COMPARING CRIME RATES:
THE INTERNATIONAL CRIME (VICTIM) SURVEY, THE
EUROPEAN SOURCEBOOK OF CRIME AND CRIMINAL JUSTICE
STATISTICS, AND INTERPOL STATISTICS

Marcelo F. Aebi,* Martin Killias,** and Cynthia Tavares***

ABSTRACT

Since its beginnings, comparative criminology has suffered from the lack of valid international data on crime and criminal justice. Over several decades, efforts were made to improve international comparisons by new data collections. In this paper, three such data sources are compared. The results show that International Crime (Victim) Survey (ICVS) data are highly correlated with police data published in the European Sourcebook of Crime and Criminal Justice Statistics, at least for offences whose survey and police definitions closely match. On the other hand, Interpol statistics are only weakly correlated with ICVS and Sourcebook data. It is argued that the careful screening process within the Sourcebook project increases considerably the comparative validity of police data by the elimination of often trivial but devastating errors. Despite such improvements, survey data may remain more valid for cross-sectional analyses on personal crime and other offences where recording by the police is uncertain and often inconsistent. In any case, unchecked police data (as those published in Interpol Statistics) should not be used for comparative purposes.

Comparative Criminology and the Search for Valid Data

Between Trivial Use and Dismissal of Comparative Data

Since the beginning of comparative analyses dating back to the first half of the nineteenth century, criminology has oscillated between naive data collection on the one hand, and, on the other, dismissal of all attempts to use comparative data to test criminological theories. Both positions were ill-designed to advance knowledge. The rejection of such data and of criminal justice data in general, which is predominant in some continental European countries, has hindered the process of theory testing in a comparative perspective, as attempted by scholars with a multilingual background like Hermann Mannheim (1965). Criminal justice policies

* Universities of Sevilla and Lausanne, IPSC-UNIL, CH-1015 Lausanne, Switzerland.
** University of Lausanne, IPSC-UNIL, CH-1015 Lausanne, Switzerland.
*** Home Office, London and University of Lausanne, Queen Anne’s Gate, London SW 1H 9AT, U.K.

© de Sitter Publications

IJCC 2(1):22-37
vary widely around the world and this stance deprives criminology of the opportunity to comparatively evaluate their merits and pitfalls. It also fails policy-makers in not providing them with relevant feedback to rethink current policies. Thus, it seems that the lack of comparative research may well have prevented policy-makers and the public from learning from significant experiences accumulated in other countries. For example, countries with high incarceration rates, such as Russia and the United States, are generally no more successful at controlling crime than countries with low or moderate incarceration rates.

From Two-Country Studies to Multi-National Research

Some studies have compared two countries with ostensibly different levels of crime. For example, studies have compared crime in Switzerland (Clinard 1978), Japan (Kühne and Miyazawa 1979), and some Latin American countries (Neapolitan 1994) with other Western countries. Recent examples include studies comparing England and Wales with the United States (Langan and Farrington 1998), England and Wales with Sweden (Farrington and Wikström 1993), Scotland with Sweden (McClintock and Wikström 1990), the former Soviet Union with Western countries (Shelley 1981), and the Nordic countries among each other (von Hofer 1997). Such studies allowed an accurate assessment of the validity of several indicators in the countries studied, due to the scholars’ familiarity with the context. However insightful and suggestive these reports may be, studies limited to comparisons between just two or three countries do not allow theories to be tested (since two cases do not constitute a sample). Archer and Gartner (1984) conducted one of those rare, truly multi-national studies, using a century’s worth of crime statistics for as many as 40 countries.

Multi-National Data Collections

Multi-national data collection initiatives are considered to be more useful than traditional cross-country studies because they provide relevant data on many countries. It is accepted that such studies will never represent an ideal random sample of all existing nations because of the fact that there is no clearly defined universe out of which a random sample could be drawn. Over the last few decades, a number of multi-nation data collections have been developed. One of the first such projects was developed by Interpol, which has published data on police-recorded crimes for a significant number of countries biennially since 1950, and annually since 1993, entitled International Crime Statistics (Interpol 2002). However, it should be noted that the relevant national authorities collect the data and Interpol publishes the information as they receive it, without too much scrutiny. Given the numerous problems associated with the measurement of crime, this type of data collection does not help to reduce the widespread skepticism towards comparative data among criminologists. The United Nations (UN) Crime Surveys,
which have been conducted every five years since 1973, collect data on police-
recorded offences and also criminal justice data on selected topics such as sen-
tencing and prison populations. As these surveys rely on official national sources, the data shares many of the shortcomings of the Interpol crime statistics. Recently, however, several United Nations Institutes (HEUNI in Helsinki, Finland, UNICRI in Torino, Italy) have devoted considerable resources and expertise in scrutinizing the raw data (Kangaspunta, Joutsen, and Ollus 1998). Thus, the UN surveys have become a major data source, particularly when interpreted using other sources of international data, such as the International Crime (Victim) Survey (ICVS) (Marshall 1998). In addition, they have provided valuable data on several criminal justice indicators whose validity is less problematic, such as prison populations (Kuhn 2000, 1998).

A related initiative came from the Council of Europe shortly after many nations in Eastern and Central Europe had joined this organization. This initiative drew on the experience of the UN crime and criminal justice surveys, but was inspired mainly by the American Sourcebook of Criminal Justice Statistics, an annual publication by the United States’ Department of Justice dating back to 1973. The working group in charge of the new Council of Europe initiative set up a network of national correspondents in each country to act as “experts” rather than official government representatives. Thus, the focus was on the “expertise” and not on the official position of those providing the data. In addition, each member of the working group was responsible for overseeing and cross-checking the data from a number of other countries. Data were not merely accepted but were scrutinized in an attempt to introduce a degree of quality control (Killias and Rau 2000). Although not yet completely satisfactory, the figures collected using this process were considered to be much more reliable than any of the past data collections. It was particularly valuable to combine data from the international crime surveys with data on police-recorded offences and offenders, convictions (including sentences imposed), and corrections (time served, number of prisoners, or serving other sanctions, as collected since the 1980s by the Council of Europe, under the responsibility of Pierre Tournier and now Marcelo Aebi). This kind of data also provides information on the “costs” of offending in the different countries and, thus, on the “outcomes” of more or less punitive policies. The European Sourcebook of Crime and Criminal Justice Statistics was first published in 1999 using data collected for 1990 to 1996 for 36 countries. The second edition, due to be published in 2003, will include more countries.

1988 saw the development of the International Crime (Victim) Survey (ICVS). The emergence of computer-assisted telephone interviews (CATI) made it possible to conduct standardized surveys with large national samples in many countries at affordable costs. The ICVS drew mainly on two methodologies. First, it drew largely on the methodology successfully tested in two national Swiss crime surveys (1984 and 1987; possibly the first such surveys ever conducted using CATI). Second, it also drew on a questionnaire that was mainly inspired by
the British Crime Surveys (whose main researcher was Pat Mayhew). The first ICVS was conducted in 1989 in 11 Western European countries, the United States, Canada, and Australia (van Dijk, Mayhew, and Killias 1990). The idea for this survey originated in the Research and Documentation Center of the Ministry of Justice of the Netherlands (whose director at that time was Jan van Dijk). The Center continued to play a decisive role in this project over the years. The survey continues to be conducted. Data collection has occurred in 1992, 1996, and 2000 (Mayhew and van Dijk 1997; van Kesteren, Mayhew, and Nieuwebeerta 2000), with the inclusion of an increasing number of countries from outside the Western hemisphere (Alvazzi del Frate 1998).

Survey or Police Data?

The validity of the ICVS, as well as the validity of data on police-recorded crime, has always been the subject of debate. Skogan (1993), in a review of the first ICVS, was quite positive on a number of issues, but was not convinced that the observed cross-sectional differences were “real.” Other researchers mentioned (in meetings and informally rather than in print) that ICVS rates for European countries “must” be wrong because they look so similar. In contrast to this skepticism about survey data, the authors of the European Sourcebook (1999:32) warn that crime levels according to police data mostly reflect differences in the recording of reported crime by the police and that, as a consequence, police data should be used for trend analyses rather than for comparative purposes.

During the first ICVS, the authors made an attempt to correlate survey data on several offences with police data gathered from Interpol statistics (van Dijk, Mayhew, and Killias 1990:180-182). The results showed some match, although discrepancies remained. Specifically, once survey-generated rates were adjusted for reporting to the police, the rank-order (Spearman rho) correlations between survey measures of crime (ICVS 1989) and Interpol’s International Crime Statistics for all the 14 countries included were: vehicle theft (.714), burglary with entry (.473), robbery (.666), assault/threats (.653), and sexual offences (.835).

At the time, the authors were more concerned with establishing the credibility of the survey data than with questioning the police data (van Dijk, Mayhew, and Killias 1990:106). Since that time, a number of studies have shown the enormous influence of different recording practices on police-measures of crime levels. In England and Wales, for example, a study has shown that less than 40 percent of all violent incidents that are brought to the attention of the police are actually recorded in police files and statistics (Burrows et al. 2000). An observational study conducted in Germany (Kürzinger 1978) more than 20 years ago yielded very similar results for violent offences (only 30 percent were recorded), whereas recording is more systematic in cases of theft, presumably because victims need a police report for insurance purposes. Therefore, it can reasonably be assumed
that, in many Western countries, minor offences of assault and other offences (where insurance coverage is unlikely), even if reported, have a lower probability of being recorded. By handling them informally, police officers may save resources since investigating violent offences is always time consuming and they may simultaneously improve clearance as well as crime rates (Loveday 1999; Burrows et al. 2000). In countries where (minor) assault is prosecuted only at the request of the victim, police data on reported offences tend to closely match police counts of suspects, suggesting that offences are recorded only once a suspect is known and whose prosecution is being demanded by the victim (European Sourcebook 1999:44, 59). Similarly, rates for convicted rapists tend to correlate inversely with the rate of recorded offences of rape (Barclay 2000). On the other hand, political pressure may also encourage the police to record all incidents in serial offence cases even if their number needs to be extrapolated. This was the case for wife abuse in Sweden where police officers assessed the number of offences by multiplying their monthly or weekly frequency by the duration of the relationship (von Hofer 2000). In summary, police counts of crime are highly affected by:

- the moment at which an offence is recorded, i.e., either at the time of reporting to the police (input), or once the police investigation has begun (output);
- the way “secondary” offences (such as forgery in a fraud case) are counted; and
- the number of incidents that are counted in cases of serial offences, such as ongoing child abuse, violent abusive relationships, or drug transactions that have lasted over time.¹

Such differences in recording practices will, of course, always affect comparisons. As long as recording by the police is not being standardized across countries, it is difficult to see how the validity of police data might be improved for comparative purposes, although such problems may affect longitudinal analyses based on police statistics to a lesser extent. On the other hand, international collections of police and justice data will suffer also from errors of a more trivial nature. As mentioned before, the distinctive characteristics of the European Sourcebook has been the establishment of a network of correspondents, and their selection according to competence rather than their official position, and also the screening of all data provided by the national correspondents by the experts in charge of the project (Killias and Rau 2000). This approach differs strongly from what had been standard practice in many other initiatives of this kind, where even blatantly implausible figures usually remained unquestioned since they were officially communicated by the nations. Despite the efforts made by the authors of the European Sourcebook to eliminate trivial errors (i.e., confusion of categories, years, and/or offences/offenders, etc.), many of them were not discovered during the first survey. For example, police-recorded rates of domestic burglary varied in
1996 between the countries with the lowest (Macedonia) and the highest rate (England and Wales) by 1:145; this is simply unbelievable. By restricting the comparison to Western Europe, England and Wales and Austria (with the lowest rates) still differ by a margin of 7:1, whereas ICVS rates (for 1996) show a ratio of only 3:1 for these two countries (Killias 2002:52; 2001:48). Obviously, there are differences in recording practices within Western Europe that cannot be eliminated by screening, but there are also many “odd” figures that had not been checked sufficiently in the 1999 version. The experts in charge of the 2003 edition of the Sourcebook have, therefore, decided to devote more time and energy in assessing the quality of the data. It remains to be seen to what extent this extra investment will improve the match between different international measures of crime.

Matching European Sourcebook Police Data with ICVS Data

In order to assess the match between police and survey findings, data from the 2000 ICVS (on incidents reported to the police in 1999) were correlated with 1999 police data from the 2003 edition of the European Sourcebook, as well as from Interpol statistics. We paid particular attention to making the crime categories as comparable as possible, a task that proved impossible to achieve to our satisfaction. The following analysis covers 12 European countries for which ICVS (2000) and Sourcebook data are available for 1999. These are Belgium, Denmark, England and Wales, Finland, France, the Netherlands, Northern Ireland, Poland, Portugal, Scotland, Sweden, and Switzerland. ICVS data on incidence rates were taken from the latest available publication (van Kesteren, Mayhew, and Nieuwbeerta 2000:180-181) and adjusted according to the percentage of offences reported to the police (van Kesteren, Mayhew, and Nieuwbeerta 2000:194-195). European Sourcebook data was taken from the preliminary version of the 1995-2000 European Sourcebook of Crime and Criminal Justice Statistics. Finally, Interpol International Crime Statistics were taken from Interpol’s Website (http://www.interpol.int/Public/Statistics/ICS/Default.asp).

Adjusting Police Categories

The comparison implied that survey data, as well as police data, required adjustment in several ways. In the case of police data, categories that most closely matched survey definitions were used, such as “domestic burglary” (Sourcebook) or “breaking and entering” (Interpol). In the case of car theft, we used “theft of motor cars” (Interpol) and “theft of motor vehicle” (Sourcebook). In the case of “rape” (according to Interpol and Sourcebook) we use the wider survey measure of “sexual incidents,” which also included more trivial forms of sexual harassment or intimidation, rather than the more restrictive categories of rape, attempted rape, and indecent assault, because the frequencies are too low (given the samples of 1,000 or 2,000 in the several countries) to allow valid cross-national comparisons.
In the case of “assault/threat” (ICVS), we used the Sourcebook measure for “assault” and the Interpol category for “serious assault.” In this context, it should be noted that in countries that have French as an official language such as Belgium, France, and Switzerland (and in Poland, a country with a strong attachment to the French culture), figures provided by Interpol for “serious assault” are almost identical to figures provided by the Sourcebook for “assault.” This may be due to a translation problem. In fact, the French category used by Interpol (coups et blessures) includes all kinds of assault, while the Spanish category (lesiones graves) and the English category refer only to serious assault. It should also be kept in mind that some forms of assault might be recorded under other crime categories in official records, such as threat, extortion, or certain forms of “harassment.” In the case of “personal theft” (ICVS), we used the category of “other thefts” in Interpol statistics, whereas in the Sourcebook, vehicle thefts and all burglaries were deducted from the category of “total theft.” The police category of robbery includes “theft with violence” (Interpol) and “bag-snatching” and mugging (Sourcebook), as did the survey definition, although the differentiation between theft and robbery may be problematic in many respects.

Adjusting Survey Categories

In order to become comparable to police data, survey data need, obviously, to be adjusted for reporting to the police. Victims are indeed differentially inclined to report incidents they experience to the police, for a number of reasons that are not of interest here. Some offences, such as car theft or burglary, are usually measured at the household level in surveys; this calls for some adjustment before levels can be compared with police data (which are per 100,000 population). This is not quite clear in the published data for the ICVS (2000), as the table that provided the data used here is labeled “Victimization in the year preceding the survey: number of offences per 100 inhabitants (incidence rates)” (van Kesteren, Mayhew, and Nieuwbeerta 2000:180-181), suggesting that all rates are calculated on the basis of population and not of households. Nevertheless, the ranking of European countries should not be affected by the choice of the denominator. It should also be noted that Interpol and Sourcebook data always provide rates per 100,000 populations. The same is true for ICVS data with the exception of rates for car theft, which are calculated according to car owners, and sexual incidents, which are calculated according to female population. Nevertheless, since we are calculating correlations and not simply comparing rates, this particularity does not affect our analysis. Finally, surveys also include offences experienced abroad, whereas they are not included in police statistics. In the present study, no such adjustment has been made, given that the fraction of crime experienced abroad is probably far too low to affect the match between survey and police data.
**Results**

**Correlational Analyses**

Table 1 gives the Pearson-Bravais (r) as well as the rank-order correlations (Spearman’s rho) between police and survey measures of comparable offences for the 12 countries included in the present study.

<table>
<thead>
<tr>
<th>Offence type</th>
<th>International Crime (Victim) Survey and European Sourcebook</th>
<th>International Crime (Victim) Survey and Interpol Statistics</th>
<th>European Sourcebook and Interpol Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pearson’s r</td>
<td>Spearman’s rho</td>
<td>Pearson’s r</td>
</tr>
<tr>
<td>Car theft</td>
<td>.67*</td>
<td>.73*</td>
<td>.12</td>
</tr>
<tr>
<td>Domestic burglary</td>
<td>.90**</td>
<td>.86**</td>
<td>.60*</td>
</tr>
<tr>
<td>Personal theft</td>
<td>.81**</td>
<td>.74**</td>
<td>.51</td>
</tr>
<tr>
<td>Robbery</td>
<td>.25</td>
<td>.45</td>
<td>.12</td>
</tr>
<tr>
<td>Assault</td>
<td>.67*</td>
<td>.76**</td>
<td>.38</td>
</tr>
<tr>
<td>Sexual offences</td>
<td>.31</td>
<td>.37</td>
<td>.35</td>
</tr>
</tbody>
</table>

* p < .05, ** p < .01

Given the many difficulties in assessing the volume of crime, the match between police (Sourcebook) and survey measures is surprising. This is particularly true for the police data from the European Sourcebook, whereas the match between survey and Interpol data is much less satisfactory. Interpol statistics often show wide variations, which are most likely due to the kind of errors the Sourcebook experts tried to eliminate. Interpol statistics also correlate rather poorly with police data from the European Sourcebook, a fact that underlines the improvements achieved through the careful data screening.

Some of the correlations in Table 1 may be low because it has not been possible to perfectly match the crime categories. This is certainly the case for car theft, sexual offences, and assault. In the case of assault, differences in recording practices, as outlined above, may play a role that even careful screening cannot overcome; besides, as we also mentioned before, differences between the French and English categories of assault in Interpol statistics could explain the poor correlation between these statistics and the European Sourcebook. In the case of sexual offences, the correlation might have been higher if only survey measures of completed and attempted rape and sexual assault could have been used (but, as we mentioned before, the frequencies are too low to do that given the samples of 1,000 or 2,000 in most countries). An interesting case is the absence of any correlation between survey and police measures of robbery, due mostly to the police-recorded robbery scores of Belgium, France, and Portugal, which are far higher.
than expected given these countries’ ICVS scores. Difficulties in differentiating robbery, violent theft, pick-pocketing, and other forms of personal theft at the recording stage might be responsible for this result.

Examination of the correlational diagrams (Figures 1-6, Appendix) further illustrates the match between police data from the Sourcebook and ICVS measures. For assault, personal theft, and especially for burglary, the match is indeed convincing since none of these correlations is inflated by extreme positions of one or two countries. On the other hand, the correlation for robbery suffers from the odd positions of three countries mentioned. In the case of sexual incidents, the Netherlands and England and Wales present quite high rates in the survey and appear as outliers in the figure, but even without these countries the correlation is not strong. Nevertheless, overall, it seems that police and survey data catch the rankings of the countries rather convincingly.

**Level Comparisons**

Despite this encouraging finding, it should be noted, however, that levels of crime measures do not necessarily match. Leaving aside “car theft” / “motor vehicle theft,” “personal theft,” and “sexual offences” / “rape,” where the match of offence definitions is more questionable, survey measures are (as one might expect) consistently higher than police figures, even after an adjustment is made for offences not reported by victims. However, police rates often are far lower than expected given survey estimates. For example, for robbery only about 1 in 10 incidents (as one would expect) appears in police statistics in some countries; it may be that attempts, particularly frequent in this domain, often go unrecorded, or that reports made by victims may be recorded under other categories (such as personal theft). Similarly odd differences can be found for assault where, again, recording practices may be responsible for the mismatch. In the case of burglary, police estimates are more in line with survey measures, although police figures are still only about half of what one would expect them to be. In this instance, attempted burglaries may explain part of the discrepancy since such incidences have a low reporting rate and may, particularly if the offender could not enter the premises and no damage occurred, be recorded rather reluctantly.

To sum up, level comparisons continue to be difficult, although offences against the person may be more affected by problems at the recording stage than burglary and other property offences where insurance-related interests of the victim offer some guarantee for a more systematic registration. However, it should be kept in mind that survey measures are not necessarily the gold standard, and that they might have, in several instances, grossly overestimated levels of crime. In the present case, however, the ICVS measures are rather valid because of the tight standardization of survey definitions, the uniform methodology, and the way problems of telescoping were addressed, for instance, by using double reference periods."
Conclusions

Several conclusions can be drawn from our analysis. First, careful screening of police data is necessary before reasonable comparisons can be made. Such efforts do not eliminate problems related to different recording practices, but may help to overcome the most trivial (but often devastating) errors concerning categories, years, offences/offenders, and so on. Second, international data collections that do not undergo such processes (such as the Interpol statistics) should not be used for comparative purposes. Third, carefully screened data (as those from the European Sourcebook) may be reasonably valid in multi-national (correlational) comparisons. Fourth, level comparisons based on police statistics remain of highly questionable validity, particularly if two or three countries are being compared, whereas such problems will be less relevant in a multi-national (correlational) analysis. Fifth, crime surveys offer more valid comparisons if definitions of measures, survey methodology, and reference periods are tightly standardized. Sixth, and beyond these specific conclusions, criminology will always have to navigate between using data whose validity might remain questionable in some respects, and using no data at all. The latter would mean simply abandoning all efforts for testing theories empirically and leaving uncontested even the most absurd, useless, or devastating conclusions.

NOTES

1 Will such incidents be counted as one offence, or will the police extrapolate their “precise” number by considering the time relationship? For example, drug offences are, in many countries, counted as one incident as long as a dealer has only sold products of one substance, no matter how many transactions were carried out over the period in question.

2 For a few countries, Interpol police data were not yet available for 1999. Therefore, 1998 data have been used for Belgium, England and Wales, the Netherlands Scotland, and Sweden; for Northern Ireland, the latest available figures (1996) were used.

3 For example, for burglary and car theft, the rates given for the highest country is more than 20 times the lowest rate; in the case of personal theft, the ratio is even 1:200. Such differences are impossible within the 12 European nations included here.

4 Of course, Interpol experts are aware of these problems and on their Website they indicate: “It should be pointed out that the Interpol General Secretariat merely reproduces the information sent to it by the NCBs. The information given is not therefore processed, but is classified according to category of offence. The data
gathered in these sets of statistics is not intended to be used as a basis for comparisons between different countries since the statistics cannot take account of the differences which exist between definitions of punishable acts in different national laws, or the diversity of statistical methods, or the changes which may occur during the reference period and affect the data collected. Police statistics reflect reported crimes, but this only represents a fraction of the real level. Furthermore, the volume of crime not reported to the police actually depends, to a certain extent, on the action of the police and can vary with time, as well as from country to country. Consequently, the data published in the current set of statistics should be interpreted with caution” (Interpol 2002).

Screening questions should always pertain to relatively long periods (e.g., five years, as in the ICVS), followed by a follow-up question concerning the exact temporal location of reported incidents (e.g., during the last 12 months or before). Many surveys in Europe have not kept to that format and may, as experiments have shown, have grossly over-estimated levels of crime (Killias 2002:75-78; 2001:71-73).

APPENDIX

Figure 1. Linear Regression for Motor vehicle (Car) Theft

![Linear Regression Graph for Motor vehicle (Car) Theft]
Figure 2. Linear Regression for Domestic Burglary (including Attempts)

Figure 3. Linear Regression for Personal Theft
Figure 4. Linear Regression for Robbery (including Attempts)

Figure 5. Linear Regression for Assault / Assault and Threats
REFERENCES


Council of Europe.


