rule that determines into which class a neighborhood falls, given its values on the explanatory variables. The objective of the supervised learning is, on the one hand, to assess the quality of the neighborhood typology and, on the other, to select the explanatory variables that have a high predictive value and to eliminate noisy variables in order to build a good parsimonious model that makes a minimum of classification errors.

As a byproduct of the training of the classifier, the random forests algorithm produces a proximity measure that indicates how distinct the different neighborhood clusters are from each other and thus allows an assessment of the quality of the neighborhood typology. The proximity measure is an indication of the probability that the random forests classifier will assign any two neighborhoods to the same cluster category based on their explanatory variables and thus how close they are to each other in original input space. As a rule, the bigger this distance measure is for neighborhoods that belong to different classes, the more distinct the neighborhood clusters are and the fewer classification errors the classifier will make. By contrast, if the neighborhood clusters overlap and the distance measure between neighborhoods of separate classes is small, the classification error rate will increase.

The second analytical tool the random forests algorithm produces is the variable importance measure. The variable importance measure indicates the predictive value of each explanatory variable and can be thought of as analogous to the p-value of a coefficient in multivariate regression. The principle of calculating the variable importance measure is rather straightforward: first, the random forests classifier predicts a class for each neighborhood in the training data set, compares the predicted values to the neighborhoods' true class, and calculates an overall classification error rate. In a second step, the algorithm randomly permutes the actual values of exactly one explanatory variable, again predicts all the neighborhoods' class, and checks by how much the overall classification error rate increases as a result of the values of that one variable being "noised up" (Breiman 2001). The importance measure for a given variable is the percentage increase in the classification error rate, if its values are changed randomly. The bigger this value, the more important that particular explanatory variable is in predicting the neighborhoods' class.

The variable importance measure can thus be used for feature selection if variables that do not reduce prediction error are removed from the training data set. This procedure is not unlike a stepwise regression procedure, in which non-significant variables are recursively removed from the regression model, until only the pertinent predictors are left.

6.1.2 Feature Selection of the Neighborhood Ecological Data

Following this brief theoretical overview, the methodological approach of the current study to develop a typology of Swiss urban neighborhoods can be summed up as follows: during the unsupervised learning phase, self-organizing maps serve to detect the spatio-temporal patterns in the neighborhood ecological data. Following the training of the SOM lattice, an agglomerative hierarchical clustering algorithm is employed to merge the prototype vectors. The resulting clustering dendrogram is then cut at the level that leads to the number of neighborhood clusters that best satisfies the twin optimization problem stated at the outset of this chapter, i.e. the typology should minimize between-cluster variance in the variables describing the neighborhood context while simultaneously accounting for a significant share of the variance in the measures of community policing impact. Once this optimum number of neighborhood clusters has been identified, the SOM lattice is partitioned into as many segments. Finally, the SOM prototype vectors are assigned to a neighborhood cluster depending on their topological position inside the lattice, as are the original data points or urban neighborhoods that take on the label of the SOM prototype vector to which they are attached.

During the supervised learning phase, this neighborhood classification system is imposed on the training data and the random forests algorithm serves to develop a classifier that predicts all neighborhoods' class based on their ecological characteristics. The random forests classifier serves to assess the quality of the neighborhood typology, i.e. how distinct the neighborhood clusters really are, and to select the explanatory variables that have a high predictive value.

The original training data set contained 89 variables to characterize the neighborhood ecology that appeared relevant according to the literature and were available for the current study. These 89 variables can be regrouped into five distinct categories: official crime rates, population demography, socio-economic status, heterogeneity and residential stability, and the built environment.

The procedure to develop the neighborhood typology consisted of two separate steps: first, the SOM-random forests algorithms described in the preceding section were run separately on four of the five categories of ecological variables in order to select the key features. The selection criterion was very simple: retain as few variables as necessary without unduly increasing the classification error rate of the model. In other words, noisy variables that do not increase the predictive accuracy of the random forests classifier were eliminated.

In a second step, the SOM-random forests procedure was run on the 24 variables that proved the most important in the preliminary runs by variable category. The rationale of this two-step approach was, on the one hand, to reduce the number of variables being analyzed in a single model in order to circumvent the "curse of dimensionality" problem (Lee & Verleysen 2007, 6). On the other hand, the idea was to retain some indicators of all five data categories to be included in the final model in order to determine which categories of explanatory variables are more important, i.e. have higher predictive value, with regard to the indicators of community policing impact.

The following section briefly describes the preliminary clustering procedures by variable category, before Chapter 6.1.3 describes the procedure and outcomes of the final model of the key 24 variables in much greater detail.

Crime rates

For the City of Basel, there were no neighborhood-level crime data available for analysis. Therefore, the clustering algorithm to develop the neighborhood typology could be run only on the neighborhoods of the four cities that provided such data: Bern, Geneva, Lausanne, and Zurich. For these four cities, local police statistics have been harmonized and standardized neighborhood-level crime rates calculated for eight different types of criminal infractions: homicides, assaults, burglaries, motor vehicle thefts, robberies, vandalism, extortion, and threats.[38]

The exploratory analysis presented in the preceding chapter has unearthed the fundamental stability of the spatio-temporal patterns of the standardized neighborhood-level crime rates. Each of the four cities has one to three neighborhoods that suffer comparatively higher crime rates,

[38] Initially, the definition according to the local police statistics of *ten* types of criminal offenses have been harmonized across the four cities (cf. Table 3.3). However, sexual assaults and narcotic crime were dropped from further analysis. Narcotics crime was dropped because no data were available for the City of Lausanne, which would have further cut the number of cities to be included in the clustering procedure down to three. Sexual assaults were dropped because the data for the City of Bern displayed an erratic pattern that appeared implausible but could not be investigated further. In order to err on the side of caution, sexual assaults were dropped as well.

whereas the remainder of the city is made up of neighborhoods that suffer much lower victimization rates per resident. What is more, these patterns appear rather stable over time.

The second key observation on the crime data is that the standardized crime rates correlate very highly. For several of the eight types of criminal offenses – notably assaults, burglaries, motor vehicle thefts, robberies, and vandalism – the correlation coefficient across all neighborhoods of the four cites is 0.9 or higher. Because of this high collinearity of the crime data, it was decided to run a principal components analysis and replace the standardized neighborhood-level crime rates for the eight criminal infractions with their component scores on the emerging principal components. Only two principal components with an eigenvalue greater than one emerged from this analysis, which accounted for 75.0 percent of the variance in the data (Fig. 3.1). Therefore, the component scores of each neighborhood on the first two principal components were retained for the complete neighborhood clustering analysis, labeled as "Crime PC1" and "Crime PC2".

Population composition

The second set of ecological variables describes the demographic composition of the neighborhood population. The list of variables includes the percentage of the total population by age group in 18 categories of 5-year intervals, from "0 to 4 years old" to "85 years old and above" as well as the composition of households, measured as the number of both single and family households as a percentage of all households in the area. All demographic data were available from the 1990 and 2000 census and included in the analysis. The resulting neighborhood typology should thus not just account for the population composition but for shifts over the intervening decade as well.

After the initial clustering procedure of the demographic data, six out of the 40 variables were retained for the final analysis: the percentage of children aged "5 to 9 years old" and "10 to 14 years old" from the 1990 census as well as the percentage of "Single households" and "Families" from both the 1990 and 2000 census.

Socio-economic status

The third set of ecological variables captures the socio-economic status (SES) of the resident population. This set of variables measures the share of the neighborhood population by level of the highest educational achievement in seven categories ranging from "Mandatory schooling" to "University" (graduates). These figures are available from the 1990 and 2000 census counts. The list of variables also includes the percentages of the professional categories of the active working population residing in a given area, subdivided into eight different groups of varying social prestige and remuneration, ranging from "Unskilled workers" to "Executives". These data were not included in the 1990 census and are available only for the year 2000.

After the initial run of the clustering algorithm on the SES data, six out of the 22 features were retained as most important for the final analysis: the percentage of residents who had completed "Mandatory schooling" or an "Apprenticeship" from both the 1990 and 2000 census, as well as the percentage of university graduates and of residents employed in a "Middle management" position from the 2000 census.

Heterogeneity and residential stability

The fourth set of ecological variables measures the degree of heterogeneity and residential stability by means of five variables: the percentage of Swiss and foreign nationals among neighborhood residents from the 1990 and 2000 census. The fifth variable was the percentage of residents who in the year 2000 still lived at the same address as five years earlier. Three of these five variables were retained for the final analysis after an initial clustering: the percentage of "Foreigners" among the resident population in 1990 and 2000 as well as the variable capturing the percentage of long-term residents in an area.

Built environment

The fifth set of ecological variables is included to characterize the built environment in a given area. A first category of variables captures the height of buildings, measuring the percentage of buildings of the total housing stock in nine subgroups from smallest to highest, i.e. "1 story", "2 stories" up to "15 and more stories". A second category of eight variables indicates the period of construction of the housing units, measuring the percentage of the total housing stock built during eight distinctive time periods ranging form "before 1900", to "1900-1920" until the most recent period "1986-1990". With these two types of variables alone a fairly good description of the housing stock is possible, allowing for the distinction of more or less densely built areas as well between neighborhoods with a housing stock dating from more than a century ago versus areas that have sprung up more recently.

Three more variables were included in the clustering analysis in order to aptly describe the housing stock in an area. These three variables indicate the functional use of a neighborhood's buildings, distinguishing between buildings that are exclusively residential, mixed housing complexes that include both apartments and offices or shops, and buildings that serve exclusively non-residential purposes such as office complexes or commercial centers. All three variables measure the share of each functional category as a percentage of the total housing stock and, as the other variables describing the built environment, were gathered from the 1990 housing census count only.

Seven variables proved important during an initial clustering procedure of the housing stock data and were retained for inclusion in the final analysis: all three variables capturing the functional use of buildings ("Residential", "Housing mixed", and "Non-housing"), three variables indicating building height ("2 stories", "6 stories", and "7-9 stories") as well as one variable indicating the construction period ("before 1900").

For easy reference, the names of the 24 key variables included in the training data set for the final clustering procedure are listed below:

```
> colnames(training)
 [1] "05-09 years 90"       "10-14 years 90"       "Single households 90"
 [4] "Families 90"          "Single households 00" "Families 00"
 [7] "Apprenticeship 90"    "Mandatory school 90"  "Apprenticeship 00"
[10] "Mandatory school 00"  "University 00"        "Middle management"
[13] "Foreigners 90"        "Foreigners 00"        "Same Address 5-yrs"
[16] "Residential"          "Housing mixed"        "Non-housing"
[19] "2 stories"            "6 stories"            "7-9 stories"
[22] "before 1900"          "Crime PC1"            "Crime PC2"
```

6.1.3 Final Clustering Procedure of the Key Variables

The initial clustering procedures of the neighborhood ecological variables served to select the most important features for each category of data. It was thus possible to pare down the original training data set of 89 variables (87 ecological variables + 2 principal components of the crime data) to a much smaller data set of 24 (22+2) key variables, without unduly increasing the classification error rate of the random forests classifier for each of the four variable categories. Once these procedures had been completed, the 24 most important variables were included in the final model. On this data set the same clustering procedure was run, relying first on the SOM algorithm and hierarchical agglomerative clustering to classify the unlabeled data and then on the random forests algorithm to assess the quality of the clustering and to identify the important features. The diagnostic plots that resulted from this final analysis were presented in Figure 3.2.

Identifying the optimum number of neighborhood clusters

In order to identify the optimum number of neighborhood clusters, the dendrogram that resulted from the hierarchical agglomerative clustering of the SOM prototype vectors is cut at different levels, dividing the pooled sample of all the neighborhoods of all four urban areas into different numbers of neighborhood clusters. The test statistics are then calculated for each number of possible neighborhood clusters ($k = 2, 3, \ldots, K$) in order to determine the optimum partitioning. As has been stated previously, the clustering algorithm has to reconcile a double objective: on the one hand, it should regroup the neighborhoods that are most similar with respect to the neighborhood ecological variables used to describe these areas. On the other hand, the resulting neighborhood typology should account for a maximum of the variance of the indicators of community policing impact, namely fear of crime.

Each optimization problem required a separate analysis. In order to choose the best partitioning of urban districts with regard to the contextual variables, some kind of clustering validity index must be used. In accord with other studies (e.g. Vesanto & Alhoniemi 2000), the current study used the Davies-Bouldin Index (DBI; Davies & Bouldin 1979) to assess the quality of clustering as a function of the number of clusters used to partition the urban districts across the four cities.

In order to chose the ideal number of neighborhood clusters with regard to the survey data measuring fear of crime, disorder, and popular attitudes towards the police, the Swiss Crime Survey respondents from the four cities were grouped by the same neighborhood clusters. The actual statistical test was a χ^2-independence test whether for the pooled sample response patterns on the threat of crime survey item varied significantly for different numbers of neighborhood clusters.

Figure 3.2 e plots in a single chart the values of both the DBI and the p-values of the χ^2-independence test as a function of the number of clusters into which the neighborhood areas across the four cities are divided. As this chart reveals, there is no unique optimal solution to the neighborhood classification problem. With regard to the neighborhood ecological variables, it would be optimal to divide the neighborhoods into only two clusters, since for two clusters the DBI is at its global minimum. This minimalist neighborhood typology of only two clusters basically pits the downtown areas of the four cities against the surrounding areas. However, with only two neighborhood clusters, the χ^2-independence test of the survey data fails a significance test by a wide margin. A neighborhood classification system that also accounts for a significant share of the variance in the fear of crime item requires a finer partitioning into at least three neighborhood clusters. It turns out that six neighborhood clusters best reconcile the twin objectives: with six clusters, the p-value of the χ^2-independence test is highly significant, while the value of the DBI is a local minimum.

The classification error rate of the final model

During supervised learning, the six neighborhood clusters just determined as optimal by the SOM clustering algorithm are imposed on the training data for the building of the random forests classifier. The random forests classifier develops a decision rule that will predict each neighborhood's class based on its ecological characteristics. Inevitably, the resulting random forests classifier will misclassify some neighborhoods. Figure 6.2 indicates the percentage of neighborhoods that are misclassified – the so-called out-of-bag error rate – if a given variable plus all the other variables to its right are included in the building of the random forests classifier. In Figure 6.2, these variables were ordered by their level of importance from left to right, from least to most important.

The classification error rate of the random forests classifier if all 24 variables retained for the clustering procedure are included is just under 10 percent. This means that a random forests classifier trained on the complete data set that was used for the final clustering procedure predicts the neighborhood type previously developed by means of the SOM clustering algorithm correctly for some 45 of the 51 neighborhoods. Some five neighborhoods are misclassified, i.e. the random forests classifier, based on an area's ecological characteristics, predicts a neighborhood type other than the area's true classification.

This misclassification rate changes only slightly, as one moves from left to right and drops the less important variables one-by-one from the training data set used to build the random forests classifier. The misclassification rate hovers between the minimum of just under 10 percent and under 20 percent as more and more variables are dropped. The misclassification rate finally exceeds 20 percent, if only the three most important variables are retained to train the random forests classifier. However, even a classifier built solely on the single most important variable – the percentage of families among a neighborhood's total households in 2000 – still classifies nearly 40 percent of all neighborhoods correctly, which, considering that there are six underlying neighborhood classes, is still far better than chance. In other words, a random forests classifiers built on the four most important ecological variables classifies some 80 percent of all neighborhoods in the sample correctly; a random forests classifier that is trained on the 10-15 most important variables is as good as it gets.

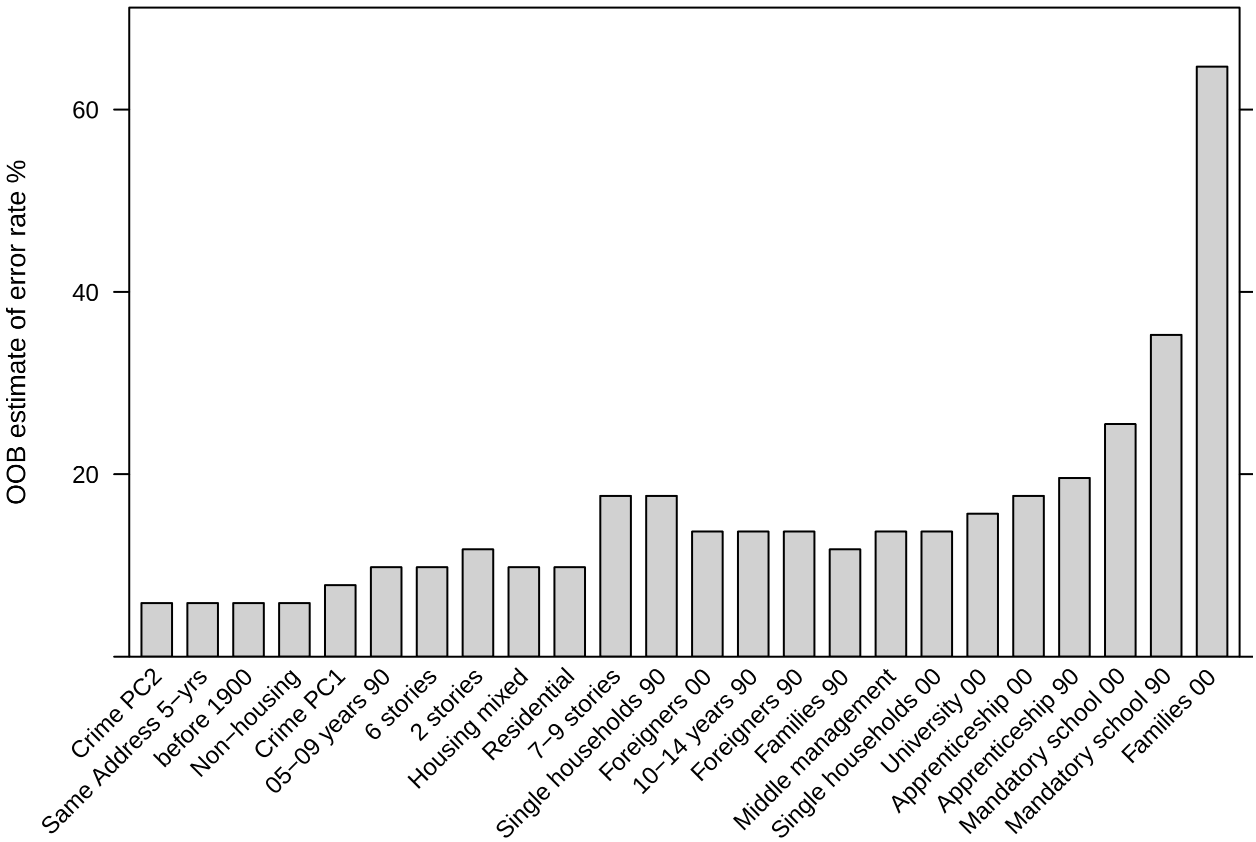


Figure 6.2 Random forests estimate of the out-of-bag error rate. Percentage of neighborhoods that are misclassified – the so-called out-of-bag error rate – if a given variable plus all the other variables to its right are included in the training of the random forests classifier. These variables were ordered by their level of importance from left to right, from least to most important.

More interesting than these crude numbers of misclassification rates from a policy perspective is which categories of ecological variables are most important in determining the neighborhoods' class. It turns out that the more important variables measure an area's demographic and socio-economic characteristics and population heterogeneity. By contrast, measures capturing neighborhood-level crime rates, the built environment, and residential stability have little impact on the accuracy of the classification. In other words, these variables can be dropped from the training data set used to build the random forests classifier without incurring much of an increase in the classification error rate.

Mapping of the neighborhood typology

Once the optimum number of neighborhood clusters has been identified, the resulting neighborhood typology can be visualized using different tools. Figure 3.2 d shows the dendrogram of the hierarchical agglomerative clustering procedure of the SOM prototype vectors cut at the level that corresponds to the optimum number of neighborhood clusters. Since the leaves of this dendrogram represent the individual prototype vectors, the results of the clustering procedure may also be visualized using the SOM lattice. Figure 3.3 plots the SOM lattice with the different segments colored according to the six neighborhood types. In addition, each neuron in the SOM lattice displays the geo-reference of the original data points, which it represents. Since these original data points correspond to ZIP code or administrative districts, it is possible to draw maps of the four cities and to visualize the results of the neighborhood clustering procedure in the original geographic space using GIS software (Fig. 6.6).

A look at these maps reveals some striking parallels between the cities. First of all, neighborhood types 1 and 2 are downtown areas located in the heart of the cities. By contrast, neighborhood types 5 and 6 are suburbs, located on the outer rim of the cities. Neighborhood types 3 and 4 lie geographically somewhere in between these two. The second striking feature of these maps is that Swiss German and Swiss French cities appear to be different. In Geneva and Lausanne, the colors yellow and green predominate, whereas Zurich and Bern are trenched more in blue and pink.

These patterns are noteworthy from at least two perspectives. First, the neighborhood typology reveals a pattern that changes from the city centers towards the outskirts, for the French- and German-speaking urban areas alike. This fact is all the more remarkable since no geographic indicator was included in the training data of the clustering algorithm. This means that the variables used to describe a neighborhood's ecology differ significantly on a center-periphery axis, so that neighborhoods fall into distinct classes according to their geographic location.

Second, the resulting pattern also highlights the *topology persevering* quality of the SOM algorithm. Topology preservation means that data points or neighborhoods that are close to each other in the high-dimensional input space remain close to each other in the projected output space. This is displayed beautifully by the plot of the trained SOM lattice in the projected 2-D output space in the top right panel of Figure 6.6. This chart clearly shows the two suburban types of neighborhoods 5 and 6 at the opposite end from the downtown neighborhood type 1, with the neighborhood types 2, 3, and 4 lying somewhere in between. In other words, without imposing *any* geographic constraints on the clustering algorithm, the spatial logic of urban development from city center to the periphery is accurately reflected by the trained SOM lattice. This result provides further evidence that self-organizing maps are an adequate clustering algorithm for the problem at hand.

6.2 Assessing the Neighborhood Classification System

6.2.1 Proximity Measure and Clustering Validity

Once the clustering algorithm has run its course, there are several ways to test the quality and assess the appropriateness of the classification system for evaluation purposes.

The first such diagnostic tool is the random forests proximity measure. As stated previously, the proximity measure is an indication of the probability that the random forests classifier will assign any two neighborhoods to the same cluster category based on their explanatory variables and thus how close they are to each other in original input space. These proximity measures can be used to calculate a distance matrix, which can then be displayed by means of a multi-dimensional scaling plot.

Figure 6.3 displays the multi-dimensional scaling plot for the current neighborhood data. In this plot, the colors correspond to the neighborhood types of Figure 6.6, whereas the four icons represent the four cities (according to the key). In this plot the neighborhoods of a given type cluster neatly according to color. Most distinct are the neighborhood types 2, 4, and 5 (yellow, light blue, and blue) that cluster neatly in the two top corners and the bottom of the plot, respectively. Less distinct are the neighborhood types 1, 3, and 6 that cluster in the middle. In particular, neighborhood type 1 (red) and type 3 (green) overlap considerably.

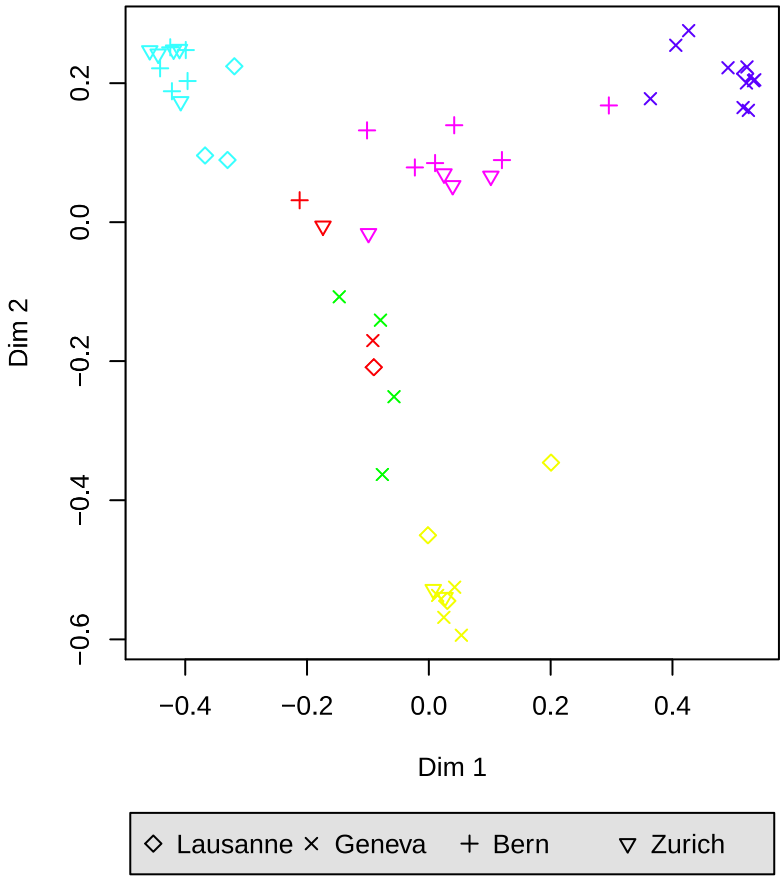


Figure 6.3 Multi-dimensional scaling plot of the random forests proximity measure. Colors correspond to the six neighborhood types while the four different icons indicate the city according to the key.

Figure 6.3 is testimony of the compromise taken to reconcile the twin optimization objectives in determining the optimum number of neighborhood clusters. If the neighborhoods had to be separated on the ecological data alone, two clusters would have been sufficient. However, in order to account for a higher share of the variance in the indicators of community policing impact, the optimum number of neighborhood clusters was found to be six. Since the random forests proximity measure is based on neighborhood ecological data alone, some types of neighborhood are less distinct than others, increasing the probability of classification errors.

6.2.2 Feature Importance and Classification Error Estimates

A second set of diagnostic plots to assess the neighborhood typology are the random forests variable importance measures by neighborhood type. The variable importance measure, it is recalled, indicates by how much the error rate of the random forests classifier increases, if the values of one variable are randomly changed. If the random forests algorithm is replicated 50 times and the values are retained for each replication, the resulting percentage increase in the classification error rate over the 50 replications can be represented as a boxplot for each of the 24 variables for each of

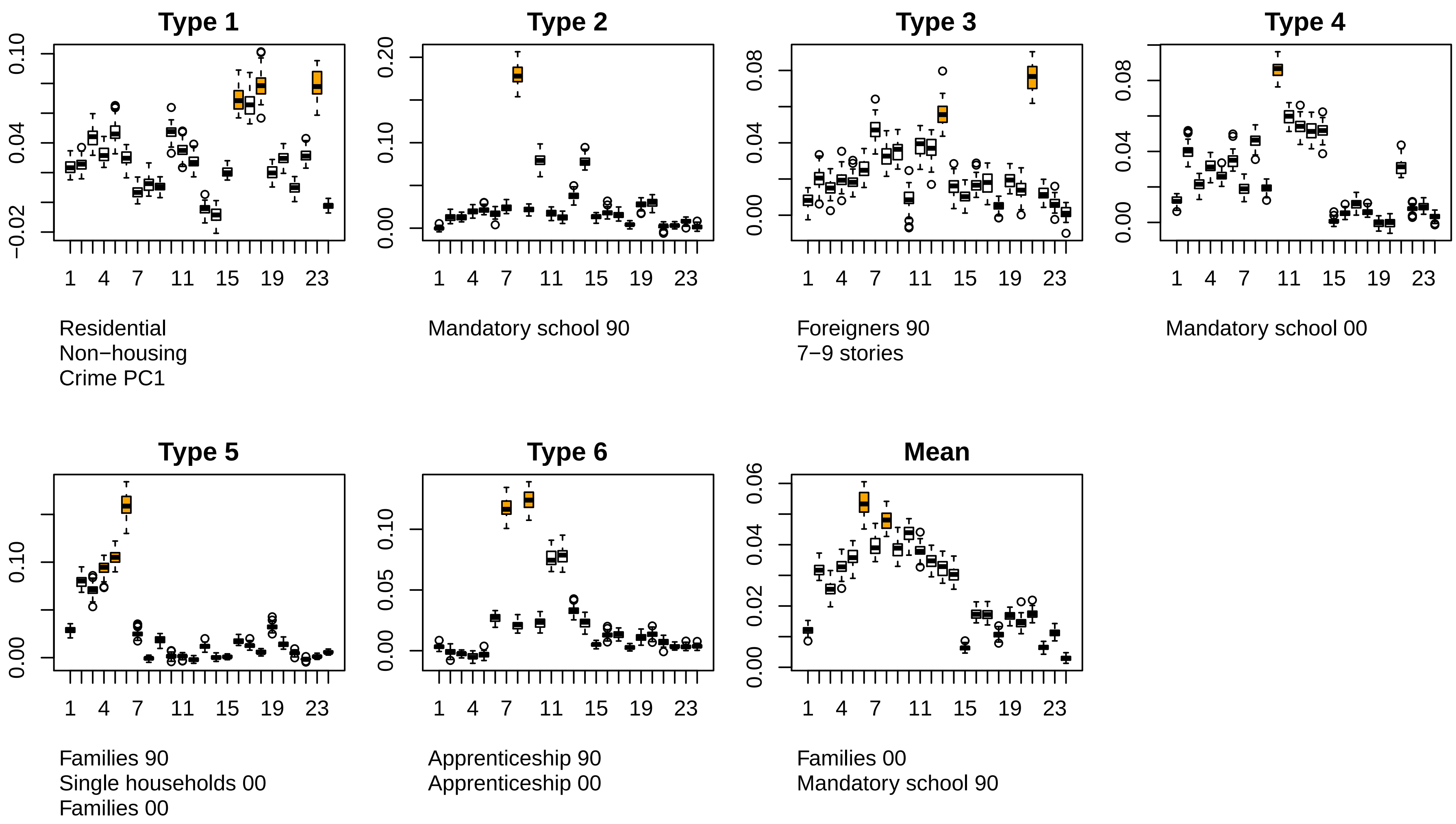


Figure 6.4 Random forests variable importance plot by neighborhood type. The boxplots display the percentage increase over the 50 replications in the classification error rate for each of the 24 variables retained in the final clustering analysis if its true values are randomly changed. The "noised up" variables for which the error rate increases by more than 1.5 times the standard deviation of all 24 variables are highlighted in yellow and their names indicated below each plot. These are the key variables in assigning a given neighborhood to that particular type.

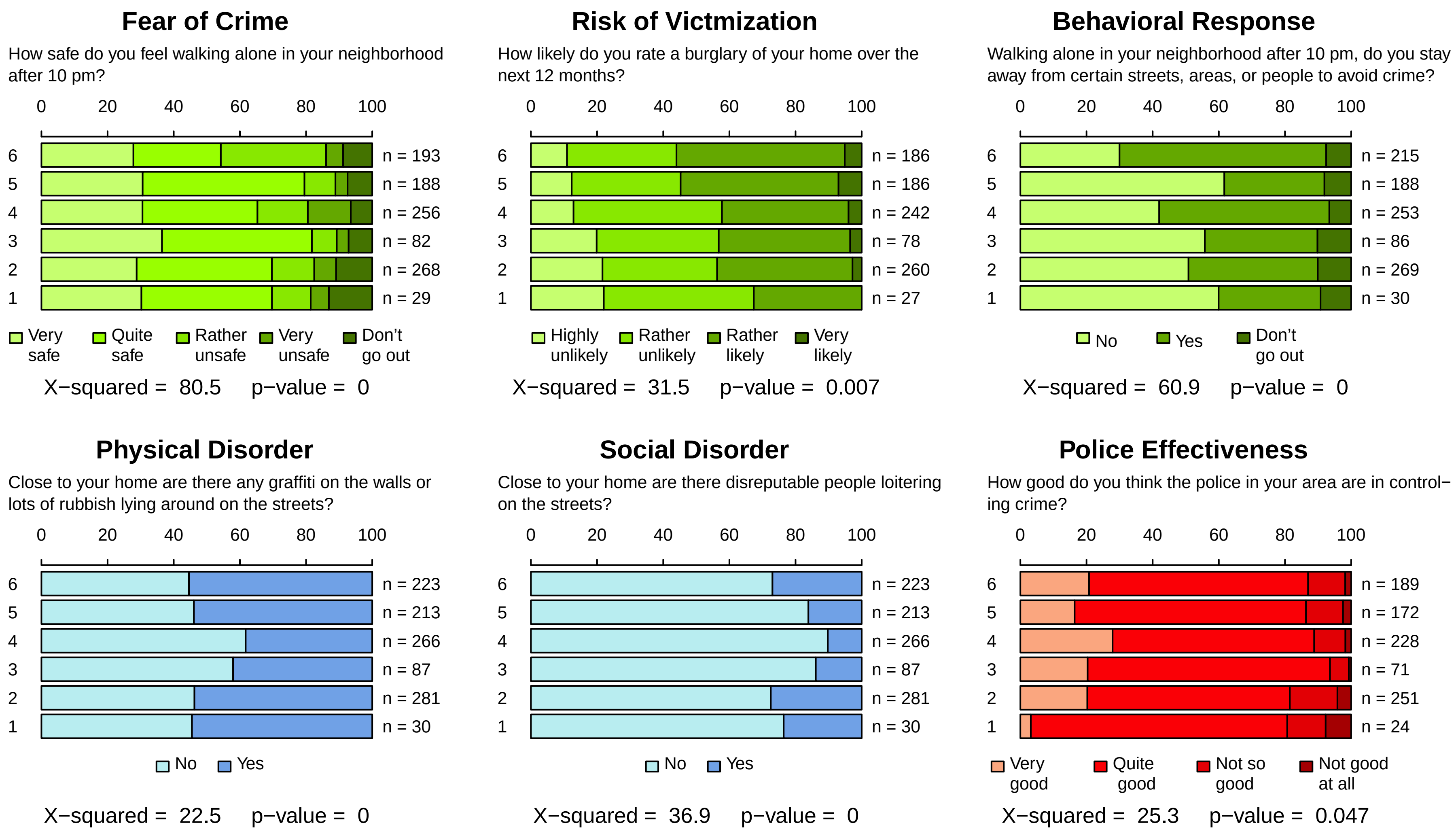


Figure 6.5 Survey response patterns by neighborhood type. Percentage of respondents by answer category for the six survey items used to assess the impact of community policing on neighborhood residents. Survey respondents were grouped together across cities by neighborhood type, excluding respondents from Basel. The χ^2-independence test statistics were calculated using Monte-Carlo simulations. The total survey sample was weighted, stratified at the neighborhood level, to correct for sampling bias in the age and gender distribution.

the 6 neighborhood types (Fig. 6.4). The 7^{th} plot ("Mean") displays the boxplots of the increase in the classification error estimate over all neighborhoods in the training sample. In each of these boxplots, the most important variables were highlighted in orange and their names are indicated below each plot.

6.2.3 Analysis of the Survey Measures by Neighborhood Type

Survey variables by neighborhood typology

The final and arguably most important test of the neighborhood typology is whether it accounts for a significant share of the survey variables that serve as indicators of community policing impact. The clustering procedure was set up so as to produce a typology that accounts for fear of crime, measured as the threat of crime. It remains to be seen to what extent the typology thus developed also accounts for a share of the variance of the other outcome indicators.

Figure 6.5 displays the response patterns of the 2000 Swiss Crime Survey on the six survey items that are being used throughout this evaluation to measure popular sentiment: three variables measuring fear of crime, the indicators capturing physical and social disorder, and one item measuring popular satisfaction with the police. For each of these survey items, the SCS sample was divided into six groups by neighborhood type in accordance with respondents' area of residence, and a χ^2-independence test was run to determine whether response patterns differ significantly by neighborhood type.

For each of the six survey items, the plot indicates the questionnaire item plus corresponding answer categories. The single digits on the left of each bar plot indicate the neighborhood type, whereas the numbers on the right indicate how many survey respondents that fall into each neighborhood class. For this analysis, survey respondents were weighted as a stratified sample at the ZIP code level to correct for sampling bias in the gender and age distribution to make the sample representative of the resident population. (The numbers "n" indicating the sample size per neighborhood cluster are rounded.) The test statistics of the χ^2-independence-tests are indicated below each bar plot. Test statistics were calculated using Monte-Carlo simulations, since the number of survey respondents for some answer categories and especially neighborhood type 6 are low.

The contingency tables underlying the charts and related χ^2-independence-tests reveal some very interesting patterns that merit further inspection. As it happens, the neighborhood types are arranged in spatial order on a center-periphery continuum. At the bottom are neighborhood types 1, 2, and 3, which are the more centrally located neighborhoods. At the top are neighborhoods 4, 5, and 6 that are located towards the outskirts of the five urban areas.

The first of the six barplots displays the survey item measuring fear of crime measured as the feeling of safety on a nightly stroll through one's neighborhood. As can be seen from this plot, the general tendency is for the percentage of respondents who feel "very safe" or "quite safe" to increase as one moves in from the periphery to the center. The only outliers from this spatial logic are neighborhood types 3 and 5, which happen to be located exclusively or predominantly in Geneva.

The general tendency for fear of crime to increase from the urban centers to the peripheries is still more pronounced for the perceived risk of a burglary. The percentage of respondents who rate the chances of a burglary of their home over the next twelve months as "likely" or "very likely" systematically increases as one moves up the typology ladder from bottom to top, i.e. from center to periphery in the real world.

For burglary risk, the pattern of subjective risk assessment can be compared to actual victimization risk as captured by official crime statistics. Figure 5.2 showed maps of the relative burglary risk of neighborhoods for the four cities for which such data were available. These maps revealed that burglary risk is highest in the city centers and tends to *decrease* towards the suburban areas. The survey response patterns by neighborhood type, however, move in exactly opposite direction.

In other words, survey respondents collectively do a poor job at assessing victimization risk: risk tends to be underrated in the city centers and to be overrated in the suburbs.

The same observation also holds for the third indicator of fear of crime, changes in behavior to avoid crime. The percentage of respondents who take avoidance strategies steadily increases from center to periphery. The only outlier is again neighborhood type 5, which are the suburbs of both Geneva and Lausanne.

For the survey item measuring physical and social disorder the spatial logic is no longer a function that increases more or less linearly from center to periphery but rather appears to be a curve-linear function. The percentage of respondents answering in the affirmative is higher for neighborhood types 1 and 2, drops slightly for types 3 and 4 but increases again for neighborhood types 5 and 6.

Popular satisfaction with the police appeared to be more attuned to actual victimization risk. As one moves from neighborhood type 1 to 6, from center to periphery, the percentage of survey respondents who think the police do a "very good" or "quite good" job both increase.

Remaining variance between cities

The goal at the outset of this clustering procedure was to develop a neighborhood typology that clusters neighborhoods of a similar type and accounts for a significant share of the variance in the outcome variables. Before this typology is being applied to evaluate community policing, especially to evaluate the impact of particular strategies not just over time but also between cities, a final test is in order to check whether the typology does indeed account for most of the variance in outcome variables between residents of different neighborhood types or if there is some remaining variance at the city level that the current neighborhood typology cannot account for.

Table 6.1 Survey response patterns by neighborhood cluster. P-values of the χ^2-independence tests of the indicators of community policing impact by city, computed separately for each of the six neighborhood clusters. (A value of $p < 0.05$ indicates that response patterns within a given neighborhood cluster still vary significantly by city.)

	Fear of Crime	Risk of Victmization	Behavioral Response	Physical Disorder	Social Disorder	Police Effectiveness
1	0.007	0.824	0.006	0.019	0.001	0.857
2	0.000	0.000	0.494	0.395	0.195	0.164
3						
4	0.001	0.130	0.137	0.000	0.007	0.034
5	0.817	0.107	0.130	0.037	0.262	0.367
6	0.301	0.000	0.590	0.734	0.060	0.052

In order to test this proposition, a second series of χ^2 tests is run. This time the χ^2-independence statistic tests whether the answering patterns vary between individual cities for all the survey respondents that reside within *any one* type of neighborhood. Table 6.1 displays all p-values of these tests for each of the six neighborhood types. The p-value of the χ^2-independence test is often not significant, meaning that the hypothesis that the response patterns of survey respondents residing in the same type of neighborhood *do not* vary between cities cannot not be rejected (which is what we were hoping for). However, for a number of the χ^2-independence tests the p-value is significant. As a matter of fact, for none of the five neighborhood types that are to be found in more than one city are all six p-values above the critical value of 0.05.

The same tests were carried out separately for respondents from the Swiss French and Swiss German cities in order to test to what extent the variance that remains in the data stems form the city or regional level. The results are presented in Tables 6.2 and 6.3. As can be seen from these two tables, the neighborhood typology works rather better for the Swiss French cities. For Geneva and Lausanne, the neighborhood typology accounts for a sufficient amount of variance in the six survey indicators for the city level no longer to wield any significant influence. The results of the χ^2-independence tests are less encouraging for the three Swiss German cities, though, where many p-values are close to or below the standard 5 percent level of significance. This means that even by regrouping survey respondents by neighborhood type, the city of residence still impinges on survey response patterns.

Table 6.2 Survey response patterns by neighborhood cluster in Geneva and Lausanne. P-values of the χ^2-independence tests of the indicators of community policing impact by city, computed separately for the neighborhood clusters found in Geneva and Lausanne.

	Fear of Crime	**Risk of Victmization**	**Behavioral Response**	**Physical Disorder**	**Social Disorder**	**Police Effectiveness**
1	0.456	0.700	0.884	0.946	0.999	0.518
2	0.536	0.711	0.459	0.358	0.416	0.520
3						
4						
5	0.817	0.107	0.130	0.037	0.262	0.367
6						

Table 6.3 Survey response patterns by neighborhood cluster in Bern and Zurich. P-values of the χ^2-independence tests of the indicators of community policing impact by city, computed separately for the neighborhood clusters found in Bern and Zurich.

	Fear of Crime	**Risk of Victmization**	**Behavioral Response**	**Physical Disorder**	**Social Disorder**	**Police Effectiveness**
1	0.102	0.340	0.141	0.381	0.027	0.789
2						
3						
4	0.033	0.081	0.490	0.000	0.003	0.035
5						
6	0.301	0.000	0.590	0.734	0.060	0.052

6.3 Characterizing the Neighborhood Clusters

6.3.1 Original Values of the Key Ecological Variables

Figure 6.7 plots the attribute values of the 51 neighborhoods in the data sample as boxplots for each of the variables retained in the complete neighborhood clustering procedure. For each of the

24 variables in the data set, the variable values of the original neighborhoods are plotted as boxplots by neighborhood type. This chart highlights the differences in the original variable values.

Figure 6.7 can be used to describe the salient characteristics of the six neighborhood types. As Figure 6.4 indicated, not every variable is equally important in determining a neighborhood's class. It is thus to be expected that for the variables that appear most important for a given neighborhood type, the original values of the neighborhoods of this class should be relatively distinct from the corresponding values for the other neighborhood types.

By now, the elements necessary to characterize the different neighborhood types have been presented. These can now be taken together to draw a relatively accurate picture of each type of neighborhood area in terms of its geographic location, salient ecological characteristics, and the popular sentiment with regard to the community policing outcome variables. It is thus possible to assess a neighborhood area prior to any evaluation of actual policing strategies that may or may not have been deployed in these areas.

Type 1

These neighborhoods are located in the city center of all cities under study save Basel. The key ecological characteristics of these areas are an unusually high share of non-housing units and a correspondingly low share of residential units of the total housing stock. In addition, these areas exhibit the highest values on the first principal component of the crime data, meaning local residents suffer higher victimization rates per resident than the cities at large.

In terms of the popular sentiment as recorded by the Swiss Crime Survey for the year 2000, residents of these areas on average were less afraid of crime, spotted comparatively more signs of physical and social disorder, and declared themselves the least satisfied with their local police.

Type 2

Type 2 neighborhoods are to be found in the western parts of Geneva and Lausanne, in the southern and northern parts of Basel, as well as the Langstrasse district in Zurich. Type 2 neighborhoods comprise economically less well-off areas. Not surprisingly, the defining ecological characteristics of these neighborhood areas are a relatively high share both of residents who have completed only mandatory schooling and of resident foreign nationals. In terms of criminal victimizations, residents of these areas are the second most afflicted after the urban centers. Considering that these areas have a considerably higher share of residential dwellings than neighborhoods of type 1 (which may artificially inflate the standardized victimization rates for the urban centers), it is arguably not far-fetched to claim that the victimization experiences of residents of these areas are comparable to residents of type 1 neighborhoods.

In terms of popular opinion as expressed through the victimization survey, type 2 neighborhoods in many ways resemble type 1 neighborhoods, with comparatively low levels of fear of crime and higher levels of signs of both physical and social disorder. However, contrary to type 1 neighborhoods, these neighborhoods have a higher share of the population that think the police do a very good job.

Type 3

These neighborhoods are to be found exclusively in Geneva, where they form a rim around the urban center on either side of the lake. The ecological characteristics that set these neighborhoods apart are a high share of resident foreigners as well as a particularly high share of buildings with between 7 and 9 floors. This finding evidences the fact that the urban center of Geneva is more densely built than other major urban areas in Switzerland. Additionally, the variable importance plot for these neighborhoods reveals that the low share of local residents whose highest level of education was an apprenticeship is another defining characteristic of type 3 neighborhoods.

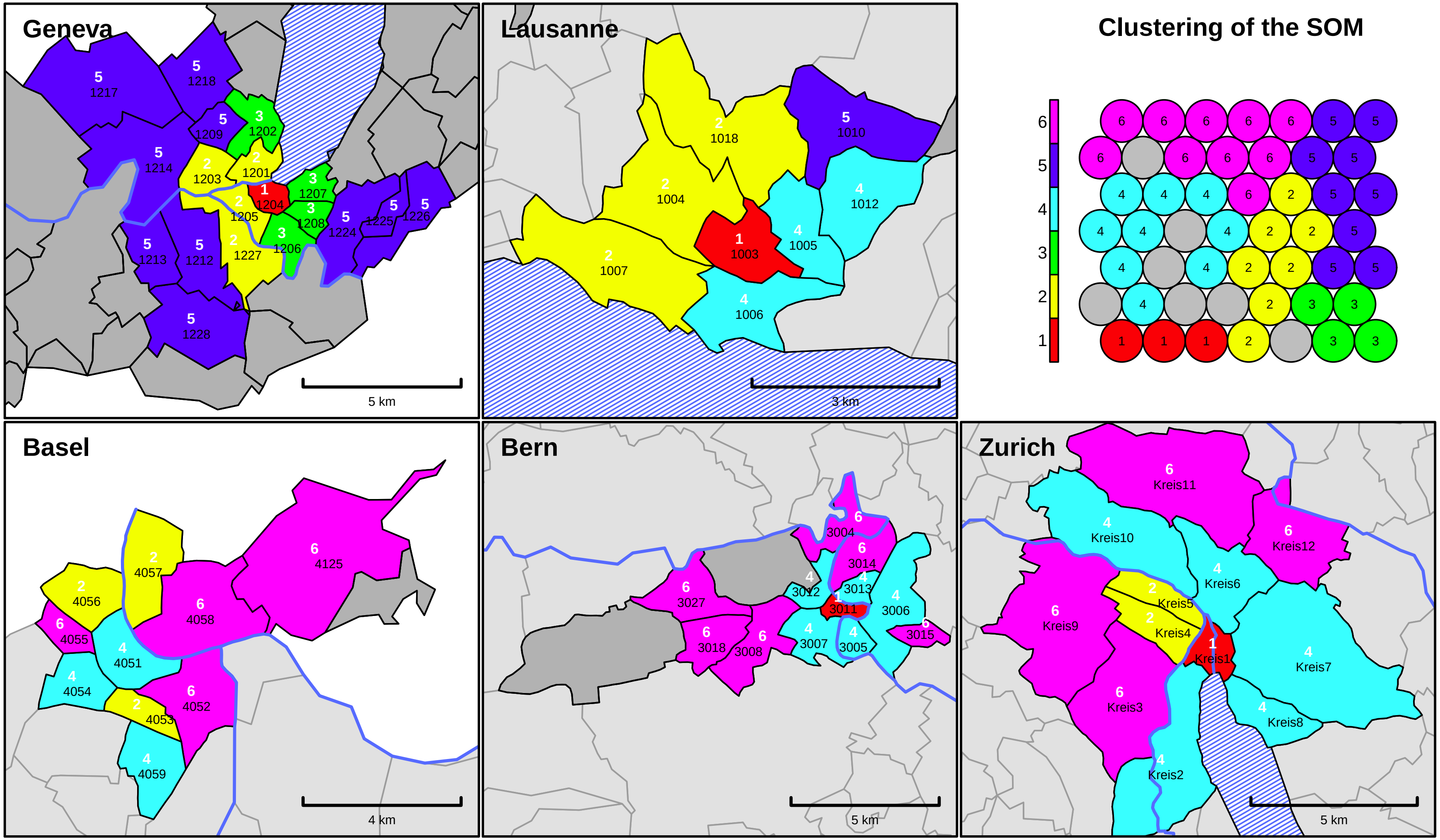


Figure 6.6 Visualization of the neighborhood typology in geographic space. Maps of the five cities with neighborhood areas colored according to type. The top right panel displays the trained SOM lattice projected into 2-D output space with the segments colored according to the best partitioning. Neighborhoods in Basel were labeled based on the random forests classifier trained on the ecological data on the other four cities since for it no neighborhood-level crime data were available.

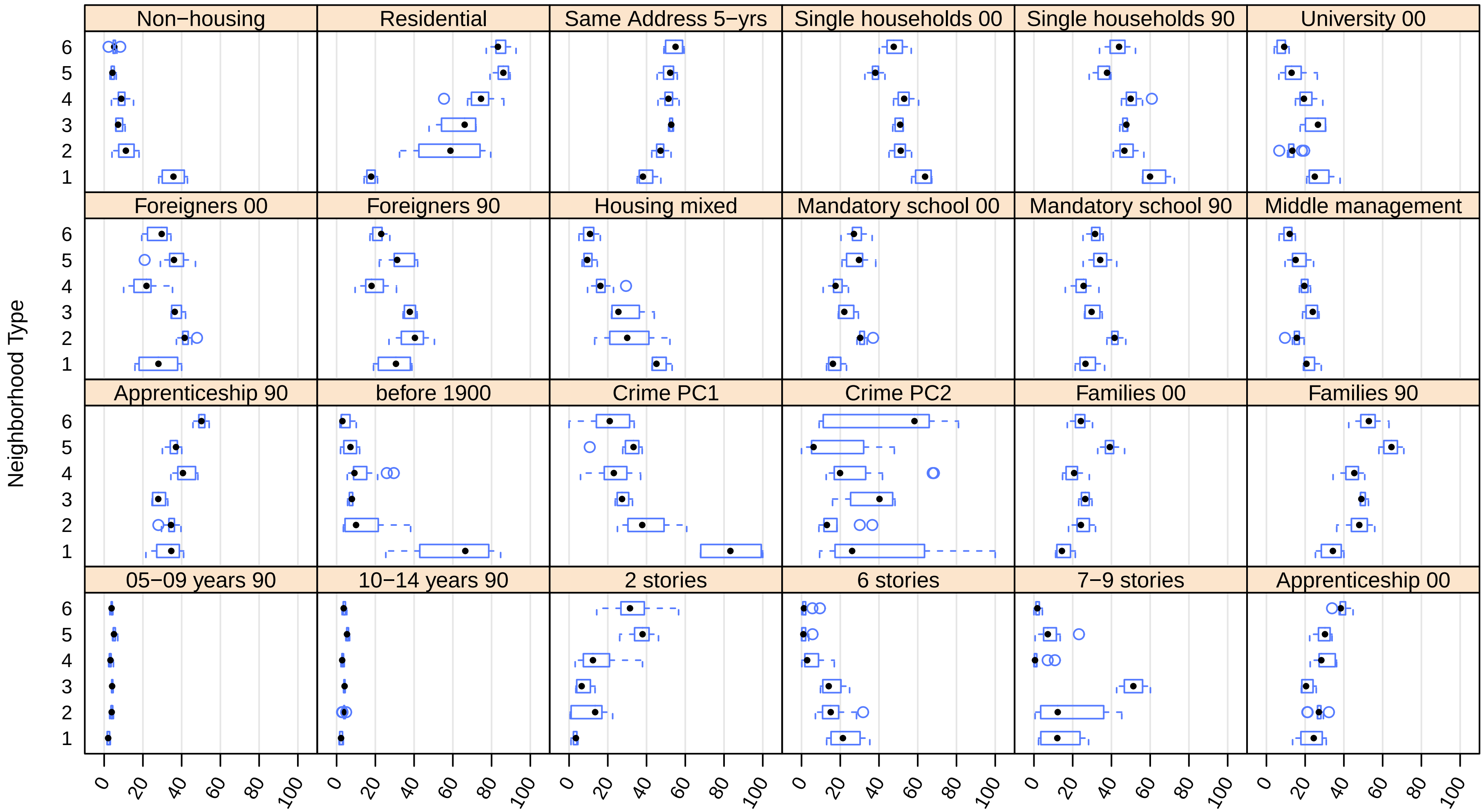


Figure 6.7 Defining characteristics of the neighborhood typology. Original values of the 24 variables retained in the final clustering analysis to describe the neighborhood ecology. All variables are percentages except for the crime principal components scores ("Crime PC1" and "Crime PC2"), whose true range was linearly transformed to a 0-100 scale.

The survey data reveal that residents of these Geneva neighborhoods score low on fear of crime measured both as the perceived threat of and behavioral responses to crime, whereas they rate as average the risk of their home being burglarized. Signs of physical and social disorder appear to be less of a concern to the residents of these areas. The percentage of respondents who are very or quite satisfied with their police is the highest among all six neighborhood types.

Type 4

Type 4 neighborhoods make up the eastern part of Lausanne, the center and two suburban areas in Basel, a rim around the center of Bern, as well as the lakefronts, Hönggerberg, and the Zürichberg in Zurich. These areas can be described as middle-class or upper middle class and, not surprisingly, the random forests variable importance plot reveals as the key defining characteristic the comparatively low share of residents who have completed only mandatory schooling. Four more ecological variables stand out for these areas that confirm the picture of relatively well-off areas: the high share of the resident population with a university degree, the comparatively high share of residents in a middle management position as well as the lower share of resident foreign nationals according to both the 1990 and 2000 census.

In terms of the survey responses, the pattern for neighborhood type 4 is not that different from type 3, except for fear measured as the threat of crime or the behavioral response, which is higher, and satisfaction with the police, which is slightly lower. In terms of disorder, residents of the two neighborhood types see virtually eye-to-eye.

Type 5

Type 5 neighborhoods are found in the north of Lausanne and make up the belt of suburban municipalities surrounding the City of Geneva. The key defining characteristics are the high share of families and the low share of single households according to both census counts. These areas are primarily residential areas where housing units predominate and buildings with mixed or non-residential functions are harder to find.

In terms of the survey responses, it is interesting to note that these areas have higher levels of fear of crime than the neighborhood types discussed so far, with one notable exception: residents here are the least likely to resort to avoidance strategies to increase their personal safety when out and about alone in their neighborhood at night. Concern about physical disorder is high, but less so for social disorder. Satisfaction with the police is rather high.

Type 6

Type 6 neighborhoods are located on the outskirts of the cities of Basel, Bern, and Zurich. The key characteristics according to the random forests variable importance plot are the share of residents whose highest educational achievement was to have completed an apprenticeship, which is the highest for all six types of neighborhood areas. Conversely, two more important variables are the share of residents with a university degree or employed in a middle management position, which is the lowest among the six neighborhood types.

In terms of survey responses, residents of these areas express the highest levels of fear of crime on all three survey items used to measure the concept. Survey respondents in these areas are more concerned than those elsewhere about crime, fear a burglary of their home, and bypass certain areas, streets or people in order to avoid crime. Perceived levels of disorder are high as well, both for the physical and the social kind. By contrast, area residents are comparatively more satisfied when it comes to judging the quality of their police.

6.4 Concluding Remarks

This chapter set out to develop a clustering procedure that, on the one hand, should minimize the between-cluster similarity of the contextual variables, which are potentially correlated with the outcome measures and may thus confound inferences about a treatment effect. On the other hand, the resulting typology of urban districts should account for a maximum of the between-cluster variance observed in the outcome measures prior to program implementation. Put simply, the overarching goal of the development of the neighborhood typology was to balance both the contextual variables or covariates and the outcome measures for the same neighborhood type across urban areas.

The diagnostic plots revealed that these twin objectives have to a large extent been met: Figure 6.7 showed that the neighborhood typology reduced between-cluster variance in the neighborhood ecological characteristics considerably for the 24 key variables retained in the final model. Also the survey response patterns varied significantly by neighborhood type across the four urban areas for all six items used as outcome measures (Fig. 6.5).

However, considering only the survey respondents of any one neighborhood type, city-level factors still affect response patterns, especially in the two Swiss German cities of Bern and Zurich. This means that in an evaluation design that compares the impact of community policing between neighborhoods of a similar type across urban areas, the current typology manages to *reduce* the threats to internal validity of selection and regression to the mean but *cannot rule them out* completely. This fact has to be borne in mind for the following analyses of the impact of community policing.[39]

[39] In order to have a better understanding of where this city-level variance comes from, the survey response patterns have been plotted by neighborhood type for individual ZIP code and administrative districts for all five urban areas (using again the labels derived from the random forests classifier for the Basel neighborhoods, cf. Appendix E). These charts reveal that within any one neighborhood type response patterns at times fluctuate considerably at the city level and indicate the most egregious outliers. These charts have to be interpreted with caution, though, because of the inordinately small sample size for some neighborhood areas.

7

The Impact of Community Policing on Neighborhood Residents

This final chapter presents the results of the analysis of the impact of community policing on neighborhood residents. The chapter builds on the insights gained in the preceding chapters, marshalling the evidence produced through the process evaluation, exploratory spatio-temporal analysis, and the development of the neighborhood typology. In order to assess the impact of community policing, the indicators of three of the four theoretical constructs of community policing impact – fear of crime, disorder, and public attitudes towards the police – will be used again. These are the constructs which have been operationalized by means of the survey data.

The preceding neighborhood typology has demonstrated that the response patterns of the survey items used to assess the impact of community policing vary significantly across different neighborhood types. It can thus not be assumed that community policing will be perceived equally by neighborhood residents all over a given city. Therefore all the analyses presented in the current chapter are put in a neighborhood context, i.e. the long-term development of the same survey indicators of community policing impact is being studied, but survey respondents are grouped by city and neighborhood type in order to embed the analysis in the neighborhood context. This step is meant to reduce the threats to internal validity such as selection or regression to the mean as alternative explanations for the observed differences in these indicators.

In the discussion in the following chapter, the results of the long-term development of the indicators of community policing impact will be compared for individual neighborhood types across cities. For this analysis, the results of the process evaluation serve for a first assessment of the plausibility that the observed changes in the indicators of treatment outcome were caused or at least partially influenced by shifts in the policing strategy. The idea is to match comparable areas that have large gaps in community policing implementation. The goal is to assess whether the observed improvements in the indicators of community policing are indeed attributable to changes in policing strategies and not just the result of a serendipitous combination of extraneous factors. If the indicators of community policing impact show a significant improvement in one area compared to an area of the same type elsewhere, and the process evaluation additionally discovered a large gap in implementation between the two, this will count as further evidence that the observed changes are at least partially explained by the shifts in policing strategy.

However, even if a perfect match could be achieved between neighborhoods of the same type – a big if since the typology in its current form failed to account for differing response patterns across the three Swiss German cities – threats to the internal validity of the current study design remain. Crucially, if the impact of treatment is assessed by comparing two random samples of the study population, the difference in response patterns is valid as an indicator of treatment impact only to the extent that the pretest and posttest samples are comparable. In a research design with repeated random samples, sampling bias is a very potent threat to internal validity (Cook & Campbell 1979, 117).

In the current study, the problem of sampling bias is further compounded by the highly unequal sample sizes of different editions of the Swiss Crime Survey. Since police departments have sometimes paid to boost the size of the survey sample for their jurisdiction, sample sizes vary substantially across cities and from one edition of the survey to the next. Weighting the data to correct for sampling bias carries the risk, especially for the smaller samples, that the weighted sample does not accurately reflect the study population or that some respondents will be attributed very large weights, or both. This is unsatisfactory because it increases the risk that the observed differences between the pretest and posttest samples are due to a non-negligible extent to sampling bias rather than to actual shifts in popular opinion. In order to circumvent this difficulty, the current study relies on propensity score matching methods to achieve balance in the demographic characteristics of the survey respondents of the pretest and posttest samples.

7.1 Methodology

7.1.1 Propensity Score Matching

Propensity score matching is a method used to balance the distribution of the observed covariates of the treated and control subjects in an observational study in order to avoid that observable differences between them confound inferences about treatment impact. The propensity score is defined as the conditional probability of a study object to receive the treatment given its observed covariates (Hansen 2007, 21). The propensity score is estimated by means of a logistic regression of the pooled sample of the treated and control subjects, whereby the dependent variable is a dummy variable with a value of 1 for treated subjects and 0 for controls and the independent variables are the baseline characteristics or observed covariates that ought to be balanced across the two groups. The estimated propensity score is the fitted value of each study subject on this regression model (Rosenbaum 2010, 240).

During the *matching* part of the method, study objects from the treatment and control groups that have similar values of the propensity scores are set up in paired groups. The rationale behind matching is that study objects with similar estimated propensity scores have similar values on the independent variables. If many such matched paired groups are formed between treated and control subjects with similar propensity scores, the two groups will show similar distributions on the observed covariates and thus be more comparable (Rosenbaum 2010, 72f.).

The act of matching can be achieved in variable ratios: in optimal pair matching, members of the treatment group are matched up pairwise with exactly one control. In the more flexible optimal full matching, treated and control subjects are set up in variable ratios, meaning that a treated subject could be set up with multiple controls and vice versa. Finally, for each paired group, the median value of the outcome variable for the treated and control subjects is retained as the response to the treatment and the control condition, respectively, and the difference between the two is considered the response to treatment for that paired group (Rosenbaum 2010, 173-183).

For the subsequent analyses of the impact of community policing, the survey sample from 2005 is considered the treatment group, and the sample from 2000 the control. For the current study, survey respondents from the treatment and control samples were matched in fully flexible ratios, to accommodate for the fact that sample sizes at times vary substantially between the two editions of the survey. For each city, matched paired groups are being formed between the two samples, whereby respondents are matched on age, gender, education level, and neighborhood type. The matching algorithm was set up so as to allow only paired groups to be formed between survey respondents that reside within the same neighborhood type as determined by the neighborhood typology. On the demographic characteristics the match needed to be close but not perfect.

Table 7.1 serves to illustrate how the propensity score matching method works. It displays the propensity scores and observed covariates for the first seven matched pair groups for the City of Zurich. The table displays the propensity score "ps" of each survey respondent in the second last column. The columns "gender", "age.cat", "educ", "educmiss", and "ntype" indicate the independent variables of the logistic regression used to estimate the propensity score, whereby "treat" indicates the treatment assignment or dependent variable of the regression model, i.e. "0" for the pretest sample from 2000 and "1" for the posttest sample from 2005. The last column "sets" displays the result of the propensity score matching process, the number of matched pair groups formed. As can be gleaned from Table 7.1, the process led to the creation of two matched pair groups for neighborhood type 1. In each matched pair group, treated subjects are set up with suitable controls. The values of the independent variables can now be inspected for the members of each matched pair group to assess the quality of the match.

Table 7.1 Propensity score matching. Propensity scores and observed covariates of the first seven matched pair groups between the 2000 and 2005 survey samples for the City of Zurich.

threat	gender	age.cat	educ	educmiss	ntype	year	treat	ps	sets
1	1	(24,44]	6	0	1	2005	1	0.20140	1.1
1	2	(44,64]	3	0	1	2005	1	0.20543	1.1
1	1	(64,99]	6	0	1	2005	1	0.15794	1.1
1	2	(14,24]	7	1	1	2005	1	0.25735	1.1
1	2	(64,99]	6	0	1	2005	1	0.13686	1.1
4	2	(64,99]	5	0	1	2000	0	0.17475	1.1
2	1	(24,44]	3	0	1	2005	1	0.26670	1.2
1	1	(24,44]	3	0	1	2005	1	0.26670	1.2
2	1	(24,44]	4	0	1	2000	0	0.38367	1.2
3	1	(24,44]	3	0	2	2005	1	0.23969	2.1
1	1	(24,44]	3	0	2	2000	0	0.23969	2.1
5	2	(64,99]	2	0	2	2005	1	0.23558	2.10
2	2	(64,99]	2	0	2	2000	0	0.23558	2.10
1	1	(14,24]	7	1	2	2005	1	0.26217	2.11
2	1	(24,44]	3	0	2	2005	1	0.23969	2.11
1	1	(24,44]	3	0	2	2005	1	0.23969	2.11
1	1	(24,44]	3	0	2	2005	1	0.23969	2.11
3	1	(24,44]	3	0	2	2005	1	0.23969	2.11
4	2	(14,24]	3	0	2	2005	1	0.24985	2.11
1	1	(24,44]	3	0	2	2000	0	0.23969	2.11
3	2	(24,44]	3	0	2	2005	1	0.21043	2.2
1	2	(24,44]	3	0	2	2000	0	0.21043	2.2
1	1	(64,99]	3	0	2	2005	1	0.18994	2.3
1	2	(24,44]	7	1	2	2005	1	0.19378	2.3
3	1	(64,99]	3	0	2	2005	1	0.18994	2.3
3	2	(44,64]	3	0	2	2005	1	0.18308	2.3
2	2	(14,24]	6	0	2	2005	1	0.18762	2.3
3	1	(44,64]	5	0	2	2005	1	0.19710	2.3
3	2	(44,64]	3	0	2	2005	1	0.18308	2.3
1	1	(44,64]	5	0	2	2000	0	0.19710	2.3

7.1.2 Statistical Test and Sensitivity Analysis

The statistical test carried out to check if the survey responses in the treatment and control group differ significantly is Wilcoxon's signed rank statistic (Wilcoxon 1945). Unlike the more common t-test, Wilcoxon's signed rank statistic is valid for ordinal and non-normally distributed variables alike. Moreover, the test is highly robust in the face of outliers and, as a non-parametric test, can be calculated correctly even for very small samples (Siegel 1997; 32f.; 73-80).[40]

Wilcoxon's signed rank statistic is based on the difference in the response between the treated and control subjects. Within each paired group, the median response of the treated subjects and the controls were retained as the response to the treatment and control condition, respectively. For example, in the first matched pair group in Table 7.1 (set 1.1), the median response to the question "How safe do you feel walking alone in your neighborhood after 10pm" ("threat") was 4 ("very unsafe") for the sample from 2000 and 1 ("very safe") for the respondents from 2005. The response to treatment of this pair group is thus $1 - 4 = -3$. By contrast, for matched pair group 2.10, the median response in 2005 was 5 ("Don't go out for security reasons") and 2 ("quite safe") in 2000, so the response to treatment of this pair group is $5 - 2 = 3$. If the median response of both treated and control subjects were identical, the response would be 0 and thus a tie.

The overall response to treatment is then calculated as the sum of the responses over all the matched pair groups $i = 1, \ldots, I$ of a given neighborhood type in a given city. If fear of crime rose in the area over the study period, respondents from 2005 will on average indicate higher values on the fear of crime item than respondents from 2000, and the sum of the responses to treatment over all the matched pairs will be positive. By contrast, if fear of crime dropped between 2000 and 2005, comparatively more matched pair groups will record negative values as the response to treatment, and the overall sum will be negative. As a rule, the categories of the six survey items have been recoded so that higher values indicate the less desirable outcomes form a policy perspective (i.e. higher levels of fear of crime, more disorder, and less satisfaction with the police). If the sign of the overall response to treatment is positive, this means that overall respondents in the 2005 sample gave higher marks to the survey item at hand (i.e. they were more afraid, more concerned about disorder, or less satisfied with their police). Conversely, if the sign of the overall response is negative, the 2005 respondents indicated lower categories on the survey items. Wilcoxon's signed rank statistic is calculated based on the list (vector) of responses to treatment over the individual matched pair groups. All statistics were computed as a one-sided test, assessing the alternative hypothesis that the overall response to treatment is negative.

If Wilcoxon's test statistic is significant, a sensitivity analysis is carried out to check how robust this finding is in the face of potential bias from some unobserved covariate. Since propensity score matching on the observed covariates does nothing also to balance the study subjects on the unobserved covariates, the findings of an observational study are always liable to bias stemming from some unmeasured characteristics. Since the treatment and control groups for the current study are random samples of the study population, it can not be ruled out that the matched samples differ in some unobserved characteristic(s) that could have produced the observed shift in response patterns between the pretest and posttest sample of the survey. The sensitivity analysis gives an indication of just how large such a bias from some unmeasured covariate would have to be, to "alter a study's conclusions." (Rosenbaum 2010, 70-76).

In order to conduct the sensitivity analysis, the Wilcoxon's signed rank statistic is first calculated under the "naïve" assumption that there is no bias stemming from some unobserved covariate(s),

[40] This last aspect is not to be underrated especially for the current study, which has to compare survey samples of varying size. The advantage of non-parametric tests is that they can prove significant effects even for very small samples. Consider the case in which there are only five matched pairs on which to base the analysis. Under the null hypothesis of no effect, the response to treatment can go either up or down, so there are $2^5 = 32$ possible combinations of positive and negative responses to treatment over the five pairs. There is, however, only one possible outcome in which all five responses to treatment are positive. The probability of this event is thus $1/32 = 0.031$, well below the common 5 percent threshold of statistical significance. In other words, five matched pairs may be sufficient for Wilcoxon's signed rank statistic to prove an effect to be significant (Rosenbaum 2010, 36-38).

i.e. that the odds of receiving the treatment are identical for two study subjects with identical observed covariates – or translated to the current context – that survey respondents from the 2000 and 2005 sample with similar propensity scores do not systematically differ due to some unmeasured characteristic. The sensitivity analysis then computes the confidence interval of the Wilcoxon T-statistic assuming that the odds of treatment may differ for two individuals with the same estimated propensity score due to some unobserved covariate. As the size of this assumed bias increases, the confidence interval of the T-statistic steadily widens, until the p-value corresponding to the upper bound will eventually exceed the 5 percent significance threshold, at which point the null hypothesis of no treatment effect can no longer be rejected. The parameter Γ (Gamma) indicates by how much the odds of treatment would have to differ to nullify a significant result (Rosenbaum 2010, 76-79).[41]

7.2 Analysis of the Impact of Community Policing

For the analysis of the impact of community policing, the samples of the Swiss Crime Survey from 2000 and 2005 are compared in order to determine whether the response patterns between these two years differ significantly. As the process evaluation has revealed, the period between 2000 and 2005 was when the implementation of community policing began in earnest in the three Swiss German cities of Basel, Bern, and Zurich. In Geneva and Lausanne, the process of community policing had formally begun already in the mid-1990s. Comparing the 2000 and 2005 survey data for these two cities does hence not qualify as a pre-post design. However, since the 1998 SCS is the oldest data set available for the French part of Switzerland, a true pre-post comparison is not possible for these two cities.

7.2.1 Fear of Crime

Threat of crime

The results of the pretest-posttest comparison for the different neighborhood types of the first SCS survey item measuring fear of crime are presented in Table 7.2. In this table, as in the ones that follow, the results from the statistical tests are arranged as follows: the first two columns indicate the city and the neighborhood type according to the typology developed in the previous chapter. The columns n_0 and n_1 indicate the number of persons interviewed living within that city and neighborhood type from the 2000 and 2005 survey, respectively. Column I indicates the number of matched pair groups that the matching algorithm set up, whereas $\mathbf{r}_T - \mathbf{r}_C$ shows the response to treatment over all the matched pairs formed through propensity score matching. The next three columns T, p-value, and Ties are related to Wilcoxon's signed rank statistic, indicating, respectively, the T statistic, related p-value assuming there is no bias stemming from some unobserved covariate, and the number of matched pair groups that recorded a tie as the response to treatment. In case the test result is significant, the final three columns present the result of the sensitivity analysis. These three columns indicate the p-value of the upper bound of the confidence interval of the T-statistic for increasing sizes of bias resulting from some unmeasured covariate ($\Gamma = 1.5, 2, 3$).

Table 7.2 presents the results of the analysis of the SCS survey item measuring fear of crime as the feeling of safety walking alone through one's neighborhood at night. In general, the matching

[41] Odds and probabilities are two different ways to indicate the likelihood of the occurrence of an event. If the probability of an event occurring is 0.5, then the odds are said to be 1-to-1; the event is as likely to occur as some other event. By contrast, if the odds are 2-to-1, then for some reason the event is twice as likely to occur as the other event. It is straightforward to convert odds ω_k into probabilities $\pi_k = \omega_k/(1 + \omega_k)$ and vice versa $\omega_k = \pi_k/(1 - \pi_k)$. The parameter Γ indicates the difference in the probability of receiving the treatment between treatment and control group. Thus if Γ takes on the value of 1.5, the odds of receiving the treatment differ by a ratio of 1.5-to-1 for subjects in the treatment and control group due to some unmeasured covariate. In other words, the probability of receiving the treatment is $1.5/(1 + 1.5) = 0.6$ rather than 0.5 as in a randomized experiment or an unbiased observational study. If $\Gamma = 2$, the probability of receiving the treatment is $2/(1 + 2) = 0.66$ (Rosenbaum 2010, 77).

algorithm succeeded rather well in forming matched pair groups without unduly reducing the sample size. For most of the sample subgroups, the number of formed pairs is virtually identical in size to the smaller of the two survey samples.

Table 7.2 Threat of crime 2000-2005. Wilcoxon's signed rank statistic of the Swiss Crime Survey item measuring the threat of crime, grouped by city and neighborhood type (ntype), where n_0 and n_1 are the sample size in 2000 and 2005, respectively, I is the number of matched pair groups, and $\mathbf{r}_T - \mathbf{r}_C$ indicates the response to treatment over all the matched pair groups. T, p-value, and Ties are the Wilcoxon test statistics for a one-sided test. If the test is significant ($p < 0.05$), the last three columns indicate the p-value of the upper bound of the confidence interval of a sensitivity analysis for increasing levels of bias resulting from some unmeasured covariate.

City	ntype	n_o	n_1	I	$\mathbf{r}_T - \mathbf{r}_C$	T	p-value	Ties	Γ=1.5	Γ=2	Γ=3
Basel	2	152	18	18	+	104.5	0.779	5			
Basel	4	99	7	7	+	17.5	0.656	1			
Basel	6	203	21	20	+	173.5	0.995	7			
Bern	1	7	4	3	−	0.5	0.125	1			
Bern	4	92	37	34	−	292.5	0.46	14			
Bern	6	88	35	33	−	221.5	0.144	9			
Geneva	1	11	2	2	−	0.5	0.25	1			
Geneva	2	120	25	25	+	175	0.625	7			
Geneva	3	93	15	15	−	55	0.381	6			
Geneva	5	157	34	29	+	234	0.633	6			
Lausanne	1	13	15	10	+	34	0.722	3			
Lausanne	2	111	126	90	−	2225	0.761	28			
Lausanne	4	45	68	37	−	350	0.488	7			
Lausanne	5	16	19	9	+	31.5	0.82	4			
Zurich	1	2	7	2	−	0	0.25	0			
Zurich	2	11	53	11	+	38	0.65	2			
Zurich	4	44	198	44	−	349	0.044	18	0.291	0.598	0.908
Zurich	6	27	204	27	+	203	0.625	10			

In terms of the actual outcome, it is notable that in Basel, all three neighborhood types witnessed an increase in fear of crime, whereas in Bern the overall response to treatment is negative four all four neighborhood types, meaning survey respondents generally reported lower levels of fear of crime. None of the drops in fear of crime in Bern is statistically significant, however, not even at the 10 percent threshold.

The picture for the other three cities is more mixed. In Geneva, Lausanne, and Zurich, which each count four different neighborhood type areas, fear of crime rose in two neighborhood types and fell in the other two. One of the observed changes is statistically significant: in Zurich in neighborhood type 4, fear of crime dropped significantly between 2000 and 2005. However, this result is not very robust as the corresponding sensitivity analysis reveals. Assuming that the pretest and posttest samples differ in some unmeasured covariate by a ratio of 1.5-to-1, the observed change ceases to be significant.

Risk of victimization

Table 7.3 presents the results of the analysis of the SCS survey item measuring the risk of victimization, i.e. the perceived risk of a burglary of one's home over the next twelve months. The matching algorithm generally worked well again, producing a number of matched pairs that is virtually as large as the smaller of the two survey samples.

Table 7.3 Risk of victimization 2000-2005. Wilcoxon's signed rank statistic of the Swiss Crime Survey item measuring the risk of victimization, grouped by city and neighborhood type (ntype), where n_0 and n_1 are the sample size in 2000 and 2005, respectively, I is the number of matched pair groups, and $\mathbf{r}_T - \mathbf{r}_C$ indicates the response to treatment over all the matched pair groups. T, p-value, and Ties are the Wilcoxon test statistics for a one-sided test. If the test is significant ($p < 0.05$), the last three columns indicate the p-value of the upper bound of the confidence interval of a sensitivity analysis for increasing levels of bias resulting from some unmeasured covariate.

City	ntype	n_o	n_1	I	$\mathbf{r}_T - \mathbf{r}_C$	T	p-value	Ties	Γ=1.5	Γ=2	Γ=3
Basel	2	143	19	19	+	114	0.767	8			
Basel	4	92	7	7	−	8.5	0.148	1			
Basel	6	196	23	22	−	122	0.437	12			
Bern	1	5	4	3	+	4	0.625	0			
Bern	4	82	38	32	−	236	0.299	13			
Bern	6	82	36	32	−	150	0.015	11	0.12	0.309	0.659
Geneva	1	11	2	2	+	2.5	0.5	1			
Geneva	2	118	24	24	+	184.5	0.828	7			
Geneva	3	89	15	15	−	42	0.151	3			
Geneva	5	149	37	34	−	192.5	0.035	14	0.213	0.462	0.802
Lausanne	1	13	15	9	+	25.5	0.59	2			
Lausanne	2	104	128	83	+	2076	0.934	35			
Lausanne	4	46	69	44	+	649	0.963	15			
Lausanne	5	17	20	13	−	41	0.368	4			
Zurich	1	3	7	3	−	1.5	0.125	2			
Zurich	2	11	51	11	−	21	0.139	2			
Zurich	4	42	195	42	−	268	0.01	15	0.116	0.337	0.733
Zurich	6	23	204	22	−	61	0.015	7	0.088	0.213	0.481

The results on victimization risk are mixed. In Basel, fear went up in the central neighborhood type 2 and fell elsewhere, but neither of the two reductions in burglary risk is statistically significant. In Bern, fear ticked up in the center but dropped elsewhere, especially in the outskirts where the observed change is statistically significant. The patterns are similar for Geneva and Lausanne, where the treatment response indicated an increase in fear in the central neighborhood types and a drop in the outskirts. In Geneva, it was the urbanized municipalities around the City of Geneva that saw a significant drop in the perceived burglary victimization risk. In Zurich, finally, all four neighborhood types witnessed a reduction in the prevalent levels of fear of crime. This fall was most pronounced in the neighborhood types 4 and 6, for which the change in response patterns is statistically significant. The related sensitivity analyses reveal, however, that these improvements are not very robust in the face of potential bias due to some unmeasured covariate. For all four neighborhood types that had witnesses a significant drop in the perceived victimization risk, the T-statistic is no longer significant for a bias due to some unmeasured covariate of $\Gamma = 1.5$.

Behavioral response

Table 7.4 presents the results of the third SCS survey item measuring fear of crime, i.e. the behavioral response to crime. What has been said so far about the performance of the matching algorithm can be repeated here: the loss in sample size due to the matching procedure is for the most part negligible.

Table 7.4 Behavioral response 2000-2005. Wilcoxon's signed rank statistic of the Swiss Crime Survey item measuring the behavioral response to crime, grouped by city and neighborhood type (ntype), where n_0 and n_1 are the sample size in 2000 and 2005, respectively, I is the number of matched pair groups, and $\mathbf{r}_T - \mathbf{r}_C$ indicates the response to treatment over all the matched pair groups. T, p-value, and Ties are the Wilcoxon test statistics for a one-sided test. If the test is significant ($p < 0.05$), the last three columns indicate the p-value of the upper bound of the confidence interval of a sensitivity analysis for increasing levels of bias resulting from some unmeasured covariate.

City	ntype	n_o	n_1	I	$\mathbf{r}_T - \mathbf{r}_C$	T	p-value	Ties	Γ=1.5	Γ=2	Γ=3
Basel	2	151	18	18	−	71	0.261	8			
Basel	4	99	7	7	−	13	0.406	3			
Basel	6	205	21	21	+	134	0.73	9			
Bern	1	7	4	3	−	1.5	0.125	2			
Bern	4	92	38	33	−	144	0.006	15	0.068	0.21	0.543
Bern	6	88	36	35	−	272	0.24	14			
Geneva	1	11	2	2	0	1.5	0.25	2			
Geneva	2	123	23	23	+	145.5	0.577	14			
Geneva	3	97	15	15	+	75	0.789	8			
Geneva	5	158	35	33	−	255.5	0.323	15			
Lausanne	1	13	14	9	−	16.5	0.213	2			
Lausanne	2	111	127	86	+	1945.5	0.624	40			
Lausanne	4	42	70	34	−	255	0.234	16			
Lausanne	5	16	21	9	+	30	0.787	5			
Zurich	1	3	7	3	0	3	0.375	3			
Zurich	2	11	54	11	+	44.5	0.817	4			
Zurich	4	44	199	44	−	256.5	0.002	22	0.045	0.181	0.557
Zurich	6	28	202	28	−	195.5	0.424	18			

In terms of actual outcomes, four cities witnessed both rising and falling levels of fear of crime measured as the behavioral response to crime; only the City of Bern saw fear of crime drop across the board. In Bern and Zurich, in neighborhood type 4 the observed drop in fear of crime is statistically significant. What is more, these findings appear somewhat more robust than the results of the analysis of the two other fear of crime measures presented so far. The sensitivity analysis reveals that for the type 4 neighborhoods in Zurich, the observed difference is statistically significant, even in the face of some bias resulting from an unobserved covariate. In Bern, the evidence is slightly weaker; there the T-statistic is no longer significant at the 5 percent threshold allowing for some unmeasured bias of $\Gamma = 1.5$.

7.2.2 Disorder

Physical disorder

Table 7.5 presents the results of the indicators of physical disorder, i.e. the presence of graffiti or litter in the streets near a respondent's home. This table confirms the general impression of the exploratory spatio-temporal analysis of this indicator, namely that physical disorder rose over the 2000-2005 period.

Table 7.5 Physical disorder 2000-2005. Wilcoxon's signed rank statistic of the Swiss Crime Survey item measuring physical disorder, grouped by city and neighborhood type (ntype), where n_0 and n_1 are the sample size in 2000 and 2005, respectively, I is the number of matched pair groups, and $\mathbf{r}_T - \mathbf{r}_C$ indicates the response to treatment over all the matched pair groups. T, p-value, and Ties are the Wilcoxon test statistics for a one-sided test. If the test is significant ($p < 0.05$), the last three columns indicate the p-value of the upper bound of the confidence interval of a sensitivity analysis for increasing levels of bias resulting from some unmeasured covariate.

City	ntype	n_o	n_1	I	$\mathbf{r}_T - \mathbf{r}_C$	T	p-value	Ties	Γ=1.5	Γ=2	Γ=3
Basel	2	154	20	20	+	129	0.806	9			
Basel	4	100	7	7	+	28	0.992	0			
Basel	6	209	23	23	−	129	0.388	10			
Bern	1	7	4	3	−	1.5	0.125	2			
Bern	4	93	39	36	+	389	0.806	22			
Bern	6	92	37	35	+	345	0.683	20			
Geneva	1	11	2	2	+	2.5	0.5	1			
Geneva	2	128	25	25	0	162.5	0.489	9			
Geneva	3	98	16	16	+	105.5	0.971	6			
Geneva	5	173	37	35	+	347.5	0.695	16			
Lausanne	1	13	15	10	+	42	0.92	4			
Lausanne	2	117	133	96	−	1799.5	0.026	44	0.41	0.83	0.995
Lausanne	4	49	74	40	+	467	0.774	17			
Lausanne	5	18	21	12	−	29.5	0.212	6			
Zurich	1	3	7	3	0	3	0.375	3			
Zurich	2	11	54	11	+	42.5	0.768	7			
Zurich	4	46	208	46	+	702.5	0.961	25			
Zurich	6	30	209	29	−	208	0.416	18			

In Basel and Bern, physical disorder rose in two neighborhood types and fell in one, but the parallels end there. Whereas in Basel disorder dropped in the outskirts and rose elsewhere, in Bern it was the city center that witnessed an improvement but not the other parts. In Geneva, disorder rose in all areas except neighborhood type area 2, whereas in Lausanne perceived disorder rose in the center and the eastern parts of the city but fell in the northern and western parts. The improvement in neighborhood type 2 in Lausanne was statistically significant. While this result is based on a relatively large sample, the effect is still not very robust. Assuming that the two samples differ by a bias of only $\Gamma = 1.5$ due to some unobserved covariate, the observed difference ceases to be statistically significant. In Zurich, the situation appears unchanged in the center but deteriorated in the adjacent neighborhood types 2 and 4. Only in the city's outskirts fewer respondents spotted signs of physical disorder in 2005 than in 2000.

Social disorder

Table 7.6 Social disorder 2000-2005. Wilcoxon's signed rank statistic of the Swiss Crime Survey item measuring social disorder, grouped by city and neighborhood type (ntype), where n_0 and n_1 are the sample size in 2000 and 2005, respectively, I is the number of matched pair groups, and $\mathbf{r}_T - \mathbf{r}_C$ indicates the response to treatment over all the matched pair groups. T, p-value, and Ties are the Wilcoxon test statistics for a one-sided test. If the test is significant ($p < 0.05$), the last three columns indicate the p-value of the upper bound of the confidence interval of a sensitivity analysis for increasing levels of bias resulting from some unmeasured covariate.

City	ntype	n_o	n_1	I	$\mathbf{r}_T - \mathbf{r}_C$	T	p-value	Ties	Γ=1.5	Γ=2	Γ=3
Basel	2	154	20	20	+	167	0.99	10			
Basel	4	100	7	7	+	20.5	0.812	5			
Basel	6	209	23	23	+	171	0.835	20			
Bern	1	7	4	3	+	4.5	0.625	2			
Bern	4	93	39	36	+	413	0.893	21			
Bern	6	92	37	35	+	439.5	0.979	20			
Geneva	1	11	2	2	+	3	0.75	0			
Geneva	2	128	25	25	+	240.5	0.982	13			
Geneva	3	98	16	16	+	68.5	0.49	14			
Geneva	5	173	37	35	+	412.5	0.943	29			
Lausanne	1	13	15	10	–	26.5	0.423	6			
Lausanne	2	117	133	96	–	2287	0.44	59			
Lausanne	4	49	74	40	+	444.5	0.672	32			
Lausanne	5	18	21	12	–	27	0.17	8			
Zurich	1	3	7	3	–	0.5	0.125	1			
Zurich	2	11	54	11	0	34	0.517	4			
Zurich	4	46	208	46	+	645.5	0.871	37			
Zurich	6	30	209	29	+	228.5	0.584	20			

Table 7.6 presents the results for social disorder. The emerging picture is comparable to the results for physical disorder. The general tendency is for social disorder to increase between 2000 and 2005.

In the cities of Basel, Bern, and Geneva, perceived social disorder rose across all neighborhood types found there. In Lausanne and Zurich, by contrast, the pattern is more nuanced. In both of these cities, there are neighborhood type areas where the perceived levels of social disorder actually dropped over the study period. In Lausanne, perceived social disorder fell everywhere save in the city's eastern neighborhoods, but none of the drops is statistically significant. In Zurich the pattern is inconclusive. While social disorder fell in the city center, the situation appears unchanged in the adjacent Langstrasse district. In the remaining two neighborhood types, social disorder had risen over the study period.

7.2.3 Public Attitudes towards the Police

Table 7.7 Police effectiveness 2000-2005. Wilcoxon's signed rank statistic of the Swiss Crime Survey item measuring popular satisfaction with police effectiveness, grouped by city and neighborhood type (ntype), where n_0 and n_1 are the sample size in 2000 and 2005, respectively, I is the number of matched pair groups, and $\mathbf{r}_T - \mathbf{r}_C$ indicates the response to treatment over all the matched pair groups. T, p-value, and Ties are the Wilcoxon test statistics for a one-sided test. If the test is significant ($p < 0.05$), the last three columns indicate the p-value of the upper bound of the confidence interval of a sensitivity analysis for increasing levels of bias resulting from some unmeasured covariate.

City	ntype	n_o	n_1	I	$\mathbf{r}_T - \mathbf{r}_C$	T	p-value	Ties	Γ=1.5	Γ=2	Γ=3
Basel	2	125	17	17	−	52.5	0.122	10			
Basel	4	82	6	6	−	5	0.109	4			
Basel	6	181	22	22	−	105	0.241	12			
Bern	1	7	3	3	+	3	0.375	0			
Bern	4	81	27	27	+	198	0.58	16			
Bern	6	71	29	26	−	144	0.211	10			
Geneva	1	8	2	2	−	0.5	0.25	1			
Geneva	2	112	20	20	0	107	0.522	8			
Geneva	3	81	13	13	−	36.5	0.249	7			
Geneva	5	142	32	31	−	209	0.222	18			
Lausanne	1	10	14	8	−	15.5	0.32	3			
Lausanne	2	102	106	77	+	1490.5	0.476	34			
Lausanne	4	41	53	34	+	378	0.913	20			
Lausanne	5	15	18	11	+	37	0.618	6			
Zurich	1	2	5	2	−	0	0.25	0			
Zurich	2	10	49	10	−	19	0.188	4			
Zurich	4	38	168	37	−	284	0.154	16			
Zurich	6	26	165	25	−	147	0.336	12			

Table 7.7 presents the results of the analyses for popular attitudes towards the police, i.e. how survey respondents rated the effectiveness of the local police in combating crime. In Basel, respondents gave the police lower (i.e. better) marks in 2005 than in 2000 across all three neighborhood types found there. The situation is comparable in Geneva and Zurich, where more respondents rated the local police as more effective across all neighborhood types, except for one type in Geneva where the situation appeared unchanged. In Bern and Lausanne, by contrast, the patterns are noticeably different. In both of these cities, satisfaction with the police dropped in a majority of neighborhood types. Only in the city center in Lausanne and the outskirts of Bern did matters improve. However, across all the five cities, none of the improvements in popular assessments of the police was statistically significant, not even at the 10% threshold.

8

Discussion

8.1 Methodology

The unusually long study period of the *process evaluation* posed serious challenges to the reliability of the data. In this regard, the data gathering process, which relied on two separate methods – combining interviews with officials from the five police departments with the analysis of written documents from the time of reform – proved advantageous for the long time-span evaluated. Moreover, in all five cities, the police officials interviewed had been with their departments for more than a decade and had been intimately involved in the internal preparation for the transition to community policing. The information on the transition process gathered through the interview process and gleaned from the written documents has been assembled in a separate community policing implementation report on each of these five cities, which have been proofread by the police department officials who gave the interviews. Finally, the process of community policing implementation was analyzed by intervals of five years. This made it possible in virtually all instances to date the onset of a given community policing element, corroborating the assumption that such a time frame corresponded to the level of accuracy that could be achieved in retrospect.

Subsequently, the information contained in the community policing implementation reports has been coded into a "Calendar of Action" (Crawley & Hope 2003) or community policing scorecard that is meant to track the chronology of implementation of each city according to six evaluative dimensions, each of which has been captured by a list of indicators. The coding work has been done by the author of the study alone. In an effort to ensure nonetheless an acceptable degree of inter-city coherence in the coding of the qualitative information, the author visited all five police departments in the study for a third time to discuss the community policing scorecard with officials from each department.

This third round of interviews notwithstanding, defining and interpreting the implementation criteria was sometimes a hard call, and much room for interpretation remains. The time line was applied as a strict benchmark, i.e. for a police department to get marks for a given indicator for a given year, the community policing element in question had to be up and running by that time (and not just be a sketch on a police department's drawing board). In deciding whether the measures taken fulfilled the requirement, however, the criteria were generally interpreted loosely, counting basically all efforts that capture the essential idea. For example, in Zurich and Lausanne, the police do not have specialized neighborhood liaison officers, and this role is incumbent on the commanders of the neighborhood police stations. In Basel, Bern, and Geneva, by contrast, in each police station there is at least one neighborhood liaison officer or community policing specialist whose sole task is problem-solving and community outreach. A one-size fits all approach in coding the process evaluation data runs into difficulties faced with such institutional differences. In this sense, the scorecards are not meant to be a definitive history of local community policing implementation but the basis for a comparative analysis of the impact of community policing strategies across the five cities in the study.

On the upside, the long study period of the current process evaluation made possible the analysis of the chronology of community policing implementation over half a generation. The findings from the present evaluation vindicate the decision to look at such a long period of time. The evidence suggests that implementation takes the better part of a decade or more to achieve. As a matter of fact, the average time period that has lapsed between a police department's formal adoption of the community policing philosophy and its implementation of most of the criteria defined as relevant according to the current process evaluation is 10 to 15 years.

Furthermore, the coding of the process evaluation data allows both cross-sectional and visual analyses of the community policing implementation data. Coding not only helps to get a grip on the wealth of information on community policing implementation and thus to highlight the peculiarities of the local implementation histories. It also makes it possible to identify the largest gaps in community policing "dosage" and hence to spot the evaluative dimension most in need of further reform.

For the *exploratory spatio-temporal analysis*, the indicators of the four theoretical constructs of community policing impact – crime, fear of crime, disorder, and public attitudes towards the police – were subjected to a long-term trend and spatio-temporal pattern analysis. Apart from the vagaries that notoriously plague official crime statistics, the study also had to grapple with measurement issues concerning the survey data. The sample size of respondents was excessively small for some ZIP code districts, and the sampling bias of the demographic variable gender at times disturbingly high, both of which could have skewed the apparent trends significantly. The study corrected for this bias by weighting the samples at the neighborhood-level so that the gender and age distribution accurately reflected the composition of the resident population.

For the long-term trend analysis of the survey data, the percentage values were calculated along with the usual margins of error of the survey samples and corresponding confidence intervals. For the spatio-temporal pattern analysis, however, the weighted mean or percentage values of each indicator were taken at face value for the calculation of Moran's I, even though they are evidently subject to a margin of error as well. To what extent ignoring the error margins affects the reliability of the value and Z-score statistic of Moran's I, is an issue that the current study could not address. An enhanced analysis of the spatial autocorrelation, however, should attempt to take into account the error margins of the underlying attribute values and indicate a confidence interval for Moran's I.

For the *neighborhood typology*, both unsupervised and supervised data mining algorithms were employed in order to build a good parsimonious model to cluster similar neighborhood areas across the five cities being studied here. Since the rationale of the clustering was to find suitable "treatment" and "control" areas, the optimization objectives for the clustering algorithm were clear: on the one hand, the neighborhood typology should cluster areas that are similar with regard to their ecological characteristics in order to reduce the risk that these covariates confound any inferences about treatment impact. On the other hand, the typology should account for a significant share of the variance in the survey items used to evaluate the impact of community policing, so that the residents of a given area collectively expressed similar views at the onset of treatment implementation.

The classification system developed for the current study goes a long way to reconcile these twin objectives despite the fact that the clustering procedure did not produce a single optimum number of clusters. Indeed, this research suggests that such a unique solution does not exist since the number of clusters that minimizes between-cluster similarity is smaller for the ecological data than for the survey data. Still, this has far-reaching ramifications: it means that the individual neighborhood categories are less distinct and that in the input space of the ecological data the clusters tend to overlap so that neighborhood areas that are on the border between two clusters could be classified either way.

This uncertainty adds to the variability inherent in the SOM clustering algorithm, the results of which vary to some extent as a function of the randomly chosen initial values or weights of the prototype vectors. In the resulting typology, there is thus an element of randomness dependent on the initial weights of the SOM clustering algorithm, which obviously affects subsequent analysis based on its results. The standard remedy to handle this aspect of the SOM is to run the clustering procedure multiple times. For the current study, the procedure was replicated 50 times, and the solution retained that best met the twin optimization criteria defined in the first place.

The random forests algorithm by contrast produced very stable results. The algorithm made it possible to identify the key indicators for each of the five categories of ecological variables and thus to limit the number of explanatory variables to build the final model down to just 24 without unduly increasing the classification error rate.

The second drawback of the neighborhood typology is that it fails to account for all or most of the variance in the survey items at the higher aggregate levels of analysis. If the SCS survey respondents are grouped by neighborhood type, there remains significant variance in the response patterns at the city level. This implies that local events or characteristics impinge on the survey data and make cross-sectional comparisons fraught with difficulties, even for survey respondents that reside within the same neighborhood class. Since the typology explicitly serves as the intelligence basis to select matching treatment and control areas to assess the impact of varying community policing strategies, this is obviously not ideal. It means that inferences about program impact are necessarily weaker, because it cannot be ruled out that some initial difference at the beginning of implementation accounts for the changes observed over subsequent periods. However, such selection bias may be less of a problem than meets the eye. First of all, this appears to be a problem only for the Swiss German cities. For Geneva and Lausanne, once survey respondents are grouped by neighborhood type, the variance at the city level ceases to be significant (cf. Fig. 6.2). More importantly, for Bern and Zurich the χ^2-independence test on the risk-of-victimization and behavioral-response-to-crime measures proved non-significant for the type 4 neighborhoods, which saw the biggest improvements over the study period (cf. Fig. 6.3). This means that on these two outcome measures at least, the views of residents of type 4 neighborhoods in Bern and Zurich did not differ significantly prior to community policing implementation and that the risk of selection bias is thus greatly reduced.

Finally, the *impact analysis* was modeled on Cochran's (1965, quoted in: Rosenbaum 2010, 4) advice that observational studies should emulate the template of a randomized trial. In a randomized trial, the probability of each study object to receive the treatment is exactly the same, and randomization balances both the *observed and unobserved* covariates of the study objects, which are thus independent of the treatment assignment. In an observational study, by contrast, the probability of a study object to receive the treatment is typically unknown and is estimated by the propensity score method. In a well-designed observational study, the treated subjects are matched up with suitable controls that have similar propensity scores. If many such matched pairs between treated subjects and controls are formed, the distribution of the *observed* covariates between the two groups will be similar and thus independent of treatment (Rosenbaum 2010). In the current study, the neighborhood typology served to embed the impact analysis in a neighborhood context and hence to achieve balance in the neighborhood-level covariates that are significantly correlated with the survey response patterns. Propensity score matching was then used to match individual survey respondents between the pretest and posttest samples and hence to match individual-level covariates within each neighborhood type area in order to keep sampling bias in check. Given the varying sample sizes of the Swiss Crime Survey and missing values on some of the observed covariates, this could not have been achieved by weighting of the data.

In propensity score matching, the standard practice is to include many covariates indiscriminately in the logistic regression model to estimate the propensity score. For this study, a slightly different approach was chosen. The propensity score of individual survey respondents was estimated using only four variables: gender, age, education, and the neighborhood type. This was done for two reasons: first, by means of the neighborhood type variable, survey respondents were matched on a series of ecological factors that are included in the neighborhood typology. Second, propensity score matching only achieves balance over a larger number of matched pairs. However, for the current analysis only a couple of matched pair groups with a handful of respondents could usually be formed per city and neighborhood type, so the goal was to achieve as good a match as possible on the few demographic variables that are known to affect fear of crime. Most prominent among the missing variables was a measure of victimization. Given the low rates for most crimes, this omission should not alter the results substantially, but it is impossible to be sure and future analysis should attempt to include it.

The statistical test performed to assess the level of significance of the observed difference in response patterns between the pretest and posttest samples was Wilcoxon's signed rank statistic. A non-parametric test was the only viable option for the task at hand. Given the small size of many of the survey samples that had to be analyzed, confidence intervals of the parameter estimates would have been too large for any meaningful analysis using a parametric test.

Finally, any significant result was subject to a sensitivity analysis. A sensitivity analysis is indispensable in an observational study design since it can never be ruled out that two individuals with similar propensity scores (i.e. that "look" similar) do not in fact differ in a systematic and relevant way in some unobserved covariate. The sensitivity analysis gives an indication of how large such a bias would have to be in order to overturn the conclusion that the null hypothesis of no effect is false (Rosenbaum 2010), i.e. by how much the pretest and posttest samples would have to differ due to some unobserved covariate in order to nullify the conclusion of a significant difference.

8.2 Hypotheses Revisited

Following this discussion of both the merits and shortcomings of the different analytical methods, it is now possible to revisit and reassess the research hypotheses formulated at the end of the literature review in light of the evidence produced by this study.

8.2.1 Neighborhood Characteristics

There were four hypotheses guiding this research on how the four theoretical constructs of this community policing evaluation – crime, fear of crime, disorder, and public attitudes towards the police – relate to each other at the neighborhood level:

- Areas with higher levels of crime have higher levels of fear of crime and lower levels of satisfaction with the police.
- Areas with higher levels of fear of crime have lower levels of satisfaction with the police.
- Areas with higher levels of disorder have higher levels of crime and of fear of crime.
- Areas with higher levels of social cohesion have lower levels of crime and of fear of crime.

Thanks to the neighborhood typology some interesting facts could be established regarding the much debated link between crime, fear of crime, disorder, and public attitudes towards the police in Swiss cities even before the evaluation of community policing strategies began. The question of whether the broken windows hypothesis is valid for Swiss urban areas is probably too simplistic by half, though. The one observation that can be made is that it all depends on the neighborhood context. In the city centers, residents generally observed higher levels of disorder, yet those same residents were relatively sanguine about victimization risk even though they were less satisfied with their police. By contrast, in the more suburban neighborhoods, residents observed similar levels of disorder as in the urban centers, but residents were more afraid of crime and comparatively more satisfied with the police. This finding vindicates both the original theoretical work by Wilson & Kelling (1982) who emphasized that disorder may play out differently in different types of neighborhoods, but also some more recent empirical work that found that neighborhood context mattered how residents perceived different levels of incivilities (Gau & Pratt 2010).

However, a big caveat is in order here since an analysis of the correlation at the aggregate level between two variables that are measured at the individual level risks falling foul of the ecological fallacy (Robinson 1950). This means that if in the neighborhoods where local residents are more concerned about disorder, they are also more afraid of crime, this does not imply that at the individual level those who perceive more disorder also worry more about crime. Since the Swiss Crime Survey data are gathered at the individual level, yet the correlations of interest here are situated at the neighborhood level, the risk of an ecological fallacy is present. However, the evidence suggests that the observed correlations may be real and not just an artifact of aggregation, since the results are comparable for varying typologies developed for the current study, i.e. different modes of aggregating individual survey respondents.

As for the relative importance of the neighborhood ecological characteristics, the random forests variable importance plot showed that not all explanatory variables are equally important in categorizing individual neighborhoods and that different variables differ in importance between the individual neighborhood types (Fig. 6.4). The random forests variable importance plot showed that the defining characteristics of the areas most afflicted by higher levels of fear of crime were not actual crime rates but demographic and socio-economic factors. This finding confirms the results of some seminal works in the fear of crime literature, namely the notion that "fear of crime is a result more of community dynamics than of crime dynamics" (Taylor & Hale 1986, 156).

8.2.2 Implementation Process

There were a total of seven hypotheses guiding the process evaluation of community policing implementation:

- If community outreach and problem-solving activities are the prerogative of a few specially trained officers, there is an increased risk of their marginalization within the police organization and the abandonment of all community-oriented policing functions.

This hypothesis is difficult to assess based on the current process evaluation because there is simply not enough variance in data: all five police departments under study practice a more or less specialized model of community policing, where community outreach and problem-solving activities are the prerogative of the few specialized neighborhood liaison officers or the specially detached patrol officers of the neighborhood police stations. Nevertheless, the present process evaluation did not discover any evidence of these "community policing specialists" being marginalized inside their police department. It is possible, however, that their elevated position within the police hierarchy together with the fact that all police officers are nowadays being instructed in community policing during their basic training at the police academy may act as countervailing forces against them being sidelined within the police organization.

- With the adoption of community policing, police departments will increase the number of strategic partnerships with other branches of the local government over time.

The evidence in support of this hypothesis is rather strong: all five police departments have over time established a growing number of strategic partnerships with other local government agencies to handle crime and disorder problems. What is more, even though such partnerships at times faced considerable initial resistance, once these cooperation agreements are concluded, they do not unravel later on.

- Involvement of the general public in community policing activities will not be representative of the overall population and be difficult to sustain in the long run.

The standing evidence that gave rise to this hypothesis stems from abroad, namely the evaluation of the CAPS policing strategy in Chicago. Assessing it for the case of Switzerland, by contrast, is more difficult because the active involvement of community members in community policing efforts only goes so far. In Switzerland, members of the community are being heard through regular citizen surveys. Business and neighborhood associations are usually more directly involved through regular round table discussions. It is difficult to assess the representativeness of such limited community involvement.

- With the adoption of community policing, the police will broaden the definition of their mission to include combating fear of crime and disorder alongside combating crime.

This hypothesis is universally true for all the police departments: all five have broadened their objectives beyond controlling crime to include the combating of fear of crime and improving police-community relations. In accord with the problem-oriented policing philosophy, they have also acknowledged that they must forge strategic partnerships with local government and private actors to achieve these goals.

- The more strategic partnerships the police have concluded with other branches of the local administration, the less law enforcement-centered strategies to deal with disorder will be.

There is more than mere anecdotal evidence in support of this hypothesis. In all five cities, the authorities have introduced novel strategies and created new government agencies to deal with signs of physical and social disorder. The campaigns to remove graffiti or clean up parks and streets are mostly multi-stakeholder initiatives, in which the police are involved but do not normally have the lead. The street worker units that the cities have established to deal with social disorder work in close cooperation with the police but remain institutionally independent from them. This is in marked contrast with some of the earlier unsuccessful strategies of the police to reign in the open air drug markets that had emerged in the Swiss German cities during the late 1980s and early 1990s.

- The more strategic partnerships the police have concluded with other branches of the local administration, the more likely they are to be involved in the planning and steering process of neighborhood regeneration projects.

The evidence from this process evaluation is also supportive of this hypothesis: the cities of Basel, Geneva, Lausanne, and Zurich, where the police have established a number of strategic partnerships with local government agencies, are also the cities where the police are involved in the planning and steering committees of neighborhood regeneration projects.

- With the adoption of community policing, the police will step up crime prevention activities.

There is also unequivocal evidence in support of the latest hypothesis: while all departments have been offering crime prevention advice for free for decades, these efforts have been stepped up since the formal adoption of community policing. Nowadays police departments target particularly vulnerable groups with their crime prevention advice, run more public campaigns, and some do send out targeted crime alerts if certain types of criminal infractions such as burglaries or shopliftings suddenly surge in an area.

8.2.3 Community Policing Impact

There were five hypotheses guiding this research on how the implementation of community policing should impact the four theoretical constructs of crime, fear of crime, disorder, and public attitudes towards the police:

- The better police can identify and attend to the proximate causes of crime patterns, the lower the levels of crime.

- Areas with a neighborhood police station and permanently assigned community liaison officers have lower levels of fear of crime and higher levels of satisfaction with the police.

- Areas with higher levels of police visibility, notably of officers on foot patrol, have lower levels of fear of crime and higher levels of satisfaction with the police.

- Areas with campaigns to remove signs of physical disorder have lower levels of crime and of fear of crime.

- Areas with intervention teams attending to social disorder have lower levels of crime and of fear of crime.

While the results of the process evaluation do not allow a detailed discussion of these hypotheses, the question remains whether community policing was indeed the cause of the drop of fear of crime in Zurich and Bern over the study period, the most important observed difference of this study? In order to assert a causal link between a treatment and an effect, a study must demonstrate three things to allow such an inference: (i) precedence in time of cause to effect, (ii) covariation between treatment and effect, and (iii) the absence of plausible alternative explanations for the observed covariance (Shadish *et al.* 2002, 53f.). These three conditions are now being discussed in turn in order to assess the quality of the evidence.

The first condition is met easily: temporal precedence of the intervention can be taken for granted. The process evaluation established that the reductions in fear of crime in both Bern and Zurich occurred over a time period during which both cities made considerable progress in community policing implementation. In particular, both cities made substantial progress on both the crucial broken windows and the problem solving axis over the period of interest here.

The second condition, covariance, can be considered as fulfilled as well, albeit on the basis of slightly weaker evidence. The study provides fair evidence of a real reduction of fear of crime in the two cities. In Bern in neighborhood type 4, all three measures indicated a drop in fear of crime, one of which was statistically significant. In Zurich in the same neighborhood type, all three fear of crime measures indicated a significant drop as well, one of them even after allowing for a bias of $\Gamma = 1.5$ due to some unmeasured covariate. This means that for this observed difference to be merely an artifact of sampling bias or an unobserved covariate rather than a sign of real improvement, less fearful respondents would need a 60 percent chance of being included in the survey sample compared to a 40 percent chance for more fearful persons due to some covariate not considered by the study. That seems a fairly wide margin. Wilcoxon's signed rank statistic and related sensitivity analysis thus established that fear of crime dropped significantly over the period of interest, even allowing for moderate sampling bias.

Still, a reduction in fear of crime after community policing efforts have been stepped up is not yet proof of covariance. Grosso modo fear of crime dropped in all three Swiss German cities and remained more or less stable in Geneva and Lausanne. Simultaneously, the five police departments made more or less steady progress towards community policing implementation. In the absence of either any instance of major backsliding in the implementation process or any significant increases in fear of crime, covariance can only partly be assessed.

The third requirement – the absence of alternative explanations – is harder to prove, and in the absence of an experimental design, doubts can never be removed completely. In an observational design, any extraneous factor that could have provoked the observed effect is a threat to internal validity (Shadish *et al.* 2002, 53f.). Still, by embedding the analyses in the neighborhood context and using propensity score methods to match survey respondents from the 2000 and 2005 SCS samples, the study made a fair attempt to address the threats to internal validity of selection, sampling bias, and regression to the mean. The neighborhood typology developed to this end allows at least tentative inferences of a causal effect. As a matter of fact, if survey respondents were grouped by the different types of neighborhood areas, the remaining variance at the city-level ceased to be significant for the outcome measures and neighborhood type that witnessed the biggest improvements over the study period. This suggests that the residents at least of this neighborhood type expressed similar views across cities prior to the onset of the treatment and that the threat of selection is hence substantially reduced, even though regression to the mean cannot not be ruled out completely.

The most plausible argument to sustain claims that progress made in implementing community policing was at least partially responsible for the reduction in fear of crime is provided by the spatial patterns of the observed reductions. In Zurich and Bern, it was the neighborhoods surrounding the city centers, where the police departments primarily focused their efforts, that saw the biggest drops in fear of crime. The fact that fear did not simultaneously fall in the urban centers themselves is not really a counter argument. First of all, the size of the 2005 survey samples in neighborhood type 1 was puny for both Zurich and Bern, making it impossible to find a statistical effect, yet the signs of the response to treatment point in the right direction for both these areas. Furthermore, as the neighborhood typology revealed, the urban centers are often the areas where residents declare themselves less afraid of crime. If people have low levels of fear of crime to begin with, it is harder still to improve matters in a statistically significant way. However, claims of an effect appear less plausible, if one also takes into account the differences on disorder. Perceived neighborhood disorder rose over the study period across those neighborhood areas that observed a drop in fear of crime. In other words, although fear dropped significantly, this did not go hand-in-hand with a drop in disorder.

In sum, the study provides fairly sound evidence that fear of crime dropped, especially in the cities of Bern and Zurich. Spatially, the perceived risk of a burglary tended to drop in the suburban neighborhood areas on the five cities' outskirts, whereas fewer respondents changed their behavior to avoid crime in the more centrally located neighborhoods. Regarding the impact of community policing, the study finds that in Bern and Zurich the changes in policing strategy preceded or coincided with the drop in fear of crime, which is robust to a moderate bias due to some unobserved covariate, and that the areas that witnessed the biggest fear reduction covary spatially with the locus of the bulk of the community policing interventions. The alternative hypothesis that the observed reductions in fear of crime were at least in part a result of community policing interventions thus appears at least as plausible as the null hypothesis of absolutely no effect, even if the observational design cannot completely rule out selection and regression to the mean as alternative explanations.

8.3 Limitations of the Study

Due to its *ex post facto* nature, the study could not measure community policing "outputs" but had to assess community policing implementation based on organizational reforms, the establishment of strategic partnerships, or declarations of intent. Moreover, the study could not assess the police internal aspects of the transition process such as how community policing implementation was received by police staff or whether police internal appraisal systems and promotion criteria have been reformed. This study makes no attempt to paper over the difficulties police departments encountered in implementing and motivating standing police officers for community policing reform. The interview partners were candid enough to admit that community policing at times met fierce resistance or even outright hostility from the rank and file. The present study could not investigate this further. However, a recent study that polled police personnel from all divisions and levels of hierarchy of the Geneva cantonal police concluded that even though for many of those interviewed community outreach was not necessarily their preferred cup of tea, a majority of them nevertheless were favorable to the idea of strengthening community policing at the Geneva police (Zorn 2008, 88f.).

The study also does not delve into such issues as the cooperation between the police and judicial authorities or special operations such as riot policing, which, to judge by the media coverage in recent years, has generated substantially more political interest than community policing. Finally, there are some factors outside the police's control that the study had to leave aside. For instance, the study could not interview representatives of the partner agencies of the police and hence cannot assess how they view the implementation of community policing. They may beg to differ with some of the assessments or the presentation of the chronology of events in this study. Likewise, the study could not touch the politics at the municipal or cantonal level that may have boosted or hindered the process of community policing adoption and implementation. Any scientific study is limited in scope and, for better or worse, these aspects must be left for future research.

The second set of limitations of the current study concern the availability, reliability, and construct validity of the data. The impact evaluation had to rely on existing data sources from standing survey instruments, which were not originally conceived for the present study. One limitation for instance concerns the study period, which is much longer for the process evaluation than for the impact evaluation, since the last sweep of the Swiss Crime Survey was conducted in 2005.[42] Individual police departments have since continued using survey instruments to gauge local levels of non-reported crime, fear, and popular satisfaction with the police. However, changing questionnaire formats (and supposedly different sampling methods) undermine the validity of these data for comparative evaluation.

For a comparative analysis of the official crime data, local police statistics first had to be harmonized. This process of data harmonization consisted of four separate steps: (i) import the different data formats into R; (ii) identify a series of criminal offenses that are numerically relevant and whose definition could conceivably be harmonized across the five cities; (iii) build a data base of the original values of these harmonized criminal offenses; and (iv) compute the standardized city-level and neighborhood-level crime rates for each type of criminal offense.

Absolute levels of criminal victimizations are obviously being lost through the data harmonization process. For the long-term trend analysis, the number of criminal infractions recorded was set at 100 for the first year for which data were available in order to track the development over the following years as percentages of the initial values. However, since absolute levels are missing, it is entirely possible that if a city shows a large relative increase in crime over the subsequent years, absolute levels of victimization per inhabitant may still be lower than in a city that experienced a smaller relative increase. Also for the spatial analysis, standardized crime rates were calculated as a multiple of the city average, so the same caveat applies to those data as well.

The fact that absolute levels of crime rates were missing could have affected the clustering procedure used to develop the typology of neighborhoods. Crime rates, it is recalled, had low importance as explanatory variables in the random forests classifier to typify Swiss urban neighborhoods. Would absolute levels of crime have fared better? While it is impossible to know for sure since the analysis cannot be done, the prima facie evidence produced by the current study suggests they would have not. First, at the neighborhood level, residents' collective subjective assessment of risk was misaligned with objective risk as measured by the official crime records. Second, at the city level over the critical 1998-2005 period, there was a noticeable convergence in the prevalent levels of fear of crime across the five cities to within each other's respective margins of error, except for burglary risk, which remained higher in Geneva and Lausanne than in the Swiss German cities. In other words, while at the neighborhood level the link between subjective and objective risk was tenuous to begin with, at the city level there has been less and less variance between cities of late that absolute level crime data could account for.

The survey data, coming from a national survey, did not suffer from the same defects of limited reliability. However, whereas crime levels were measured using eight different types of offences, the other three theoretical constructs of community policing outcome had to be operationalized with a limited number of variables. This problem was less acute for fear of crime, which was measured using three different questionnaire items, but more so for disorder and popular attitudes towards the police, which were operationalized by means of only one survey question (the indicators on

[42] The data from the most recent edition of the Swiss Crime Survey, polling for which took place in early summer of 2011, became available too late to be included in the current study.

physical and social disorder were recoded from a single questionnaire item). Measuring latent constructs by means of a single survey item poses a threat to the construct validity of the analysis. Reassuringly, over the crucial period between 2000 and 2005, the question and answer categories of the survey items used as indicators of the treatment outcome remained the same.

The third set of limitations concerns shortcomings in the data analysis, namely the risk of overfitting the survey data and the use of an inappropriate unit of analysis. When developing the neighborhood typology, the optimum number of neighborhood clusters was determined by a χ^2-independence test of the survey data. For this test, the data were weighted, but for each replication of the clustering procedure the complete survey sample was used to calculate the test statistics. Since any survey sample includes noise or sampling bias, the current procedure risks modeling the noise and thus to overfit the survey data.

Still, dealing with the risk of overfitting the survey data is not straightforward. Sampling bias is a perennial fact of all survey samples and must be corrected using weights. Bootstrapping survey respondents would require recalculating the weights for each iteration lest a biased sample alters the result, and this would get computationally rather intensive very quickly. In other words, the problem is being acknowledged but must be left for future research.

Finally, for an evaluation that aims to match comparable treatment and control areas, the choice of the appropriate spatial unit of analysis is a critical part of the research design (Rosenbaum 2002, 194). The current study had to deal with the frequent problem of varying units of analysis for different types of data. Whereas the census data are based on the OFS' "statistical neighborhoods" (census districts), the Swiss Crime Survey retained a respondent's ZIP code number and were thus regrouped by ZIP code districts. For the crime data, the unit of analysis varied even across the five urban areas. However, shifting boundaries of the units of analysis influence the results of the statistical analyses based on them, a phenomenon that is well-established in geography as the modifiable areal unit problem (MAUP; Openshaw 1984). If treatment impact is assessed based on the response to treatment of matched comparison areas, the MAUP risks affecting the inferences from the evaluation.

The issue of the appropriate unit of analysis has recently generated interest in the field of environmental criminology (Weisburd *et al.* 2009). The little evidence that is available so far suggests that the impact of varying shapes of the boundaries of the higher level aggregates are not inordinate, though. In their literature review, Bridenball & Jesilow (2008, 13) concluded that the basic link between an area's ecological characteristics such as crime rates or income levels and individual attitudes of local residents remained relatively stable for different higher-level spatial aggregates, which is consistent with Sampson *et al.*'s (2002, 446f.) findings that patterns of "social-ecological differentiation" of American communities are relatively robust across different units of analysis. A first indication that the potential effect of the MAUP on the current study may not be as dramatic as might be feared was provided by Figure 5.2. For Geneva and Lausanne, neighborhood burglary victimization rates could be calculated at the level of individual census tracts. These maps show that the crime rates for neighboring census tracts were for the most part rather comparable. In other words, aggregating census tract crime rates by ZIP code districts does not appear to distort the data that much.

9

Conclusion

The first conclusion to draw from the *process evaluation* is that all five police departments have made substantial progress over time towards the goal of community policing implementation. Over the last fifteen years, a paradigm shift has silently occurred in the policing of Switzerland's major urban areas. Whereas such notions as strategic partnerships, broken windows, or problem solving policing were anathema to all but the most progressive police forces as recently as 1995, a decade and a half later they have become commonplace. What is more, the implementation process for the most part has been a one-way street. Although progress has come by fits and starts, none of the five police departments has seen instances of serious backsliding along the way towards full implementation of community policing.

Second, a more detailed look at individual implementation histories reveals that there is no single best or ideal way to implement community policing, and each police department has moved forward in its own particular manner. No two track records look alike: while some departments like the Basel police made a formal transition to community policing and overhauled their organization if not exactly over night then still in a matter of a few years, others like the Zurich municipal police have been practicing many elements of community policing for years in all but name. Local events play a role; mergers of police agencies or personnel changes at a department's helm can slow down or halt reform efforts temporarily. At the end of the day, however, it appears that in the world of community policing all roads lead to Rome.

Third, even if there are no signs of major setbacks, this does by no means imply that community policing implementation is plain sailing. On the contrary, the evidence from this process evaluation suggests that determined leadership at the top is vital and that without it, the momentum of reform can quickly dissipate. This is perhaps best illustrated by the case of the *gendarmerie* of the Geneva police, which as long as Guy Baer served as its commander, used to lead the other police departments in Switzerland in community policing reform during the 1990s. After his retirement and a series of rapid successions at the top of the Geneva cantonal police, reform efforts stalled and the erstwhile leader was overtaken by its former pupils. More recently, with renewed continuity at their helm and a re-invigorated community policing program, the Geneva police are catching up with their peers in other major urban areas.

Finally, community policing is here to stay. The introduction of the federal examination that police recruits now take upon graduation from the local academy in order to obtain a federal diploma as police officer, which firmly established community policing in the curriculum for the basic training, marks a watershed. As the share of police officers that have studied the concept during basic training steadily increases over the coming years, internal resistance to reform is likely to dissipate. What is more, the demand for increased cooperation in matters of security is likely to grow stronger from outside the ranks of the police. This trend is evidenced by the long and growing list of partnerships that the five police departments have forged with other agencies of the local government and partners from private enterprise and civil society groups. The evidence so far suggests that such partnerships, once concluded, endure. Considering the cultural battles that first had to be fought before these partnerships were forged, that is by itself a remarkable achievement.

As Skogan (2006, 35) remarked: "If community policing is the police department's program, it may fail. Community policing must be the city's program." In Switzerland's five major urban areas, it is indeed increasingly the city's program.

The key insight from the *exploratory data analysis* of the indicators of the four theoretical constructs of community policing impact was the striking parallels between the spatio-temporal patterns across the five cities. Trends in official crime records showed an increase in threats and assaults, whereas crimes against property showed stagnant or even falling trend lines. Also the spatial distribution of crime showed noticeable parallels from one city to another. The same goes for the survey data which also exhibited very similar patterns between cities, at least within each of the two linguistic regions. Whereas fear of crime tended to drop over the crucial period between 1998 and 2005, especially in the three Swiss German cities, the neighborhoods worst afflicted tended to shift from the city centers to the outskirts of the five urban areas. Disorder appeared to rise across the board without any sign of major spatial displacement, and popular satisfaction with the police was either high to begin with or rising, with few spatial patterns readily discernible. Despite the apparent shortcomings in the measurement of the data, the exploratory spatio-temporal analysis has thus demonstrated clearly the importance of the spatial dimension for any more sophisticated analysis of the subject at hand. Most importantly, the exploratory analysis demonstrated that changes in the neighborhood context imply differences in the links between the outcome variables, underlining the need to take the ecological context into account, especially for an observational study of a policing intervention.

As a consequence, the current research applied geospatial data mining techniques to develop a *neighborhood typology* that attempts to do exactly that. These methods made it possible to take data on the demographic and socio-economic structure as well as the built environment of urban neighborhoods into account, which influence the survey response patterns of the indicators of community policing impact and may affect how neighborhood residents perceive different community policing strategies. This approach hence addressed some of the criticism that much recent criminological research in the area of crime prevention has been too focused on the question of "what works?" while neglecting the influence of contextual and environmental factors on the success of such initiatives (Williamson *et al.* 2006, 199f.).

Despite the shortcomings previously discussed, the neighborhood typology goes a long way in meeting the twin objectives of achieving a degree of homogeneity in the ecological data while accounting for a significant share of the variance in the survey data. The neighborhoods within each cluster thus resemble each other not only with regard to the contextual variables that might confound inferences about the effect of treatment but also with regard to the variables of treatment outcome at the onset of program implementation, which the latter is trying to influence. The typology thus increases the validity of an observational study by comparing the treatment response between matched areas within the same neighborhood type.

Apart from reducing the threat of selection bias for the impact evaluation, the neighborhood typology unveiled some interesting facts in its own right. First of all, the typology revealed that, collectively, neighborhood residents do a rather poor job at assessing the objective risk of crime. Whereas the objective risk of crime is comparatively higher in the urban centers and a few nearby areas, subjective assessments of risk tend to increase as one moves from the centers to the peripheries. The study thus found evidence that also in Swiss cities there are neighborhoods that display a phenomenon that has come to be known as the "reassurance gap" (Tuffin *et al.* 2006), meaning low or falling levels of actual victimization paired with above average levels of fear. By contrast, popular attitudes towards the police appear more attuned with reality. Popular assessments of whether the police do a good job of combating crime tend to covary spatially with actual victimization risk.

The second extra benefit of the neighborhood classification system is to highlight the key characteristics of each neighborhood type. The random forests variable importance plot is highly informative in this regard, indicating for each neighborhood class the most salient characteristics compared to the other categories. Needless to say, this opens avenues for targeted interventions, especially for a police department solidly committed to community policing.

Finally, the *impact analysis* suggests that fear of crime dropped over the study period between 2000 and 2005, particularly in the cities of Bern and Zurich. In these two cities, the relatively more affluent neighborhoods around the urban centers observed a drop in the levels of fear of crime that was statistically significant. Furthermore, these improvements appear fairly robust even in the face of bias due to some unobserved covariate that can never be ruled out in an observational study design. In Basel, Geneva, and Lausanne, the trends in fear of crime were less clear-cut. For these three cities, the observed difference is mixed for different neighborhood types. Only one area, the suburban communities surrounding the City of Geneva witnessed a significant improvement in the perceived risk of a burglary victimization, but this effect is not robust against potential bias introduced by some unobserved covariate.

Regarding disorder, the study finds signs of an increase, especially in the cities of Basel, Bern, and Geneva. In Lausanne and Zurich, the emerging picture is more mixed with different neighborhood types experiencing both rising and falling levels of perceived physical and social disorder.

Finally, regarding public attitudes towards the police, the data suggest that survey respondents were more favorably disposed to their police, most notably in Basel, Geneva, and Zurich. In these three cities, residents in all neighborhood types rated the police as more effective at the end of the study period than five years earlier, but none of the changes is statistically significant. In Bern and Lausanne, by contrast, local residents deemed the police less effective in 2005 than in 2000 in a majority of neighborhood types.

Does community policing work? In this era of "evidence-based" policy that is the question. Marshalling the evidence from the process and impact evaluation, the study finds that the reductions in fear of crime covary both temporarily and spatially with progress made in the implementation of community policing in the cities of Bern and Zurich. These improvements are fairly robust in the face of potential bias due to some unobserved covariate. It thus appears plausible that the observed reductions in fear of crime are at least in part a result of community policing interventions, even if the observational design cannot completely rule out selection and regression to the mean as alternative explanations.

9.1 Implications for Policy

Keep up the pace of reform. Whereas the past decade was one of experimentation and the gradual adoption of community policing, the next decade must become a period of consolidation. According to the current study, the area most in need of progress is training, which is still lagging beyond the basic training stage at the police academy. Continued training of standing police personnel, especially the highly exposed neighborhood liaison officers and police cadres, should – resources permitting – be stepped up.

Decentralize the police organization. The second recommendation is a confirmation of the perennial demand of the community policing literature for police departments to decentralize their organization. The current study revealed that also in the five major urban areas in Switzerland, the socio-economic dynamics of different parts of a city vary noticeably, and the problems the police face change accordingly. Police need to fine tune their strategies to the characteristics of the neighborhoods in which they operate. This call for the decentralization of police organizations says nothing about the desirability of neighborhood police stations or the spatial organization of the police. Decentralization paired with long-term assignments to police beats will simply make it easier for individual officers to become familiar with and better understand their turf.

Exchange best practices between cities. A related policy implication of this research is that individual police departments may benefit from increased cooperation at the regional level, especially between police officials responsible for community outreach in comparable neighborhood types. As the socio-dynamic processes appear to repeat themselves across cities and the problems these officials face in their respective areas of activity are similar from one city to another, regular get-togethers to exchange experiences and best practices could be of mutual interest.

9.2 Implications for Research

Extend the analysis to include the data from the 2011 Swiss Crime Survey. In 2011, the latest sweep of the Swiss Crime Survey was conducted. Alas, the data of this most recent edition of the survey became available too late to be included in the current study. As a consequence, the study period of the impact evaluation is considerably shorter than the process evaluation. Whereas the process evaluation covered the years from the early 1990s to 2010, the impact evaluation could so far only examine the 2000-2005 period.

It thus only makes sense to extend the current study to include the latest survey data as well and conduct the impact analysis over the 2000-2011 period. This extended impact analysis could first be done relying on the neighborhood typology developed on the basis of the survey and census data from the year 2000. In a second step, however, it would also be interesting to redo the clustering procedure on the basis of the 2005 survey data, and use it as the baseline for the subsequent impact evaluation of the 2005-2011 period.

Such an analysis, supposedly, would not only provide insights into whether the community policing benefits of reduced fear of crime achieved over the 2000-2005 period could be sustained over the longer run. It would also constitute an important step towards the development of a monitoring system that allowed criminologists and police practitioners to make better use of regular public surveys to formulate community-oriented policing strategies and to assess their effectiveness on a permanent basis.

Refine the geospatial clustering procedure to match neighborhood areas. The implications of the neighborhood typology for future research flow directly from its two shortcomings. First, the clustering procedure, due to the conflicting optimization objectives, did not produce a unique solution. Since the results of the subsequent impact analysis depend on its results, however, methods that achieve a higher degree of reliability of the resulting neighborhood typology should be used. For instance, consensus clustering methods (Monti *et al.* 2003), which analyze the degree of consensus over multiple runs of clustering procedures that depend on randomized initial values such as the SOM, could be employed.

Second, the neighborhood typology could not account for all the variance at the higher levels of aggregation. The second, more radical proposition is thus that future research be carried out at a finer scale of spatial granularity than the large units of postal ZIP code or administrative districts. Ideally, such an analysis would do away completely with administrative boundaries, which are always liable to affect the results of a study due to the aggregation effects they induce. In a study for which all data are available in raster format, larger units of analysis with fluid boundaries could be created by aggregating individual raster cells depending on the results of the clustering procedure. Such a research design would combine several advantages: first of all, by using smaller, more homogeneous units of analysis the chances of finding a significant effect of treatment would increase. As Sherman (Sherman 2010, 603f.) noted, evaluations of policing intervention that found significant effects mostly used smaller units of analysis such as places or hotspots rather than larger patrol beats or neighborhoods. Second, if such "neighborhoods" were aggregated by flexible geometry according to a study's requirements, it should be possible to match treatment and control areas that are similar enough to allow valid comparisons. This would address the potent threat of selection in an observational study, which the neighborhood typology developed for the current study could not achieve. Finally, such an approach would also elegantly bypass the conceptual question of where to draw neighborhood boundaries, the definition of which remains an open research question (cf. Sampson *et al.* 2002, 470).

Invest in high-quality observational study designs. An observational study is the empirical analysis of the effects of a treatment "when randomized experimentation is unethical or infeasible." (Rosenbaum 2010, ix). There are at least three reasons why high-quality observational designs deserve to be held in higher regard in police research than they are today. First, in the fields of crime prevention and policing, randomized designs have seldom been implemented, because field experiments are perceived as politically risky, or ethically questionable, or both. The need for valid alternative methods of program evaluation is particularly acute regarding community policing. Advocates have never denied that its extraordinary variability and flexibility make community policing hard to evaluate and inherently unsuited for randomized experiments and have long pressed for a special evaluation agenda (e.g. Kennedy & Moore 1995). Observational designs, by contrast, are easier to carry out. Moreover, as data bases are being built up in this era of "big data", possibilities for high-quality observational study designs are likely to grow over the coming years, whereas randomization controlled trials in policing will always depend on the political will of police administrators and their political masters and hence be much more difficult to implement.

Second, community policing takes years to implement. As the process evaluation unveiled, the five Swiss police departments took ten to fifteen years to reform their organization, institutionalize training, and establish the necessary partnerships with local government. Sustaining an experimental research design over such a long period of time would be a tall order, yet the police literature often alludes to the "glacial" pace of police strategic reforms, the impact of which may take years to materialize (Greene 2000, 309). Furthermore, experience shows that community policing is inevitably scaled up to the city-level over such long time-periods, precluding experimental designs beyond the micro- and meso-levels. In other words, there is a need for high-quality observational studies to assess what police departments actually do.

Third, by combining a statistical test with a sensitivity analysis, covariance between treatment and outcome can be established with reasonable confidence in an observational design as well. If the current study permitted only tentative inferences about a treatment effect, this was in large part due to the fact that (i) treatment had to be assessed *ex post*, (ii) the process evaluation could thus only capture changes at the organizational and strategic level but not measure actual outputs, and (iii) at this highly abstract level the treatment showed relatively little variance between different cities to boot. In an observational design, in which the treatment output is measured contemporaneously with the data on treatment outcome, and there are larger implementation gaps between the matched treatment and control areas, stronger inferences about a treatment effect become possible.

Bibliography

Arend, Michal. 2007. *Integration und Quartierentwicklung in mittelgrossen und kleineren Städten der Schweiz*. Zollikon: Bundesamt für Migration/Bundesamt für Wohnungswesen.

Arend, Michal, Baur, Martin, & Schuler, Martin. 2005. *Bevölkerungszusammensetzung, Integration und Ausgrenzeng in urbanen Zonen*. Neuchâtel: Bundesamt für Statistik.

Ashby, David I. 2005. Policing neighbourhoods: Exploring the geographies of crime, policing and performance assessment. *Policing and Society*, **15**(4), 413–447.

Ashby, David I., & Longley, Paul A. 2005. Geocomputation, geodemographics and resource allocation for local policing. *Transactions in GIS*, **9**(1), 53–72.

Atkins, Stephen, Husain, Sohail, & Storey, Angele. 1991. *The Influence of Street Lighting on Crime and Fear of Crime*. Crime Prevention Unit Paper, vol. 28. London: Home Office.

Baer, Guy (Commandant de la gendarmerie). 1993. PEGASE: Redéfinition des missions prioritaires et nouvelle organisation de la gendarmerie. *Le Journal: publication trimestrielle de la Police genevoise*, **26**, 3–6.

Bação, Fernando, Lobo, Victor, & Painho, Marco. 2008. Applications of different self-organizing map variants to geographical information science problems. *Pages 21–44 of:* Agarwal, Pragya, & Skupin, André (eds), *Self-Organising Maps: Applications in Geographic Information Science*. Chichester: Wiley.

Bartholomew, David J., Steele, Fiona, Moustaki, Irini, & Galbraith, Jane I. 2008. *Analysis of Multivariate Social Science Data*. 2nd edn. Statistics in the Social and Behavioral Sciences Series. Boca Raton: CRC Press.

Bennett, Trevor. 1991. The effectiveness of a police-initiated fear-reducing strategy. *British Journal of Criminology, Delinquency and Deviant Social Behaviour*, **31**(1), 1–14.

Bennett, Trevor, Holloway, Katy, & Farrington, David. 2006. Does neighborhood watch reduce crime? A systematic review and meta-analysis. *Journal of Experimental Criminology*, **2**(4), 437–458.

Bennett, Trevor, Holloway, Katy, & Farrington, David. 2008. *The Effectiveness of Neighborhood Watch*. Campbell Systematic Reviews. Oslo: The Campbell Collaboration.

Bivand, Roger S., Pebesma, Edzer J., & Gómez-Rubio, Virgilio. 2008. *Applied Spatial Data Analysis with R*. New York: Springer.

Bolle, Pierre-Henri, & Knoepfler, Julien. 2000a. *Eléments pour un bilan de connaissances sur les expériences de police de proximité en Suisse. Bilan et perspectives au 31 décembre 1999. Sur mandat du Ministère français de l'Intérieur par l'Institut des Hautes Etudes de la Sécurité Intérieure (IHESI), à Paris.* Polices de proximité en Europe Occidentale : Bilans de connaissances en Allemagne, Angleterre, Belgique, Espagne, Pays-Bas, Suisse. Paris: Institut des Hautes Etudes de la Sécurité Intérieure.

Bolle, Pierre-Henri, & Knoepfler, Julien. 2000b. La police de proximité en Suisse: Cinq modèles pour une définition. *Pages 103–122 of:* Institut des Hautes Etudes de la Sécurité Intérieure (ed), *Les dilemmes de la proximité*. Les Cahiers de la sécurité intérieure, vol. 39. Paris: IHESI.

Boruch, Robert F. 1975. On common contentions about randomized field experiments. *In:* Boruch, Robert F., & Reicken, Henry L. (eds), *Experimental Testing of Public Policy: The Proceedings of the 1974 Social Sciences Research Council Conference on Social Experimentation*. Boulder: Westview Press.

Bowers, William, & Hirsch, Jon H. 1987. The impact of foot patrol staffing on crime and disorder in Boston: An unmet promise. *American Journal of Policing*, **6**(1), 17–44.

Braga, Anthony A. 2001. The effects of hot spots policing on crime. *The ANNALS of the American Academy of Political and Social Science*, **578**(1), 104–125.

Braga, Anthony A. 2005. Hot spots policing and crime prevention: A systematic review of randomized controlled trials. *Journal of Experimental Criminology*, **1**(3), 317–342.

Braga, Anthony A. 2007. *The Effects of Hot Spots Policing on Crime*. Campbell Systematic Reviews. Oslo: The Campbell Collaboration.

Braga, Anthony A. 2010. Setting a higher standard for the evaluation of problem-oriented policing initiatives. *Criminology and Public Policy*, **9**(1), 173–182.

Braga, Anthony A., & Bond, Brenda J. 2008. Policing crime and disorder hot spots: A randomized controlled trial. *Criminology*, **46**(3), 577–607.

Braga, Anthony A., Weisburd, David L., Waring, Elin J., Mazerolle, Lorraine Green, Spelman, William, & Gajewski, Francis. 1999. Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology*, **37**(3), 541–580.

Brantingham, Paul, & Brantingham, Patricia. 1991. *Environmental Criminology*. 2nd edn. Prospect Heights: Waveland Press.

Breiman, Leo. 2001. Random forests. *Machine Learning*, **45**(1), 5–32.

Breiman, Leo, Friedman, Jerome H., Olshen, Richard A., & Stone, Charles J. 1984. *Classification and Regression Trees*. New York: Chapman & Hall.

Bridenball, Blaine, & Jesilow, Paul. 2008. What matters: The formation of attitudes toward the police. *Police Quarterly*, **11**(2), 151–181.

Brown, Barbara B., Perkins, Douglas D., & Brown, Graham. 2004. Crime, new housing, and housing incivilities in a first-ring suburb: Multilevel relationships across time. *Housing Policy Debate*, **15**(2), 301–345.

Bässmann, Jörg, & Vogt, Sabine. 1997. *Community Policing: Projektbericht des Bundeskriminalamtes zu den Erfahrungen in den USA*. Bundeskriminalamt Wiesbaden.

Bureau of Justice Assistance. 1994. *Understanding Community Policing: A Framework for Action*. NIJ Monograph, vol. 148457. Washington, DC: U.S. Department of Justice, Bureau of Justice Assistance.

Bursik, Robert J. 1988. Social disorganization and theories of crime and delinquency - problems and prospects. *Criminology*, **26**(4), 519–551.

Campbell, Donald T., & Stanley, Julian C. 1963. *Experimental and Quasi-experimental Designs for Research*. Boston: Houghton Mifflin Company.

Cao, Liqun, Frank, James, & Cullen, Francis T. 1996. Race, community context, and confidence in the police. *American Journal of Police*, **15**(1), 3–22.

Cerdá, Magdalena, Tracy, Melissa, Messner, Steven F., Vlahov, David, Tardiff, Kenneth, & Galea, Sandro. 2009. Misdemeanor policing, physical disorder, and gun-related homicide: A spatial analytic test of Broken-Windows Theory. *Epidemiology*, **20**(4), 533–541.

Clarke, Ronald V. G. 1980. "Situational" crime prevention: Theory and practice. *British Journal of Criminology*, **20**, 136–47.

Clarke, Ronald V. G., & Cornish, Derek Blaikie. 1986. *The Reasoning Criminal: Rational Choice Perspectives on Offending.* Research in criminology. New York: Springer.

Clerici, Christian, & Killias, Martin. 1999a. Unsicherheit im öffentlichen Raum: Eine Folge der Kriminalität? *Crimiscope*, **6**, 1–9.

Clerici, Christian, & Killias, Martin. 1999b. Zum Bild der Polizei in der Öffentlichkeit. *Crimiscope*, **5**, 1–7.

Cochran, William G. 1965. The planning of observational studies of human populations. *Journal of the Royal Statistical Society. Series A (General)*, **128**(2), 234–266.

Cohen, Lawrence E., & Felson, Marcus. 1979. Social change and crime rate trends: A routine activity approach. *American Sociological Review*, **44**(4), 588–608.

Cook, Thomas D., & Campbell, Donald T. 1979. *Quasi-experimentation: Design and Analysis Issues for Field Settings.* Chicago: Rand McNally College Publishing Company.

Cordner, Gary. 2010. *Reducing Fear of Crime: Strategies for Police.* Washington, DC: U.S. Department of Justice, Office of Community Oriented Policing Services.

Corman, Hope, & Mocan, Naci. 2002. *Carrots, Sticks and Broken Windows.* National Bureau of Economic Research Working Paper Series, vol. 9061. Cambridge: National Bureau of Economic Research.

Crawley, Michael J. 2007. *The R Book.* Chichester: Wiley.

Crawley, Peter, & Hope, Tim. 2003. Measuring the outputs of crime reduction projects. *Community Safety Journal*, **2**(3), 4–12.

Dalgleish, David, & Myhill, Andy. 2004. *Reassuring the Public: A Review of International Policing Interventions.* Home Office Research Study, vol. 284. London: Home Office.

Davies, David L., & Bouldin, Donald W. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **1**(2), 224–227.

Dessemontet, Pierre, Schuler, Martin, & Walser, Olivier. 2008. Eine Typologie der Quartiere der Schweizer Städte / Une typologie des quartiers des villes suisses. *In: Journées suisses de la statistique publique, Davos, 10.09.08.*

Ditton, Jason, & Farrall, Stephen. 2000. *The Fear of Crime.* The International Library of Criminology, Criminal Justice and Penology. Aldershot: Ashgate.

DuBow, F, McCabe, E, & Kaplan, G. 1979. *Reactions to crime: A critical review of the literature.* Washington, DC: Government Printing Office.

Eck, John E. 2002. Preventing crime at places. *Pages 241–294 of:* Sherman, Lawrence W., Farrington, David P., Welsh, Brandon C., & MacKenzie, Doris Layton (eds), *Evidence-Based Crime Prevention.* New York: Routledge.

Eck, John E., & Rosenbaum, Dennis P. 1994. The new police order: Effectiveness, equity, and efficiency in community policing. *Pages 3–23 of:* Rosenbaum, Dennis P. (ed), *The Challenge of Community Policing: Testing the Promises*. Thousand Oaks: Sage Publications.

Eisner, Manuel, & Manzoni, Patrik. 2000. *Lebensqualität und Sicherheit im Wohnquartier*. Chur, Zürich: Rüegger.

Ernst, Marie-Louise, Aeschbacher, Monique, & Spencer, Brenda. 1999a. *Gesundheitliche Aspekte der Polizeiarbeit im Drogenbereich*. Raisons de santé 37a. Université de Lausanne: Institut universitaire de médecine sociale et préventive.

Ernst, Marie-Louise, Aeschbacher, Monique, & Spencer, Brenda. 1999b. *Les aspects sanitaires du travail de la police avec les consommateurs de drogues*. Raisons de santé 37a. Université de Lausanne: Institut universitaire de médecine sociale et préventive.

Farrall, Stephen, & Gadd, David. 2004. Evaluating crime fears: A research note on a pilot study to improve the measurement of the 'fear of crime' as a performance indicator. *Evaluation*, **10**(4), 493–502.

Farrington, David P. 2003. Methodological quality standards for evaluation research. *The ANNALS of the American Academy of Political and Social Science*, **587**(1), 49–68.

Farrington, David P., & Petrosino, Anthony. 2001. The Campbell Collaboration Crime and Justice Group. *The ANNALS of the American Academy of Political and Social Science*, **578**, 35–49.

Farrington, David P., Gottfredson, Denise C., Sherman, Lawrence, & Welsh, Brandon C. 2002. The Maryland Scientific Methods Scale. *Pages 13–21 of:* Sherman, Lawrence W., Farrington, David P., Welsh, Brandon C., & MacKenzie, Doris Layton (eds), *Evidence-Based Crime Prevention*. New York: Routledge.

Fielding, Nigel, & Innes, Martin. 2006. Reassurance policing, community policing and measuring police performance. *Policing and Society*, **16**(2), 127–145.

Fridell, Lorie. 2004. The defining characteristics of community policing. *Pages 3–12 of:* Fridell, Lorie, & Wycoff, Mary Ann (eds), *Community Policing: The Past, Present, and Future*. Washington, DC: Annie E. Casey Foundation and Police Executive Research Forum.

Furstenberg, Frank F., Jr. 1971. Public reaction to crime in the streets. *The American Scholar*, **40**(4), 601–610.

Garofalo, James. 1979. Victimization and the fear of crime. *Journal of Research in Crime and Delinquency*, **16**(1), 80–97.

Garofalo, James, & Laub, John H. 1978. Fear of crime: Broadening our perspective. *Victimology*, **3**, 242–253.

Gau, Jacinta M., & Pratt, Travis C. 2010. Revisiting broken windows theory: Examining the sources of the discriminant validity of perceived disorder and crime. *Journal of Criminal Justice*, **38**(4), 758–766.

Genuer, Robin, Poggi, Jean-Michel, & Tuleau-Malot, Christine. 2010. Variable selection using random forests. *Pattern Recognition Letters*, **31**(14), 2225–2236.

Gianakis, Gerasimos A., & Davis, G. John. 1998. Reinventing or repackaging public services? The case of community-oriented policing. *Public Administration Review*, **58**(6), 485–498.

Goldstein, Herman. 1977. *Policing a Free Society*. Cambridge: Ballinger.

Goldstein, Herman. 1979. Improving policing: A problem-oriented approach. *Crime & Delinquency*, **25**(2), 236–258.

Goldstein, Herman. 1990. *Problem-Oriented Policing.* New York: McGraw-Hill.

Greene, Jack R. 1998. Evaluating planned change strategies in modern law enforcement: Implementing community-based policing. *Pages 141–160 of:* Brodeur, Jean-Paul (ed), *How to Recognize Good Policing: Problems and Issues.* Washington, DC; Thousand Oaks: Police Executive Research Forum and Sage Publications.

Greene, Jack R. 2000. Community policing in America: Changing the nature, structure, and function of the police. *Pages 299–370 of:* Horney, Julie (ed), *Policies, Processes, and Decisions of the Criminal Justice System.* Criminal Justice 2000, vol. 3. Washington, DC: U.S. Department of Justice, National Institute of Justice.

Greene, Jack R., & Mastrofski, Stephen D. 1988. *Community Policing: Rhetoric or Reality?* New York: Praeger.

Greene, Jack R., & Taylor, Ralph B. 1988. Community-based policing and foot patrol. Issues of theory and evaluation. *Pages 195–223 of:* Greene, Jack R., & Mastrofski, Stephen D. (eds), *Community Policing: Rhetoric or Reality.* New York: Praeger.

Guerette, Rob T., & Bowers, Kate J. 2009. Assessing the extent of crime displacement and diffusion of benefits: A review of situational crime prevention evaluations. *Criminology*, **47**(4), 1331–1368.

Hagenlocher, Gérald (Commandant de la Police municipale de Lausanne). 1996. *Police de proximité: l'approche lausannoise.* Polycopié interne à la Police municipale.

Hale, C. 1996. Fear of crime: A review of the literature. *International Review of Victimology*, **4**(2), 79–150.

Hansen, Ben B. 2007. Optmatch: Flexible, optimal matching for observational studies. *R News*, **7**(2), 18–24.

Harcourt, Bernard E. 1998. Reflecting on the subject: A critique of the social influence conception of deterrence, the broken windows theory, and order-maintenance policing New York style. *Michigan Law Review*, **97**(2), 291–389.

Harcourt, Bernard E., & Ludwig, Jens. 2006. Broken windows: New evidence from New York City and a five-city social experiment. *The University of Chicago Law Review*, **73**(1), 271.

Harper, Gareth, Williamson, Ian, Clarke, Graham, & See, Linda. 2002. *Family Origins: Developing Groups of Crime and Disorder Reduction Partnerships and Police Basic Command Units for Comparative Purposes.* London: Home Office.

Haymoz, Sandrine, Aebi, Marcelo F., Killias, Martin, & Lamon, Philippe. 2009. Comparisons of crime data based on general population surveys with criminal statistics in Switzerland. *Pages 127–140 of:* Robert, Philippe (ed), *Comparing Crime Data in Europe: Official Crime Statistics and Survey Based Data.* Criminological Studies. Brussels: VUBPRESS Brussels University Press.

Henig, Jeffrey, & Maxfield, Michael G. 1978. Reducing fear of crime: Strategies for intervention. *Victimology*, **3**(3/4), 297–313.

Heye, Corinna, & Leuthold, Heiri. 2006. *Segregation und Umzüge in der Stadt und Agglomeration Zürich.* Stadt Zürich: Präsidialdepartement, Statistik Stadt Zürich.

Hinkle, Joshua. C., & Weisburd, David. 2008. The irony of broken windows policing: A micro-place study of the relationship between disorder, focused police crackdowns and fear of crime. *Journal of Criminal Justice*, **36**(6), 503–512.

Hope, Tim, & Hough, Mike. 1988. Area crime and incivility: A profile from the British Crime Survey. *Pages 30–47 of:* Shaw, Margaret, Hope, Tim, Research, Great Britain Home Office, & Unit, Planning (eds), *Communities and Crime Reduction*. London: Her Majesty's Stationery Office.

Hope, Tim, Bryan, Jane, Crawley, Elaine, Crawley, Peter, Russell, Noel, & Trickett, Alan. 2004. *Strategic Development Projects in the Yorkshire and the Humber, East Midlands and Eastern regions*. Home Office Online Report 41/04. London: Home Office.

Hunter, Albert. 1978. Symbols of incivility. *In: Paper presented to the annual meeting of the American Society of Criminology*.

Innes, Martin. 2004. Signal crimes and signal disorders: notes on deviance as communicative action. *The British Journal of Sociology*, **55**(3), 335–355.

Jackson, Jonathan, Bradford, Ben, Hohl, Katrin, & Farrall, Stephen. 2009. Does the fear of crime erode public confidence in policing? *Policing*, **3**(1), 100–111.

Jacobs, Jane. 1961. *The Death and Life of Great American Cities*. New York: Random House.

Jang, Sung Joon, & Johnson, Byron R. 2001. Neighborhood disorder, individual religiosity, and adolescent use of illicit drugs: A test of multilevel hypotheses. *Criminology*, **39**(1), 109–144.

Kanevski, Mikhail, Pozdnoukhov, Alexei, & Timonin, Vadim. 2009. *Machine Learning for Spatial Environmental Data: Theory, Applications and Software*. Lausanne: EPFL Press.

Kansas City Police Department. 1977. *Response Time Analysis*. Kansas City: Kansas City Police Department.

Kantonspolizei Basel-Stadt. 1998. *Community Policing*. [Basel]: [Kantonspolizei Basel-Stadt].

Kantonspolizei Basel-Stadt. 1999. *Die EMRK und der Polizeidienst: Ein Unterrichts-Hilfsmittel und Behelf für den praktischen Dienst*. Basel: Kantonspolizei Basel-Stadt.

Kantonspolizei Basel-Stadt. 2003. *Community Policing*. Basel: Kantonspolizei Basel-Stadt.

Keizer, Kees, Lindenberg, Siegwart, & Steg, Linda. 2008. The spreading of disorder. *Science*, **322**(5908), 1681–1685.

Kelling, George L., & Coles, Catherine M. 1996. *Fixing Broken Windows: Restoring Order and Reducing Crime in Our Communities*. New York: Martin Kessler Books.

Kelling, George L., & Sousa, William H. 2001. *Do police matter? An analysis of the impact of New York City's Police Reforms*. Civic Report, vol. 22. New York: Manhattan Institute for Policy Research.

Kelling, George L., Pate, Tony, Dieckman, Duane, & Brown, Charles. 1974. *The Kansas City Preventive Patrol Experiment: A Summary Report*. Washington, DC: Police Foundation.

Kennedy, David. M., & Moore, Mark. H. 1995. Underwriting the risky investment in community policing: What social science should be doing to evaluate community policing. *Justice System Journal*, **17**(3), 271–289.

Killias, Martin. 2003. Objektive und subjektive Sicherheit in Städten. *Pages 7–24 of:* Konferenz der Städtischen Polizeidirektorinnen und Polizeidirektoren and Stadt Zürich (ed), *Urbane Sicherheit im 21. Jahrhundert: Referate und Podiumsdiskussion der Tagung vom 12. September 2003.* Schriftenreihe der KSPD. Zürich: Schulthess.

Killias, Martin, Chevalier, Claude, & Kuhn, André. 1989. *Les Suisses face au crime: leurs expériences et attitudes à la lumière des enquêtes suisses de victimisation.* Collection Criminologie, vol. 5. Grüsch: Rüegger.

Killias, Martin, Haymoz, Sandrine, & Lamon, Philippe. 2007. *Swiss Crime Survey: Die Kriminalität in der Schweiz im Lichte der Opferbefragungen von 1985 bis 2005.* Bern: Stämpfli.

Killias, Martin, Kuhn, André, & Aebi, Marcelo F. 2011. *Grundriss der Kriminologie: Eine europäische Perspektive.* 2nd edn. Bern: Stämpfli.

Killias, Martin, Aebi, Marcelo Fernando, & Kuhn, André. 2012. *Précis de criminologie.* 3e edn. Berne: Stämpfli.

Kohonen, Teuvo. 1990. The self-organizing map. *Proceedings of the IEEE*, **78**(9), 1464–1480.

Kohonen, Teuvo. 2001. *Self-Organizing Maps.* 3rd edn. Berlin: Springer.

Kreis, Christian. 1999. *The Impact of Community Policing on Neighborhood Residents: An Evaluation Study of the Community Policing Strategy of the Police in Portland, Oregon (USA).* Unpublished M.A. Thesis. Universtiät Zürich: Institut für Politikwissenschaft.

Kreis, Christian. 2009. *The Spatio-Temporal Patterns of Fear of Crime in Major Swiss Cities: A quantitative analysis of the spatial distribution of fear of crime, disorder, and public satisfaction with the police among neighborhood residents in Switzerland's five metropolitan areas between 1987 and 2005.* Travail de mémoire, Master en études avancées en urbanisme durable. Université de Lausanne: Institut de géographie.

Kreis, Christian, Kuhn, André, & Killias, Martin. 2010. *Community Policing in Switzerland: A Comparative Evaluation Study of the Community Policing Strategies of the Police in Major Swiss Urban Areas.* SNSF Research Project no. 100015-122463 Final Report. Université de Lausanne: Institut de criminologie et de droit pénal.

Kuhn, André, & Viredaz, Baptiste. 2004. *Comment lutter contre le sentiment d'insécurité à Lausanne ?* Rapport élaboré sur la base du sondage 2003 de l'Observatoire de la sécurité de la Ville de Lausanne. Université de Lausanne: Institut de criminologie et de droit pénal.

Kuhn, André, & Viredaz, Baptiste. 2007a. *Sondage ≪convivialité≫ (2007).* Observatoire de la sécurité de la Ville de Lausanne - Compte rendu des résultats. Université de Lausanne: Institut de criminologie et de droit pénal.

Kuhn, André, & Viredaz, Baptiste. 2007b. *Sondage ≪street lighting≫ (2006-2007).* Observatoire de la sécurité de la Ville de Lausanne - Compte rendu des résultats. Université de Lausanne: Institut de criminologie et de droit pénal.

Kuhn-Roux, Viviane, & Kuhn, André. 1994. L'image de la police en Valais. *Bulletin de criminologie*, **20**(2), 41–59.

LaGrange, Randy L., & Ferraro, Kenneth F. 1989. Assessing age and gender differences in perceived risk and fear of crime. *Criminology*, **27**(4), 697–719.

Lee, John A., & Verleysen, Michel. 2007. *Nonlinear Dimensionality Reduction.* New York: Springer.

Lewis, Dan A., & Maxfield, Michael G. 1980. Fear in the neighborhoods: An investigation of the impact of crime. *Journal of Research in Crime and Delinquency*, **17**(2), 160–189.

Lewis, Dan A., & Salem, Greta. 1986. *Fear of Crime: Incivility and the Production of a Social Problem*. New Brunswick: Transaction Books.

Li, Yan, & Shanmuganathan, Subana. 2007. *Social Area Analysis Using SOM and GIS: A Preliminary Research*. RCAPS Working Paper, vol 07-3. Ritsumeikan Asia Pacific University, Ritsumeikan Center for Asia Pacific Studies.

Liaw, Andy, & Wiener, Matthew. 2002. Classification and regression by randomForest. *R News*, **2**(3), 18–22.

Lipsey, Mark, Petrie, Carol, Weisburd, David, & Gottfredson, Denise. 2006. Improving evaluation of anti-crime programs: Summary of a National Research Council report. *Journal of Experimental Criminology*, **2**(3), 271–307.

Lum, Cynthia, Koper, Christopher, & Telep, Cody. 2011. The evidence-based policing matrix. *Journal of Experimental Criminology*, **7**(1), 3–26.

Lumley, Thomas. 2010. *Complex Surveys: A Guide to Analysis Using R*. Hoboken: Wiley.

Lurigio, Arthur J., & Rosenbaum, Dennis P. 1986. Evaluation research in community crime prevention: A critical look at the field. *Pages 19–44 of:* Rosenbaum, Dennis P. (ed), *Community Crime Prevention: Does It Work?* Sage Criminal Justice System Annuals, vol. 22. Beverly Hills: Sage Publications.

Lurigio, Arthur J., & Rosenbaum, Dennis P. 1994. The impact of community policing on police personnel: A review of the literature. *Pages 147–163 of:* Rosenbaum, Dennis P. (ed), *The Challenge of Community Policing: Testing the Promises*. Thousand Oaks: Sage Publications.

Maguire, Edward, & Gantley, Megan. 2009a. Decentralization and geographic accountability. *Pages 35–43 of:* Maguire, Edward, & Wells, William (eds), *Implementing Community Policing: Lessons from 12 Agencies*. Washington, DC: U.S. Department of Justice, Office of Community Oriented Policing Services.

Maguire, Edward, & Gantley, Megan. 2009b. Specialist and generalist models. *Pages 45–55 of:* Maguire, Edward, & Wells, William (eds), *Implementing Community Policing: Lessons from 12 Agencies*. Washington, DC: U.S. Department of Justice, Office of Community Oriented Policing Services.

Maguire, Edward, & Wells, William. 2009. *Implementing Community Policing: Lessons from 12 Agencies*. Washington, DC: U.S. Department of Justice, Office of Community Oriented Policing Services.

Maguire, Edward R. 1997. Structural change in large municipal police organizations during the community policing era. *Justice Quarterly*, **14**(3), 547–576.

Maltz, Michael D. 1972. *Evaluation of Crime Control Programs*. Washington, DC: U.S. Department of Justice, National Institute of Law Enforcement and Criminal Justice.

Manzoni, Patrik, & Thalmann, Urs. 2008. Crime Mapping als Instrument der Einbruchsprävention. *Pages 155–172 of:* Schwarzenegger, Christian, & Müller, Jürg (eds), *Erstes Zürcher Präventionsforum: Kommunale Kriminalprävention, Crime Mapping, Einbruchskriminalität*. Zürich, Basel, Genf: Schulthess.

Mason, Mark. 2009. *Findings from the Second Year of the National Neighbourhood Policing Programme Evaluation*. Research Report 14. London: Home Office.

Mastrofski, Stephen D. 2006. Community policing: A skeptical view. *Pages 44–73 of:* Weisburd, David, & Braga, Anthony A. (eds), *Police Innovation: Contrasting Perspectives*. Cambridge Studies in Criminology. Cambridge, MA: Cambridge University Press.

Mastrofski, Stephen D., Snipes, Jeffrey B., & Supina, Anne E. 1996. Compliance on demand: The public response to specific police requests. *Journal of Research in Crime and Delinquency*, **33**(3), 269–305.

Mastrofski, Stephen D., Reisig, Michael D., & McCluskey, John D. 2002. Police disrespect toward the public: An encounter-based analysis. *Criminology*, **40**(3), 519–552.

Matthews, Roger. 1992. Replacing "Broken Windows": Crime, incivilities and urban change. *Pages 19–50 of:* Matthews, Roger, & Young, Jock (eds), *Issues in Realist Criminology*. Sage contemporary criminology. London: Sage Publications.

Maxfield, Michael G. 1988. The London Metropolitan Police and their clients: Victim and suspect attitudes. *Journal of Research in Crime and Delinquency*, **25**(2), 188–206.

Mazerolle, Lorraine, Soole, David, & Rombouts, Sacha. 2006. Street-level drug law enforcement: A meta-analytical review. *Journal of Experimental Criminology*, **2**(4), 409–435.

Mazerolle, Lorraine, Soole, David, & Rombouts, Sacha. 2007. Drug Law Enforcement. *Police Quarterly*, **10**(2), 115–153.

Mazerolle, Lorraine Green, Price, James F., & Roehl, Jan. 2000. Civil remedies and drug control: A randomized field trial in Oakland, California. *Evaluation Review*, **24**(2), 212–241.

McElroy, Jerome E. 1998. Evaluating service delivery of police agencies: Suggestions for focus and strategy. *Pages 73–87 of:* Brodeur, Jean-Paul (ed), *How to Recognize Good Policing: Problems and Issues*. Washington, DC; Thousand Oaks: Police Executive Research Forum and Sage Publications.

Messner, Steven F., Galea, Sandro, Tardiff, Kenneth J., Tracy, Melissa, Bucciarelli, Angela, Piper, Tinka Markham, Frye, Victoria, & Vlahov, David. 2007. Policing, drugs, and the homicide decline in New York City in the 1990s. *Criminology*, **45**(2), 385–414.

Mitchell, Andy. 1999. *The ESRI© Guide to GIS-Analysis. Volume 1: Geographic Patterns & Relationships*. Redlands: ESRI Press.

Mitchell, Andy. 2005. *The ESRI© Guide to GIS-Analysis. Volume 2: Spatial Measurements & Statistics*. Redlands: ESRI Press.

Müller, Jürg. 2006. *Community Policing*. Zürich: Stadtpolizei Zürich.

Müller, Jürg. 2008. Kommunale Kriminalprävention in der Stadt Zürich: Massnahmen und Erfahrungen der Stadtpolizei Zürich. *Pages 5–20 of:* Schwarzenegger, Christian, & Müller, Jürg (eds), *Erstes Zürcher Präventionsforum: Kommunale Kriminalprävention, Crime Mapping, Einbruchskriminalität*. Zürich, Basel, Genf: Schulthess.

Monti, Stefano, Tamayo, Pablo, Mesirov, Jill, & Golub, Todd. 2003. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning*, **52**(1), 91–118.

Moore, Mark H., & Kelling, George L. 1983. To serve and protect: Learning from police history. *The Public Interest*, **70**, 49–65.

Moore, Mark H., & Kelling, George L. 1988. *The Evolving Strategy of Policing.* Perspectives on Policing, vol. 4. Washington, DC: U.S. Department of Justice, National Institute of Justice, and John F. Kennedy School of Government, Harvard University.

Moore, Mark H., & Trojanowicz, Robert C. 1988. *Policing and the Fear of Crime.* Perspectives on Policing, vol. 3. Washington, DC: U.S. Department of Justice, National Institute of Justice, and John F. Kennedy School of Government, Harvard University.

Moore, Mark H., Trojanowicz, Robert C., & Kelling, George L. 1988. *Crime and Policing.* Perspectives on Policing, vol. 2. Washington, DC: U.S. Department of Justice, National Institute of Justice, and John F. Kennedy School of Government, Harvard University.

Municipalité de Lausanne. 2001. *Sécurité à Lausanne: état des lieux - actions entreprises - perspectives. Création d'un observatoire de la sécurité. Réponse aux motions de MM. Olivier Français et Jean-Christophe Bourquin ainsi qu'à la pétition de M. Carl Kyril Gossweiler.* Rapport-préavis N° 239 du 20 septembre 2001. Lausanne: Direction de la sécurité publique et des affaires sportives.

Murtagh, F. 1995. Interpreting the Kohonen self-organizing feature map using contiguity-constrained clustering. *Pattern Recognition Letters*, **16**(4), 399–408.

Myhill, Andy. 2006. *Community Engagement in Policing: Lessons from the Literature.* London: Home Office.

Newman, Oscar. 1972. *Defensible Space: Crime Prevention Through Urban Design.* New York: Macmillan.

Nolan, James J., Conti, Norman, & McDevitt, Jack. 2004. Situational policing: Neighbourhood development and crime control. *Policing and Society*, **14**(2), 99–117.

Openshaw, Stan. 1984. *The Modifiable Areal Unit Problem.* Norwich: Geo Books.

Painter, Kate. 1996. The influence of street lighting improvements on crime, fear and pedestrian street use, after dark. *Landscape and Urban Planning*, **35**(2-3), 193–201.

Pate, Antony M., Brown, Charles E., Kelling, George L., Ferrara, Amy, & Utne, Mary. 1981. *The Newark Foot Patrol Experiment.* Washington, DC: Police Foundation.

Pate, Antony M., Wycoff, Mary Ann, Skogan, Wesley G., & Sherman, Lawrence W. 1986. *Reducing Fear of Crime in Houston and Newark: A Summary Report.* Washington, DC: Police Foundation and National Institute of Justice.

Patil, G. P., Modarres, R., Myers, W. L., & Patankar, P. 2006. Spatially constrained clustering and upper level set scan hotspot detection in surveillance geoinformatics. *Environmental and Ecological Statistics*, **13**(4), 365–377.

Police Executive Research Forum. 1996. *Themes and Variations in Community Policing: Case Studies of Communtiy Policing.* Washington, DC: Police Executive Research Forum.

Police municipale de Lausanne. 1986. *Police de Lausanne.* Brochure edité par la Direction de police à propos de l'inauguration de l'Hôtel de police, 1er octobre 1986.

Quinton, Paul, & Morris, Julia. 2008. *Neighbourhood Policing: The Impact of Piloting and Early National Implementation.* Home Office On-Line Report, vol. 01/08. London: Home Office.

Quinton, Paul, & Tuffin, Rachel. 2007. Neighbourhood change: the impact of the National Reassurance Policing Programme. *Policing: A Journal of Policy and Practice*, **1**(2), 149–160.

R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing.

Ramsay, Malcolm, & Newton, Rosemary. 1991. *The Effect of Better Street Lighting on Crime and Fear: A Review.* Crime Prevention Unit Paper, vol. 29. London: Home Office.

Ratcliffe, Jerry H., Taniguchi, Travis, Groff, Elizabeth R., & Wood, Jennifer D. 2011. The Philadelphia foot patrol experiment: A randomized controlled trial of police patrol effectiveness in violent crime hotspots. *Criminology*, **49**(3), 795–831.

Reisig, Michael D., & Parks, Roger B. 2000. Experience, quality of life, and neighborhood context: A hierarchical analysis of satisfaction with police. *Justice Quarterly*, **17**(3), 607–630.

Reisig, Michael D., & Parks, Roger B. 2004. Can community policing help the truly disadvantaged? *Crime & Delinquency*, **50**(2), 139–167.

Riva, Gilles. 1988. *Morphologie de l'espace urbain et délits contre le patrimoine à Lausanne en 1980.* Collection Criminologie, vol. 4. Grüsch: Rüegger.

Rix, Andrew, Joshua, Faye, Maguire, Mike, & Morton, Sarah. 2009. *Improving public confidence in the police: A review of the evidence. 2nd Edition.* Research Report 28. London: Home Office.

Robinson, Jennifer B., Lawton, Brian A., Taylor, Ralph B., & Perkins, Douglas D. 2003. Multilevel longitudinal impacts of incivilities: Fear of crime, expected safety, and block satisfaction. *Journal of Quantitative Criminology*, **19**(3), 237–274.

Robinson, W. S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review*, **15**(3), 351–357.

Robitaille, Eric, & Séguin, Anne-Marie. 2007. Effets de quartier et délinquance juvénile: Le défi posé par l'élaboration d'indicateurs territoriaux de désorganisation sociale. *In:* Sénécal, G (ed), *Les indicateurs socio-territoriaux et les métropoles.* Université Laval, Québec: Presses de l'Université Laval.

Roh, Christiane. 2010. La criminalité à Lausanne, dans ses quartiers et dans l'agglomération. *Lausanne déchiffrée - Courrier statistique*, **6**(2), 1–4.

Rosenbaum, Dennis. 2002. Evaluating multi-agency anti-crime partnerships: Theory, design, and measurement issues. *Pages 171–225 of:* Tilley, Nick (ed), *Evaluation for Crime Prevention.* Crime Prevention Studies, vol. 14. Monsey: Criminal Justice Press.

Rosenbaum, Dennis P. 1986. *Community Crime Prevention: Does It Work?* Sage Criminal Justice Annuals. Beverly Hills: Sage.

Rosenbaum, Dennis P. 1988. Community crime prevention: A review and synthesis of the literature. *Justice Quarterly*, **5**(3), 323–395.

Rosenbaum, Dennis P. 1994. *The Challenge of Community Policing: Testing the Promises.* Thousand Oaks: Sage Publications.

Rosenbaum, Dennis P. 1998. The changing role of the police: Assessing the current transition to community policing. *Pages 3–29 of:* Brodeur, Jean-Paul (ed), *How to Recognize Good Policing: Problems and Issues.* Washington, DC; Thousand Oaks: Police Executive Research Forum and Sage Publications.

Rosenbaum, Paul R. 2010. *Design of Observational Studies.* New York: Springer.

Rosenbaum, Paul R., & Rubin, Donald B. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70**(1), 41–55.

Rosenfeld, Richard, Fornango, Robert, & Rengifo, Andres F. 2007. The impact of order-maintenance policing on New York City homicide and robbery rates: 1988-2001. *Criminology*, **45**(2), 355–384.

Rossi, Peter Henry. 1987. The iron law of evaluation and other metallic rules. *Pages 3–20 of:* Miller, J., & Lewis, M. (eds), *Research in Social Problems and Public Policy*. Greenwich, UK: JAI Press.

Sacco, Vincent F. 1998. Evaluating satisfaction. *Pages 123–137 of:* Brodeur, Jean-Paul (ed), *How to Recognize Good Policing: Problems and Issues*. Washington, DC; Thousand Oaks: Police Executive Research Forum and Sage Publications.

Sadd, Susan, & Grinc, Randolph M. 1994. Innovative neighborhood oriented policing: An evaluation of community policing programs in eight cities. *Pages 27–52 of:* Rosenbaum, Dennis P. (ed), *The Challenge of Community Policing: Testing the Promises*. Thousand Oaks: Sage Publications.

Sadd, Susan, & Grinc, Randolph M. 1996. *Implementation Challenges in Community Policing: Innovative Neighbourhood-Oriented Policing in Eight Cities*. Washington, DC: National Institute of Justice.

Sampson, Robert J. 2009. Disparity and diversity in the contemporary city: Social (dis)order revisited. *The British Journal of Sociology*, **60**(1), 1–31.

Sampson, Robert J., & Bartusch, Dawn Jeglum. 1998. Legal cynicism and (subcultural?) tolerance of deviance: The neighborhood context of racial differences. *Law & Society Review*, **32**(4), 777–804.

Sampson, Robert J., & Groves, W. Byron. 1989. Community structure and crime: Testing social disorganization theory. *American Journal of Sociology*, **94**(4), 774–802.

Sampson, Robert J., & Raudenbush, Stephen W. 1999. Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology*, **105**(3), 603–651.

Sampson, Robert J., Raudenbush, Stephen W., & Earls, Felton. 1997. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, **277**(5328), 918–924.

Sampson, Robert J., Jeffrey, D. Morenoff, & Thomas, Gannon-Rowley. 2002. Assessing neighborhood "effects": Social processes and new directions in research. *Annual Review of Sociology*, **28**, 443.

Sarkar, Deepayan. 2008. *Lattice: Multivariate Data Visualization with R*. New York: Springer.

Sauter, Regine, Schedler, Kuno, & Schäfer, Werner. 2005. Forschungsprojekt "Urbane Sicherheit". *Pages 73–182 of:* Konferenz der Städtischen Polizeidirektorinnen und Polizeidirektoren (KSPD), & Polizeidepartement Zürich (eds), *Community Policing: Modelle für eine vernetzte Polizeiarbeit in der Schweiz*. Schriftenreihe der KSPD, vol. 2. Zürich: Schulthess.

Schafer, Joseph A., Huebner, Beth M., & Bynum, Timothy S. 2003. Citizen perceptions of police services: Race, neighborhood context, and community policing. *Police Quarterly*, **6**(4), 440–468.

Schwarzenegger, Christian. 1992. *Die Einstellungen der Bevölkerung zur Kriminalität und Verbrechenskontrolle: Ergebnisse einer repräsentativen Befragung der Zürcher Kantonsbevölkerung im internationalen Vergleich*. Freiburg im Breisgau: Max-Planck-Institut für Ausländisches und Internationales Strafrecht.

Schwarzenegger, Christian, Schmohl, Denise, Thalmann, Urs, Vertone, Leonardo, & Zanolini, Veio. 2009. *Kriminalität und kommunale Kriminalprävention in Zürich: Eine kriminologische Regionalanalyse mit besonderer Ausrichtung auf das Zürcher Langstrassenquartier (Projekt "Langstrasse PLUS")*. Zürich: Kriminologisches Institut der Universität Zürich.

Segmüller, Pius. 2005. Community Policing - Wie die Stadtpolizei Luzern zur bürgerorientierten Polizeiarbeit findet. *Pages 37–41 of:* Konferenz der Städtischen Polizeidirektorinnen und Polizeidirektoren (KSPD), & Polizeidepartement Zürich (eds), *Community Policing: Modelle für eine vernetzte Polizeiarbeit in der Schweiz. Referate der Tagung vom 10. September 2004*. Schriftenreihe der KSPD, vol. 2. Zürich: Schulthess.

Shadish, William R., Cook, Thomas D., & Campbell, Donald T. 2002. *Experimental and Quasi-experimental designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Shaw, Clifford R., & McKay, Henry D. 1942/1972. *Juvenile Delinquency and Urban Areas: A Study of Rates of Delinquency in Relation to Differential Characteristics of Local Communities in American Cities*. Rev. edn. Chicago: The University of Chicago Press.

Sheldon, Gerard, Hall, Rebecca, Brunsdon, Chris, Charlton, Martin, Alvanides, Seraphim, & Mostratos, Nikos. 2002. *Maintaining basic command unit and crime and disorder partnership families for comparative purposes*. RDS Online Report.

Sherman, Lawrence W. 1997. Policing for crime prevention. *In:* Sherman, Lawrence W., Gottfredson, Denise, MacKenzie, Doris, Eck, John, Reuter, Peter, & Bushway, Shawn (eds), *Preventing Crime: What Works, What Doesn't, What's Promising. A report to the United States Congress*. Washington, DC: US Department of Justice, Office of Justice Programs.

Sherman, Lawrence W. 2010. Democratic policing on the evidence. *Pages 589–618 of:* Wilson, James Q., & Petersilia, Joan (eds), *Crime and Public Policy*, 2nd edn. Oxford: Oxford University Press.

Sherman, Lawrence W., & Eck, John E. 2002. Policing for crime prevention. *Pages 295–329 of:* Sherman, Lawrence W., Farrington, David P., Welsh, Brandon C., & MacKenzie, Doris Layton (eds), *Evidence-Based Crime Prevention*. London: Routledge.

Sherman, Lawrence W., & Weisburd, David. 1995. General deterrent effects of police patrol in crime hot spots: A randomized, controlled trial. *Justice Quarterly*, **12**(4), 625–648.

Sherman, Lawrence W., Gottfredson, Denise, MacKenzie, Doris, Eck, John, Reuter, Peter, & Bushway, Shawn. 1997. *Preventing Crime: What Works, What Doesn't, What's Promising. A Report to the United States Congress*. Washington, DC: US Deptartment of Justice, Office of Justice Programs.

Sherman, Lawrence W., Farrington, David P., Welsh, Brandon C., & MacKenzie, Doris Layton. 2002a. *Evidence-Based Crime Prevention*. London: Routledge.

Sherman, Lawrence W., Farrington, David P., Welsh, Brandon C., & MacKenzie, Doris Layton. 2002b. Preventing crime. *Pages 1–12 of:* Sherman, Lawrence W., Farrington, David P., Welsh, Brandon C., & MacKenzie, Doris Layton (eds), *Evidence-Based Crime Prevention*. London: Routledge.

Siegel, Sidney. 1997. *Nichtparametrische statistische Methoden*. 4th edn. Eschborn bei Frankfurt am Main: Klotz.

Silverman, Robert A., & Kennedy, Leslie W. 1985. Loneliness, satisfaction and fear of crime: A test for non-recursive effects. *Canadian Journal of Criminology*, **27**(1), 1–13.

Skogan, Wesley G. 1986. Fear of crime and neighborhood change. *Pages 203–229 of:* Reiss, Albert J., & Tonry, Michael (eds), *Communities and Crime*. Crime and Justice: A Review of Research. Chicago: University of Chicago Press.

Skogan, Wesley G. 1987. The impact of victimization on fear. *Crime & Delinquency*, **33**(1), 135–154.

Skogan, Wesley G. 1990. *Disorder and Decline: Crime and the Spiral of Decay in American Neighborhoods*. New York, Toronto: The Free Press.

Skogan, Wesley G. 1994. The impact of community policing on neighborhood residents. A cross-site analysis. *Pages 167–181 of:* Rosenbaum, Dennis P. (ed), *The Challenge of Community Policing: Testing the Promises*. Thousand Oaks: Sage Publications.

Skogan, Wesley G. 1999. Measuring what matters: Crime, disorder, and fear. *Pages 37–53 of:* Langworthy, Robert H. (ed), *Measuring What Matters: Proceedings from the Policing Research Institute Meetings*. Washington, DC: US Deptartment of Justice, National Institute of Justice and the Office of Community Oriented Policing Services.

Skogan, Wesley G. 2006a. Asymmetry in the impact of encounters with police. *Policing and Society: An International Journal of Research and Policy*, **16**(2), 99–126.

Skogan, Wesley G. 2006b. *Police and Community in Chicago: A Tale of Three Cities*. Studies in crime and public policy. Oxford: Oxford University Press.

Skogan, Wesley G. 2006c. The promise of community policing. *Pages 27–43 of:* Weisburd, David, & Braga, Anthony A. (eds), *Police Innovation: Contrasting Perspectives*. Cambridge Studies in Criminology. Cambridge: Cambridge University Press.

Skogan, Wesley G. 2009. Concern about crime and confidence in the police. *Police Quarterly*, **12**(3), 301–318.

Skogan, Wesley G., & Frydl, Kathleen. 2004. *Fairness and Effectiveness in Policing: The Evidence*. Washington, DC: National Academies Press.

Skogan, Wesley G., & Hartnett, Susan M. 1997. *Community Policing: Chicago Style*. New York; Oxford: Oxford University Press.

Skogan, Wesley G., & Maxfield, Michael G. 1981. *Coping with Crime: Individual and Neighborhood Reactions*. Sage Library of Social Research, vol. 124. Beverly Hills; London: Sage Publications.

Skogan, Wesley G., & Roth, Jeffrey A. 2004. Introduction. *Pages xvii–xxxiv of:* Skogan, Wesley G. (ed), *Community Policing: Can It Work?* The Wadsworth Professionalism in Policing Series. Belmont: Wadsworth/Thomson Learning.

Skogan, Wesley G., Steiner, Lynn, Benitez, Claudia, Bennis, Jason, Borchers, Sarah, DuBois, Jill, Gondocs, Rita, Hartnett, Susan, Kim, So Young, & Rosenbaum, Sarah. 2004. *Community Policing in Chicago, Year Ten: An Evaluation of Chicago's Alternative Policing Strategy*. Chicago: Illinois Criminal Justice Information Authority.

Skupin, André, & Agarwal, Pragya. 2008. Introduction: What is a self-organizing map? *Pages 1–20 of:* Agarwal, Pragya, & Skupin, André (eds), *Self-Organising Maps: Applications in Geographic Information Science*. Chichester: Wiley.

Sousa, William H., & Kelling, George L. 2006. Of "broken windows," criminology, and criminal justice. *Pages 77–97 of:* Weisburd, David, & Braga, Anthony A. (eds), *Police Innovation: Contrasting Perspectives*. Cambridge: Cambridge University Press.

Spaar, Rudolf, Roubaty, Jean-Pierre, Hänni, Urs, Krebs, Stefan, Scherrer, Andreas, Müller, Jürg, & Böss, Martin. 2007. *Community Policing: Grundlagendokument für die eidgenössische Berufsprüfung Polizist/Polizistin*. Neuchâtel: Schweizerisches Polizei-Institut.

Sparrow, Malcolm K. 1988. *Implementing Community Policing*. Perspectives on Policing, vol. 9. Washington, DC: US Department of Justice, National Institute of Justice and John F. Kennedy School of Government, Harvard University.

Spielman, Seth E., & Thill, Jean-Claude. 2008. Social area analysis, data mining, and GIS. *Computers, Environment and Urban Systems*, **32**(2), 110–122.

Stadt Bern. 2005. *PINTO - Prävention, Intervention, Toleranz: Detailkonzept*. Bern: Direktion für Bildung Soziales und Sport.

Stadtpolizei Bern. 1999. *Gemeinsam für Sicherheit – Strategie zur Umsetzung des Projekts APOLLO*. Stadtpolizei Bern.

Sunshine, Jason, & Tyler, Tom R. 2003. The role of procedural justice and legitimacy in shaping public support for policing. *Law & Society Review*, **37**(3), 513–548.

Taylor, Ralph B. 1999. The incivilities thesis: Theory, measurement, and policy. *In:* Langworthy, Robert H. (ed), *Measuring What Matters: Proceedings from the Policing Research Institute Meetings*. Washington, DC: US Deptartment of Justice, National Institute of Justice and the Office of Community Oriented Policing Services.

Taylor, Ralph B. 2001. *Breaking Away from Broken Windows: Baltimore Neighborhoods and the Nationwide Fight against Crime, Grime, Fear, and Decline*. Crime & Society. Boulder: Westview Press.

Taylor, Ralph B. 2002. Fear of crime, social ties, and collective efficacy: Maybe masquerading measurement, maybe déjà vu all over again. *Justice Quarterly*, **19**(4), 773–792.

Taylor, Ralph B. 2006. Incivilities reduction policing, zero tolerance, and the retreat from coproduction: Weak foundations and strong pressures. *Pages 98–114 of:* Weisburd, David, & Braga, Anthony A. (eds), *Police Innovation: Contrasting Perspectives*. Cambridge: Cambridge University Press.

Taylor, Ralph B., & Hale, Margaret. 1986. Testing alternative models of fear of crime. *The Journal of Criminal Law and Criminology*, **77**(1), 151–189.

Taylor, Ralph B., Gottfredson, Stephen D., & Brower, Sidney. 1984. Block crime and fear: Defensible space, local social ties, and territorial functioning. *Journal of Research in Crime and Delinquency*, **21**(4), 303–331.

Terrill, William, & Reisig, Michael D. 2003. Neighborhood context and police use of force. *Journal of Research in Crime and Delinquency*, **40**(3), 291–321.

Thacher, David. 2004. Order maintenance reconsidered: Moving beyond strong causal reasoning. *The Journal of Criminal Law and Criminology*, **94**(2), 381–414.

Tien, James M., O'Donnell, Vincent F., Bamett, Arnold, & Mirchandani, Pitu B. 1979. *Street Lighting Projects*. Washington, DC: US Department of Justice, National Institute of Law Enforcement and Criminal Justice.

Trojanowicz, Robert, Kappeler, Victor E., & Gaines, Larry K. 2002. *Community Policing: A Contemporary Perspective*. 3 edn. Cincinnati: Anderson Publishing.

Trojanowicz, Robert C. 1983. Evaluation of a neighborhood foot patrol program. *Journal of Police Science and Administration*, **11**(4), 410–419.

Trojanowicz, Robert C., & Bucqueroux, Bonnie. 1994. *Community Policing: How to Get Started*. Cincinnati: Anderson Publishing.

Tseloni, Andromachi, & Zarafonitou, Christina. 2008. Fear of crime and victimization: A multivariate multilevel analysis of competing measurements. *European Journal of Criminology*, **5**(4), 387–409.

Tuffin, Rachel, Morris, Julia, & Poole, Alexis. 2006. *An Evaluation of the Impact of the National Reassurance Policing Programme*. Home Office Research Study, vol. 296. London: Home Office.

Tuia, Devis, Kaiser, Christian, Da Cunha, Antonio, & Kanevski, Mikhail. 2009. Detection of urban socio-economic patterns using clustering techniques. *Pages 19–36 of:* Murgante, Beniamino, Borruso, Giuseppe, & Lapucci, Alessandra (eds), *Geocomputation and Urban Planning*. Berlin: Springer.

Tyler, Tom R. 1990. *Why People Obey the Law*. New Haven: Yale University Press.

Tyler, Tom R. 2004. Enhancing police legitimacy. *The ANNALS of the American Academy of Political and Social Science*, **593**, 84–99.

Tyler, Tom R., & Huo, Yuen J. 2002. *Trust in the Law: Encouraging Public Cooperation With the Police and Courts*. New York: Russell Sage Foundation.

Ultsch, A., & Siemon, H. P. 1990. Kohonen's self organizing feature maps for exploratory data analysis. *Pages 305–308 of: Proceedings of the INNC'90, International Neural Network Conference*. Dordrecht: Kluwer.

van der Knaap, Leontien M., Leeuw, Frans L., Bogaerts, Stefan, & Nijssen, Laura T. J. 2008. Combining Campbell standards and the realist evaluation approach. The best of two worlds? *American Journal of Evaluation*, **29**(1), 48–57.

Vesanto, Juha, & Alhoniemi, Esa. 2000. Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, **11**(3), 586 – 600.

Vieli, Rolf. 2006. *Die Langgasse, die Langstrasse und die „Longstreet“*. Zürich: Polizeidepartement der Stadt Zürich.

Wakefield, Alison. 2007. Carry on constable? Revaluing foot patrol. *Policing*, **1**(3), 342–355.

Waller, Lance A., & Gotway, Carol A. 2004. *Applied Spatial Statistics for Public Health Data*. Wiley Series in Probability and Statistics. Hoboken: Wiley.

Walpen, Laurent. 1996. How to introduce humanity into a sometimes inhuman profession. Paper presented at the International Police Executive Symposium in Geneva (Switzerland) in May 1994. *The Police Journal*, **69**(1), 21–29.

Warr, Mark. 1984. Fear of victimization: Why are women and the elderly more afraid. *Social Science Quarterly*, **65**(3), 681–702.

Wehrens, Ron, & Buydens, Lutgarde M. C. 2007. Self- and super-organizing maps in R: the kohonen package. *Journal of Statistical Software*, **21**(5), 1–19.

Weisburd, David, & Eck, John E. 2004. What can police do to reduce crime, disorder, and fear? *The ANNALS of the American Academy of Political and Social Science*, **593**(1), 42–65.

Weisburd, David, & Green, Lorraine. 1995. Policing drug hot spots: The Jersey City drug market analysis experiment. *Justice Quarterly*, **12**(4), 711 – 735.

Weisburd, David, Lum, Cynthia, & Petrosino, Anthony. 2001. Does research design affect study outcomes in criminal justice? *The ANNALS of the American Academy of Political and Social Science*, **578**, 50–70.

Weisburd, David, Eck, John E., Hinkle, Joshua C., & Telep, Cody W. 2008. *The effects of problem-oriented policing on crime and disorder*. Campbell Systematic Reviews. Oslo: The Campbell Collaboration.

Weisburd, David, Bernasco, Wim, & Bruinsma, Gerben J.N. 2009. *Putting Crime in its Place: Units of Analysis in Geographic Criminology*. New York: Springer.

Weisburd, David, Telep, Cody W., Hinkle, Joshua C., & Eck, John E. 2010. Is problem-oriented policing effective in reducing crime and disorder? Findings from a Campbell systematic review. *Criminology and Public Policy*, **9**(1), 139–172.

Weiss, Carol H. 1972. *Evaluation Research: Methods for Assessing Program Effectiveness*. Prentice-Hall Methods of Social Science Series. Englewood Cliffs: Prentice-Hall.

Weitzer, Ronald, & Tuch, Steven A. 2004. Race and perceptions of police misconduct. *Social Problems*, **51**(3), 305–325.

Wells, William. 2009a. Community partnerships. *Pages 3–12 of:* Maguire, Edward, & Wells, William (eds), *Implementing Community Policing: Lessons from 12 Agencies*. Washington, DC: US Department of Justice, Office of Community Oriented Policing Services.

Wells, William. 2009b. Performance appraisal systems. *Pages 71–78 of:* Maguire, Edward, & Wells, William (eds), *Implementing Community Policing: Lessons from 12 Agencies*. Washington, DC: US Department of Justice, Office of Community Oriented Policing Services.

Wells, William. 2009c. Problem solving. *Pages 13–34 of:* Maguire, Edward, & Wells, William (eds), *Implementing Community Policing: Lessons from 12 Agencies*. Washington, DC: US Department of Justice, Office of Community Oriented Policing Services.

Welsh, Brandon C., & Farrington, David P. 2004. Surveillance for crime prevention in public space: Results and policy choices in Britain and America. *Criminology & Public Policy*, **3**(3), 497–526.

Welsh, Brandon C., & Farrington, David P. 2008. *Effects of Improved Street Lighting on Crime*. Campbell Systematic Reviews. Oslo: The Campbell Collaboration.

Welsh, Brandon C., & Farrington, David P. 2009. Public area CCTV and crime prevention: An updated systematic review and meta-analysis. *Justice Quarterly*, **26**(4), 716–745.

Welsh, Brandon C., & Hoshi, Akemi. 2002. Communities and crime prevention. *Pages 165–197 of:* Sherman, Lawrence W., Farrington, David P., Welsh, Brandon C., & MacKenzie, Doris Layton (eds), *Evidence-Based Crime Prevention*. London: Routledge.

Wilcoxon, Frank. 1945. Individual comparisons by ranking methods. *Biometrics Bulletin*, **1**(6), 80–83.

Williamson, Tom, Ashby, David I., & Webber, Richard. 2006. Classifying neighbourhoods for reassurance policing. *Policing and Society*, **16**(2), 189–218.

Wilson, James Q. 1977. *Thinking About Crime*. New York: Vintage Books.

Wilson, James Q., & Kelling, George L. 1982. Broken Windows: The police and neighborhood safety. *The Atlantic Monthly*, **249**(3), 29–38.

Wisler, Dominique, Froidevaux, Didier, Faggioni, André, Genier, Didier, & Tcherniaeva, Ioulia (sur mandat de la Police cantonale de Genève). 2004. *Diagnostic local de sécurité 2004 du canton de Genève*. Genève: Police cantonale de Genève, Service des études stratégiques.

Wisler, Dominique, Froidevaux, Didier, Faggioni, André, & Fasola, Ioulia (sur mandat de la Police cantonale de Genève). 2008a. *Diagnostic local de sécurité 2007*. Genève: Police cantonale de Genève, Service des études stratégiques.

Wisler, Dominique, Froidevaux, Didier, Faggioni, André, & Fasola, Ioulia (sur mandat de la Police cantonale de Genève). 2008b. *Diagnostic local de sécurité 2007: Annexes*. Genève: Police cantonale de Genève, Service des études stratégiques.

Wisler, Dominique, Froidevaux, Didier, Faggioni, André, Fasola, Ioulia, Moret, Melody, & Ohannessian, Silva (sur mandat de la Police cantonale de Genève). 2011. *Diagnostic local de sécurité 2010*. Genève: Police cantonale de Genève, Service des études stratégiques.

Wyant, Brian R. 2008. Multilevel impacts of perceived incivilities and perceptions of crime risk on fear of crime: Isolating endogenous impacts. *Journal of Research in Crime and Delinquency*, **45**(1), 39–64.

Wycoff, Mary Ann, & Skogan, Wesley G. 1994. Community policing in Madison: An analysis of implementation and impact. *Pages 75–91 of:* Rosenbaum, Dennis P. (ed), *The Challenge of Community Policing: Testing the Promises*. Thousand Oaks: Sage Publications.

Xu, Yili, Fiedler, Mora L., & Flaming, Karl. H. 2005. Discovering the impact of community policing: The broken windows thesis, collective efficacy, and citizens' judgment. *Journal of Research in Crime and Delinquency*, **42**(2), 147–186.

Zhao, Jihong. H., Lovrich, Nicholas. P., & Robinson, T. Hank. 2001. Community policing - Is it changing the basic functions of policing? Findings from a longitudinal study of 200+ municipal police agencies. *Journal of Criminal Justice*, **29**(5), 365–377.

Zhao, Jihong "Solomon", Matthew, Scheider, & Quint, Thurman. 2002. The effect of police presence on public fear reduction and satisfaction: A review of the literature. *The Justice Professional*, **15**(3), 273–299.

Zhao, Jihong "Solomon", He, Ni, & Lovrich, Nicholas P. 2003. Community policing: Did it change the basic functions of policing in the 1990s? A national follow-up study. *Justice Quarterly*, **20**(4), 697–724.

Zorn, Jehan. 2008. *La police de proximité dans une perspective genevoise*. Travail de mémoire en criminologie. Université de Lausanne: Institut de criminologie et de droit pénal.

A

Technical Appendices

A.1 Weights

A salient feature of the administration of the Swiss Crime Survey in its most recent editions has been that a number of cantonal and urban police departments financed extra interviews in their jurisdiction. These cash infusions boosted the samples of the cities of interest for the current study to between 200 and 500 interviews, which are thus substantially overrepresented relative to their actual population size. For example, in the 2005 survey, the sample of the City of Lausanne counts 244 cases out of a total of 3898 completed interviews nationwide, which is equivalent to roughly four times its true share of the Swiss resident population (Killias *et al.* 2007, 161).

The survey samples were thus weighted on the basis of official census data. The data sets from 1998, 2000, and 2005 were adjusted using census data from the year 2000. The 1987 survey was weighted on the basis of the 1990 census count.

Upon our request, the OFS prepared summary statistics of the total population, gender and age distribution as well as educational achievements for all neighborhoods of the five cities under study (plus the surrounding municipalities in the Cantons of Basel-City and Geneva). Where the area covered by a ZIP code district encompasses more than one OFS neighborhood (census tract), the census data were added up accordingly. The same aggregating procedure of OFS neighborhoods had previously been performed to create polygons that best resembled the five cities' ZIP code districts, which served as the basic unit for the spatial analyses (cf. Chapter 3.2.3).

Weights were computed independently for the two demographic variables gender and age in order to make the survey sample proportional to the census count at the ZIP code level. The sub-sample of each ZIP code district was further weighted according to the area's total population in order to reflect its relative population size.

These individual weights were multiplied to compute overall weights. Overall weights were adjusted so that their sum corresponded to the actual size of the survey sample for all analyses. This procedure resulted in three separate sets of weights for each data set, in accord with the three levels of analysis of this study – ZIP code districts, cities, and the combined total survey sample of all five urban areas. For all computations, weights were used at the level appropriate for the basic unit of analysis of the inquiry at hand, i.e. ZIP code-level weights for analyses of neighborhood attributes, city-level weights for city attributes, and national-level weights for the combined analyses of the respondents from all five urban areas.

Weights were not adjusted for non-response rates, because such data were not available for all four data sets.

The variables gender and age were complete in all four data sets. The variable education level, by contrast, contained a substantial number of missing values. The problem of missing values was particularly acute for respondents aged 25 and younger, supposedly because they had not yet completed their education at the time of the interview. The survey samples were therefore not weighted regarding the level of education, in order not to lose these cases for subsequent analyses.

All final weight scores were tested for validity. Frequency distributions of both the raw and weighted data were calculated by the variables gender, age, and total population at each of the three levels of analysis for all four data sets. These distributions were subjected to a χ^2-homogeneity test against the census data to check whether the weighted samples better reflected the true distribution of the attribute classes. A success rate of between 73 and 92 percent for the neighborhood-level weights indicated that the weighting procedure greatly enhanced the validity of subsequent data analyses. Crucially, the weighting procedure managed to adjust the distribution of those few ZIP code areas for which the p-value of the χ^2-test of the raw survey sample was below 0.1. At the city and overall level, the rate at which the p-value of the χ^2-homogeneity test was higher for the weighted data fluctuated more widely (due to the smaller denominator). Nevertheless, the weighting procedure represented a net improvement over the raw survey samples.

A.2 Data Checks

As the primary geographic indicator used for the present study, the ZIP code numbers of all survey respondents in each of the four data sets were checked for their accuracy and geographic location. Based on the OFS' county number, which was recorded for each case in the three most recent data sets, all survey respondents from the cities of Bern, Lausanne, and Zurich as well as the cantons of Basel City and Geneva were selected and their ZIP code number verified. The 1987 data set did not yet record the OFS county number, so individual ZIP code numbers were scrutinized directly.

A.2.1 Invalid ZIP Code Numbers

In the 1987 data set, the following few cases have been set to missing because the recorded ZIP code number did either not exist or was a generic value that does not indicate a specific geographic area within a city:

Table A.1 Invalid ZIP code numbers

Year	Original ZIP	N Cases
1987	3000	2
	3022	1
	4000	9
	4045	1
	8000	7
	8018	1

A.2.2 Recoded ZIP Code Numbers

For each of the four data sets, the ZIP code number of a few individual respondents had to be recoded because their ZIP code district has no unique corresponding feature in the GIS shape file of all counties and urban neighborhoods in Switzerland used as basic spatial units of analysis of this study. Table A.2 indicates the ZIP code numbers that were recoded.

Table A.2 Recoded ZIP code numbers

Year	Original	Recoded	N Cases
2005	1219	1214	2
	1220	1214	1
	1222	1245	1
	3020	3019	1
	8064	8048	3
2000	1216	1217	2
	1219	1214	23
	1220	1214	7
	1222	1245	5
	1231	1224	1
	1234	1255	7
	3020	3019	1
	4001	4051	1
	8064	8048	1
1998	1216	1217	3
	1219	1214	29
	1220	1214	7
	1234	1255	2
	1282	1283	1
	8064	8048	5
1987	3032	3027	3
	3048	3014	1
	4027	4054	1
	8026	8004	1
	8034	8008	1
	8035	8006	1
	8036	8003	1
	8042	8006	1

A.3 Data Recoding

In order to facilitate the data handling and interpretation of results, various variables in the four data sets were recoded. The guiding principles of this process were to uphold the ordinal character of answer categories wherever possible and to ensure that higher category values consistently indicate higher levels of fear, disorder, or dissatisfaction with police services (i.e. the less desirable outcome). Following is an exhaustive list of all the data recodings made to individual variables. Variables not explicitly mentioned here were used in their original format. The data recoding log is grouped according to the three theoretical constructs fear, disorder, and policing and then subdivided by survey year.

A.3.1 Fear

2005

Variable Q30010 (Feeling safe at night) was used in its original format but category 6 was set to missing to exclude respondents who do not leave their home at night for other reasons than security concerns.

Variable Q302 (Likelihood of burglary) was recoded into variable q302rec in reversed order so that higher category values stand for increased probability of a burglary.

Variable Q303 (Avoiding alleys/people) was recoded into variable q303rec. Category 1 was recoded as 2 and category 2 was recoded as 1. Category 4 was set to missing to exclude respondents who do not leave their home at night for other reasons than security concerns. This was done to ensure that higher category values stand for more drastic avoidance strategies.

2000

The exact same transformations described above were made to variables q30010, q302, and q303 in the 2000 data set.

1998

The exact same transformations described above were also made to variables q30010, q302, and q303 in the 1998 data set.

1987

Variables v45, v46, and v47 (Questionnaire items 489, 490, and 491) were transformed into the single variable q489_91. This made it possible to combine the variance of three survey items in a single variable. This new variable takes on the value of 1 for respondents who go out at night and do not experience fear, 2 for respondents who do go out at night but are afraid at some place within a short range from their home, 3 for respondents who do not go out at night but would be afraid if they did, and 4 for respondents who do not go out at night but would not be afraid in their neighborhood if they did. This latter category was set to missing on the grounds that these respondents stay put in their homes for different reasons than concerns about their personal safety.

Variable q489_91 has been compared with the questionnaire item on fear of crime of the following years. However, both the wording and format of the three underlying questionnaire items in 1987 (filter question and the two follow up questions) are rather different from the single question probing for "feeling of safety while walking in one's neighborhood at night" used subsequently. In fact, the wording of the 1987 question is more reminiscent of the question about "avoiding certain types of areas or people" while out and about in one's neighborhood at night, which was added in subsequent years. At any rate, comparisons of the variable q489_91 with fear indicators of other survey years must be made carefully.

Variables v56, vv56, vvv56, vvvv56, and vvvvv56 (Questionnaire item 500) were transformed into the six dummy variables q500_1 to q500_6, indicating five different precautionary measures people take to reduce the risk of falling victim to a crime.

A.3.2 Disorder

2005

Variables Q30916_1, Q30916_2, and Q30916_3 were transformed into the four dummy variables q30916GR, q30916LI, q30916DP, and q30916AF, indicating graffiti, litter, disorderly people, or all of the above forms of disorder, respectively.

Variables Q30916_1, Q30916_2, and Q30916_3 were further transformed into the two dummy variables q30916PD and q30916SD, indicating physical or social disorder, respectively. q30916PD takes on the value 1 if either of the Variables q30916GR q30916LI or q30916AF has a value of 1. q30916SD takes on the value 1 if either of the Variables q30916DP q30916AF has a value of 1. The six cases that had missing values for Variable Q30916_1 were also set to missing for variables q30916PD and q30916SD.

2000

Variable Q30916 was transformed into the two dummy variables q30916PD and q30916SD, indicating physical or social disorder, respectively. q30916PD takes on the value 1 if variable Q30916 had a value of 1, 2, or 4 and 0 otherwise. q30916SD takes on the value 1 if Q30916 had a value of 3 or 4 and 0 otherwise. Values 8 and 9 were copied and set as missing values for both dummy variables.

1998

The dummy variable q30916 (Questionnaire item 309p) was recoded into the variable q30916rec. A category value of 1 implies the presence of physical or social disorder. The categories of the questionnaire do not distinguish between social and physical forms of disorder.

Variable q299 (Questionnaire item 299) was recoded into variable q299rec. Categories 2 and 3 were interchanged so that a higher category value implies less mutual help among neighbors.

1987

Variable v58 (Questionnaire item 502) was recoded into variable q502rec. Categories 2 and 3 were interchanged so that a higher category value implies less mutual help among neighbors.

A.3.3 Policing

2005

Variable Q31016 was recoded into variable q31016rec to ensure that higher categories indicate the wish for more frequent police patrols.

1987

Variable v65 (Questionnaire item 505; overall appreciation of the police) was recoded into the variable q505rec with the order of categories being reversed. This is in keeping with previous coding practice that a higher category value implies greater dissatisfaction with the police.

B

Community Policing Scorecards

Table B.1 Community policing scorecard: Basel

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Organization and Partnerships							
Organizational reform							
Decentralization	4plus/Optima				√	√	
Flat hierarchy	4plus/Optima				√	√	√
Intelligence sharing	4plus/Optima				√	√	√
Citizen contacts	4plus/Optima				√	√	√
Training							
Basic training							
Soft skills	Transaktionsanalyse	√	√	√	√	√	√
Community policing	Kurstage/Berufsprüfung					√	√
Continued training							
Community policing						√	√
Police administrators							
Community policing	Kaderkurse B/C					√	√
Neighborhood liaison officers	CP-Verantwortliche						
Soft skills	Selbstständige Weiterbildung				√	√	√
Community policing	Selbstständige Weiterbildung				√	√	√
Strategic partnerships							
Social services	Polizeilicher Sozialdienst	√	√	√	√	√	√
Street workers	Schwarzer Peter/Mittler im öffentlichen Raum				√	√	√
Public works	Amt für Stadtreinigung/Industrielle Werke Basel				√	√	√
Parks services	Amt für Stadtgärtnerei				√	√	√

Table B.1 Community policing scorecard: Basel (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Schools	Besondere Präventions-abteilung						√
Neighborhood associations	Runde Tische Quartiere				√	√	√
Business associations/retailers	Community Information System					√	√
Objectives							
Community policing manual	CP Broschüren 1998/2003				√	√	√
Theoretical basis	CP Broschüre 2003/ Lehrmittel SPI					√	√
Crime reduction	CP Broschüren 1998/2003				√	√	√
Fear reduction	CP Broschüren 1998/2003				√	√	√
Disorder reduction	CP Broschüren 1998/2003				√	√	√
Strategic partnerships	CP Broschüren 1998/2003				√	√	√
Police-community relations	CP Broschüren 1998/2003				√	√	√
Decentralization and Deployment							
Patrol officers	Sicherheitspolizei						
Long-term assignment					√	√	√
Foot/bike patrol					√	√	√
Police Assistants	Polizeidienstangestellte						
Foot/bike patrol		√	√	√	√	√	√
Neighborhood liaison officers	CP-Verantwortliche						
Autonomy					√	√	√
Position of authority					√	√	√
Foot/bike patrol						√	√
Round tables					√	√	√
Primary task					√	√	√
Broken Windows Approach							

Table B.1 Community policing scorecard: Basel (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Clean-up campaigns							
Graffiti	Baudepartement/Malermeisterverband					√	√
Parks	Amt für Stadtgärtnerei				√	√	√
Streets	Amt für Stadtreinigung				√	√	√
Management of public spaces							
Reducing social disorder	Interventionen CP-Verantwortliche/Platzhirsch					√	√
Anti-loitering laws	Wegweisungsartikel						√
Problem-solving							
SARA method							
Neighborhood liaison officers	CP-Verantwortliche					√	√
Best practices	CP-Plattform				√	√	√
Hotspot policing							
Operational data analysis	Lagezentrum/Kriminalanalysestelle				√	√	√
Special intervention units	Schwerpunktbildung	√	√	√	√	√	√
Homeless/Alcoholics							
Shelters/drop-in centers	Schwarzer Peter				√	√	√
Drug addicts							
Shelters/drop-in centers	Suchthilfe Region Basel				√	√	√
Referrals/deportation	Mittler im öffentlichen Raum				√	√	√

Table B.1 Community policing scorecard: Basel (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Prostitution							
Interagency cooperation	Aids-Hilfe beider Basel, Frauen-Oase				√	√	√
Youth violence							
Interagency cooperation	Präventionsabteilung Gewaltprävention					√	√
Neighborhood regeneration projects	Ripa forte/Integrale Aufwertung Basel Nord				4057, 4058	4057, 4058	4057, 4058
Crime prevention							
Counseling services	CP-Verantworliche/ Abteilung Prävention	√	√	√	√	√	√
Pro-active campaigns	Senioren			?	√	√	√
Public campaigns	Zelte/Infomobil				√	√	√
Crime alerts	CIS/Sicherheitspartner-schaft Detailhandel				√	√	√
Performance Appraisal Systems							
Evaluation/Monitoring							
Controlling of effectiveness	Jahresplanung/Qualitäts-zirkel/Monitoring	√	√	√	√	√	√
Project evaluations	Zielvereinbarungen				√	√	√
Citizen surveys							
Police department	Konso-Studien/ICVS	√	√	√	√	√	√

Table B.1 Community policing scorecard: Basel (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
City government	Kantonales statistisches Amt					√	√
Equity and Legitimacy							
Victim assistance							
Police officer training	Grundausbildung			√	√	√	√
Referral to counseling centers	Opferhilfe-Beratungsstellen			√	√	√	√
Ethics and human rights							
Police complaints board	Beschwerdestelle Justiz- und Sicherheitsdepartement, kantonale Ombudsstelle	√	√	√	√	√	√
Police ethics code	Werte- und Bekenntnissystem, Lehrmittel EMRK und Polizeidienst			?	√	√	√
Minorities							
Cultural diversity training	Grundausbildung/Berufsprüfung					√	√
Meetings ethnic/religious groups	CP-Spezialisten					√	√
Minority police officers	Niedergelassene (C-Bewilligung)				√	√	√

Table B.2 Community policing scorecard: Bern

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Organization and Partnerships							
Organizational reform							
Decentralization	APOLLO				√	√	√
Flat hierarchy	APOLLO				√	√	√
Intelligence sharing	APOLLO				√	√	√
Citizen contacts	APOLLO				√	√	√
Training							
Basic training							
Soft skills	Reform 1983/84	√	√	√	√	√	√
Community policing	Pilotprojekt SPI/ Berufsprüfung					√	√
Continued training							
Community policing						√	√
Police administrators							
Community policing	Modulausbildung CP						√
Neighborhood liaison officers	GfS/CP Spezialisten						
Soft skills	Polizeilehrerkurs					√	√
Community policing	Selbstständige Weiterbildung					√	√
Strategic partnerships							
Social services	Koordinationsgruppe Drogen, Drogenkonferenz				√	√	√
Street workers	PINTO					√	√
Public works	Tiefbauamt					√	√
Parks services	Stadtgärtnerei			√	√	√	√
Schools	Koordinationsgruppe Drogen, Drogenkonferenz				√	√	√

Table B.2 Community policing scorecard: Bern (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Neighborhood associations	Quartierleisten/-vereine, Beschwerdemanagement Stadtteil sechs					√	√
Business associations/retailers	Sicherheitspartnerschaften mit Dienstleistungs- und Gewerbetrieben					√	√
Objectives							
Community policing manual	Grundsatzpapier GfS				√	√	√
Theoretical basis	Lehrmittel SPI					√	√
Crime reduction	Grundsatzpapier/Strategisches Konzept GfS/CP				√	√	√
Fear reduction	Grundsatzpapier/Strategisches Konzept GfS/CP				√	√	√
Disorder reduction							
Strategic partnerships	Grundsatzpapier/Strategisches Konzept GfS/CP				√	√	√
Police-community relations	Grundsatzpapier/Strategisches Konzept GfS/CP				√	√	√
Decentralization and Deployment							
Patrol officers	Sicherheitspolizei, Grundversorgung						
Long-term assignment		√	√	√			
Foot/bike patrol		√	√	√	√	√	√
Police Assistants							
Foot/bike patrol	APOLLO				√	√	√
Neighborhood liaison officers	GfS/CP Spezialisten						
Autonomy						√	√
Position of authority						√	√

Table B.2 Community policing scorecard: Bern (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Foot/bike patrol						√	√
Round tables						√	√
Primary task						√	√
Broken Windows Approach							
Clean-up campaigns							
Graffiti	Casa Blanca					3011	√
Parks	City Pflege/Stadtgärtnerei				√	√	√
Streets	Tiefbauamt, Betrieb und Unterhalt					√	√
Management of public spaces							
Reducing social disorder	PINTO					3011	√
Anti-loitering laws	Wegweisungsartikel				3011, 3012	3011, 3012	3011, 3012
Problem-solving							
SARA method							
Neighborhood liaison officers	GfS/CP Spezialisten					√	√
Best practices	Wochenrapport GfS					√	√
Hotspot policing							
Operational data analysis	Lagezentrum/Kriminalanalysestelle				√	√	√
Special intervention units	Einsatzgruppe Krokus/ Schwerpunktbildung			√	√	√	√
Homeless/Alcoholics							
Shelters/drop-in centers	Aufenthaltsraum Christoffelgasse					√	√

Table B.2 Community policing scorecard: Bern (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Drug addicts							
Shelters/drop-in centers	Münstergasse/Hodlergasse		√	√	√	√	√
Referrals/deportation	Ambulante Vermittlung und Rückführung/PINTO				√	√	√
Prostitution							
Interagency cooperation	Beratungsstelle Xenia					?	√
Youth violence							
Interagency cooperation	Projektgruppe Jugendgewalt, Stadtteilkonferenzen					√	√
Neighborhood regeneration projects							
Crime prevention							
Counseling services	Sicherheitsberatungen, GfS/CP Beratungen	√	√	√	√	√	√
Pro-active campaigns	Senioren					√	√
Public campaigns	GfS/CP Auftritte					√	√
Crime alerts	Sicherheitspartnerschaften					√	√
Performance Appraisal Systems							
Evaluation/Monitoring							
Controlling of effectiveness	Jahresbericht GfS-Koordinator					√	√

Table B.2 Community policing scorecard: Bern (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Project evaluations	Projektauswertungen GfS-Spezialisten, Dienstchef Prävention Bern					√	√
Citizen surveys							
Police department	ICVS				√	√	√
City government	Statistikdienste der Stadt Bern			√	√	√	√
Equity and Legitimacy							
Victim assistance							
Police officer training	Grundausbildung		√	√	√	√	√
Referral to counseling centers	Opferhilfe-Beratungsstellen			√	√	√	√
Ethics and human rights							
Police complaints board	Städtische/kantonale Ombudsstelle				√	√	√
Police ethics code	Leitfaden Polizeiethik	√	√	√	√	√	√
Minorities							
Cultural diversity training	Grundausbildung/Berufsprüfung				√	√	√
Meetings ethnic/religious groups	Polizeiwachen				√	√	√
Minority police officers							

Table B.3 Community policing scorecard: Geneva

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Organization and Partnerships							
Organizational reform							
Decentralization	Plan Pégase/Proxipol			√	√	√	√
Flat hierarchy							
Intelligence sharing	Plan Pégase/Proxipol			√	√	√	√
Citizen contacts	Plan Pégase/Proxipol			√	√	√	√
Training							
Basic training							
Soft skills	L'analyse transactionnelle		√	√	√	√	√
Community policing	Brevet fédéral					√	√
Continued training							
Community policing	Travail en partenariat et médiation				√	√	√
Police administrators							
Community policing							
Neighborhood liaison officers	Îlotiers de quartier						
Soft skills	Formation psychologique en négociation et résolution de conflits			√	√	√	√
Community policing							
Strategic partnerships							
Social services	Collaboration lutte contre toxicomanie			√	√	√	√
Street workers	FASe/Travailleurs sociaux hors murs				√	√	√
Public works	Voiries municipales				√	√	√
Parks services	Services des espaces verts municipaux				√	√	√

Table B.3 Community policing scorecard: Geneva (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Schools	Coordinations de quartier, Protocole DIP-DI			√	√	√	√
Neighborhood associations	Associations de quartier			√	√	√	√
Business associations/retailers	Commerçants			√	√	√	√
Objectives							
Community policing manual	Plan Pégase/Proxipol			√	√	√	√
Theoretical basis				√	√	√	√
Crime reduction	Plan Pégase/Proxipol			√	√	√	√
Fear reduction	Plan Pégase/Proxipol			√	√	√	√
Disorder reduction	Proxipol					√	√
Strategic partnerships	Plan Pégase/Proxipol			√	√	√	√
Police-community relations	Plan Pégase/Proxipol			√	√	√	√
Decentralization and Deployment							
Patrol officers	Gendarmerie						
Long-term assignment		√	√	√	√	√	√
Foot/bike patrol	Plan Pégase			√	√	√	√
Police Assistants	ASM/APM						
Foot/bike patrol					√	√	√
Neighborhood liaison officers	Îlotiers de quartier						
Autonomy				√	√	√	√
Position of authority				√	√	√	√
Foot/bike patrol				√	√	√	√
Round tables				√	√	√	√
Primary task				√	√	√	√
Broken Windows Approach							

Table B.3 Community policing scorecard: Geneva (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Clean-up campaigns							
Graffiti	Assurance anti-tag						√
Parks							√
Streets	Plan Propriété canton-communes, campagnes de sensibilisation					√	√
Management of public spaces							
Reducing social disorder	Plan Pégase/ ASM/APM/Opérations de police intensives			√	√	√	√
Anti-loitering laws							√
Problem-solving							
SARA method							
Neighborhood liaison officers	Îlotiers de quartier						√
Best practices	Réunions des îlotiers, Rapport Sécurité-Tranquillité-Ordre Public			√	√	√	√
Hotspot policing							
Operational data analysis	Centre de situation/Service des études stratégiques			√	√	√	√
Special intervention units	Peloton de Gendarmerie Mobile/Opération Figaro					√	√
Homeless/Alcoholics							

Table B.3 Community policing scorecard: Geneva (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Shelters/drop-in centers	Centre-Espoir	√	√	√	√	√	√
Drug addicts							
Shelters/drop-in centers	Quai 9					√	√
Referrals/deportation	Unité Mobile d'Urgences Sociales, Fondation des services d'aide et de soins à domicile					√	√
Prostitution							
Interagency cooperation	Aspasie, Departement de police et justice		√	√	√	√	√
Youth violence							
Interagency cooperation	Coordinations de quartier, Protocole DIP-DI			√	√	√	√
Neighborhood regeneration projects	Vernier-Le Lignon, Les Avanchets					1214, 1219, 1220	1214, 1219, 1220
Crime prevention							
Counseling services	POL-SHOP, Postes de quartier	√	√	√	√	√	√
Pro-active campaigns	Îlotiers de quartier/BRNP			√	√	√	√
Public campaigns	Foires, fêtes	√	√	√	√	√	√
Crime alerts	Commerçants					√	√
Performance Appraisal Systems							

Table B.3 Community policing scorecard: Geneva (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Evaluation/Monitoring							
Controlling of effectiveness	Cellule Pégase/Diagnostic local de sécurité, Datapol					√	√
Project evaluations	Opération Figaro						√
Citizen surveys							
Police department	Diagnostic local de sécurité/ICVS				√	√	√
City government							
Equity and Legitimacy							
Victim assistance							
Police officer training	Formation de base		√	√	√	√	√
Referral to counseling centers	Centre de consultation LAVI			√	√	√	√
Ethics and human rights							
Police complaints board	Commissariat à la déontologie				√	√	√
Police ethics code	Code de déontologie				√	√	√
Minorities							
Cultural diversity training	Formation de base/Brevet fédéral				√	√	√
Meetings ethnic/religious groups	Mondial Contact/ réunions BIE					√	√
Minority police officers							

Table B.4 Community policing scorecard: Lausanne

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Organization and Partnerships							
Organizational reform							
Decentralization							
Flat hierarchy	Réforme du statut du policier						√
Intelligence sharing	Réunions chefs des postes de quartier et direction de police secours et police judiciaire						√
Citizen contacts	Concept PdP 1996				√	√	√
Training							
Basic training							
Soft skills	Formation psychologique			√	√	√	√
Community policing	Formation de base/Brevet fédéral				√	√	√
Continued training							
Community policing							
Police administrators							
Community policing							
Neighborhood liaison officers	Chefs des postes de quartier						
Soft skills							
Community policing							
Strategic partnerships							
Social services	Groupe de travail de la municipalité					√	√
Street workers	UnISET/Travailleurs sociaux hors murs					√	√

Table B.4 Community policing scorecard: Lausanne (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Public works	Services des travaux/ Services industriels					√	√
Parks services	Services des parcs et promenades					√	√
Schools	Brigade jeunesse		√	√	√	√	√
Neighborhood associations	Associations de quartier				√	√	√
Business associations/retailers	Commerçants				√	√	√
Objectives							
Community policing manual	Concept PdP 1996				√	√	√
Theoretical basis	Jean-Pierre Harvin				√	√	√
Crime reduction	Concept PdP 1996						
Fear reduction	Concept PdP 1996				√	√	√
Disorder reduction	Concept PdP 1996				√	√	√
Strategic partnerships	Concept PdP 1996				√	√	√
Police-community relations	Concept PdP 1996				√	√	√
Decentralization and Deployment							
Patrol officers	Agents des postes de quartier						
Long-term assignment	Concept PdP 1996				√	√	√
Foot/bike patrol	Concept PdP 1996				√	√	√
Police Assistants	Assistants de police						
Foot/bike patrol		√	√	√	√	√	√
Neighborhood liaison officers	Chef des postes de quartier						
Autonomy					√	√	√
Position of authority					√	√	√
Foot/bike patrol						√	√
Round tables					√	√	√
Primary task					√	√	√

Table B.4 Community policing scorecard: Lausanne (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Broken Windows Approach							
Clean-up campaigns							
Graffiti	Services des travaux					√	√
Parks	Services des parcs et promenades				√	√	√
Streets	Eclairages publics				√	√	√
Management of public spaces							
Reducing social disorder	Plantons de place/ patrouille pédestres/ UnISET	1003, 1005, 1006	1003, 1005, 1006	1003, 1005, 1006	√	√	√
Anti-loitering laws							
Problem-solving							
SARA method							
Neighborhood liaison officers	Cadres dirigeants formés en PdP						√
Best practices							
Hotspot policing							
Operational data analysis	Rapport de synthèse du coordinateur judiciaire				√	√	√
Special intervention units	Groupe d'intervention	√	√	√	√	√	√
Homeless/Alcoholics							
Shelters/drop-in centers	Le Passage						√
Drug addicts							

Table B.4 Community policing scorecard: Lausanne (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Shelters/drop-in centers							
Referrals/deportation	Centre Saint-Martin						√
Prostitution							
Interagency cooperation	Fleur de Pavé					√	√
Youth violence							
Interagency cooperation	Brigade de jeunesse, Groupe de travail interdépartemental, Plate-forme FASL		√	√	√	√	√
Neighborhood regeneration projects	Flon					1003	1003
Crime prevention							
Counseling services	Visites à domicile, renseignements en prévention			√	√	√	√
Pro-active campaigns	Personnes âgées			√	√	√	√
Public campaigns	Foires, fêtes					√	√
Crime alerts	Commerçants, Bijouteries					√	√
Performance Appraisal Systems							
Evaluation/Monitoring							
Controlling of effectiveness	Evaluation des sondages					√	√
Project evaluations	Evaluation de la démarche éthique						√

Table B.4 Community policing scorecard: Lausanne (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Citizen surveys							
Police department	ICVS				√	√	√
City government	Observatoire de la sécurité					√	√
Equity and Legitimacy							
Victim assistance							
Police officer training	Formation de base			√	√	√	√
Referral to counseling centers	Centre de consultation LAVI			√	√	√	√
Ethics and human rights							
Police complaints board	Commission de déontologie						√
Police ethics code	Code de déontologie/ Charte des valeurs						√
Minorities							
Cultural diversity training	Formation de base/Brevet fédéral					√	√
Meetings ethnic/religious groups	Chef des postes de quartier					√	√
Minority police officers							

Table B.5 Community policing scorecard: Zurich

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Organization and Partnerships							
Organizational reform							
Decentralization	Stapo 200X					√	√
Flat hierarchy	Phönix Q				√	√	√
Intelligence sharing	Stapo 200X					√	√
Citizen contacts	Phönix Q				√	√	√
Training							
Basic training							
Soft skills	Psychologische Schulung/Transaktionsanalyse	√	√	√	√	√	√
Community policing	Berufsprüfung					√	√
Continued training							
Community policing							
Police administrators							
Community policing	Weiterbildung Stadtpolizei						√
Neighborhood liaison officers	Kreischefs						
Soft skills	Seminar Psychologie				√	√	√
Community policing							
Strategic partnerships							
Social services	Drogendelegation Stadtrat			√	√	√	√
Street workers	sip züri				√	√	√
Public works	Entsorgung und Recycling Zürich, Tiefbauamt				√	√	√
Parks services	Grün Stadt Zürich				√	√	√
Schools	Task Force Jugendgewalt						√
Neighborhood associations	Gemischtes Drogenforum, Runde Tische Quartiervereine				√	√	√

Table B.5 Community policing scorecard: Zurich (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Business associations/retailers	Gewerbevereine					√	√
Objectives							
Community policing manual	Müller 2006						√
Theoretical basis	Bässmann & Vogt						√
Crime reduction	Müller 2006						√
Fear reduction	Müller 2006						√
Disorder reduction	Müller 2006						√
Strategic partnerships	Müller 2006						√
Police-community relations	Phönix Q/Müller 2006				√	√	√
Decentralization and Deployment							
Patrol officers	Sicherheitspolizei, Grundversorgung						
Long-term assignment	Phönix Q				8032, 8038	√	√
Foot/bike patrol	Phönix Q				8032, 8038	√	√
Police Assistants	Polizeilicher Assistenzdienst						
Foot/bike patrol		√	√	√	√	√	√
Neighborhood liaison officers	Kreischefs/Quartierpolizisten						
Autonomy					√	√	√
Position of authority					√	√	√
Foot/bike patrol					8032, 8038	√	√
Round tables					?	√	√
Primary task					√	√	√
Broken Windows Approach							
Clean-up campaigns							

Table B.5 Community policing scorecard: Zurich (continued)

Dimension	Instrument	1985	1990	1995	2000	2005	2010
Graffiti	Graffiti-Reinigungsabon-nement					√	√
Parks	SiSa, Langstrasse PLUS				8008	√	√
Streets	SiSa, Langstrasse PLUS				8008	√	√
Management of public spaces							
Reducing social disorder	sip züri				√	√	√
Anti-loitering laws	Wegweisungsartikel						√
Problem-solving							
SARA method							
Neighborhood liaison officers	Kreischefs					√	√
Best practices	Monatliche Treffen						√
	Kreischefs "SARA-Review"						
Hotspot policing							
Operational data analysis	Tageslage Fachbereich Analyse, Planung und Einsatzunterstützung					√	√
Special intervention units	Turicum			√	√	√	√
	SMER/Brennpunkt						
Homeless/Alcoholics							
Shelters/drop-in centers	t-alk				√	√	√
Drug addicts							
Shelters/drop-in centers	Kontakt-/Anlaufstellen		√	√	√	√	√
Referrals/deportation	Rückführungen, sip züri			√	√	√	√
Prostitution							

Table B.5 Community policing scorecard: Zurich (continued)

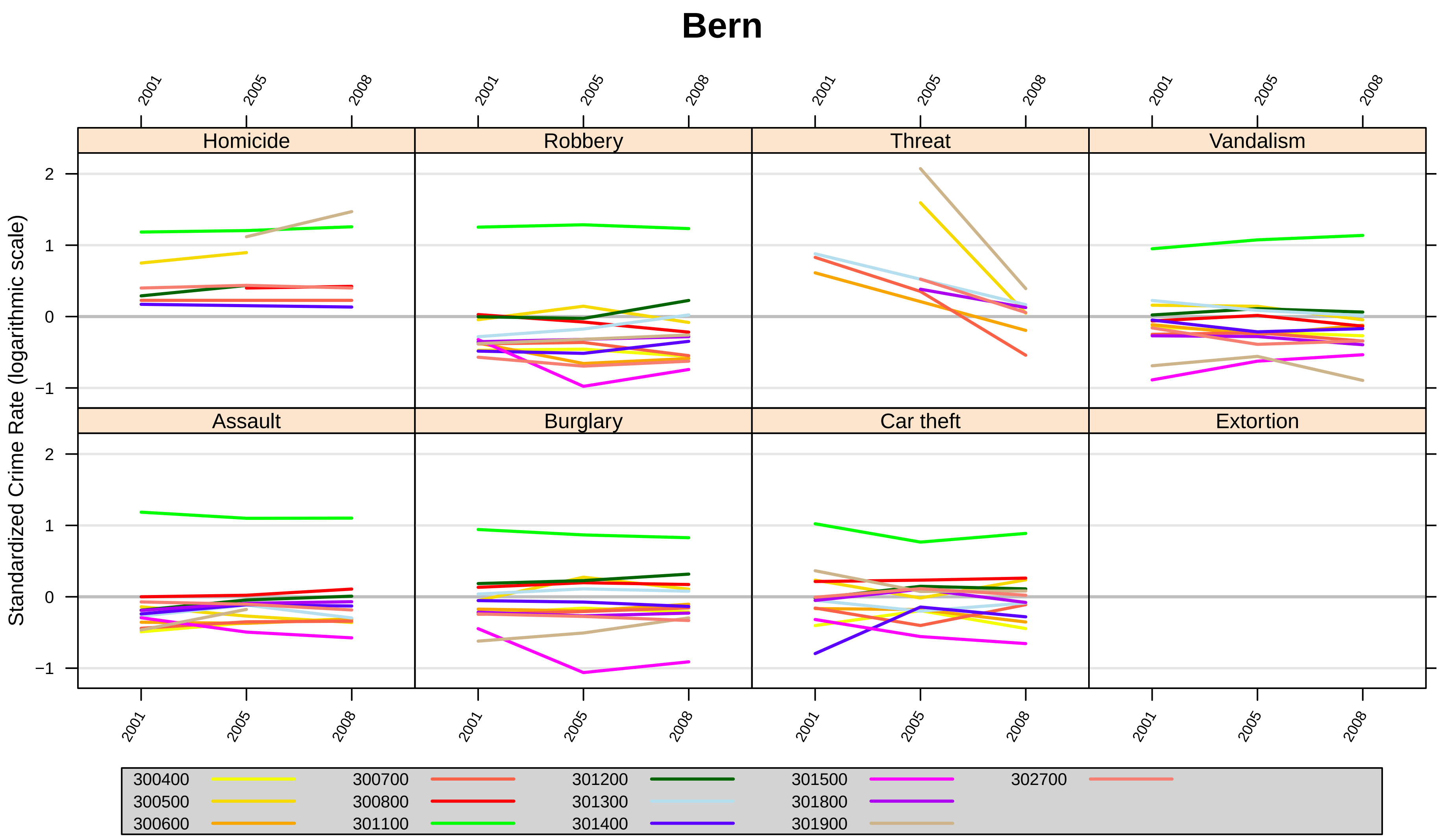
Dimension	Instrument	1985	1990	1995	2000	2005	2010
Interagency cooperation	Langstrasse PLUS, Interdepartementale Arbeitsgruppe					8004, 8005	√
Youth violence							
Interagency cooperation	Task force Jugendgewalt, Kriminalprävention Schulen						√
Neighborhood regeneration projects	Langstrasse PLUS, Seebach, Schwamendingen, Grünau, Hardquartier				8050	8004, 8005, 8050, 8051, 8064	8004, 8005, 8050, 8051
Crime prevention							
Counseling services	Sicherheitsberatungen, Abteilung Prävention	√	√	√	√	√	√
Pro-active campaigns	Abteilung Prävention					√	√
Public campaigns	Abteilung Prävention	√	√	√	√	√	√
Crime alerts	Banken, Bijouterien				√	√	√
Performance Appraisal Systems							
Evaluation/Monitoring							
Controlling of effectiveness	Tageslage APE, LageFIS, POLIS					√	√
Project evaluations	Abteilung Prävention, SiSa, Langstrasse PLUS				√	√	√
Citizen surveys							
Police department	Isopublic/ICVS				√	√	√

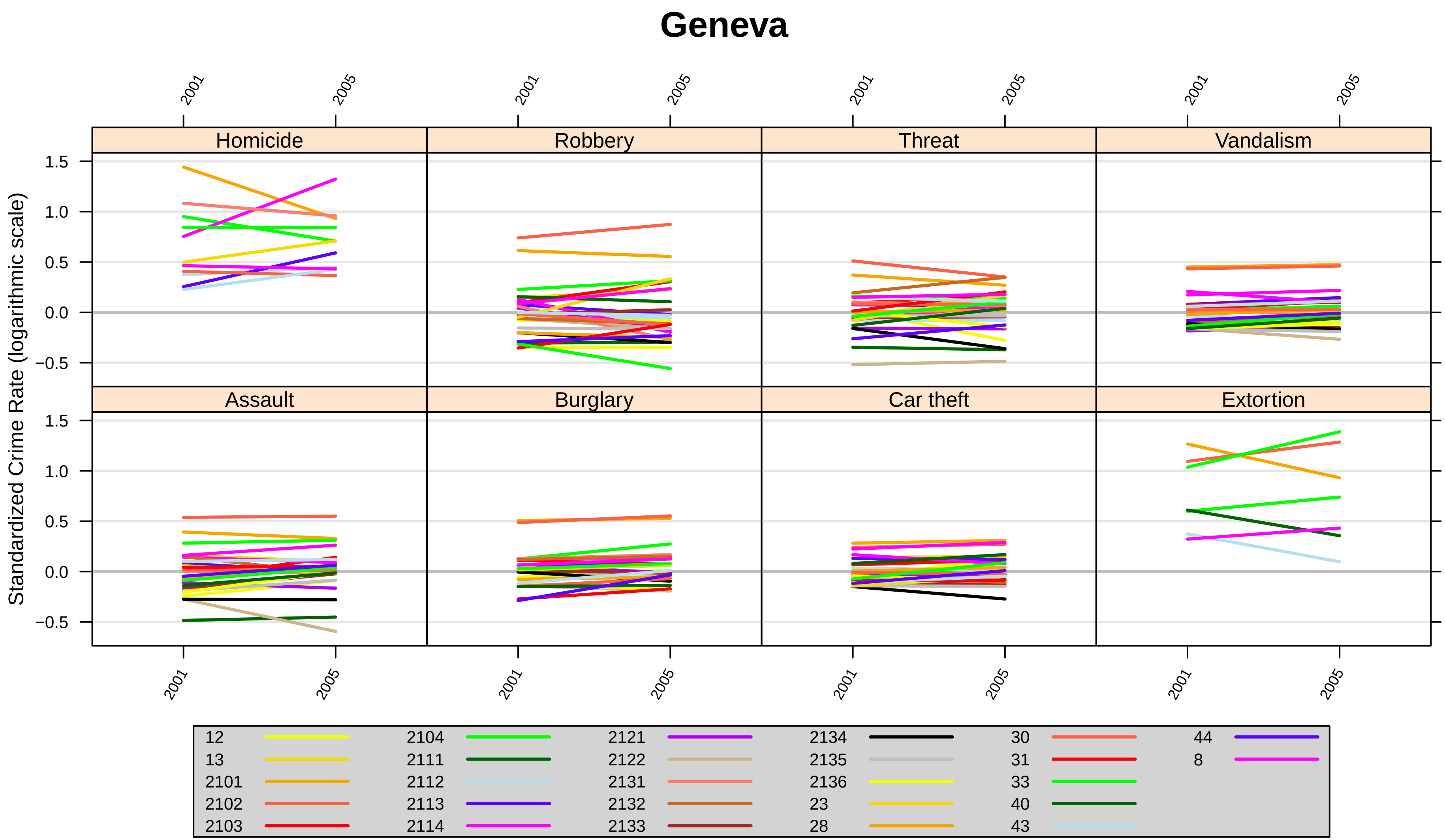
Table B.5 Community policing scorecard: Zurich (continued)

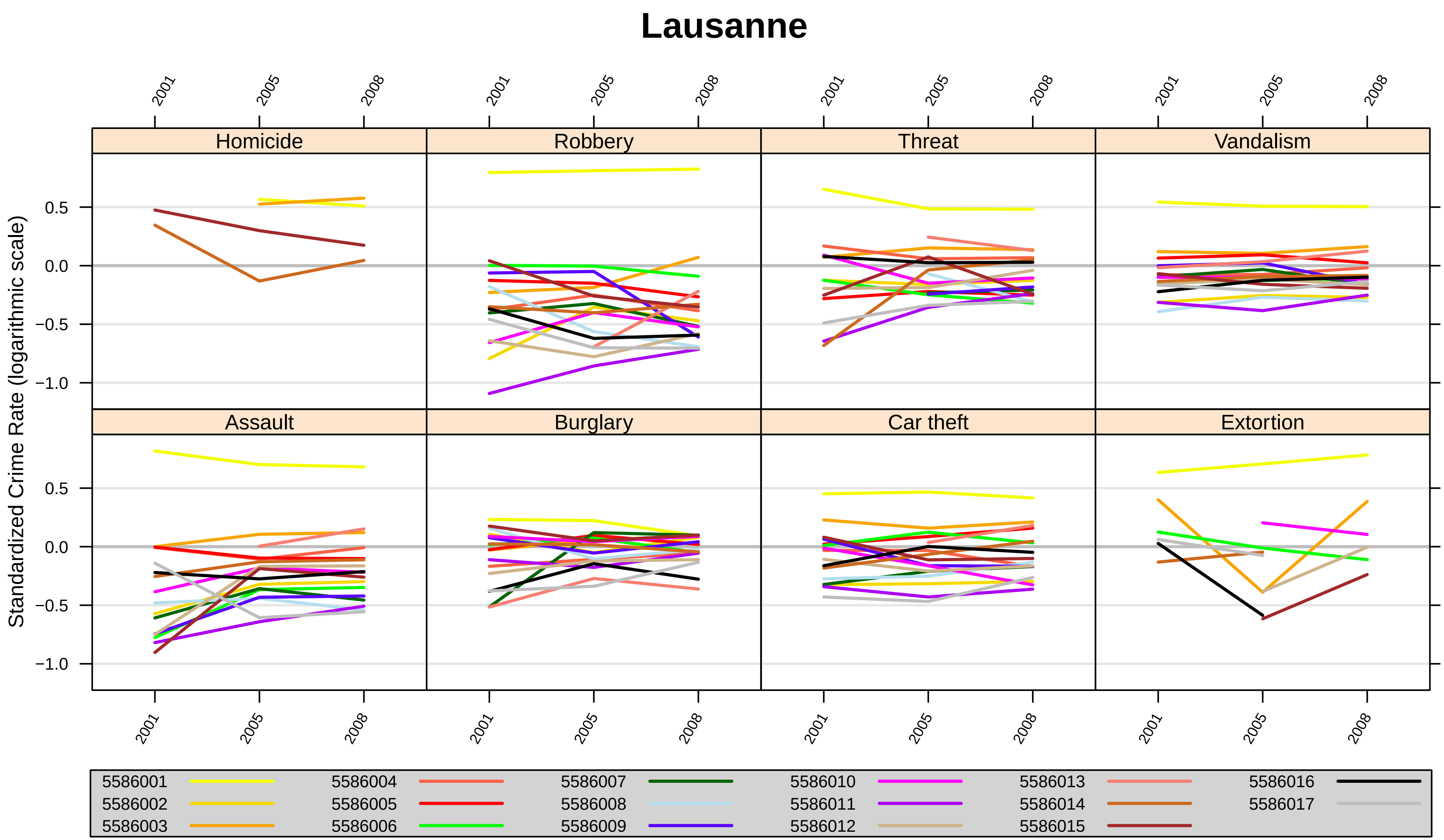
Dimension	Instrument	1985	1990	1995	2000	2005	2010
City government	Stadtentwicklung Zürich				√	√	√
Equity and Legitimacy							
Victim assistance							
Police officer training	Grundausbildung			√	√	√	√
Referral to counseling centers	Opferhilfe-Beratungsstellen			√	√	√	√
Ethics and human rights							
Police complaints board	Städtische Ombudsstelle	√	√	√	√	√	√
Police ethics code	Leitbild Stadtpolizei	√	√	√	√	√	√
Minorities							
Cultural diversity training	Grundausbildung/Berufs-prüfung	√	√	√	√	√	√
Meetings ethnic/religious groups	Kreischefs/Prävention					√	√
Minority police officers							

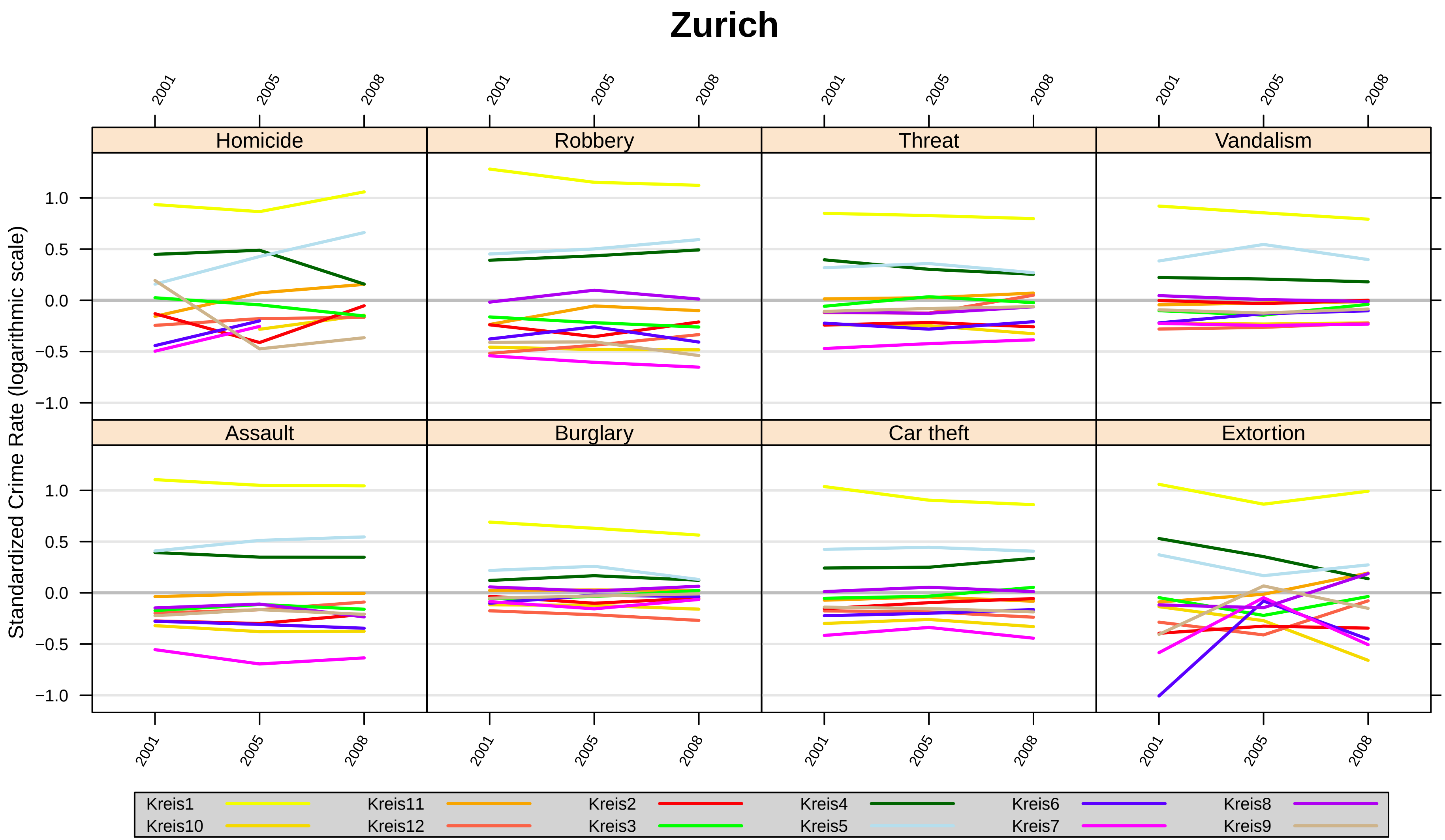
C

Standardized Area-Level Crime Rates


Bern
Homicide
Robbery
Threat
Vandalism
Assault
Burglary
Car theft
Extortion
Standardized Crime Rate (logarithmic scale)
2001
2005
2008
2
1
0
−1
300400
300500
300600
300700
300800
301100
301200
301300
301400
301500
301800
301900
302700


Geneva
Homicide
Robbery
Threat
Vandalism
Assault
Burglary
Car theft
Extortion
Standardized Crime Rate (logarithmic scale)
1.5
1.0
0.5
0.0
−0.5
2001
2005
12
13
2101
2102
2103
2104
2111
2112
2113
2114
2121
2122
2131
2132
2133
2134
2135
2136
23
28
30
31
33
40
43
44
8


Lausanne
Homicide
Robbery
Threat
Vandalism
Assault
Burglary
Car theft
Extortion
Standardized Crime Rate (logarithmic scale)
0.5
0.0
−0.5
−1.0
2001
2005
2008
5586001
5586002
5586003
5586004
5586005
5586006
5586007
5586008
5586009
5586010
5586011
5586012
5586013
5586014
5586015
5586016
5586017

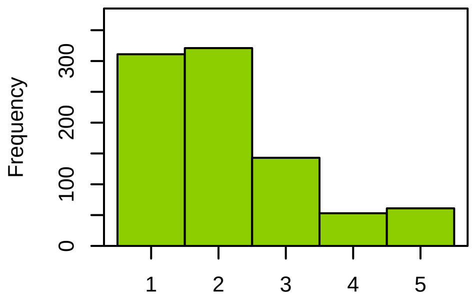

Zurich
Homicide
Robbery
Threat
Vandalism
Assault
Burglary
Car theft
Extortion
Standardized Crime Rate (logarithmic scale)
1.0
0.5
0.0
−0.5
−1.0
2001
2005
2008
Kreis1
Kreis10
Kreis11
Kreis12
Kreis2
Kreis3
Kreis4
Kreis5
Kreis6
Kreis7
Kreis8
Kreis9

D

Analysis of the Outcome Indicators by the Demographic Characteristics of Survey Respondents

Fear of Crime 2005

Histogram



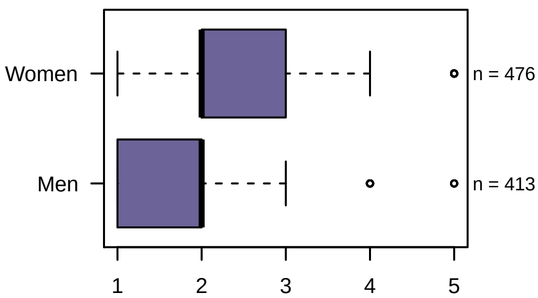
Summary Statistics

Mean = 2.136
Std. Dev. = 1.162
N = 889

Skewness = 1.012
Std. Err. Skewness = 0.08215
p−value = 0

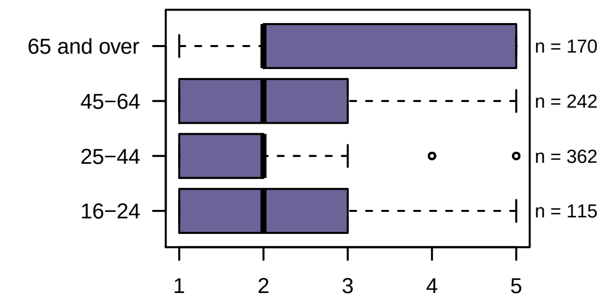
Kurtosis = 0.298
Std. Err. Kurtosis = 0.1643
p−value = 0.03502

Gender



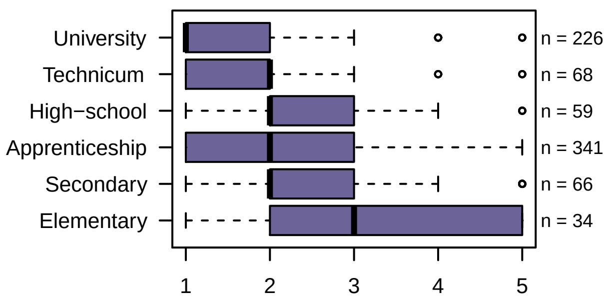
Wilcoxon rank sum test
W = 57802; p = 7.62e−29

Age



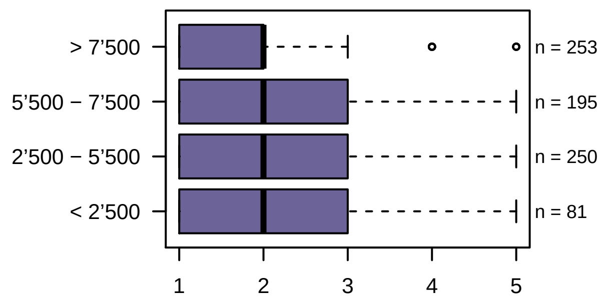
Kruskal−Wallis rank sum test
chi−squared = 64.6; df = 3; p = 6.05e−14

Education



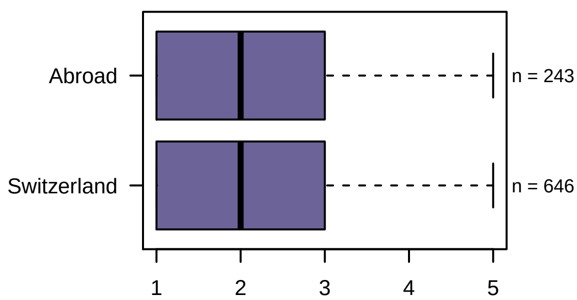
Kruskal−Wallis rank sum test
chi−squared = 76.5; df = 5; p = 4.44e−15

Income



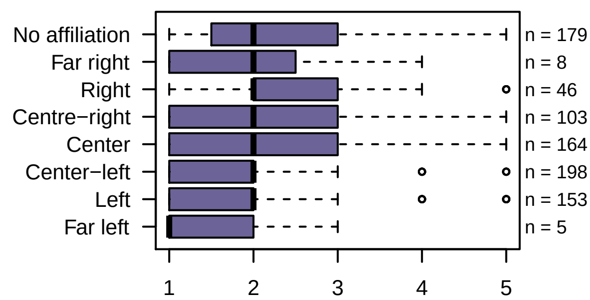
Kruskal−Wallis rank sum test
chi−squared = 31.3; df = 3; p = 7.49e−07

Place of birth



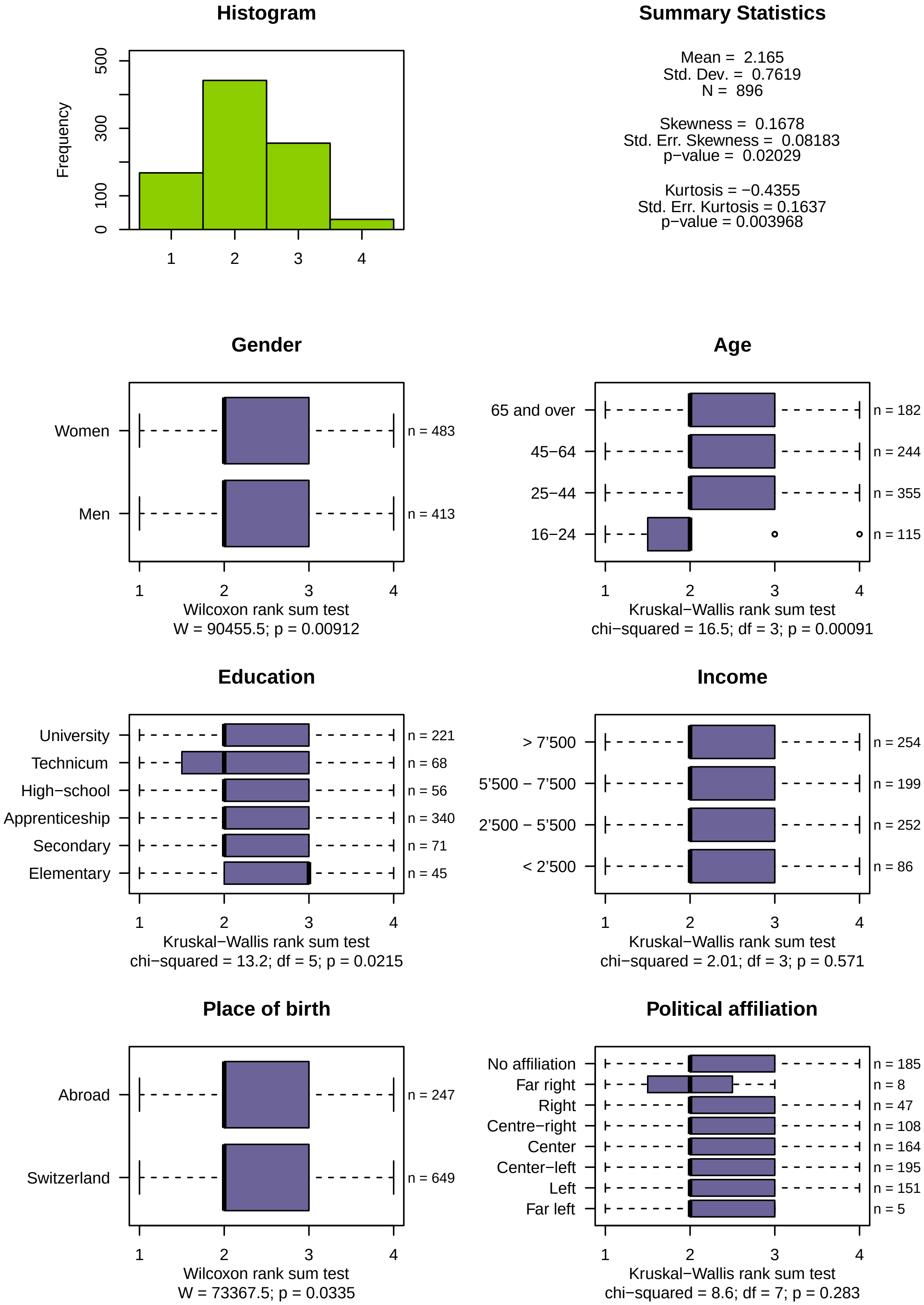
Wilcoxon rank sum test
W = 76903; p = 0.625

Political affiliation

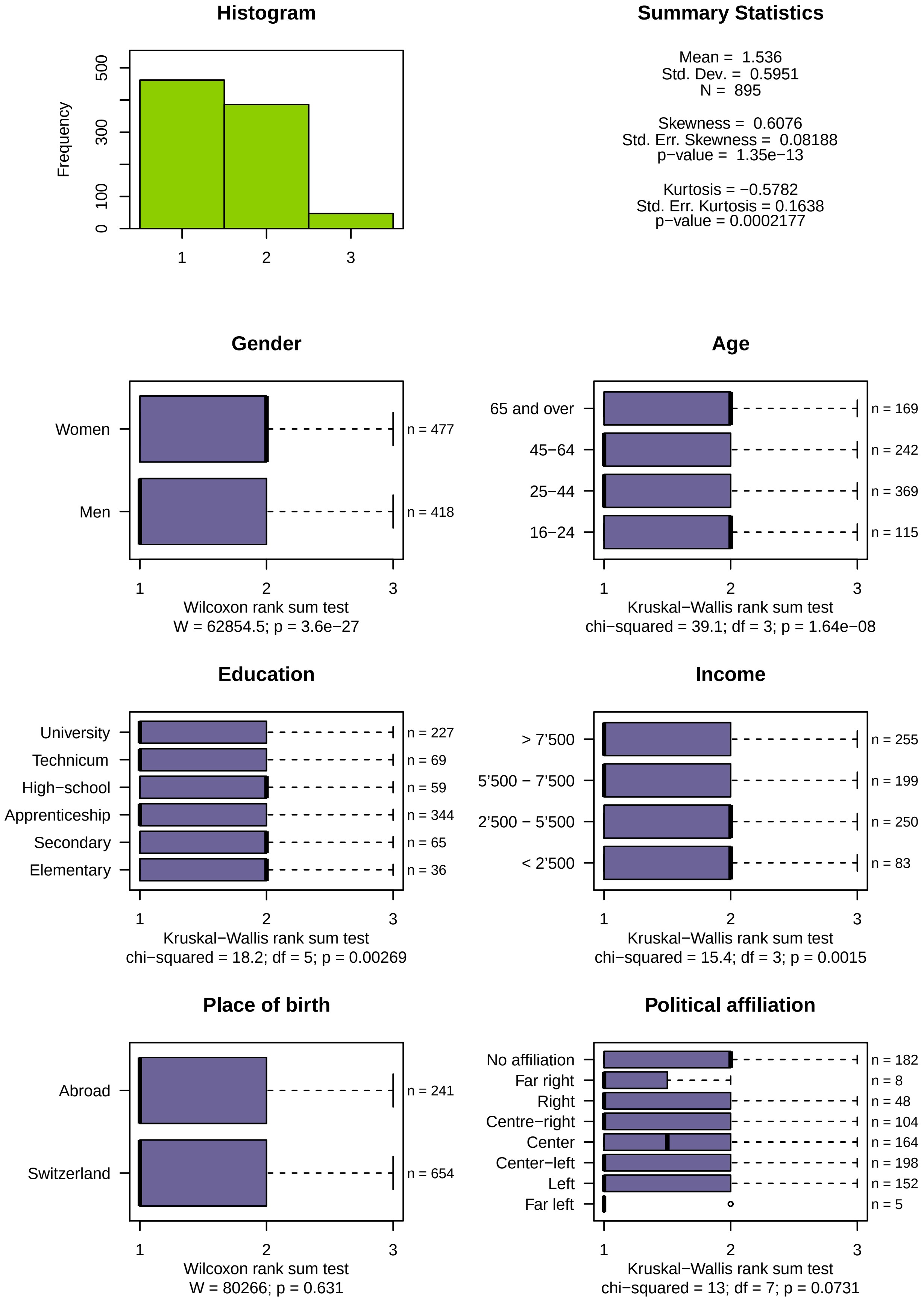


Kruskal−Wallis rank sum test
chi−squared = 44.3; df = 7; p = 1.84e−07

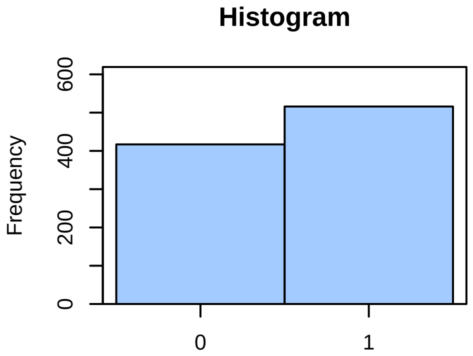
Risk of Victmization 2005



Behavioral Response 2005



Physical Disorder 2005

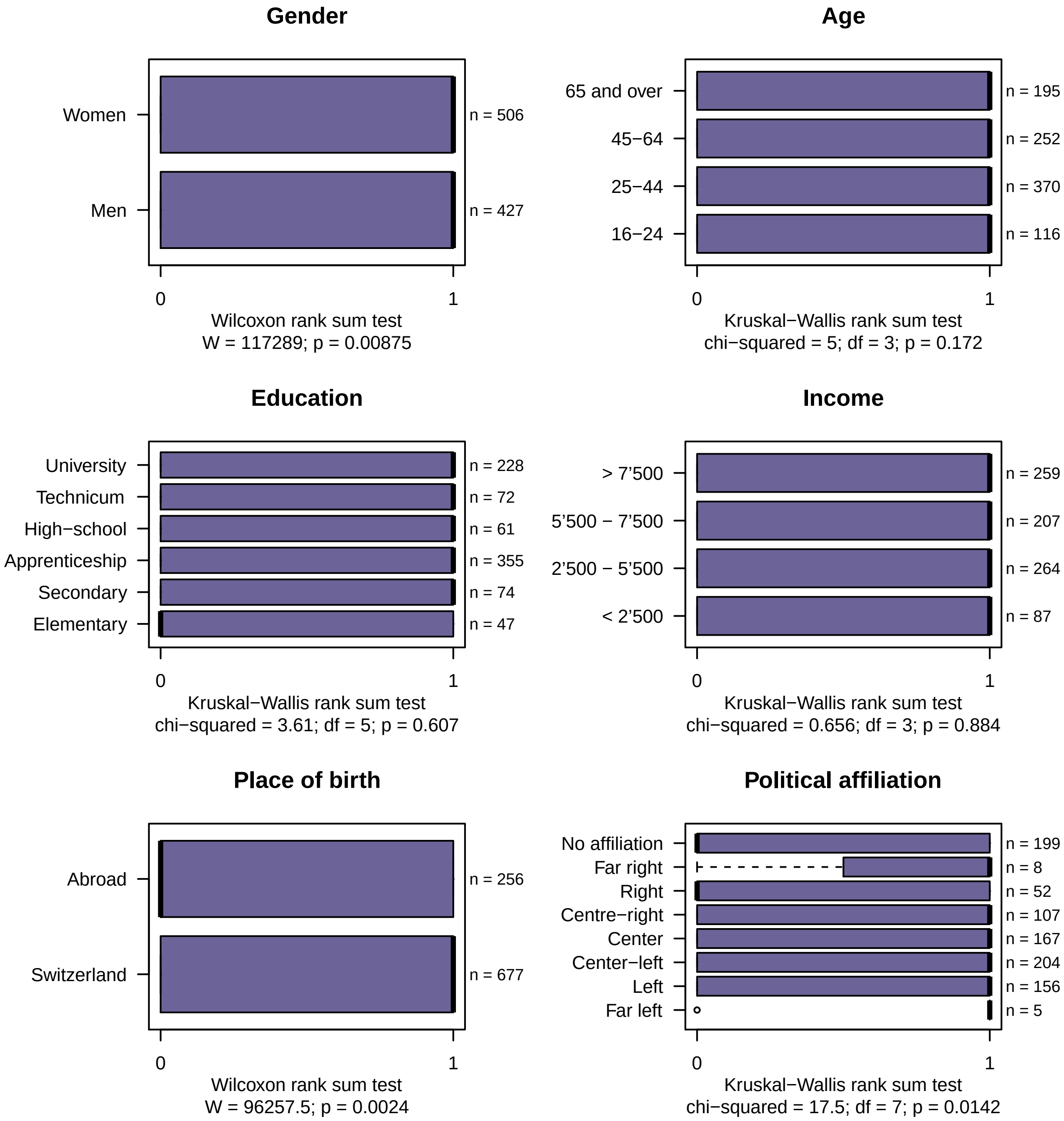


Summary Statistics

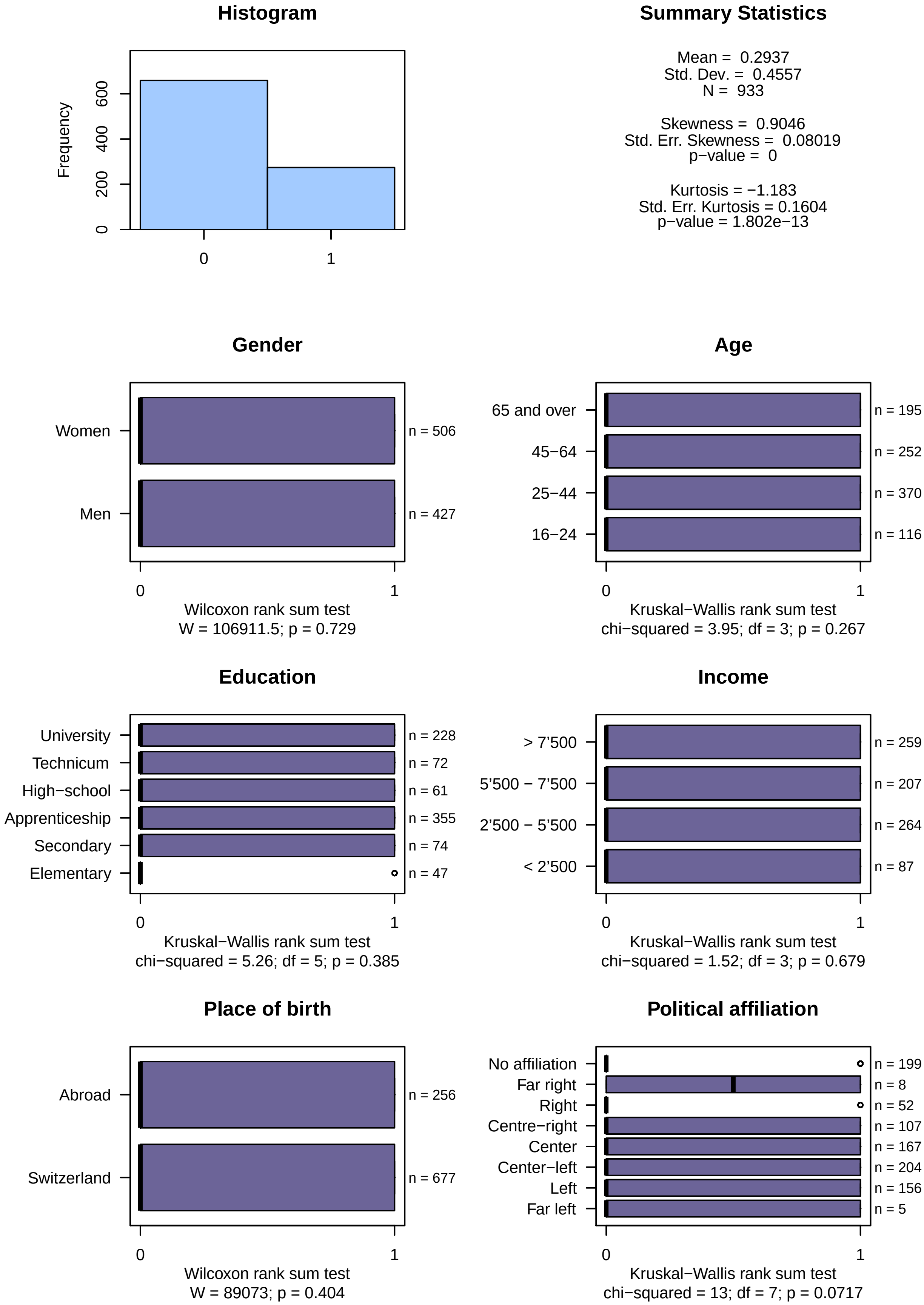
Mean = 0.5531
Std. Dev. = 0.4974
N = 933

Skewness = −0.2131
Std. Err. Skewness = 0.08019
p−value = 0.004008

Kurtosis = −1.957
Std. Err. Kurtosis = 0.1604
p−value = 0

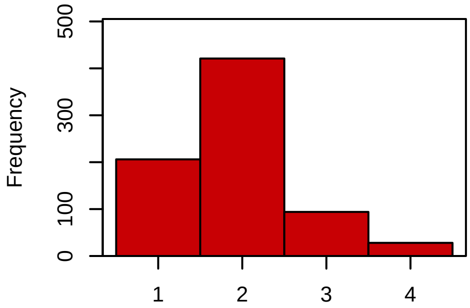


Social Disorder 2005



Police Effectiveness 2005

Histogram



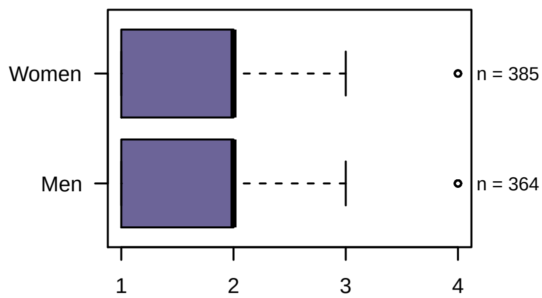
Summary Statistics

Mean = 1.925
Std. Dev. = 0.7384
N = 749

Skewness = 0.6759
Std. Err. Skewness = 0.0895
p−value = 6.295e−14

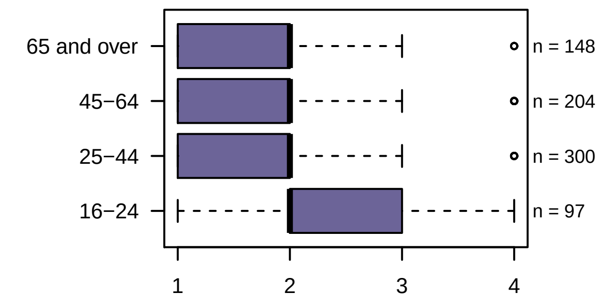
Kurtosis = 0.5719
Std. Err. Kurtosis = 0.179
p−value = 0.0007288

Gender



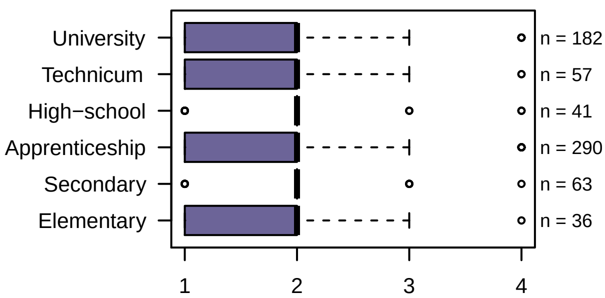
Wilcoxon rank sum test
W = 70373; p = 0.909

Age



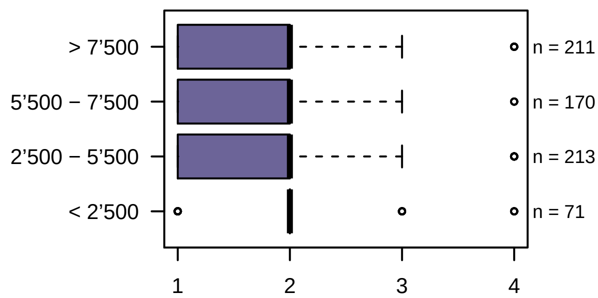
Kruskal−Wallis rank sum test
chi−squared = 12.6; df = 3; p = 0.00557

Education



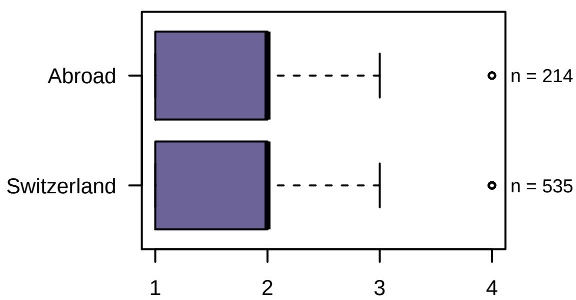
Kruskal−Wallis rank sum test
chi−squared = 1.23; df = 5; p = 0.942

Income



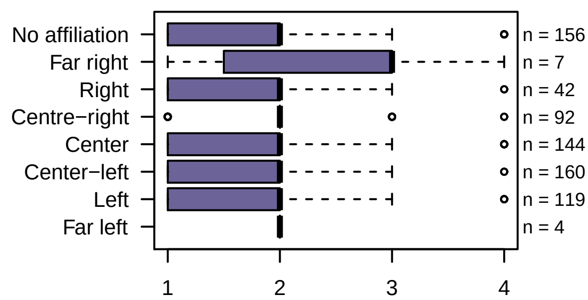
Kruskal−Wallis rank sum test
chi−squared = 2.17; df = 3; p = 0.539

Place of birth



Wilcoxon rank sum test
W = 57778; p = 0.824

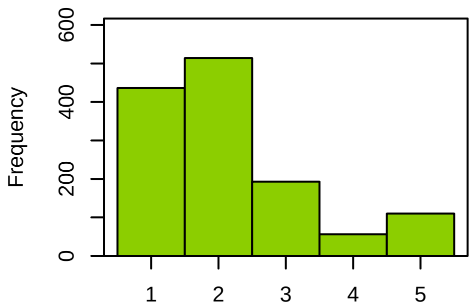
Political affiliation



Kruskal−Wallis rank sum test
chi−squared = 5.11; df = 7; p = 0.647

Fear of Crime 2000

Histogram



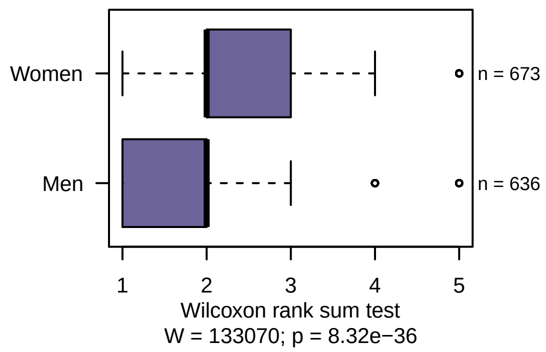
Summary Statistics

Mean = 2.152
Std. Dev. = 1.177
N = 1309

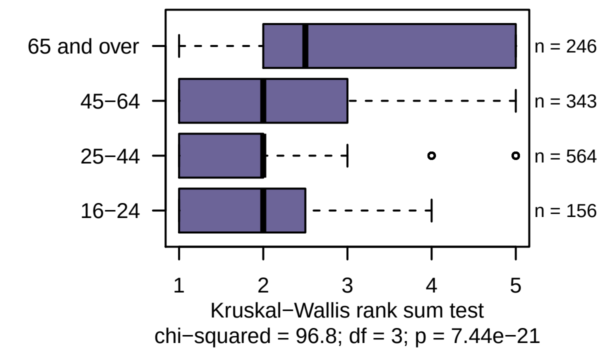
Skewness = 1.097
Std. Err. Skewness = 0.0677
p−value = 0

Kurtosis = 0.4833
Std. Err. Kurtosis = 0.1354
p−value = 0.0001853

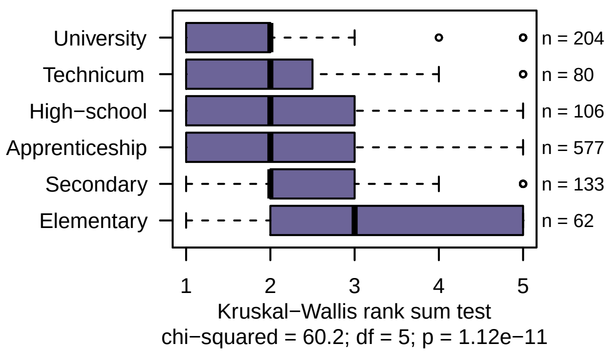
Gender



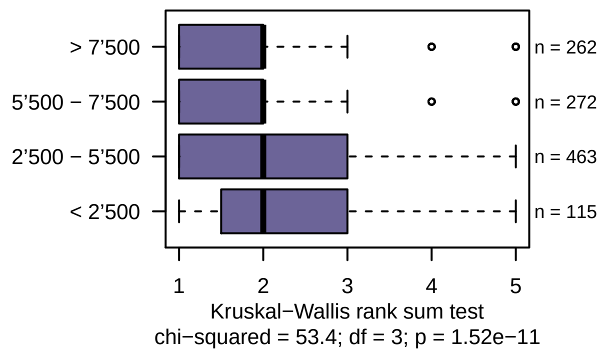
Age



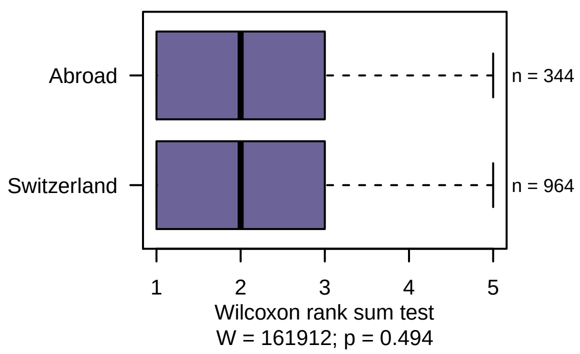
Education



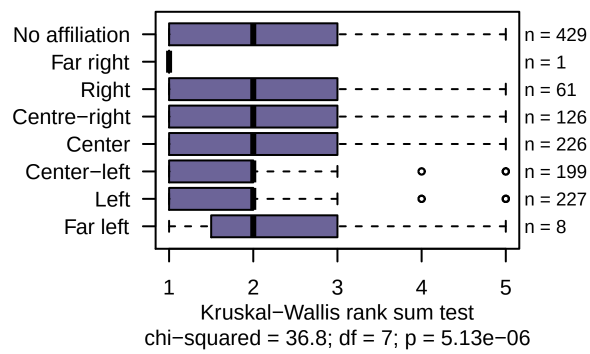
Income



Place of birth

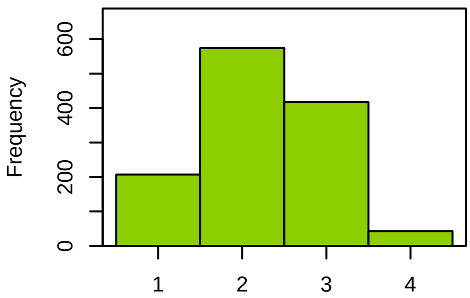


Political affiliation



Risk of Victmization 2000

Histogram



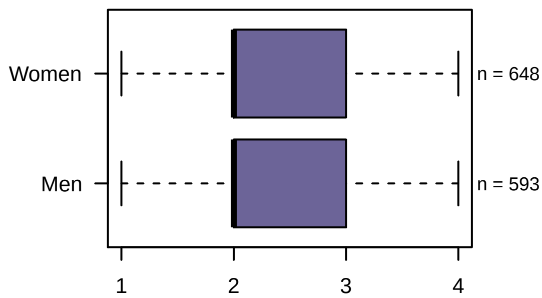
Summary Statistics

Mean = 2.239
Std. Dev. = 0.7649
N = 1241

Skewness = 0.0326
Std. Err. Skewness = 0.06953
p−value = 0.3197

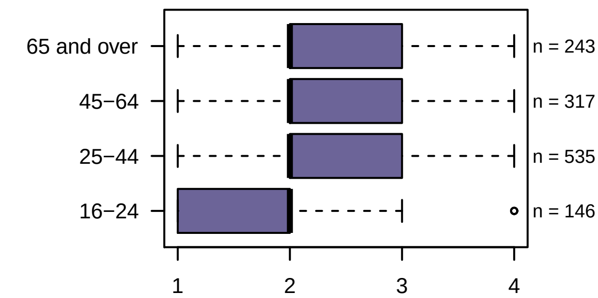
Kurtosis = −0.5439
Std. Err. Kurtosis = 0.1391
p−value = 4.843e−05

Gender



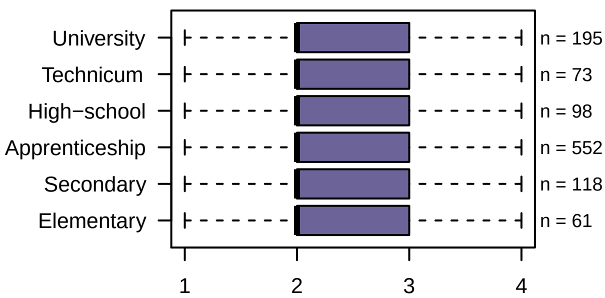
Wilcoxon rank sum test
W = 181297; p = 0.0637

Age



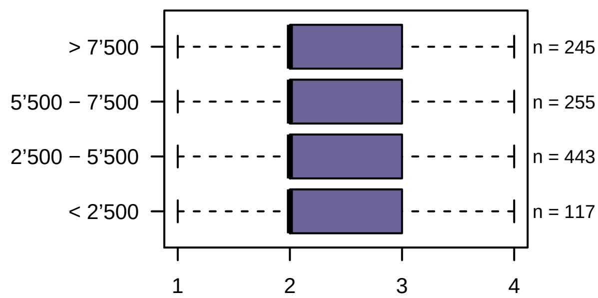
Kruskal−Wallis rank sum test
chi−squared = 22.9; df = 3; p = 4.28e−05

Education



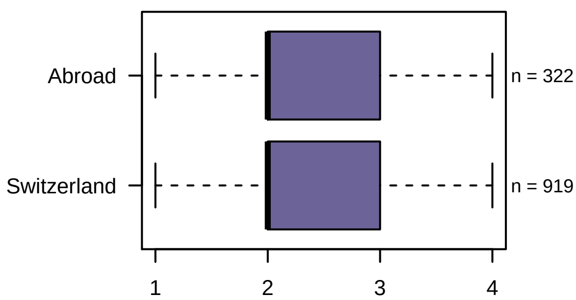
Kruskal−Wallis rank sum test
chi−squared = 10.6; df = 5; p = 0.0597

Income



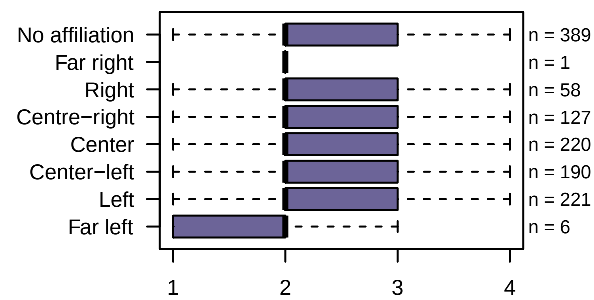
Kruskal−Wallis rank sum test
chi−squared = 2.4; df = 3; p = 0.494

Place of birth



Wilcoxon rank sum test
W = 136947; p = 0.0318

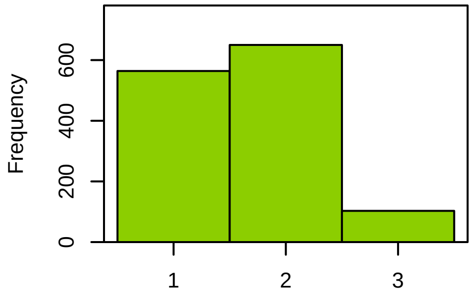
Political affiliation



Kruskal−Wallis rank sum test
chi−squared = 18.5; df = 7; p = 0.00994

Behavioral Response 2000

Histogram



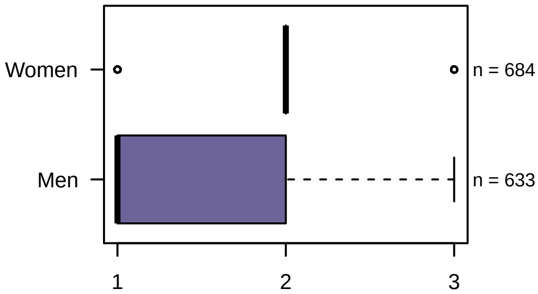
Summary Statistics

Mean = 1.65
Std. Dev. = 0.6199
N = 1317

Skewness = 0.4032
Std. Err. Skewness = 0.0675
p−value = 1.495e−09

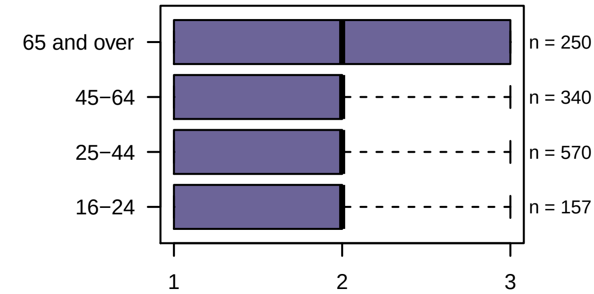
Kurtosis = −0.6722
Std. Err. Kurtosis = 0.135
p−value = 3.605e−07

Gender



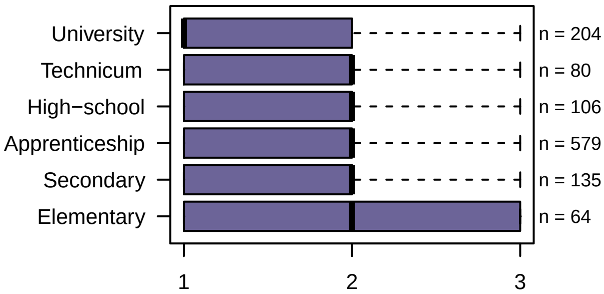
Wilcoxon rank sum test
W = 128651.5; p = 5.67e−46

Age



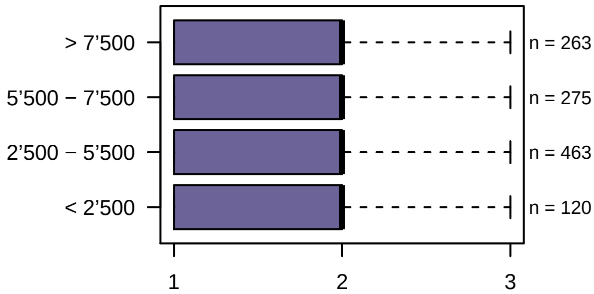
Kruskal−Wallis rank sum test
chi−squared = 67.5; df = 3; p = 1.45e−14

Education



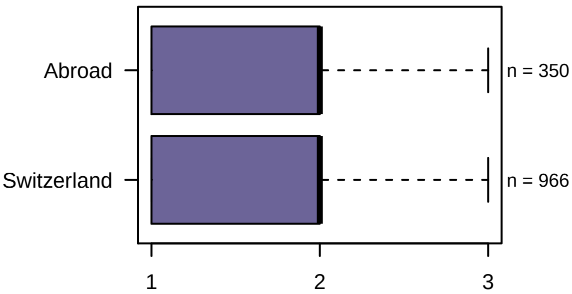
Kruskal−Wallis rank sum test
chi−squared = 30.3; df = 5; p = 1.29e−05

Income



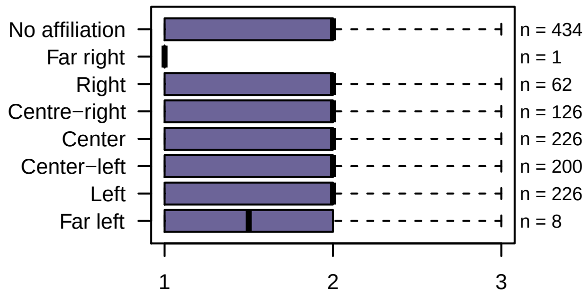
Kruskal−Wallis rank sum test
chi−squared = 29.4; df = 3; p = 1.89e−06

Place of birth



Wilcoxon rank sum test
W = 178987; p = 0.0683

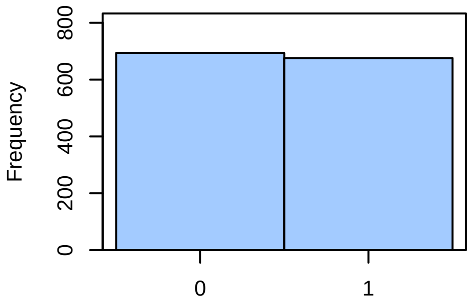
Political affiliation



Kruskal−Wallis rank sum test
chi−squared = 13.2; df = 7; p = 0.0666

Physical Disorder 2000

Histogram



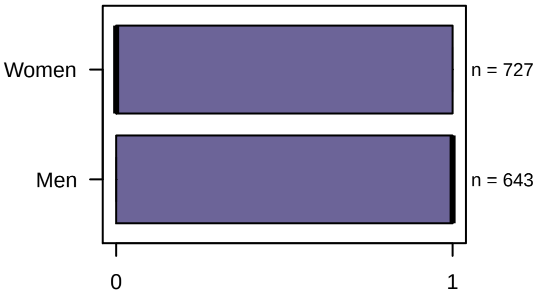
Summary Statistics

Mean = 0.4934
Std. Dev. = 0.5001
N = 1370

Skewness = 0.02625
Std. Err. Skewness = 0.06618
p−value = 0.3458

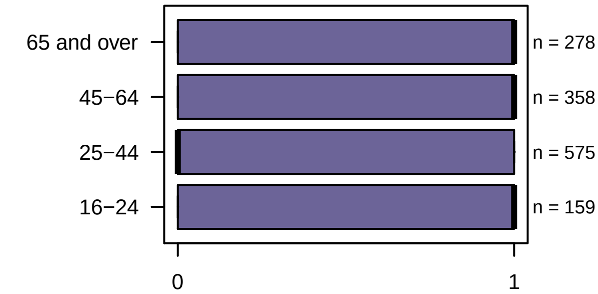
Kurtosis = −2.001
Std. Err. Kurtosis = 0.1324
p−value = 0

Gender



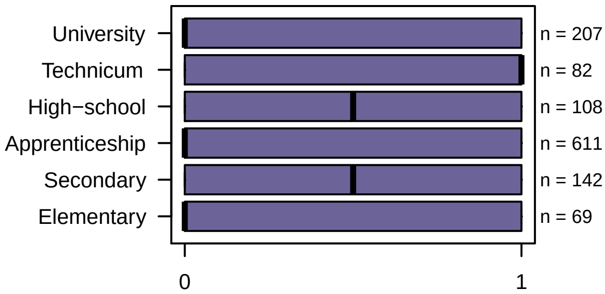
Wilcoxon rank sum test
W = 254776.5; p = 0.000882

Age



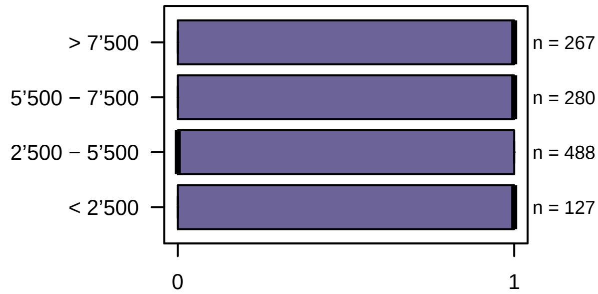
Kruskal−Wallis rank sum test
chi−squared = 13.2; df = 3; p = 0.00425

Education



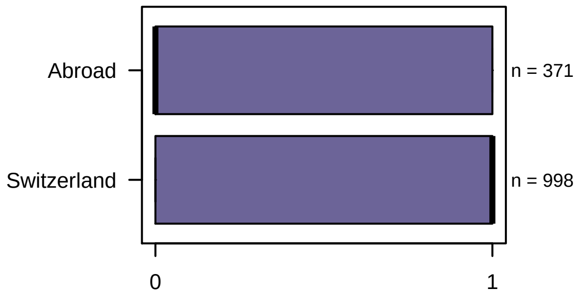
Kruskal−Wallis rank sum test
chi−squared = 0.562; df = 5; p = 0.99

Income



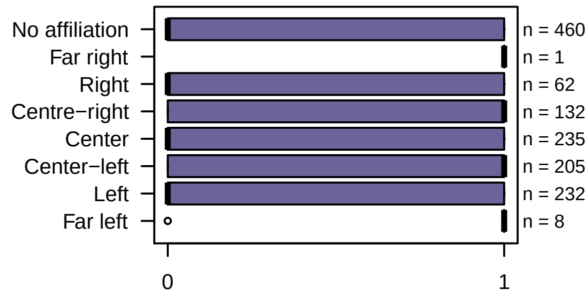
Kruskal−Wallis rank sum test
chi−squared = 7.53; df = 3; p = 0.0568

Place of birth



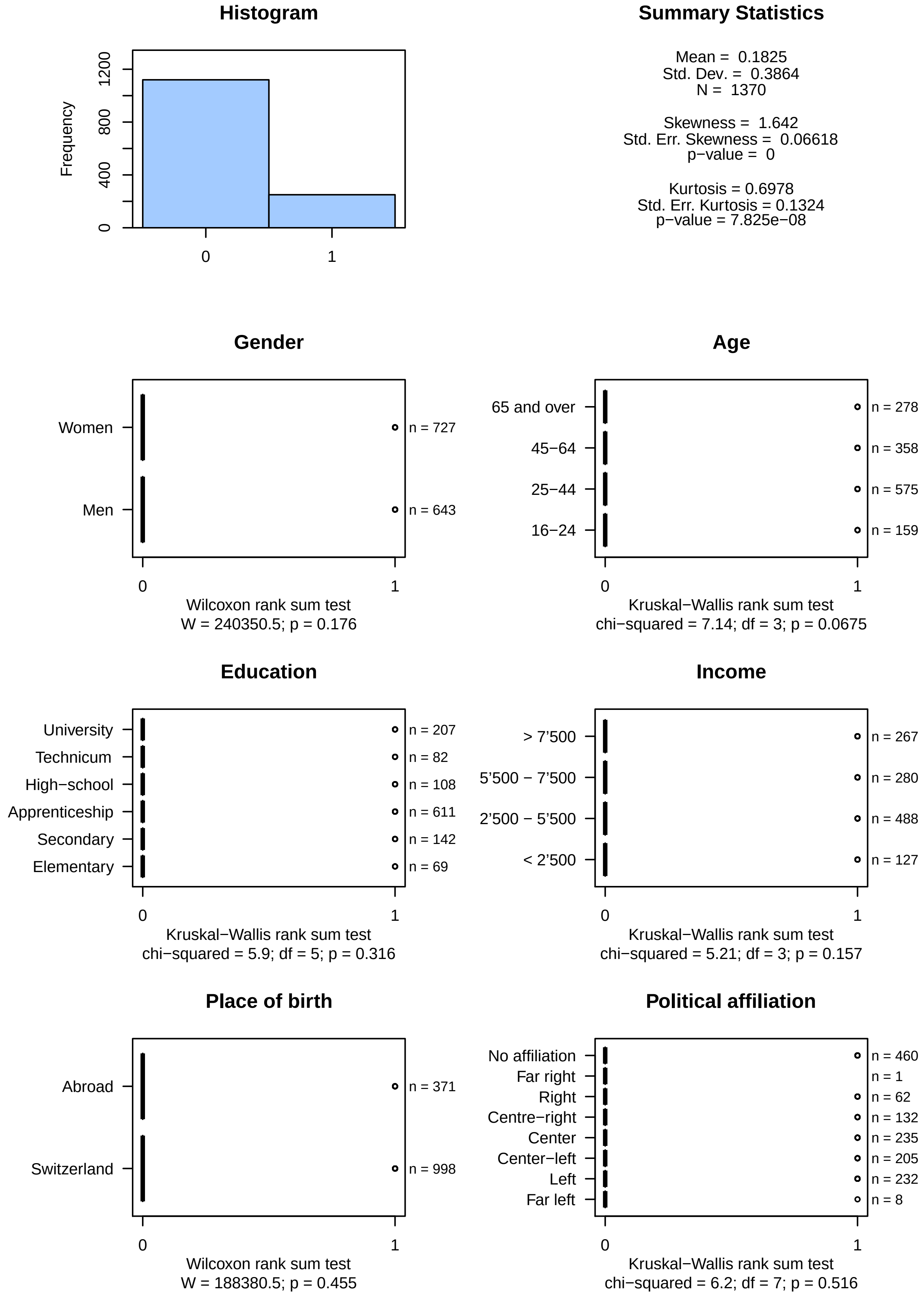
Wilcoxon rank sum test
W = 203559.5; p = 0.00106

Political affiliation



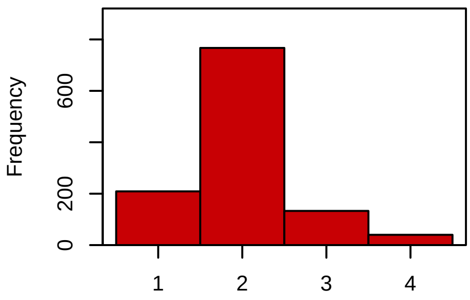
Kruskal−Wallis rank sum test
chi−squared = 15.8; df = 7; p = 0.0269

Social Disorder 2000



Police Effectiveness 2000

Histogram



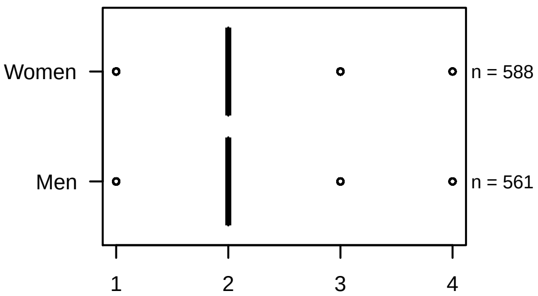
Summary Statistics

Mean = 2.003
Std. Dev. = 0.6613
N = 1149

Skewness = 0.7186
Std. Err. Skewness = 0.07226
p−value = 0

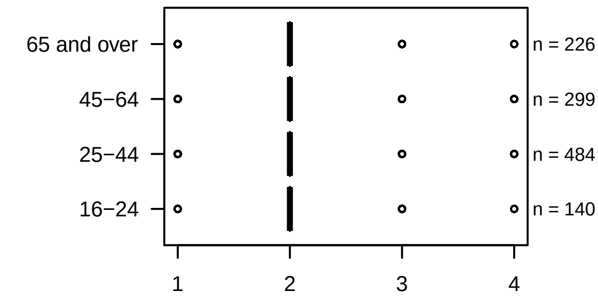
Kurtosis = 1.455
Std. Err. Kurtosis = 0.1445
p−value = 0

Gender



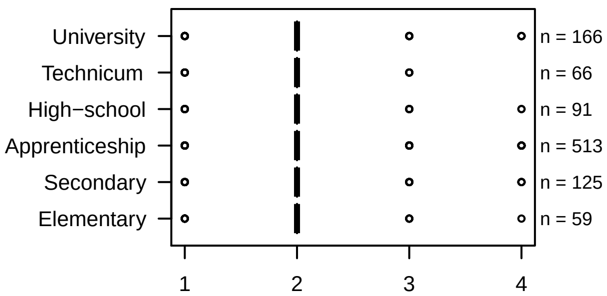
Wilcoxon rank sum test
W = 167745.5; p = 0.549

Age



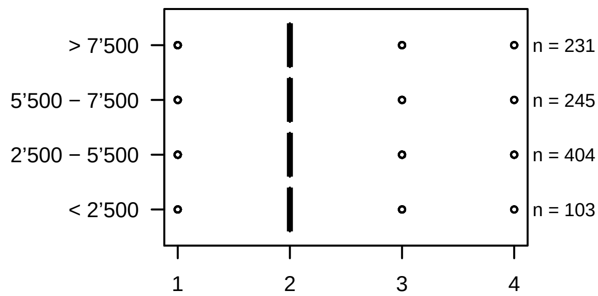
Kruskal−Wallis rank sum test
chi−squared = 10.5; df = 3; p = 0.0147

Education



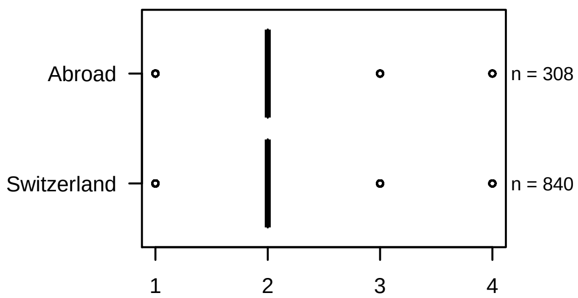
Kruskal−Wallis rank sum test
chi−squared = 4.47; df = 5; p = 0.484

Income



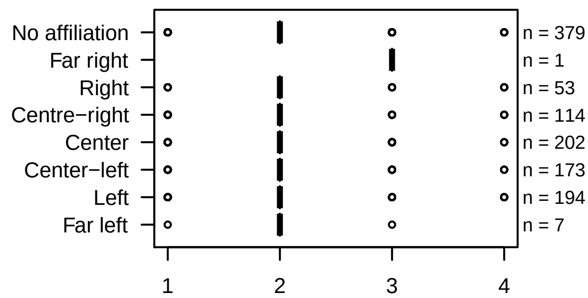
Kruskal−Wallis rank sum test
chi−squared = 1.66; df = 3; p = 0.645

Place of birth



Wilcoxon rank sum test
W = 143017.5; p = 0.00099

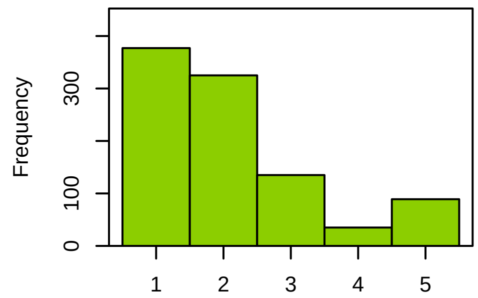
Political affiliation



Kruskal−Wallis rank sum test
chi−squared = 7.1; df = 7; p = 0.419

Fear of Crime 1998

Histogram



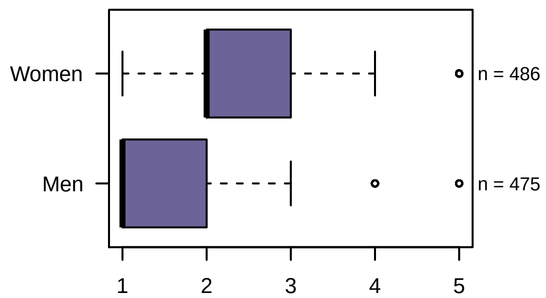
Summary Statistics

Mean = 2.099
Std. Dev. = 1.226
N = 961

Skewness = 1.135
Std. Err. Skewness = 0.07902
p−value = 0

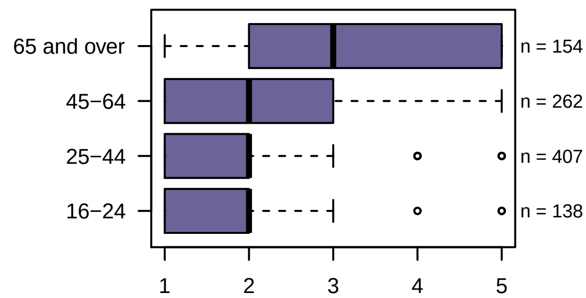
Kurtosis = 0.4056
Std. Err. Kurtosis = 0.158
p−value = 0.005212

Gender



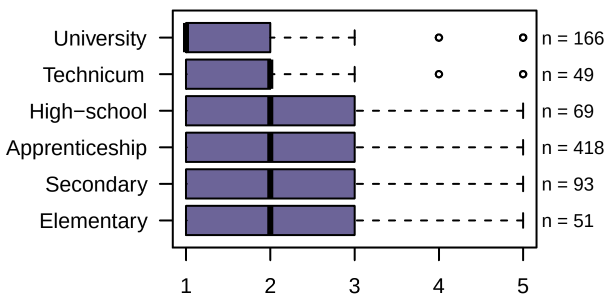
Wilcoxon rank sum test
W = 65187; p = 6.41e−35

Age



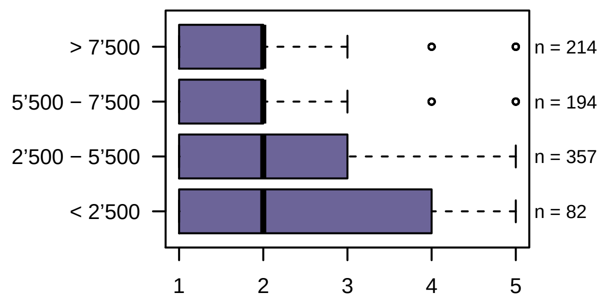
Kruskal−Wallis rank sum test
chi−squared = 111; df = 3; p = 6.64e−24

Education



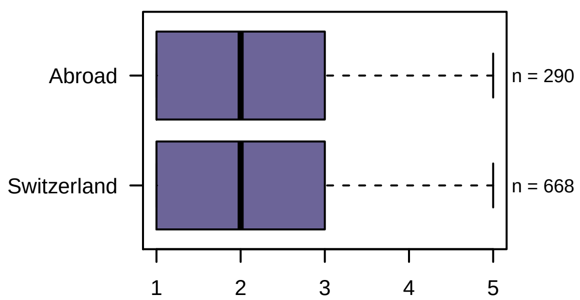
Kruskal−Wallis rank sum test
chi−squared = 36.4; df = 5; p = 7.9e−07

Income



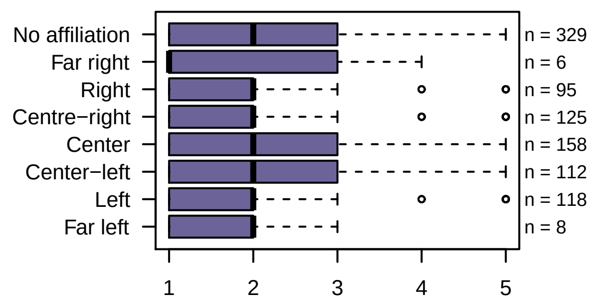
Kruskal−Wallis rank sum test
chi−squared = 22.3; df = 3; p = 5.6e−05

Place of birth



Wilcoxon rank sum test
W = 94655; p = 0.554

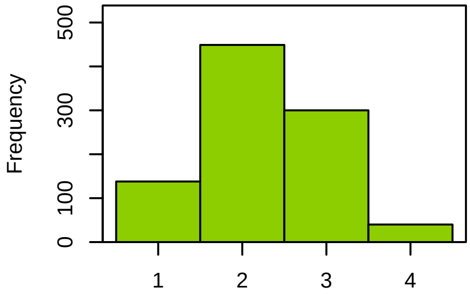
Political affiliation



Kruskal−Wallis rank sum test
chi−squared = 18.5; df = 7; p = 0.01

Risk of Victmization 1998

Histogram



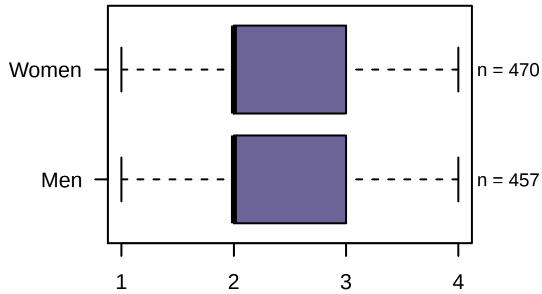
Summary Statistics

Mean = 2.261
Std. Dev. = 0.76
N = 927

Skewness = 0.1146
Std. Err. Skewness = 0.08045
p−value = 0.07724

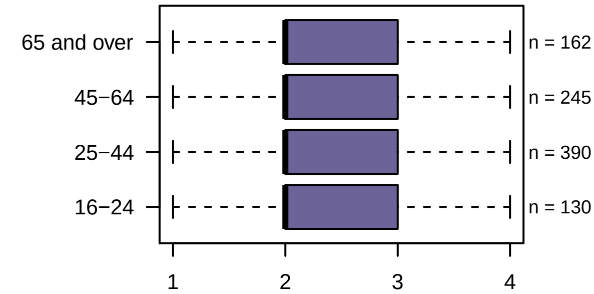
Kurtosis = −0.3926
Std. Err. Kurtosis = 0.1609
p−value = 0.007441

Gender



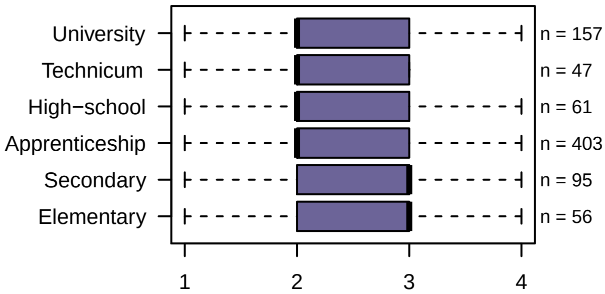
Wilcoxon rank sum test
W = 105294.5; p = 0.576

Age



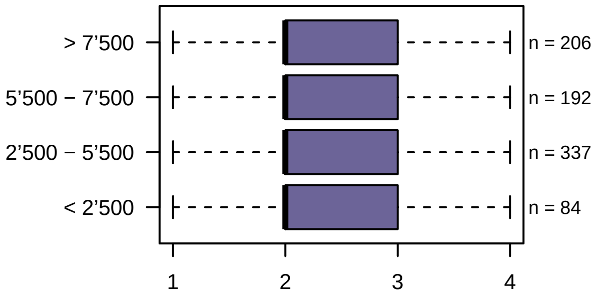
Kruskal−Wallis rank sum test
chi−squared = 21.2; df = 3; p = 9.65e−05

Education



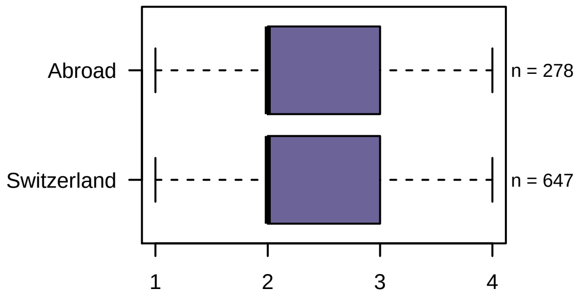
Kruskal−Wallis rank sum test
chi−squared = 19.5; df = 5; p = 0.00154

Income



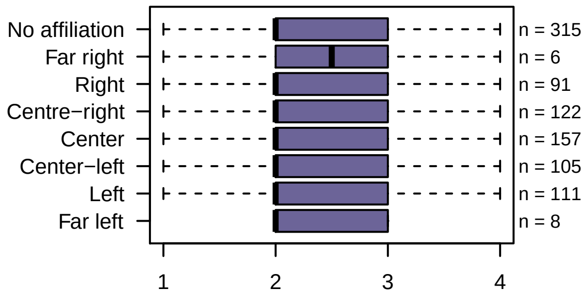
Kruskal−Wallis rank sum test
chi−squared = 1.11; df = 3; p = 0.774

Place of birth



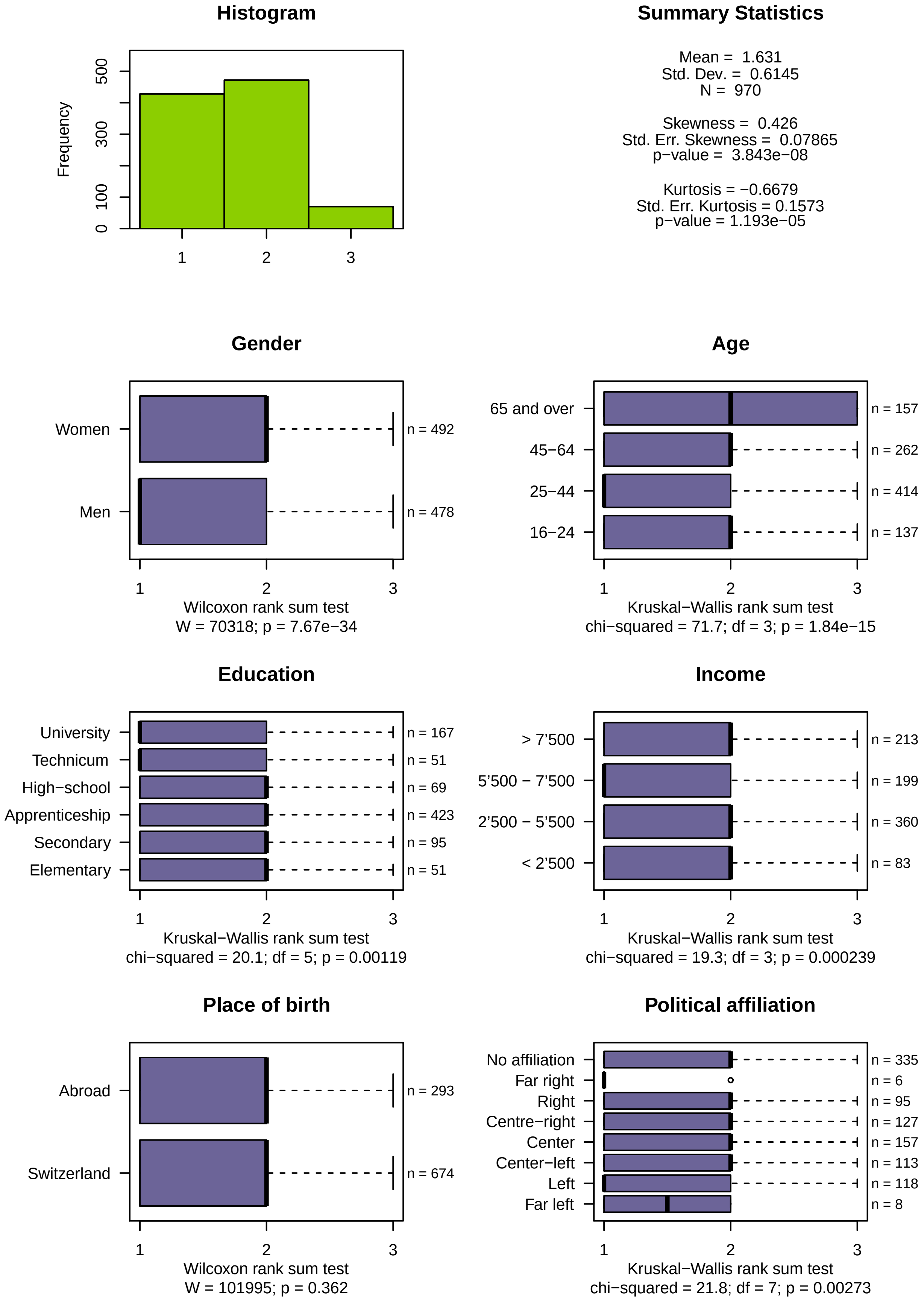
Wilcoxon rank sum test
W = 84248.5; p = 0.0977

Political affiliation



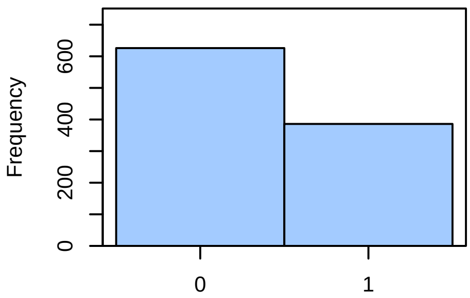
Kruskal−Wallis rank sum test
chi−squared = 6.09; df = 7; p = 0.529

Behavioral Response 1998



Physical Disorder 1998

Histogram



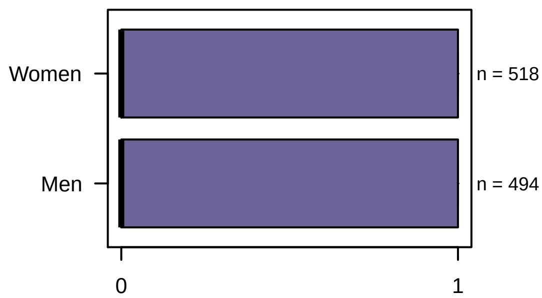
Summary Statistics

Mean = 0.3814
Std. Dev. = 0.486
N = 1012

Skewness = 0.4875
Std. Err. Skewness = 0.077
p−value = 1.824e−10

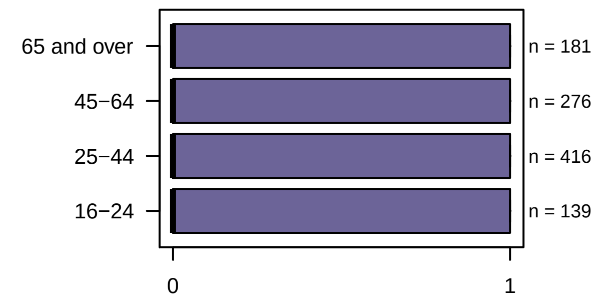
Kurtosis = −1.764
Std. Err. Kurtosis = 0.154
p−value = 0

Gender



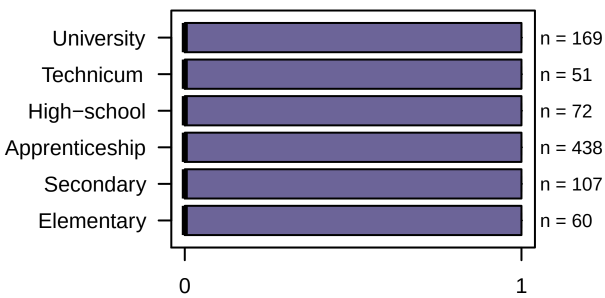
Wilcoxon rank sum test
W = 131780; p = 0.327

Age



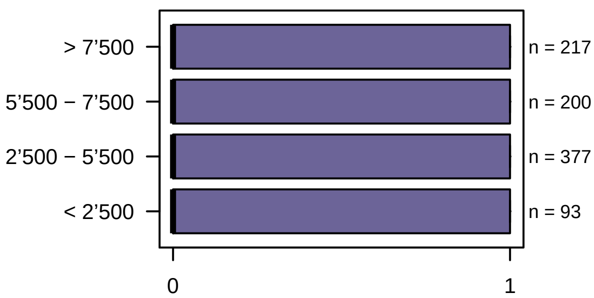
Kruskal−Wallis rank sum test
chi−squared = 9.58; df = 3; p = 0.0225

Education



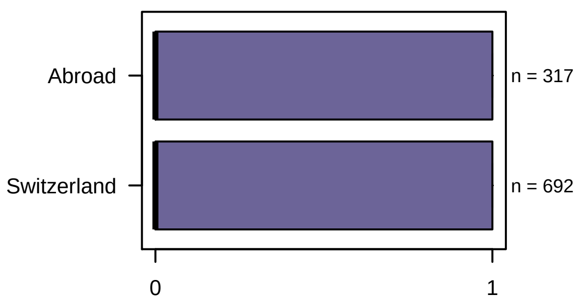
Kruskal−Wallis rank sum test
chi−squared = 7.66; df = 5; p = 0.176

Income



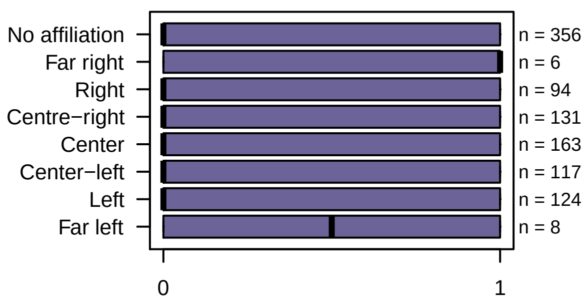
Kruskal−Wallis rank sum test
chi−squared = 6.15; df = 3; p = 0.105

Place of birth



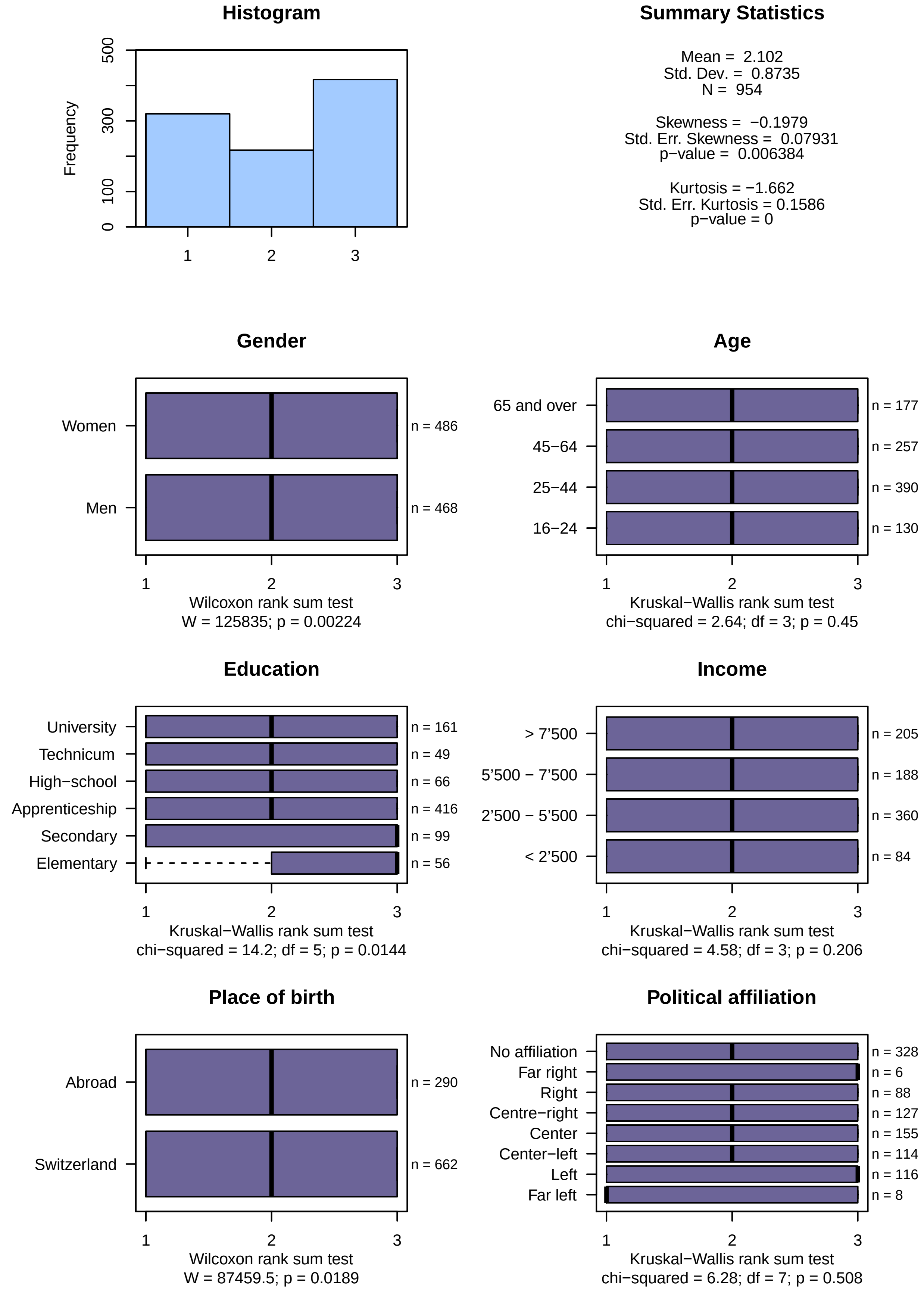
Wilcoxon rank sum test
W = 116881.5; p = 0.0466

Political affiliation



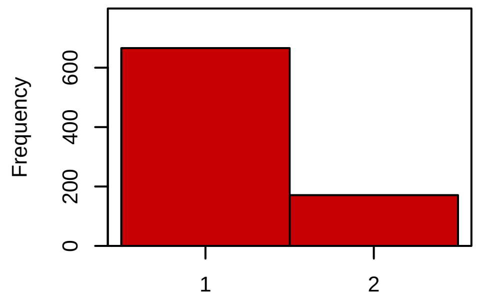
Kruskal−Wallis rank sum test
chi−squared = 6.41; df = 7; p = 0.493

Social Cohesion 1998



Police Effectiveness 1998

Histogram



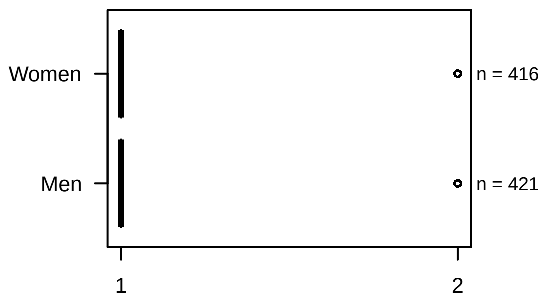
Summary Statistics

Mean = 1.204
Std. Dev. = 0.4034
N = 837

Skewness = 1.464
Std. Err. Skewness = 0.08467
p−value = 0

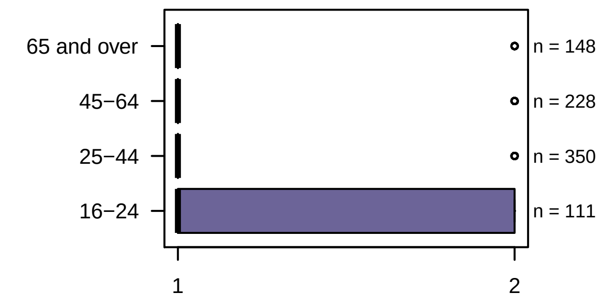
Kurtosis = 0.144
Std. Err. Kurtosis = 0.1693
p−value = 0.1977

Gender



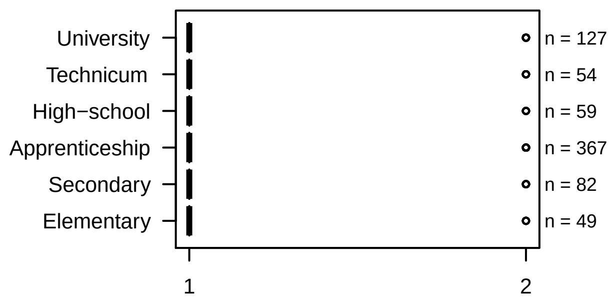
Wilcoxon rank sum test
W = 91330; p = 0.124

Age



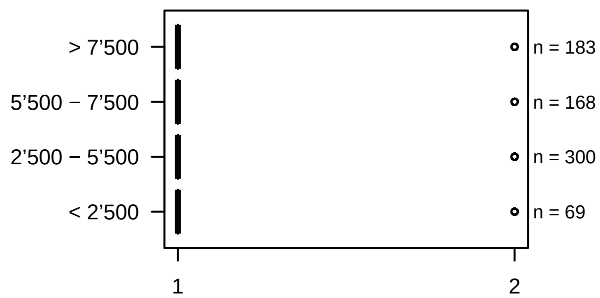
Kruskal−Wallis rank sum test
chi−squared = 4.39; df = 3; p = 0.222

Education



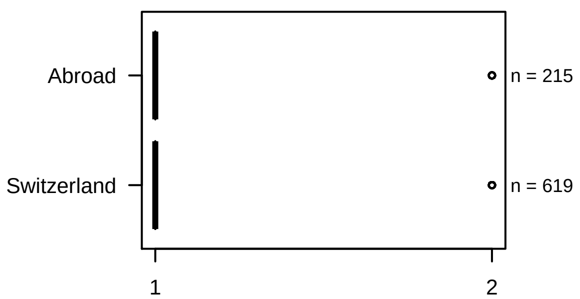
Kruskal−Wallis rank sum test
chi−squared = 4.37; df = 5; p = 0.498

Income



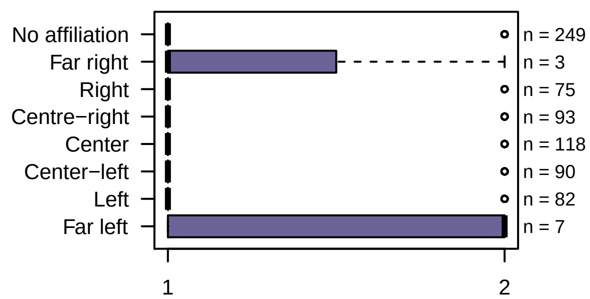
Kruskal−Wallis rank sum test
chi−squared = 6.2; df = 3; p = 0.102

Place of birth



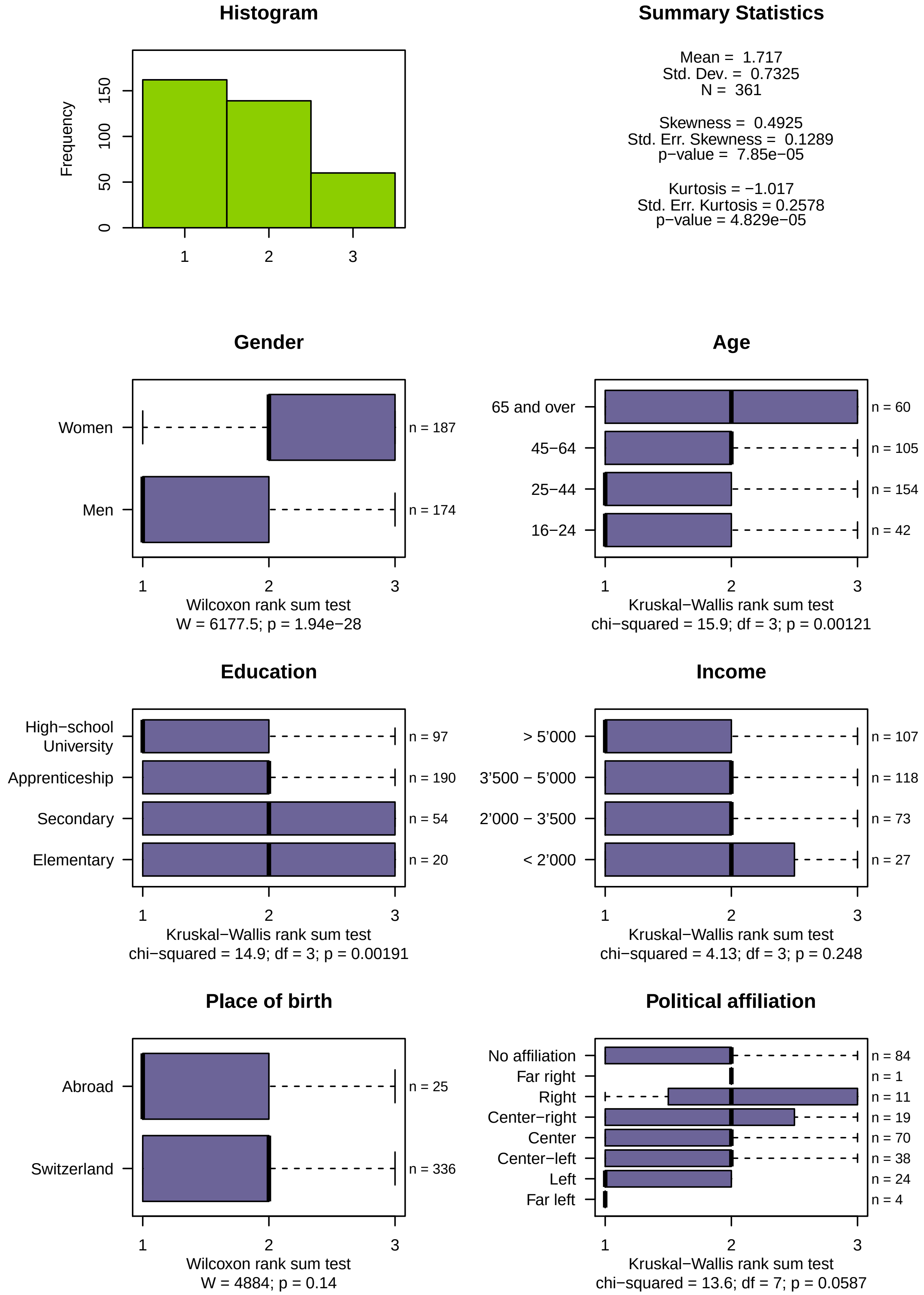
Wilcoxon rank sum test
W = 70747; p = 0.0482

Political affiliation

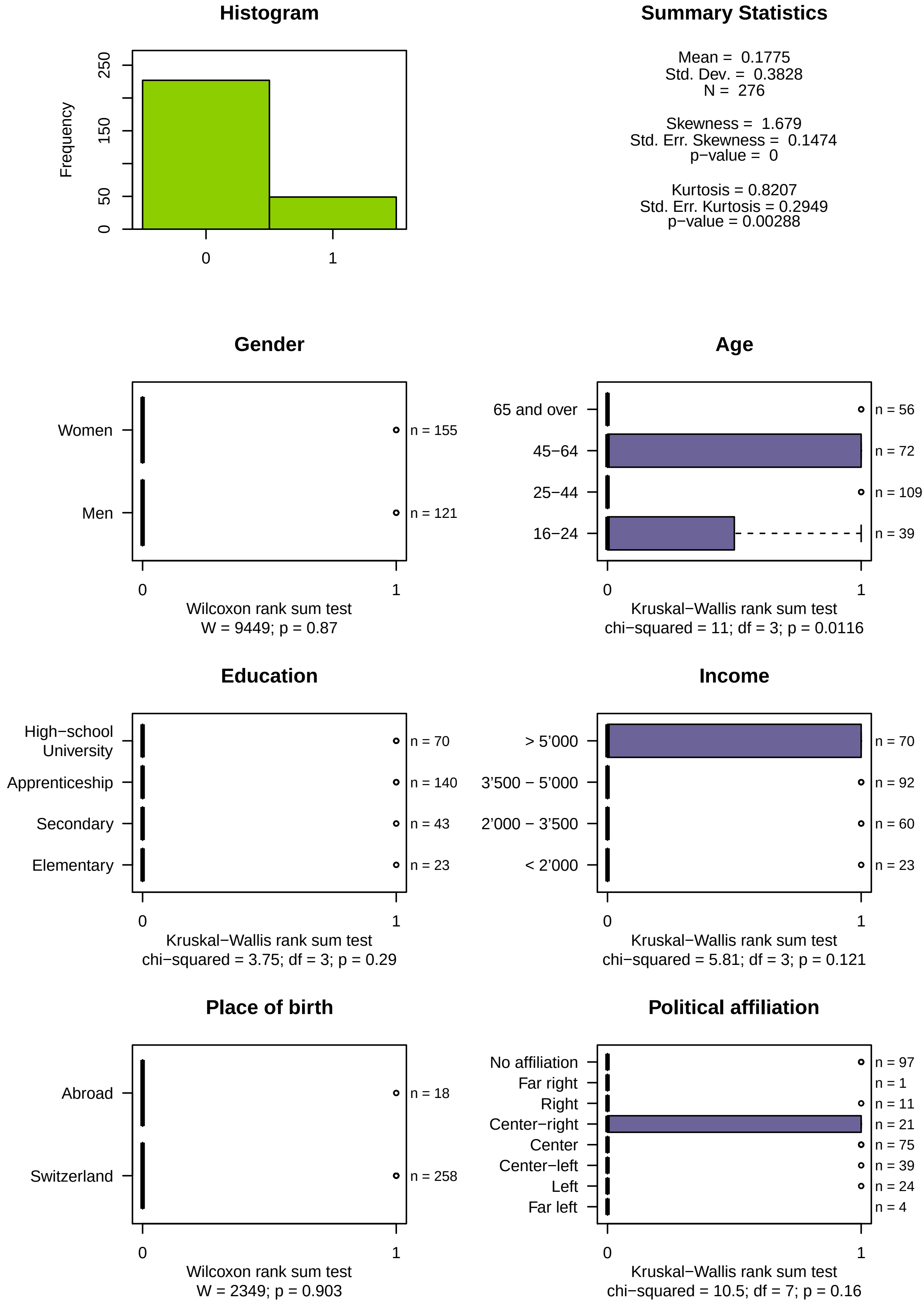


Kruskal−Wallis rank sum test
chi−squared = 7.82; df = 7; p = 0.348

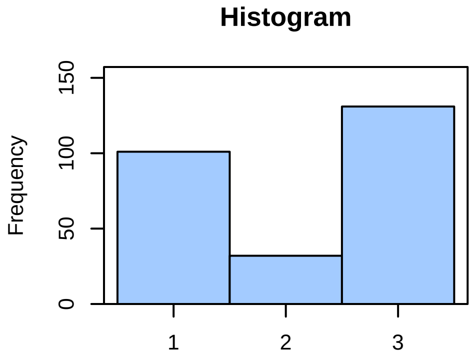
Fear of Crime 1987



Behavioral Response 1987



Social Cohesion 1987

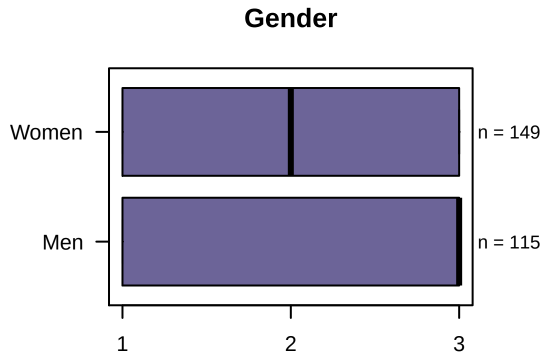


Summary Statistics

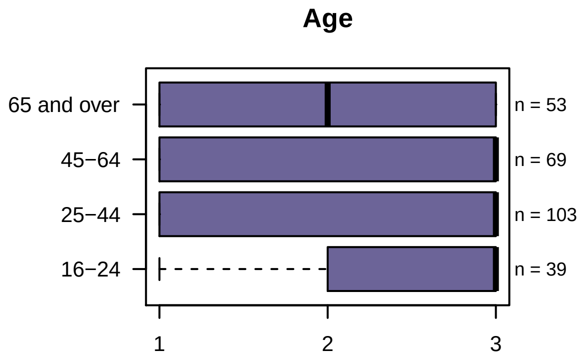
Mean = 2.114
Std. Dev. = 0.9323
N = 264

Skewness = −0.2259
Std. Err. Skewness = 0.1508
p−value = 0.06765

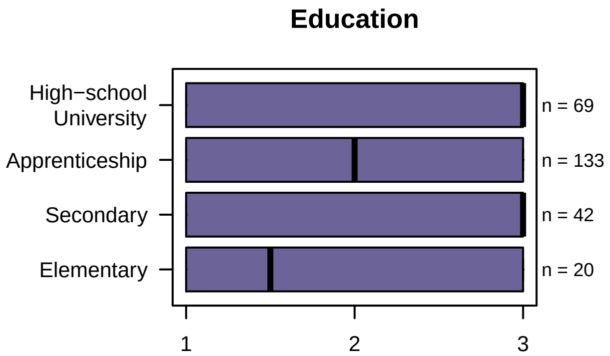
Kurtosis = −1.816
Std. Err. Kurtosis = 0.3015
p−value = 2.903e−09



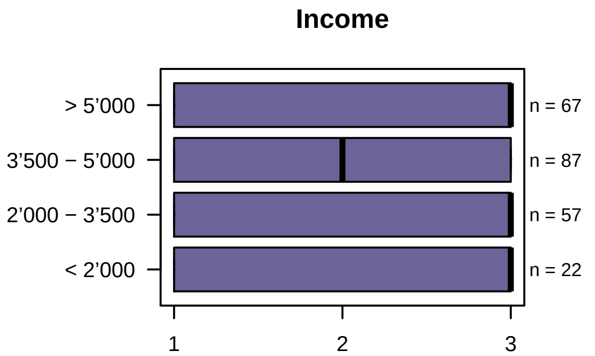
Wilcoxon rank sum test
W = 9169; p = 0.281



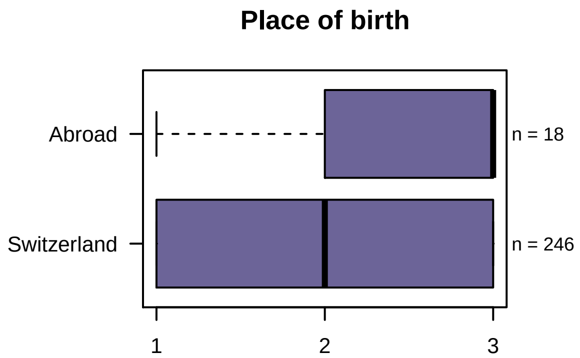
Kruskal−Wallis rank sum test
chi−squared = 5.91; df = 3; p = 0.116



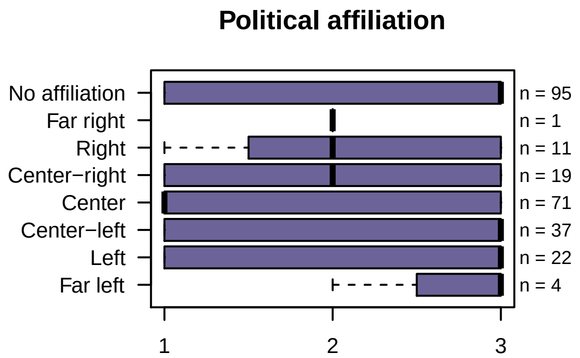
Kruskal−Wallis rank sum test
chi−squared = 4.41; df = 3; p = 0.22



Kruskal−Wallis rank sum test
chi−squared = 1.94; df = 3; p = 0.586



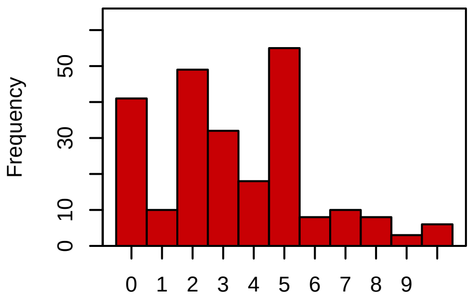
Wilcoxon rank sum test
W = 1853.5; p = 0.204



Kruskal−Wallis rank sum test
chi−squared = 9.68; df = 7; p = 0.207

Police Effectiveness 1987

Histogram



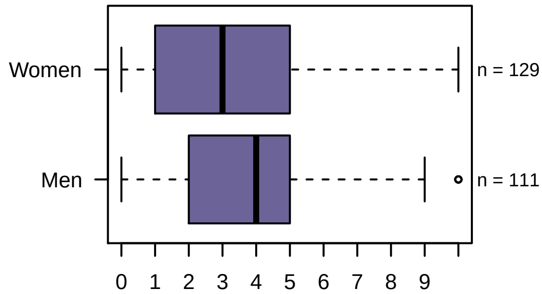
Summary Statistics

Mean = 3.417
Std. Dev. = 2.495
N = 240

Skewness = 0.5111
Std. Err. Skewness = 0.1581
p−value = 0.0007004

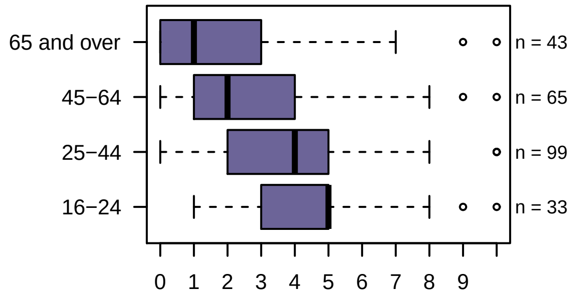
Kurtosis = −0.1855
Std. Err. Kurtosis = 0.3162
p−value = 0.2791

Gender



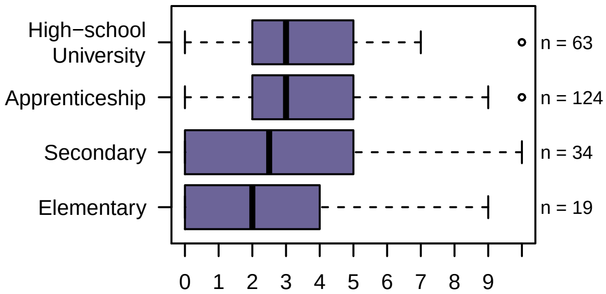
Wilcoxon rank sum test
W = 8605; p = 0.00626

Age



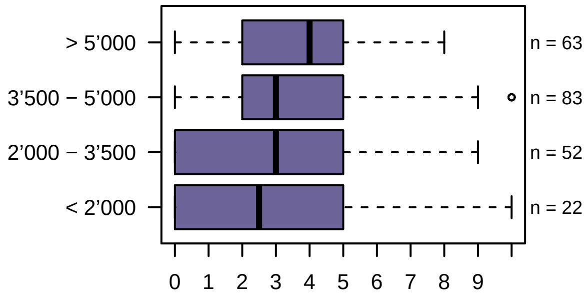
Kruskal−Wallis rank sum test
chi−squared = 47.1; df = 3; p = 3.23e−10

Education



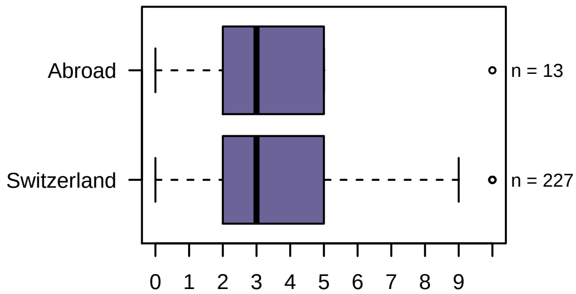
Kruskal−Wallis rank sum test
chi−squared = 7.73; df = 3; p = 0.052

Income



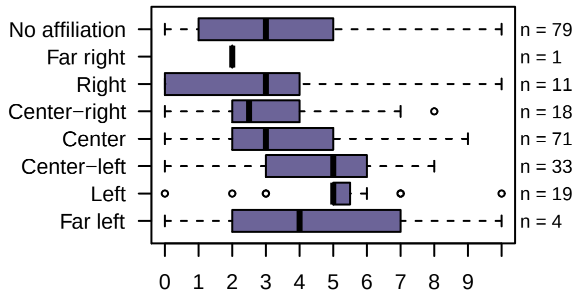
Kruskal−Wallis rank sum test
chi−squared = 3.38; df = 3; p = 0.337

Place of birth



Wilcoxon rank sum test
W = 1437; p = 0.874

Political affiliation



Kruskal−Wallis rank sum test
chi−squared = 20.1; df = 7; p = 0.0054

E

Analysis of the Outcome Indicators by Neighborhood Cluster

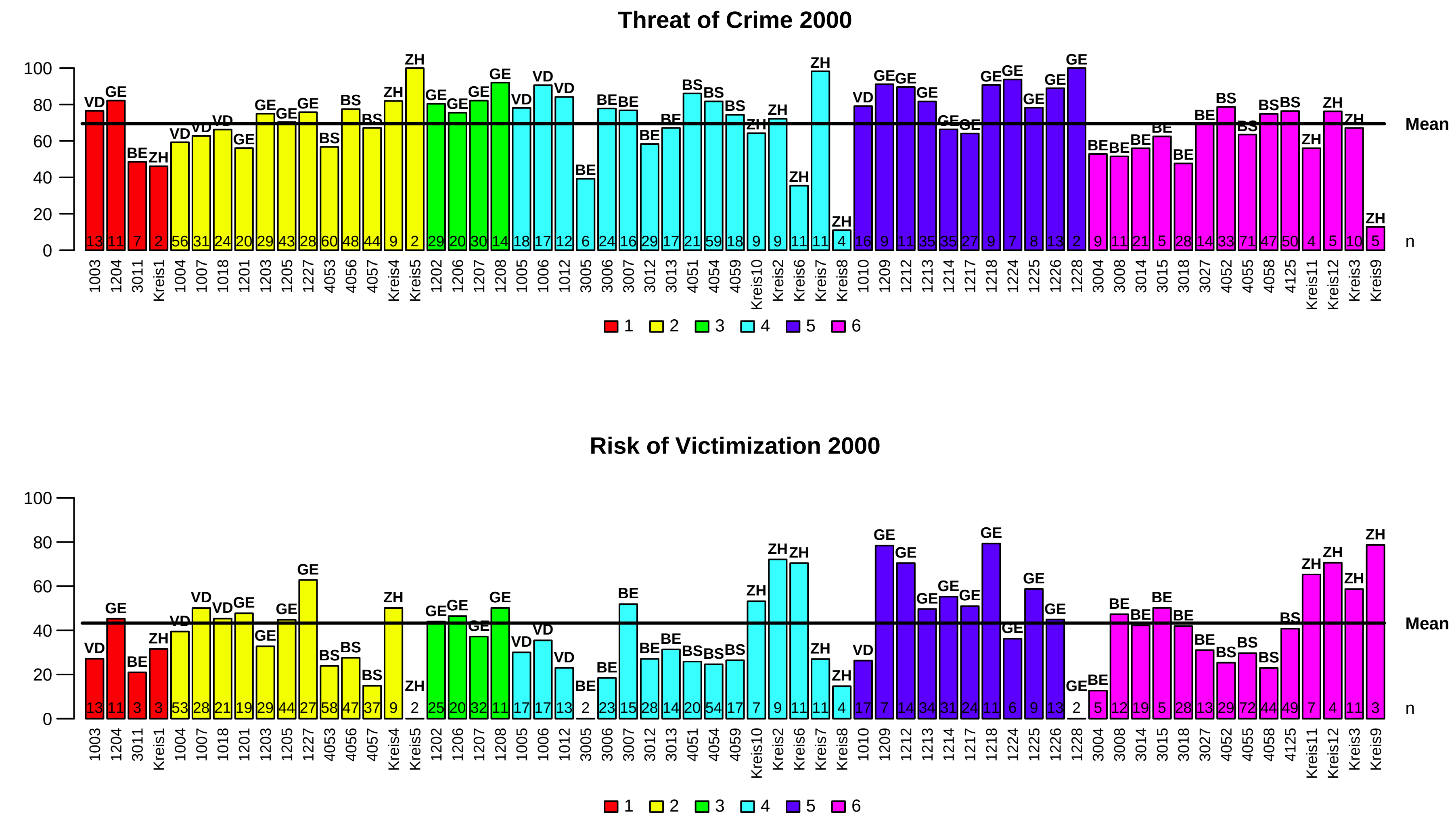
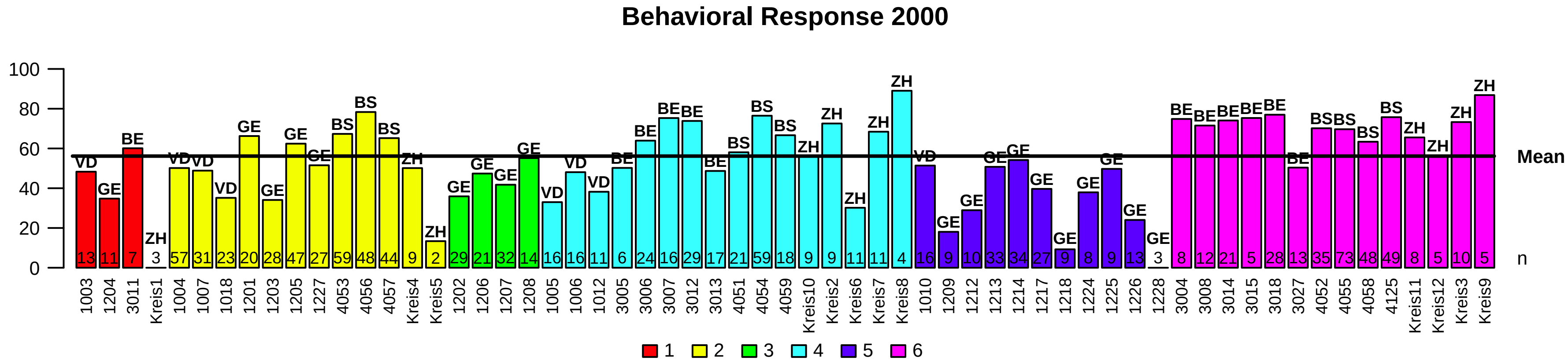


Figure E.1 Threat of crime by neighborhood cluster. Percentage of respondents who feel "very safe" or "quite safe" walking alone in their neighborhood at night (top). Risk of victimization by neighborhood cluster. Percentage of respondents who think chances are "rather likely" or "very likely" that someone will try to break into their home over the course of the next twelve months (bottom). The data were weighted at the neighborhood level to correct for sampling bias in the age and gender distribution.



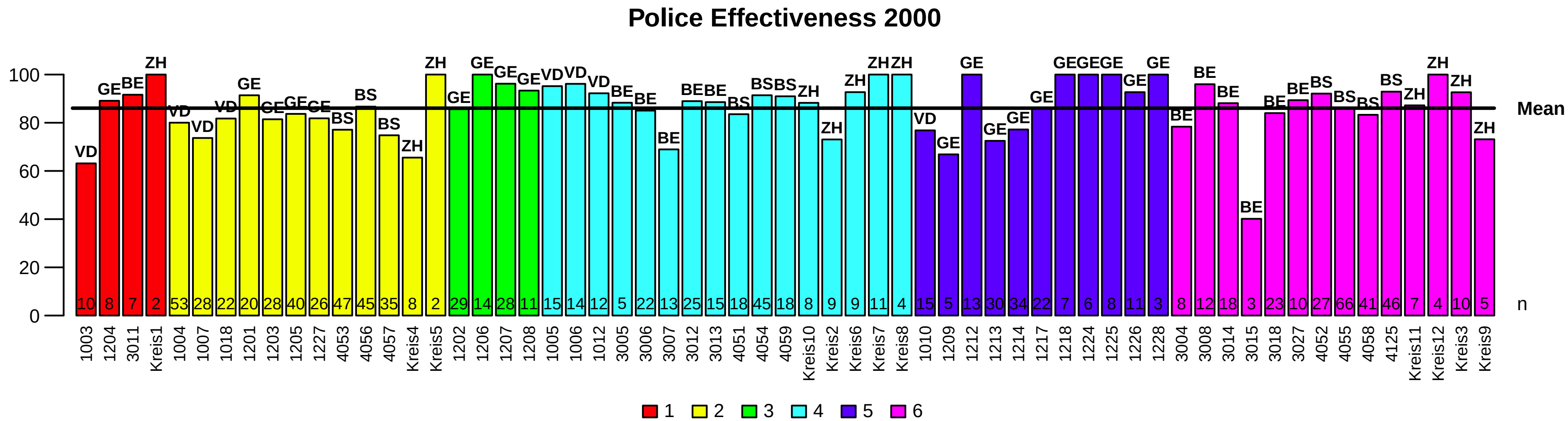


Figure E.2 Behavioral response to crime by neighborhood cluster. Percentage of respondents who – walking alone in their neighborhood after 10 pm – stay away from certain streets, areas, or people to avoid crime (top). Public assessment of police effectiveness by neighborhood cluster. Percentage of respondents who rate the effectiveness of the police in controling crime in their area as "very good" or "rather good" (bottom). The data were weighted at the neighborhood level to correct for sampling bias in the age and gender distribution.

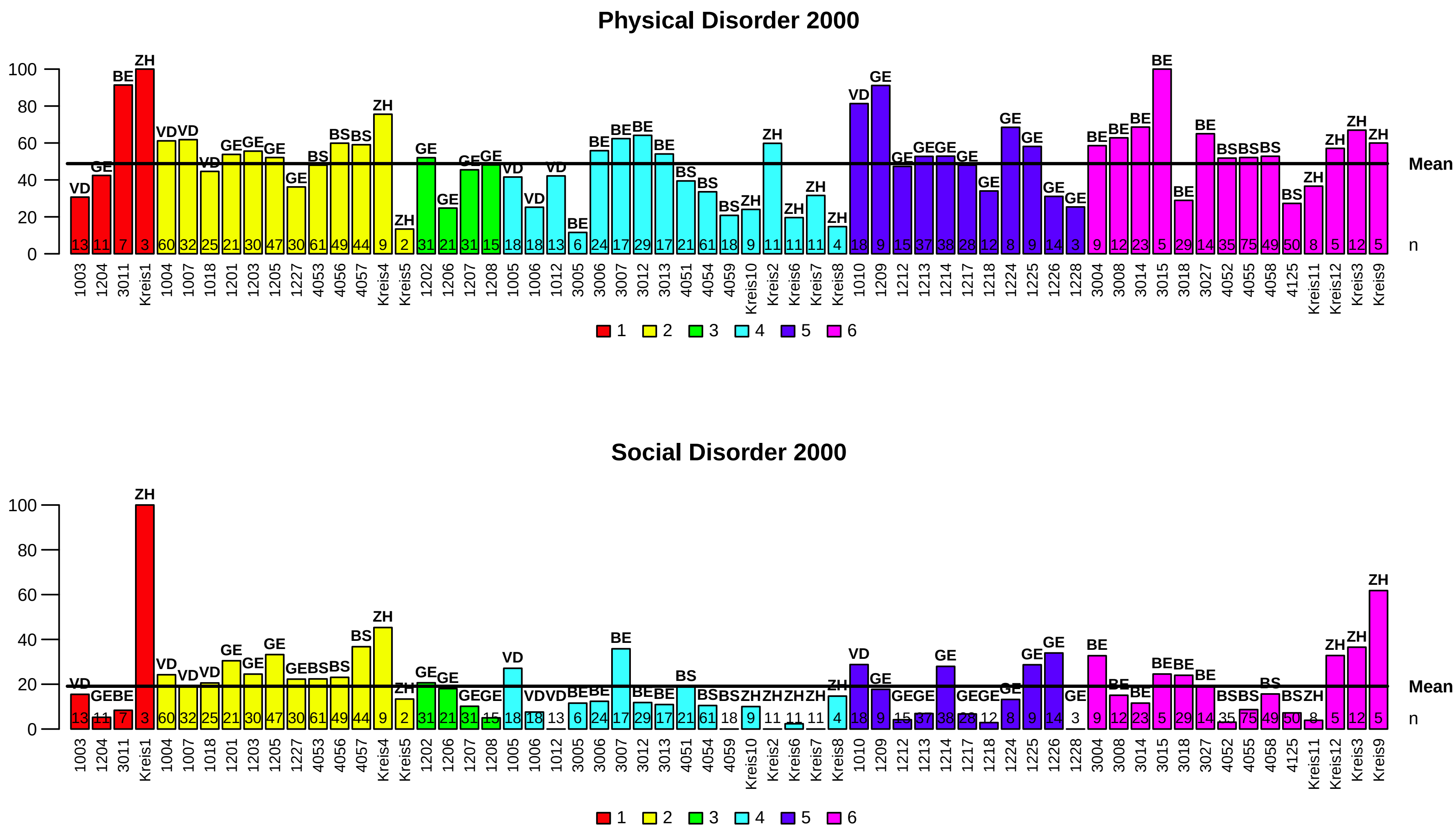


Figure E.3 Physical disorder by neighborhood cluster. Percentage of respondents who spotted graffiti on the walls or lots of rubbish lying around in the streets close to their home (top). Social disorder by neighborhood cluster. Percentage of respondents who spotted loitering groups of disreputable people close to their home (bottom). The data were weighted at the neighborhood level to correct for sampling bias in the age and gender distribution.

F

Survey Questionnaire

Table F.1 SCS survey questionnaire

Survey	Wording	Coding
	Fear of Crime – Threat of Crime	
CH 1987	• *Gehen Sie manchmal allein in der Nacht aus Ihrem Haus?* ——— 1 = ja 2 = nein	v45
CH 1987	• *Gibt es im Umkreis von 1 km von Ihrer Wohnung einen Ort, wo Sie Angst hätten, in der Nacht allein spazieren zu gehen?* ——— 1 = ja 2 = nein 3 = weiss nicht	v46
CH 1987	• *Aber wenn Sie in der Nacht alleine spazieren gehen müssten, gibt es im Umkresi von 1 km von Ihrer Wohnung einen Ort wo Sie Angst hätten?* ——— 1 = ja 2 = nein 3 = weiss nicht	v47
CH 1998	• *Wie sicher fühlen Sie sich, wenn Sie nach 22 Uhr am Abend allein zu Fuss in Ihrer Wohngegend unterwegs sind?* • *Comment vous sentez-vous lorsque vous vous promenez seul après 22 heures dans votre quartier ? Vous sentez-vous très sécurisé, assez sécurisé, pas très sécurisé ou pas du tout sécurisé ?* ——— 1 = sehr sicher/très sécurisé 2 = ziemlich sicher/assez sécurisé 3 = ein wenig unsicher/pas très sécurisé 4 = sehr unsicher/pas du tout sécurisé 5 = aus Sicherheitsgründen gehe ich nie nach 22 Uhr am Abend aus dem Haus/ne sort jamais après 22 heures pour des raisons de sécurité 6 = aus anderen Gründen gehe ich nie nach 22 Uhr am Abend aus dem Haus/ne sort jamais après 22 heures pour d'autres raisons 8 = weiss nicht/ne sait pas 9 = keine Antwort/sans réponse	q30010
ICVS 2000	idem	q30010
ICVS 2005	idem	Q30010
	Fear of Crime – Risk of Victimization	
CH 1998	• *Wie hoch würden Sie sagen ist die Wahrscheinlichkeit, dass jemand im Laufe von den nächsten 12 Monate in Ihre Wohnung (ev. Ihr Haus) einbricht?* • *Quel est d'après vous le degré de probabilité que quelqu'un s'introduise dans votre habitation par effraction au cours des 12 prochains mois ? Pensez-vous que cela soit très probable, probable, improbable ou très improbable ?*	

Table F.1 SCS survey questionnaire (continued)

Survey	Wording	Coding
	1 = sehr wahrscheinlich/très probable 2 = wahrscheinlich/probable 3 = unwahrscheinlich/improbable 4 = sehr unwahrscheinlich/très improbable 8 = weiss nicht/ne sait pas 9 = keine Antwort/sans réponse	q302
ICVS 2000	idem	q302
ICVS 2005	idem	Q302
	Fear of Crime – Behavioral Changes	
CH 1998	• *Wenn Sie nach 22 Uhr am Abend alleine zu Fuss in Ihrem Quartier unterwegs sind, meiden Sie da absichtlich gewisse Strassen, Orte oder gewisse Leute aus Sicherheitsgründen?* • *Lorsque vous vous promenez après 22 heures dans votre quartier, évitez-vous volontairement certaines rues, certains endroits ou certaines personnes pour des raisons de sécurité ?* 1 = ja/oui 2 = nein/non 3 = aus Sicherheitsgründen gehe ich nie nach 22 Uhr am Abend aus dem Haus/ne sort jamais après 22 heures pour des raisons de sécurité 4 = aus anderen Gründen gehe ich nie nach 22 Uhr am Abend aus dem Haus/ne sort jamais après 22 heures pour d'autres raisons que la sécurité 8 = weiss nicht/ne sait pas 9 = keine Antwort/sans réponse	q303
ICVS 2000	idem	q303
ICVS 2005	idem	Q303
	Neighborhood Development – Disorder	
CH 1998	• *Hat es in der Nähe von Ihrem Wohnort Graffiti-Zeichnungen an den Wänden, viel Abfall, wo herumliegt oder zweifelhafte Personen, wo oft in Gruppen herumstehen und zusammen diskutieren?* • *Y a-t-il proche de chez vous des graffitis sur les murs, beaucoup d'ordures qui traînent ou des groupes de personnes douteuses qui discutent souvent ensemble ?* 1 = ja/oui 2 = nein/non 8 = weiss nicht/ne sait pas 9 = keine Antwort/sans réponse	q30916
ICVS 2000	• *Hat es in der Nähe von Ihrem Wohnort Graffiti-Zeichnungen an den Wänden, viel Abfall, wo herumliegt oder zweifelhafte Personen, wo oft in Gruppen herumstehen und zusammen diskutieren?*	

Table F.1 SCS survey questionnaire (continued)

Survey	Wording	Coding
	• *Y a-t-il proche de chez vous des graffiti sur les murs, beaucoup d'ordures qui traînent ou des groupes de personnes douteuses qui discutent souvent ensemble ?* ——— 1 = ja, Graffiti/oui, des graffiti 2 = ja, Abfälle/oui, des ordures 3 = ja, zweifelhafte Personen/oui, des personnes douteuses 4 = ja, ein wenig von allem/oui, un peu de tout 5 = nein/non 8 = weiss nicht/ne sait pas 9 = keine Antwort/sans réponse	q30916
ICVS 2005	idem	Q30916_1 Q30916_2 Q30916_3
	Neighborhood Development – Social Cohesion	
CH 1987	• *In gewissen Quartieren oder Dörfern versuchen die Leute sich gegenseitig zu helfen, während anderswo jeder sein eigenes Leben führt ohne sich um die anderen zu kümmern. Wie ist es dort wo Sie wohnen?* ——— 1 = die Leute helfen sich gegenseitig 2 = jeder für sich 3 = halb-halb 4 = weiss nicht	v58
CH 1998	• *Nun möchte ich Sie noch fragen, wie Sie das Klima in Ihrem Quartier und das Problem von der Kriminalität dort beurteilen. Es gibt Quartiere und Gegenden, wo die Leute sich um einander kümmern und versuchen, einander zu helfen, während in anderen Gegenden und Quartieren jeder nur für sich schaut. Wie sehen Sie in dieser Hinsicht Ihren Wohnort oder Ihr Quartier? Helfen sich hier die Leute meistens gegenseitig aus, oder ist jeder eher für sich allein?* • *Maintenant, j'aimerais vous poser quelques questions concernant le climat de votre quartier et le problème de la criminalité dans ce dernier. En effet, dans certains quartiers ou certaines régions, les gens se soucient les uns des autres et essaient de s'aider, tandis qu'ailleurs, c'est plutôt "chacun pour soi". Dans votre lieu d'habitation ou votre quartier, les gens s'entraident-ils ou est-ce plutôt chacun pour soi ?* ——— 1 = die Leute helfen sich hier meistens gegenseitig/plutôt l'entraide 2 = jeder ist eher für sich allein/plutôt chacun pour soi 3 = gemischt, so zwischendurch/un mélange des deux 8 = weiss nicht/ne sait pas 9 = keine Antwort/sans réponse	q299
	Public Attitudes towards the Police – Overall Satisfaction	

Table F.1 SCS survey questionnaire (continued)

Survey	Wording	Coding
CH 1987	• *Welche Note würden Sie der Polizei Ihres Kantons (oder Ihrer Stadt) im Allgemeinen geben?* (0 = sehr schlecht bis 10 = sehr gut) ——— 0 = 0 1 = 1 2 = 2 3 = 3 4 = 4 5 = 5 6 = 6 7 = 7 8 = 8 9 = 9 10 = 10 99 = weiss nicht/keine Antwort	v65
CH 1998	• *Wie gut kommt – alles in allem – Ihrer Ansicht nach die Polizei der Aufgabe der Verbrechensbekämpfung in Ihrer Wohngegend nach? Finden Sie, dass sie das gut macht, oder finden Sie das eher nicht?* • *Tout bien considéré, comment trouvez-vous l'action de la police dans votre quartier en matière de lutte contre la criminalité ? Trouvez-vous qu'elle fait plutôt du bon travail ou plutôt du mauvais travail ?* ——— 1 = sie macht das eher gut/elle fait plutôt du bon travail 2 = sie macht das eher schlecht/elle fait plutôt du mauvais travail 8 = weiss nicht/ne sait pas 9 = keine Antwort/sans réponse	q310
ICVS 2000	• *Alles in allem, wie finden Sie macht die Polizei in Ihrem Quartier/ Wohnort ihre Arbeit in Bezug auf die Bekämpfung der Kriminalität?* • *Tout bien considéré, comment trouvez-vous l'action de la police dans votre quartier/lieu de domicile en matière de lutte contre la criminalité ? Trouvez-vous qu'elle fait du très bon travail, de l'assez bon travail, du travail plutôt mauvais ou du travail très mauvais ?* ——— 1 = sehr gut/très bon travail 2 = ziemlich gut/assez bon travail 3 = nicht so gut/travail plutôt mauvais 4 = überhaupt nicht gut/très mauvais travail 8 = weiss nicht/ne sait pas 9 = keine Antwort/sans réponse	q31010
ICVS 2005	idem	Q31010

	Public Attitudes towards the Police – Wish for more patrols	
ICVS 2005	• *Würden Sie es wünschen sie öfters oder weniger oft zu sehen?* • *Souhaiteriez-vous les voir plus souvent ou moins souvent ?* ——— 1 = öfter/plus souvent	

Table F.1 SCS survey questionnaire (continued)

Survey	Wording	Coding
	2 = weniger oft/moins souvent 3 = gleich viele Male/la même chose 8 = weiss es nicht/ne sait pas 9 = keine Antwort/sans réponse	Q31016
	Public Attitudes towards the Police – Police Equity	
CH 1998	• *Finden Sie, dass die Polizei alle Leute gleich behandelt oder werden bestimmte Personen besser oder schlechter als andere behandelt?* • *En général, trouvez-vous que la police traite tout le monde de la même façon ou pensez-vous qu'elle traite certaines personnes plus ou moins favorablement que d'autres ?* ——— 1 = immer gleiche Behandlung/toujours le même traitement 2 = sehr oft gleiche Behandlung/plutôt le même traitement 3 = manchmal ungleiche Behandlung/parfois un traitement différent 4 = immer ungleiche Behandlung/toujours un traitement différent 8 = weiss nicht/ne sait pas 9 = keine Antwort/sans réponse	q311
ICVS 2000	idem	q311
ICVS 2005	idem	Q311