The Endogenous Analysis of Flows, with Applications to Migrations, Social Mobility and Opinion Shifts

FRANÇOIS BAVAUD [a]

[a] University of Lausanne, Switzerland

PLEASE SCROLL DOWN FOR ARTICLE

Routledge
Taylor & Francis Group

The Endogenous Analysis of Flows, with Applications to Migrations, Social Mobility and Opinion Shifts

François Bavaud

University of Lausanne, Switzerland

Two major traditions coexist in the statistical analysis of spatial and social change: mobility analysis (e.g., Boudon, 1973) initiated by Prais (1953) on one hand, and gravity modelling for geographic flows (e.g., Fotheringham and O'Kelly, 1989; Sen and Smith, 1995), initiated by Reilly (1931) on the other hand. This paper focuses on the formal content of both traditions and explores their relationship. Three small data sets (migrations, occupational mobility and opinion shifts) illustrate the issues. In particular, we show how the sojourn times and jump matrices of mobility analysis are related to the accessibilities and expansivities of gravity modelling. We also address in detail the issues of time-dependence and categories aggregation. Two well-behaved indices of mobility are defined and discussed, namely the average sojourn time, which measures short-time mobility, and the second eigenvalue of the transition matrix measuring long-time mobility.

1. INTRODUCTION

Spatial motion (inter-regional migrations, international communications, etc.) and social change (social mobility, opinion shifts, etc.) are quantified by a matrix of flows, counting the number of units (people, goods, news, opinions, etc.) initially located in some category and lying in some other category (possibly identical) after some reference time. Section 2 introduces the notations as well as three typical data sets, chosen for convenience and illustrative purposes.

Exogenous modelling consists in fitting or "explaining" the observed flows by tuning parameters associated to independent (i.e., differing

Address correspondence to François Bavaud, Section d'informatique et methodes mathematiques, Quartier UNIL-Dorigny, Batiment Anthropole, CH-1015, Lausanne, Switzerland. E-mail: Francois.Bavaud@unil.ch

from the flow data themselves) predictors, such that the effective distance between regions or the employment level (for migrations) or the socio-economic structures and government policies (for social mobility). By contrast, *endogenous* modelling investigates the configuration of flows itself, that is, the relationship between their components, typically by adjusting the observed flows to some low-dimensional theoretical flow families of interest. In this paper, we focus on the two most investigated endogenous models of flows (whose relationship is discussed in 5), namely

- the *quasi-symmetric* models (section 3), which correspond to the endogenous content of the gravity models of geographers, as, for example, further developed by Fotheringham and O'Kelly (1989), Sen and Smith (1995), Weidlich and Haag (1988) or Bavaud (2002).
- the *Markov embeddable* models (section 4), which postulate that the Markov transition matrix associated to the observed flows can be extended to continuous times, or equivalently that the dynamics of flows is governed by a *transition rate* matrix.

In the scientific literature of the last 50 years or so, exogenous approachs overwhelmingly outnumber endogenous studies. This historical circumstance seems hard to justify: rushing into exogenous modelling without *first* paying enough attention to endogenous modelling is questionable, for the origin (endogenous or exogenous) of a possible misfit cannot be identified anymore (confounding), and makes inter-author and inter-data comparisons problematic, despite the enormous amount of empirical research done ever since the pioneering works of Reilly (1931) and Prais (1955).

Also, the basic and fundamental issue of *mobility indices*, measuring the importance of the change; that is, the speed of spatial or social motion, exclusively depend on endogenous considerations by construction. It would of course be desirable to be able to compare the mobilities of two systems differing by their granularity (the number of categories) or their reference time, that is to be able to construct *aggregation-invariant* and *time-invariant* mobility indices.

The question of categories aggregation is explored in sections 3.3 and 4.5. While the aggregation of embeddable flows is automatically embeddable, it is not the case for quasi-symmetric flows, unless a special condition (apparently not documented so far) we call *equi-expansivity* holds.

The issue of the time dependence of parameters is addressed in sections 4.2, 5 and 5.3. Section 5.3 introduces a presumably new family of information-theoretical mobility indices and emphasizes the

dynamical significance of the (harmonic, weighted) average sojourn time, measuring the short time mobility, as well as of the second eigenvalue of the transition matrix, measuring the long-time mobility.

This paper present original results (Theorems 4, 8, 9, and 10) or extensions of previously known results (Theorems 1, 5, and 7). It also recalls relevant lesser-known results or otherwise scattered in the litterature (Theorems 2, 3, and 6). Familiar indices are put in a presumably original perspective, aimed at elucidating and emphasizing their formal properties and relationships.

2. NOTATIONS AND DATA

Let $n_{jk}(T)$ be a square $(m \times m)$ flow matrix, that is a contingency table counting the number of units being in state j $(j = 1, \ldots, m)$ at some initial time and in state k $(k = 1, \ldots, m)$ after some observation time $T > 0$. Tables of migration (Table 1), social mobility or opinion shifts (Table 2) are examples of flow matrices.

2.1. Markov Transitions

The j-th row marginal $n_{j\bullet} := \sum_{k=1}^{m} n_{jk}$ counts the number of units initially in j. The k-th column marginal $n_{\bullet k} := \sum_{j=1}^{m} n_{jk}$ counts the

TABLE 1 Observed Inter-Regional Migration Flows $n_{jk}(T)$ Relatively to 7 Swiss Regions for the Period 1985–1990. The Regions are Built up from the Aggregation of the 26 Cantons, Namely Region 1 = "ZH, ZG, AG"; 2 = "BE, FR, NE"; 3 = "LU, UR, SZ, OW, NW, GL"; 4 = "SO, BS, BL, JU"; 5 = "SH, AR, AI, SG, GR, TG"; 6 = "TI"; 7 = "VD, VS, GE". Row Margins Give Total Counts $n_{j\bullet}$ in 1985, and Column Margins Give Total Counts $n_{\bullet k}$ in 1990. Diagonal Values Dominate: Most People Did Not Move to Another Region in the Period Under Consideration

To region →	1	2	3	4	5	6	7	Total
From Region ↓								
1	1′488′047	11′446	14′641	11′799	25′674	3′716	5′130	1′560′453
2	11′856	1′131′965	3′427	11′684	4′346	1′353	14′774	1′179′405
3	14′857	4′410	478′442	2′626	5′125	1′028	1′401	507′889
4	13′836	12′764	2′750	621′088	3′321	986	3′691	658′436
5	26′139	5′783	5′339	3′175	785′784	2′072	1′999	830′291
6	4′605	1′649	690	863	1′393	242′380	1′813	253′393
7	6′001	14′314	1′189	2′531	1′548	1′151	1′022′712	1′049′446
Total	1′565′341	1′182′331	506′478	653′766	827′191	252′686	1′051′520	6′039′313

TABLE 2 Left: British Occupational Mobility Data (Glass, 1954, Cited in Bishop et al. [1975]), Where Row j Determines the Father's Occupation, and Column k the Son's Occupation, with $n = n_{\bullet\bullet} = 3'500$. The Categories are $1 =$ "Professional Occupations"; $2 =$ "Managerial and Technical Occupations"; $3 =$ "Skilled Non-Manual and Manual Occupations"; $4 =$ "Partly-Skilled Occupations"; $5 =$ "Unskilled Occupations". Right: Attitude Towards Belonging to the EEC (Anderson, 1990) in October 1971 and in December 1973 for a Random Sample of $n = 493$ Danes. The Categories are $1 =$ "Yes"; $2 =$ "No"; $3 =$ "Undecided"

$$(n_{jk}) = \begin{pmatrix} 50 & 45 & 8 & 18 & 8 \\ 28 & 174 & 84 & 154 & 55 \\ 11 & 78 & 110 & 223 & 96 \\ 14 & 150 & 185 & 714 & 447 \\ 3 & 42 & 72 & 320 & 411 \end{pmatrix} \qquad (n_{jk}) = \begin{pmatrix} 167 & 36 & 15 \\ 19 & 131 & 10 \\ 45 & 50 & 20 \end{pmatrix}$$

number of units finally in k. The total number of units is $n = n_{\bullet\bullet} := \sum_{jk} n_{jk}$. The matrix

$$W = (w_{jk}) \qquad w_{jk} := \frac{n_{jk}}{n_{j\bullet}} \tag{1}$$

is non-negative with rows normalized to unity: W is thus, formally, the transition matrix of a Markov chain. We assume the chain to be *regular* (see, e.g., Karlin and Taylor (1975)), that is, to be

(a) *irreducible*, namely each category can be reached from each other category by following the random walk during a finite number l of steps.

(b) *aperiodic*, namely the greatest common divisors of all possible l above is one.

Condition (a) is violated in the presence of migration regulations forbidding citizens from a country to ever reach another country, or of tight cultural barriers ruling out specific socio-economic transitions. Condition (b) is violated for strict commuters flows between housing and working places, of period two.

Regular Markov chains possess a unique stationary or long-term distribution π obeying $\sum_j \pi_j w_{jk} = \pi_k$. By contrast, $\sum_j f_j(0) w_{jk} = f_k(T)$.

2.2. Shorrocks' Index

Three special classes of Markov chains will be considered further, namely

- *perfect mobility*: $w_{jk} = \pi_k$
- *frozen flows*: $w_{jk} = \delta_{jk}$
- *off-diagonal flows*: $w_{jj} = 0$.

Shorrocks (1978) defines a *mobility index s* as

$$s := \frac{m-d}{m-1} \quad \text{where } d := \text{trace}(W) = \sum_j w_{jj} \tag{2}$$

whose minimum value $s = 0$ is attained for frozen flows, while $s = 1$ for perfect mobility. Its maximum value $s = \frac{m}{m-1} > 1$ is attained for off-diagonal flows—see, however, section 5.1.

2.3. Attractivities

Size-independent attractivities between categories j and k can be measured by the *quotients* q_{jk}, or by their symmetrized versions, the *accessibilities* $g_{jk} = g_{kj}$:

$$q_{jk} := \frac{n_{jk}n}{n_{j\bullet}n_{\bullet k}} \qquad g_{jk} := \sqrt{q_{jk}q_{kj}}$$

By construction, $g_{jk} > 1$ (respectively $g_{jk} < 1$) if spatial interaction or social transitions between j and k is higher (respectively, smaller) than on average.

2.4. Flows Aggregation

Let $J = 1, \ldots, M$ denote coarse-grained categories, obtained by aggregating the finer categories $j = 1, \ldots, m$ belonging to the same J ($M \leq m$). The aggregated flows are

$$n_{JK} = \sum_{j \in J} \sum_{k \in K} n_{jk}$$

3. QUASI-SYMMETRIC MODELS

Some simple model candidates for expected flows associated to Tables 1 and 2 can be quickly ruled out by inspection:

- *Independence models* of the form $n_{jk}^{\text{theo}} = \alpha_j \beta_k$ are unsuitable, when considering the large numerical values of the diagonal. The latter are actually strongly related to the observation time $T > 0$: the smaller T, the less changes can occur; in the limit $T \to 0$, the flow matrix n_{jk} becomes diagonal (see section 5).

- Independence models for off-diagonal components of the form $n_{jk}^{\text{theo}} = \alpha_j \beta_k$ for $j \neq k$ (quasi-independence models) are unsuitable: the transition rates is smaller between far away social categories than close ones; the distance deterrence makes the relative exchange of migrants between near regions more frequent than between distant regions.
- *Symmetric models* of the form $n_{jk}^{\text{theo}} = \gamma_{jk}$ with $\gamma_{jk} = \gamma_{kj}$ are unsuitable as well: during migration, some regions might grow and other shrink; a political opinion can gain more voices during the transition; some occupational categories can become more or less represented.

The simplest satisfactory model for data sets of Tables 1 and 2 (and many others) is the *quasi-symmetric model* (QS), first proposed and investigated by Caussinus (1965). The model is defined by $n_{jk}^{\text{theo}} = \alpha_j \beta_k \gamma_{jk}$ with $\gamma_{jk} = \gamma_{kj}$, but possesses many equivalent definitions (Bavaud, 2002). Bishop et al. (1975) describe the QS model in some detail; applications of the latter can be found at http://www.lsp. ups-tlse.fr/Projet_QS.

The QS model is nothing but the endogenous formulation of the *gravity model* of geographers (see, e.g., Sen and Smith (1995) or Fotheringham and O'Kelly (1989)), we shall henceforth describe in its original context, namely population migrations:

$$n_{jk}^{\text{theo}} = \alpha_j \begin{pmatrix} \text{origin socio-} \\ \text{economic conditions} \end{pmatrix} \beta_k \begin{pmatrix} \text{destination socio-} \\ \text{economic conditions} \end{pmatrix}$$
$$\gamma_{jk} \begin{pmatrix} \text{effective origin-} \\ \text{destination distance} \end{pmatrix} \tag{3}$$

where α_j is the *push effect*, β_k the *pull effect* and γ_{jk} the *distance deterrence* effect; by contrast, the exogenous part of gravity modelling, historically much more investigated, consists in seeking satisfactorily functional forms for α_j, β_k and γ_{jk} in terms of exogenous predictors.

Paradoxically, the quasi-symmetric nature of gravity flows has hardly been recognized as such by geographers (Willekens (1983) constitutes an exception); this state of things might partially be explained by the fact that typical constraints entering into the popular maximum-entropy flow modelling (Wilson, 1971) produce counts that are automatically quasi-symmetric. We develop in section 3 a parameterization of quasi-symmetric flows in terms of accessibilities and expansivities, initiated by Bavaud (2002).

Equation (3) can be canonically normalized, allowing a fairly intuitive behavioral interpretation (in the Bradley-Terry-Luce decision formalism) through disentanglement of size, distance deterrence and push-pull effects (Bavaud, 2002):

Theorem 1 (proof in appendix). *Quasi-symmetric flows n_{jk}^{theo} decompose in a unique way as*

$$n_{jk}^{\text{theo}} = \underbrace{n}_{absolute\ size} \cdot \underbrace{f_j^m \cdot f_k^m}_{relative\ size} \cdot \underbrace{g_{jk}}_{distance\ deterrence} \cdot \underbrace{\exp(u_k - u_j)}_{push-pull} \qquad (4)$$

where[1]

$$n := n_{\bullet\bullet}^{\text{theo}}\ total\ flow \qquad (5)$$

$$f_j(0) := \frac{n_{j\bullet}^{\text{theo}}}{n}\ initial\ population\ distribution \qquad (6)$$

$$f_k(T) := \frac{n_{\bullet k}^{\text{theo}}}{n}\ final\ population\ distribution \qquad (7)$$

$$f_j^m := \sqrt{f_j(0)f_j(T)}\ mixed\ profile \qquad (8)$$

$$g_{jk} := \sqrt{\frac{n_{jk}^{\text{theo}} n_{kj}^{\text{theo}}}{n f_j^m f_k^m}}\ pairwise\ accessibility \qquad (9)$$

$$u_j := \frac{1}{2} \sum_{l=1}^{m} a_l \ln \frac{n_{lj}^{\text{theo}}}{n_{jl}^{\text{theo}}}\ regional\ expansivity \qquad (10)$$

The accessibilities $g_{jk} = g_{kj}$, reflecting the easiness in reaching j from k and vice-versa, express the origin-destination attractivities (see section 2). While accessibilities have been constructed so as to constitute size-independent indices, their diagonal components exhibit in general an effective size dependence due to the nature of the migratory dynamics itself (see section 5).

Expansivities u_j, supposed to reflect the socio-economic attractivity of the j-th region, have been normalized to $\sum_{j=1}^{m} a_j u_j = 0$, where the

---

[1]The parameters in (4) are in one-to-one correspondence with the parameters of the log-linear formulation (see, e.g., Bishop et al., 1975) $n_{jk}^{\text{theo}} = \exp(\kappa + \lambda_j + \mu_k + \nu_{jk})$ with $\nu_{jk} = \nu_{kj}$, $\lambda_\bullet = 0$, $\mu_\bullet = 0$ and $\nu_{\bullet k} = 0$; however, the correspondence is not particularly enlightening. Although formally convenient, log-linear parameters possess no easy interpretation nor good aggregation properties.

weights a_j (with $a_j \geq 0$ and $\sum_{j=1}^m a_j = 1$) quantify the relative impor-
tance of the regions. As such, weights are arbitrary.[2] Under time
reversal, that is under transposition of the flow matrix n_{jk}, expansiv-
ities transform as $u_j \to -u_j$, provided weights are invariant under
time reversal. Examples of such weights are the uniform weights,
$a_j = 1/m$, the standard weights $a_j = \frac{1}{2}(f_j(0) + f_j(T))$, or any exogenous
weights a_j independent of the flow under consideration.

During the transition time T, the number of units in category j
expands (respectively, shrinks) if $u_j > 0$ (respectively $u_j < 0$), at least
on average, in virtue of the inequality $\sum_j u_j(f_j(T) - f_j(0)) \geq 0$, proved
in Theorem 9.

Theorem 2 (Bavaud, 2002). *Quasi-symmetric flows n_{jk}^{theo}*

- *are symmetric iff $u_j = 0$ for all j*
- *are independent iff $g_{jk} = 1$ for all j,k.*

Symmetric flows characterize a system in *equilibrium*, with con-
stant occupation of each category, that is, $f_j(0) = f_j(T)$. Independent
flows $n_{jk}^{\text{theo}} = \alpha_j \beta_k$ characterize the perfect mobility case (Prais, 1955).

Among competing estimation procedures for n_{jk}^{theo}, the maximum-
likelihood method is arguably the most convenient (Bavaud, 2002),
and results in obtaining quasi-symmetric estimates n_{jk}^{theo} obeying

$$n_{j\bullet}^{\text{theo}} = n_{j\bullet} \qquad n_{jk}^{\text{theo}} + n_{kj}^{\text{theo}} = n_{jk} + n_{kj} \tag{11}$$

Solution $n_{jk}^{\text{theo}} = b_j^{(\infty)} c_{jk}^{(\infty)}$ to (11) can be obtained numerically by iteration

$$b_j^{(r)} := \frac{n_{j\bullet}}{c_{j\bullet}^{(r)}} \quad c_{jk}^{(r+1)} := \frac{n_{jk} + n_{kj}}{b_j^{(r)} + b_k^{(r)}} \tag{12}$$

with some initial condition such as $b_j^{(0)} := 1/m$. Table 3 shows the result
of the procedure applied on the migration data. The resulting expansiv-
ities (10) and accessibilities (9) (with standard weights) are

$$(u_j) = \begin{pmatrix} .014 \\ .021 \\ -.022 \\ -.054 \\ -.029 \\ -.028 \\ .030 \end{pmatrix} \quad (g_{jk}) = \begin{pmatrix} 3.68 & .04 & .11 & .08 & .12 & .06 & .02 \\ .04 & 4.90 & .04 & .10 & .03 & .03 & .07 \\ .11 & .04 & 11.23 & .05 & .08 & .04 & .01 \\ .08 & .10 & .05 & 8.71 & .04 & .03 & .03 \\ .12 & .03 & .08 & .04 & 6.91 & .05 & .01 \\ .06 & .03 & .04 & .03 & .05 & 22.86 & .03 \\ .02 & .07 & .01 & .03 & .01 & .03 & 5.60 \end{pmatrix} \tag{13}$$

[2]Defining u_j^a (respectively u_j^b) as the left-hand side of (10) with weights a_j (respect-
ively b_j) and substituting (4) yields $u_j^a = u_j - \sum_l a_l u_l$ and $u_j^b = u_j - \sum_l b_l u_l$, that is
$u_k^a - u_j^a = u_k^b - u_j^b$, as required from (4).

TABLE 3 Quasi-symmetric Estimates n_{jk}^{theo} Associated to Table 1, Obtained Iteratively from Eq. (12). Observe that $n_{j\bullet}^{theo} = n_{j\bullet}$, $n_{\bullet k}^{theo} = n_{\bullet k}$ and $n_{jj}^{theo} = n_{jj}$

To region →	1	2	3	4	5	6	7	Total
From region ↓								
1	1'488'047	11'800.76	14'199.44	11'949.00	24'746.69	3'979.87	5'730.25	1'560'453
2	11'501.24	1'131'965	3'722.22	11'239.44	4'772.84	1'416.59	14'787.68	1'179'405
3	15'298.56	4'114.78	478'442	2'605.81	5'192.65	853.70	1'381.49	507'889
4	13'686.00	13'208.56	2'770.19	621'088	3'322.90	947.07	3'413.28	658'436
5	27'066.31	5'356.16	5'271.35	3'173.10	785'784	1'734.85	1'905.22	830'291
6	4'341.13	1'585.41	864.30	901.93	1'730.15	242'380	1'590.07	253'393
7	5'400.75	14'300.32	1'208.51	2'808.72	1'641.78	1'373.93	1'022'712	1'049'446
Total	1'565'341	1'182'331	506'478	653'766	827'191	252'686	1'051'520	6'039'313

247

The off-diagonal accessibilities quantify the importance of *spatial interaction* between regions, that is the number of exchanged migrants. Spatial interaction is here maximum between region 1 and region 5.

The same procedure applied on British occupational mobility data (Table 2, left) yields expansivities and accessibilities (14)

$$(u_j) = \begin{pmatrix} -.203 \\ -.045 \\ -.078 \\ -.024 \\ .127 \end{pmatrix} \quad (g_{jk}) = \begin{pmatrix} 12.80 & 2.19 & .58 & .32 & .17 \\ 2.19 & 2.52 & 1.18 & .74 & .37 \\ .58 & 1.18 & 1.62 & 1.00 & .64 \\ .32 & .74 & 1.00 & 1.16 & .97 \\ .17 & .37 & .64 & .97 & 1.67 \end{pmatrix} \quad (14)$$

Corresponding results for the Danish opinion shifts data (Table 2, right) read

$$(u_j) = \begin{pmatrix} -.029 \\ .272 \\ -.559 \end{pmatrix} \quad (g_{jk}) = \begin{pmatrix} 1.63 & .31 & .80 \\ .31 & 1.86 & .81 \\ .80 & .81 & 1.91 \end{pmatrix} \quad (15)$$

The empirical fit of the QS model (obeying a total of $m(m-1)/2$ constraints: see, e.g., Bishop et al. (1975)) is satisfactory for the latter data (with a log-likelihood chi-square of $L2 = 4.67$ with df=6 for the British occupational mobility; $L2 = 0.05$ with df $= 1$ for the Danish opinion shifts), but much less for the Swiss migration data ($L2 = 785$ with df $= 21$). Recall, however, $L2$ to be proportional to the size n of the data, which is huge in this case and therefore bound to amplify any deviation from the model beyond any chance level: In *relative* terms, the empirical and theoretical proportions hardly differ.

3.1. Quasi-symmetry and Reversibility

A Markov transition matrix (1) satisfying the local balance condition $\pi_j w_{jk}(T) = \pi_k w_{kj}(T)$ (for all j, k) is said to be *reversible*. Berger and Snell (1957) refer to it as a "strict exchange system" in the social mobility context.

Theorem 3. $n_{jk}(T)$ *is quasi-symmetric* $\Leftrightarrow$ $w_{jk}(T)$ *is reversible.*

Theorem 3, whose proof is straightforward, seems to have been rediscovered independently by many researchers during the last decades. Under reversibility, the stationary distribution obtains as $\pi_j = (\sum_l \frac{w_{jl}}{w_{lj}})^{-1}$. Expansivities (10) are alternatively expressible as

$$u_j = \frac{1}{2} \ln \frac{\pi_j}{f_j(0)} - \text{cst} \quad \text{cst} := \frac{1}{2} \sum_{l=1}^{m} a_l \ln \frac{\pi_l}{f_l(0)} \quad (16)$$

which shows u_j to measure the ratio "long-range occupation" over "initial occupation:" In accordance with Theorem 2, expansivities are all zero iff the system is in equilibrium.

Also, accessibilities (9) express as

$$g_{jk}(T) = \frac{\pi_j w_{jk}(T)}{\sqrt{\pi_j \pi_k f_j(T) f_k(T)}} \tag{17}$$

which further simplifies as $g_{jk} = w_{jk}(T)/\pi_k$ for a system in equilibrium.

3.2. Courgeau's Index

Courgeau (1980) defines an *index of migratory stress* $c_j \in [-1, 1]$ as

$$c_j := \frac{I_j - E_j}{I_j + E_j} \tag{18}$$

where $I_j := n_{\bullet j} - n_{jj}$ is the number of immigrants in region j, and $E_j := n_{j\bullet} - n_{jj}$ is the number of emigrants. In the quasi-symmetric approximation, (4) yields $n_{kj} \cong n_{jk} \exp(2u_j - 2u_k) = n_{jk}(1 + 2u_j - 2u_k) + 0(u^2)$, and hence, in the limit of low expansivities

$$c_j \cong \frac{u_j - \sum_k w_{jk} u_k}{1 - w_{jj}} \cong \frac{\hat{\tau}_j}{T} \left(u_j - \sum_{k=1}^{m} w_{jk} u_k \right) \tag{19}$$

the latter expression being valid in the small time approximation. Hence, for a system not too far from equilibrium, Courgeau's migratory stress is proportional to the so-called *Laplacian* of the expansivities, that is, to the difference between the expansivity of the region and the average expansivity of its neighbors - neighbors in the sense of the Markov transition matrix.

3.3. Aggregation Properties

Under aggregation of categories $j \to J$ (section 2), initial densities, final densities and transition matrices transform as

$$f_J(0) = \sum_{j \in J} f_j(0) \quad f_K(T) = \sum_{k \in K} f_k(0) \quad w_{JK}(T) = \sum_{j \in J} \sum_{k \in K} \frac{f_j(0)}{f_J(0)} w_{jk}(T) \tag{20}$$

Let π_J be the stationary distribution associated to $w_{JK}(T)$. A flow is said to be *aggregation invariant* if $\pi_J = \sum_{j \in J} \pi_j$. It turns out that a flow

is aggregation invariant iff for any three distinct categories j, j' and k, one (or both) of the following conditions holds (Bavaud, 1998):

(a) $n_{jk} n_{j'\bullet} = n_{j'k} n_{j\bullet}$
(b) $n_{j\bullet} \pi_{j'} = n_{j'\bullet} \pi_j$

Independent flows satisfy (a); marginally homogeneous flows (that is obeying $n_{j\bullet} = n_{\bullet j}$), in particular symmetric flows, satisfy (b). Unfortunately, QS flows *do not* in general satisfy a) or b); even more disturbingly, flows n_{JK} resulting from the aggregation of quasi-symmetric flows n_{jk} are generally *not* quasi-symmetric.[3] However, restricting aggregations to *equi-expansive* ones preserves quasi-symmetry and aggregation invariance:

Theorem 4   (proof in appendix). *Let n_{jk} be quasi-symmetric. Consider an* equi-expansive *aggregation, that is, such that $u_j = constant =: c_J$ for all $j \in J$. Then n_{JK} is quasi-symmetric, with expansivities (defined with weights $a_J := \sum_{j \in J} a_j$), accessibilities, mixed profiles and stationary distributions given by*

$$u_J = c_J \quad g_{JK} = \frac{\sum_{j \in J} \sum_{k \in K} f_j^m f_k^m g_{jk}}{f_J^m f_K^m} \quad f_J^m = \sqrt{f_J(0) f_J(T)} \quad \pi_J = \sum_{j \in J} \pi_j$$

In view of (16), the condition of equi-expansivity equivalently reads $f_j(0)/f_J(0) = \pi_j/\pi_J$; under equilibrium, any aggregation is equi-expansive by definition since $u_j = 0$. The aggregated expansivities are normalized as $\sum_J a_J u_J = 0$.

The methodological consequences of theorem 4 are fairly straightforward: In gravity modelling, avoid aggregating regions with strongly differing expansivities - and don't worry if the flows are close to equilibrium.

4. MARKOV EMBEDDABLE MODELS

A common hypothesis in studies on social mobility (Singer and Spilerman, 1976) is to assume the Markov transition matrix (1) to be *embeddable*, that is to be generated in continuous time by a *proper*

---

[3]Note that flows resulting from the aggregation of independent flows are not independent either. This well know "defect" actually justifies factorial approaches such that *correspondence analysis*, aiming at decomposing dependent joint counts into a sum of independent components.

transition rate matrix R (proper meaning that off-diagonal elements are positive and rows sum to zero), such that $W(T) = \exp(TR)$. Then $W(T)$ can be extended to continuous times $t > 0$ as $W(t) = \exp(tR)$.

Requiring transition matrices $W(T)$ to be embeddable (which seems natural for migration or opinion shifts data, but more questionable for intergenerational mobility data) severely restricts the class of admissible flows, and solves a major difficulty, first addressed by Shorrocks (1978), in defining an all-purposes, single mobility index: namely, it is not clear whether the (unembeddable) transition matrix U

$$U := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad V := \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix} \quad (0 < p < 1) \tag{21}$$

should be qualified as more mobile than the perfectly mobile (embeddable) matrix V, or if one should distinguish between one kind of mobility expressed by U from another kind of mobility expressed by V and construct accordingly two distinct mobility indices. The embeddability requirement simply rules out possibilities such as U and rehabilitates the quest for a single mobility index.

In general, for T fixed, to a given transition rate matrix $W(T)$ might correspond

- many distinct transition proper rate matrices R (aliasing), in which case it is not immediate to determine which R should be selected
- a unique proper transition rate matrix R
- no proper transition rate matrix, in which case the problem should be cast in terms of determining a proper transition rate matrix R^{theo} such that $W^{\text{theo}}(T) = \exp(TR^{\text{theo}})$, where the theoretical parameters $n_{jk}^{\text{theo}}(T) := n_{j\bullet}(T)w_{jk}^{\text{theo}}(T)$ should be optimally close to the empirical counts $n_{jk}(T)$ (see Geweke et al., 1986 for a maximum likelihood approach restricted on a particular low-parameter family R^{theo}).

In full generality, the question of the embeddability of a given $W(T)$ is rather involved (Singer and Spilerman, 1976; Israel et al., 2001). We can attempt to express $R = \frac{1}{T} \ln W(T)$ as a power series expansion

$$R := \frac{1}{T} \sum_{l=1}^{\infty} \frac{1}{l} (-1)^l (I - W(T))^l \tag{22}$$

which turns out to converge for the three examples. The resulting R is proper for the Swiss migration and Danish opinion shifts data (which are therefore embeddable) but not for the British mobility data. In the latter case, the eigenvalues of $W(T)$ are all real, making R defined in

(22) unique (Singer and Spilerman, 1976), which definitely establishes the nonembeddability of British mobility data.

4.1. Sojourn Times and Jump Matrices

Suppose the transition rate matrix R (obeying $r_{jk} \geq 0$ for $j \neq k$ and $\sum_k r_{jk} = 0$) to solve the embedding problem for some $W(T)$. We can further parameterize the rate matrix in terms of *sojourn times* (τ_j) and *jump matrix* (κ_{jk}) by defining the quantities

$$\tau_j := -\frac{1}{r_{jj}} \geq 0 \quad \kappa_{jk} := \begin{cases} \tau_j r_{jk} & \text{if } j \neq k \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

Hence, time-continuous Markov transition matrices $W(t)$ can be generated by *jump processes*, where a unit initially in region j remains in this state for some random time t_j, exponentially distributed with mean $E(t_j) = \tau_j > 0$. The unit then jumps into a *different* region $k \neq j$ with probability transition κ_{jk} (satisfying $\kappa_{jk} \geq 0$, $\kappa_{jj} = 0$ and $\kappa_{j\bullet} = 1$): the jump matrix (κ_{jk}) determines the pattern of future destinations, while the sojourn times τ_j control the mean stay duration in each state $j = 1, \ldots, m$. The latter are $\tau_j = 0$ under perfect mobility, and $\tau_j = \infty$ for frozen flows.

4.2. Small Period Approximations

Suppose the observation period T to be small, so that the *small period approximation* $W(T) = I + RT + 0(T^2)$ makes sense. That is, the transition matrix rate is approximatively $\hat{r}_{jk} = (w_{jk}(T) - \delta_{jk})/T = (\frac{n_{jk}(T)}{n_{j\bullet}(T)} - \delta_{jk})/T$. The resulting sojourn times and jump matrix (23) are

$$\hat{\tau}_j := \frac{T}{1 - w_{jj}} = \frac{T}{1 - \frac{n_{jj}}{n_{j\bullet}}} \quad \hat{\kappa}_{jk} := \begin{cases} \frac{w_{jk}}{1 - w_{jj}} = \frac{n_{jk}}{n_{j\bullet} - n_{jj}} & \text{if } j \neq k \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

$\hat{\tau}_j$ is referred to as the *Prais local immobility index* (Prais, 1955), increasing with the propensity w_{jj} to stay in the same state, that its *autarchy*. Approximate sojourn times and jump matrix (24) are always defined, in contrast to their exact counterparts (23) which exist for embeddable flows only.

Jarrow et al. (1997) propose another small time approximate transition rate matrix $\overline{R}$ by assuming that an individual changes at most once during the observation period T. Their result is

$$\bar{r}_{jk} := \begin{cases} -\frac{1}{T} \frac{\ln w_{jj}}{1 - w_{jj}} w_{jk} & \text{if } j \neq k \\ \frac{1}{T} \ln w_{jj} & \text{otherwise} \end{cases} \tag{25}$$

yielding sojourn times

$$\bar{\tau}_j := -\frac{T}{\ln w_{jj}} = -\frac{T}{\ln \frac{n_{jj}}{n_{j\bullet}}} \tag{26}$$

and jump matrices $\bar{\kappa}$ identical to $\hat{\kappa}$ as given in (24).

Theorem 5   (proof in appendix). *Define*

- π *as the stationary distribution of* $W(T)$ *(or equivalently of* $W(t)$ *for any* $t > 0$)
- σ *as the stationary distribution of* κ
- $\hat{\sigma}$ *as the stationary distribution of* $\hat{\kappa}$.

  Then

$$\sigma_j = \tau_{\mathrm{Av}} \frac{\pi_j}{\tau_j} \quad where \quad \frac{1}{\tau_{\mathrm{Av}}} := \sum_{j=1}^{m} \frac{\pi_j}{\tau_j} = \frac{1}{\sum_{j=1}^{m} \sigma_j \tau_j} \tag{27}$$

$$\hat{\sigma}_j = \hat{\tau}_{\mathrm{Av}} \frac{\pi_j}{\hat{\tau}_j} \quad where \quad \frac{1}{\hat{\tau}_{\mathrm{Av}}} := \sum_{j=1}^{m} \frac{\pi_j}{\hat{\tau}_j} = \frac{1}{\sum_{j=1}^{m} \hat{\sigma}_j \hat{\tau}_j} \tag{28}$$

Theorem 5 shows the stationary distribution π_j of the transition matrix to be proportional to the relative visit frequency σ_j times the sojourn time τ_j of the state. Note that (27) also holds for the more general case of the *semi-Markov models* (see, e.g., Çinlar, 1969), generalizing the jump processes of section 4.1 to non-exponentially distributed residence times.

Theorem 5 also emphasizes the role of the average sojourn time τ_{Av}, defined as the weighted harmonic average of the individual sojourn times (in terms of the stationary distribution), or alternatively as the ordinary weighted average of the individual sojourn times (in terms of the relative visit frequency).

Exact and approximative sojourn times for the Swiss migrations data are (in years):

$$\tau = \begin{pmatrix} 103.9 \\ 121.0 \\ 83.4 \\ 85.3 \\ 90.2 \\ 112.4 \\ 192.9 \end{pmatrix} \quad \hat{\tau} = \begin{pmatrix} 107.8 \\ 124.3 \\ 86.2 \\ 88.1 \\ 93.3 \\ 115.0 \\ 196.3 \end{pmatrix} \quad \hat{\tau} = \begin{pmatrix} 105.2 \\ 121.8 \\ 83.7 \\ 85.6 \\ 90.8 \\ 112.5 \\ 193.8 \end{pmatrix} \tag{29}$$

with $T = 5 \ll \tau_{\mathrm{Av}} = 110.3$. The small time approximations are satisfactory and produce exact and approximate jump matrices which are numerically close:

$$\kappa \cong \hat{\kappa} = \bar{\kappa} = \begin{pmatrix} 0 & .16 & .20 & .16 & .35 & .05 & .07 \\ .25 & 0 & .07 & .25 & .09 & .03 & .31 \\ .50 & .15 & 0 & .09 & .17 & .03 & .05 \\ .37 & .34 & .07 & 0 & .09 & .03 & .10 \\ .59 & .13 & .12 & .07 & 0 & .05 & .04 \\ .42 & .15 & .06 & .08 & .13 & 0 & .16 \\ .22 & .54 & .04 & .09 & .06 & .04 & 0 \end{pmatrix} \tag{30}$$

Exact and approximative sojourn times for the Danish opinion shifts data are (in months):

$$\tau = \begin{pmatrix} 78.0 \\ 99.6 \\ 12.6 \end{pmatrix} \quad \hat{\tau} = \begin{pmatrix} 111.1 \\ 143.4 \\ 31.5 \end{pmatrix} \quad \bar{\tau} = \begin{pmatrix} 97.6 \\ 130.0 \\ 14.9 \end{pmatrix} \tag{31}$$

with $T = 26$ and $\tau_{\mathrm{Av}} = 61.8$. The small time approximations are less satisfactory, as indicated by the large discrepancy between exact and approximate jump matrices:

$$\kappa = \begin{pmatrix} 0 & .48 & .52 \\ .41 & 0 & .59 \\ .49 & .51 & 0 \end{pmatrix} \quad \hat{\kappa} = \bar{\kappa} = \begin{pmatrix} 0 & .71 & .29 \\ .66 & 0 & .34 \\ .47 & .53 & 0 \end{pmatrix} \tag{32}$$

4.3. Off-diagonal Flows

In administrative reports and scientific literature, the diagonal of flows data n_{jk} is frequently not available, that is, formally set to $n_{jj} = 0$. Those off-diagonal flows are not embeddable (see section 5.1), and the whole modelling enterprise is compromised. In the short period regime, a remediation could consist in considering $n_{jk}/n_{j\bullet}$ as defining a *jump* transition matrix $\hat{\kappa}_{jk}$ (rather than an ordinary transition matrix), and, under additional assumptions about sojourn times $\hat{\tau}_j$, to reconstruct the rate matrix $\hat{r}_{jk}$ by (24) as well as $\hat{W}(T) := \exp(T\hat{R})$.

For instance, setting the unknown diagonals to $\hat{n}_{jj} = \lambda n_{j\bullet}$ for some $\lambda \in (0, 1)$ (where $n_{j\bullet} := \sum_{k|k \neq j} n_{jk}$) yields constant sojourn times $\hat{\tau}_j = T/(1 - \lambda)$ with rate matrix

$$\hat{r}_{jk} := \begin{cases} \frac{(1-\lambda)}{T} \frac{n_{jk}}{n_{j\bullet}} & \text{if } j \neq k \\ -\frac{(1-\lambda)}{T} & \text{otherwise} \end{cases} \tag{33}$$

4.4. Mobility Indices

Sojourn times τ_j, their small-period versions $\hat{\tau}_j$ as well as their harmonic averages τ_{Av} and $\hat{\tau}_{\mathrm{Av}}$ constitute evident (im-)mobility indices.

Equation (24) shows Shorrocks' index to be related to the *uniform* harmonic average sojourn time $\hat{\tau}_{\mathrm{uniform}}$ as

$$s = \frac{m}{m-1}\frac{T}{\hat{\tau}_{\mathrm{uniform}}} \qquad \frac{1}{\hat{\tau}_{\mathrm{uniform}}} := \frac{1}{m}\sum_{j=1}^{m}\frac{1}{\hat{\tau}_j} = \frac{\sum_{j=1}^{m}(1-w_{jj})}{mT} \qquad (34)$$

Under quasi-symmetry, one intuitively expects strong diagonal (respectively, off-diagonal) accessibilites to correspond to low (respectively, high) mobility. This is indeed the case, in view of the *exact* identities (see Eqs. (9), (24), and (28))

$$g_{jk}(T) = \begin{cases} \left(1-\frac{T}{\hat{\tau}_j}\right)/f_j(T) & \text{if } j=k \\ \frac{T}{\hat{\tau}_j}\frac{\pi_j\hat{\kappa}_{jk}}{\sqrt{\pi_j\pi_k f_j(T)f_k(T)}} = \frac{T}{\hat{\tau}_{\mathrm{Av}}}\frac{\hat{\sigma}_j\hat{\kappa}_{jk}}{\sqrt{\pi_j\pi_k f_j(T)f_k(T)}} & \text{if } j\neq k \end{cases} \qquad (35)$$

which in turn imply the following sum rules:

$$\sum_j \pi_j f_j(T)g_{jj}(T) = 1 - \frac{T}{\hat{\tau}_{\mathrm{Av}}} \qquad\qquad \sum_{j\neq k}\sqrt{\pi_j\pi_k f_j(T)f_k(T)}\,g_{jk}(T) = \frac{T}{\hat{\tau}_{\mathrm{Av}}}$$

4.5. Aggregation Properties

Let $w_{jk}(T)$ be an embeddable transition matrix, transforming under aggregation as described in (20). By construction, the (proper) transition rate matrix r_{jk} follows the same behaviour, namely

$$r_{JK} = \sum_{j\in J}\sum_{k\in K}\frac{f_j(0)}{f_J(0)}r_{jk} \qquad (36)$$

Applying definitions (23) on the aggregated system yields sojourn times and jump matrices of the form

$$\frac{1}{\tau_J} = \sum_{j\in J}\frac{f_j(0)}{f_J(0)}(1-\kappa_{jJ})\frac{1}{\tau_j} \quad \kappa_{JK} = \begin{cases} \frac{\sum_{j\in J}f_j(0)\frac{\kappa_{jK}}{\tau_j}}{\sum_{l\in J}f_l(0)\frac{(1-\kappa_{lJ})}{\tau_l}} & \text{if } J\neq K \\ 0 & \text{otherwise} \end{cases} \qquad (37)$$

where $\kappa_{jK} := \sum_{k\in K}\kappa_{jk}$. The aggregated sojourn time τ_J is the harmonic average of the sojourn times τ_j of the categories $j\in J$, corrected by the term $1-\kappa_{jJ}$, which is the probability that an unit initially in $j\in J$ will jump outside J: Since units can jump from a fine-grained category to another while staying in the same coarse-grained category,

the coarse-grained sojourn time is larger than the harmonic average of the fine-grained sojourn times. For instance, the sojourn time of the region $[1 \cup 2]$ resulting of the aggregation of region 1 (with sojourn time $\tau_1 = 103.9$ years) and region 2 ($\tau_2 = 121.0$ years) reads $\tau_{[1 \cup 2]} = 137.5$ years from formula (37).

In the limit of a complete aggregation into a single class J (with $M = 1$), one gets $\kappa_{jJ} = 1$ and thus $\tau_J = \infty$, as expected.

Aggregation rules for small time parameters (24) $\hat{\tau}$ and $\hat{\kappa}(= \bar{\kappa})$ are identical to (37). The logarithmic sojourn time $\bar{\tau}$ (26) aggregates as

$$\frac{1}{\bar{\tau}_J} = -\frac{1}{T} \ln \sum_{j \in J} \frac{f_j(0)}{f_J(0)} \left\{ \bar{\kappa}_{jK} + (1 - \bar{\kappa}_{jK}) \exp\left(-\frac{T}{\bar{\tau}_j}\right) \right\}$$

where $\bar{\kappa}_{jK} := \sum_{k \in K} \bar{\kappa}_{jk}$. Again, $\bar{\tau}_J = \infty$ when $M = 1$.

Finally, observe r_{JK} as defined in (36) to obey $r_{J\bullet} = 0$; also, $r_{JK} \geq 0$ for $J \neq K$ (since $r_{jk} \geq 0$ if $j \in J$ and $k \in K$ by hypothesis). Thus the aggregation of a proper rate matrix yields another proper rate matrix, which shows the embeddability condition to be automatically conserved under aggregation.

5. QUASI-SYMMETRIC EMBEDDABLE MODELS

In this section, we assume flows to be quasi-symmetric *and* embeddable. By theorem (7) below, the reversibility property of the Markov transition holds at any times, and is inherited by the (exact or small time) jump matrices.

5.1. Spectral Decomposition

Theorem 6   (proof in appendix). *The transition matrix of a reversible and embeddable regular Markov chain can be spectrally decomposed as*

$$w_{jk} = \pi_k \sum_{\alpha=1}^{m} \lambda_\alpha x_{j\alpha} x_{k\alpha} = \pi_k \left(1 + \sum_{\alpha \geq 2} \lambda_\alpha x_{j\alpha} x_{k\alpha} \right) \tag{38}$$

where the eigenvalues λ_α verify $1 = \lambda_1 > \lambda_2 \geq \ldots \lambda_m \geq 0$, and the eigen-coordinates $x_{j\alpha}$ satisfy

$$x_{j1} \equiv 1 \qquad \sum_{j=1}^{m} \pi_j x_{j\alpha} x_{j\beta} = \delta_{\alpha\beta} \qquad \sum_{\alpha=1}^{m} x_{j\alpha} x_{k\alpha} = \frac{\delta_{jk}}{\pi_j} \tag{39}$$

For instance, the quasi-symmetric transitions associated to the Swiss migrations, Danish opinion shifts and matrix V in (21) satisfy

Theorem 6, in constrast to the British occupational mobility data and matrix U in (21). The latter (with $\lambda_2 = -1$) exhibits an oscillatory behaviour compatible with pure commuters flows, but not with the diffusive nature of embeddable processes.

Theorem 6 entails the inequalities

$$1 \leq \text{trace}(W) = 1 + \sum_{\alpha \geq 2} \lambda_\alpha \leq m \tag{40}$$

which demonstrate Shorrock's index to satisfy $s \in [0,1]$, and rule out the case $\text{trace}(W) = 0$ of off-diagonal flows. The lower bound in (40) is attained under perfect mobility (corresponding to $\lambda_2 = \cdots \lambda_m = 0$), and the upper bound is attained for frozen flows (corresponding to $\lambda_2 = \cdots \lambda_m = 1$).

5.2. Time Extensions

Following section 4, embeddable flows $n_{jk}(T)$ collected during the observation period T can be formally extended to arbitrary times $t \geq 0$ as

$$n_{jk}(t) := n_{j\bullet}(T)w_{jk}(t) \quad W(t) := W^{\frac{t}{T}}(T) \quad w_{jk}(T) := \frac{n_{jk}(T)}{n_{j\bullet}(T)} \tag{41}$$

One should certainly not expect (41), based upon the fiction of a constant rate together with never-altered, "immortal" units, to provide a realistic model for the long-term flow dynamics $t \gg T$. It is, however, important, from the endogenous point of view, to understand how the various quantities encountered so far are transformed (or not) under the time extensions $T \to t$.

First, suppose that the observed flows $n_{jk}(T)$ are quasi-symmetric. Then so are their time extensions $n_{jk}(t)$ (see theorem 3):

Theorem 7   (proof in appendix). *Let $w_{jk}(T) := n_{jk}(T)/n_{j\bullet}(T)$ be embeddable. Then $W(T) \Leftrightarrow W(t)$ reversible for any $t > 0 \Leftrightarrow \kappa$ reversible $\Leftrightarrow \hat\kappa = \bar\kappa$ reversible.*

In particular, the quasi-symmetric decomposition (theorem 1) holds for time-extended flows, permitting to define time-extended accessibilities and expansivities. Theorem 8 (whose proof in straightforward in view of equation 16 and theorem 6) describes the behavior of the latter as well as the other quantities of interest.

Theorem 8.   *Under time extension (41), the initial population distri-bution, $f_j(0)$, the total flow n, the stationary distribution π_j, the rate matrix r_{jk}, the expansivities u_j, the eigencoordinates $x_{j\alpha}$, the sojourn times τ_j and jumps components κ_{jk} are constant. Also*

$$f_k(t) = \sum_{j=1}^{m} f_j(0)w_{jk}(t) \; \textit{final distribution} \tag{42}$$

$$q_{jk}(t) = \frac{w_{jk}(t)}{f_k(t)} \; \textit{quotients} \tag{43}$$

$$g_{jk}(t) = \frac{\sqrt{\pi_j}}{\sqrt{\pi_k}} \frac{w_{jk}(t)}{\sqrt{f_j(t)f_k(t)}} \; \textit{accessibilities} \tag{44}$$

$$u_j(t) = u_j(T) \; \textit{expansivities} \tag{45}$$

$$\hat{\tau}_j(t) = \frac{t}{1 - w_{jj}(t)} \; \textit{small period sojurn times} \tag{46}$$

$$\hat{\kappa}_{jk}(t) = \frac{w_{jk}(t)}{1 - w_{jj}(t)} \; \textit{small period jump components} \tag{47}$$

$$c_j(t) = \frac{u_j - \sum_k w_{jk}(t)u_k}{1 - w_{jj}(t)} \; \textit{Courgeau migratory stress} \tag{48}$$

$$s(t) = \frac{m - \sum_j w_{jj}(t)}{m - 1} \; \textit{Shorrocks mobility index} \tag{49}$$

$$\lambda_\alpha(t) = \lambda_\alpha^{\frac{t}{T}}(T) \; \textit{eigenvalues} \tag{50}$$

In particular, the short-time limit $t \to 0$ yields frozen flows $w_{jk}(0) = \delta_{jk}$. Diagonal accessibilities are $g_{jj}(0) = 1/f_j(0)$, decreasing linearly with slope $g'_{jj}(0) = -1/(\hat{\tau}_j f_j(0))$, while off-diagonal accessibil-ities $g_{jk}(0)$ are zero, and increase linearly with t (see Figures 1 and 2). Also, the small period sojourn times and jump components reach their exact values (23). Courgeau index tends to $c_j = u_j - \sum_k \kappa_{jk}u_k$, Shorrock index to zero, and all eigenvalues to one.

The long-term limit $t \to \infty$ yields perfect mobility $w_{jk}(\infty) = \pi_k$. Diagonal and off-diagonal accessibilities tend (slowly) to unity (see Figure 2). Small period sojourn times are infinite. Small period
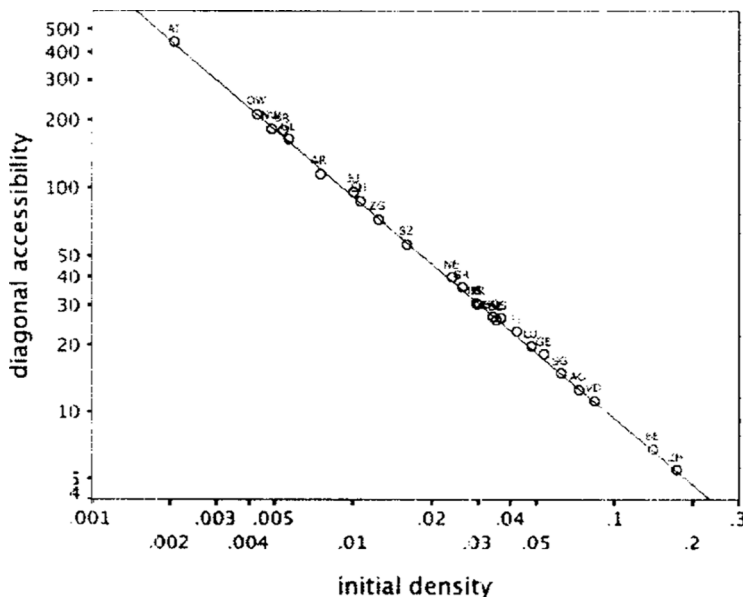
FIGURE 1 Dependence of the diagonal accessibilites $g_{jj}(T)$ on the initial densities $f_j(0)$ (logarithmic scale), for the complete Swiss migrations data on 26 cantons. A weighted regression yields $g_{jj}(T) \cong 0.97 \cdot f_j(0)^{-0.986}$, practically identical to the small time limit $g_{jj}(0) = f_j(0)^{-1}$.

jump components are $\kappa_{jk} = \pi_k/(1 - \pi_j)$. Courgeau index is $c_j = (u_j - \sum_k \pi_k u_k)/(1 - \pi_j)$, Shorrock index tends to unity, and all nontrivial eigenvalues tend to zero.

5.3. Information-Theoretical Mobility Indices

Under perfect mobility, the final destination is independent of the initial category; under frozen flows, the final destination is entirely determined by the initial category. The conditional information-theoretical uncertainty $\breve{h}(t)$ of the position at time t given the initial position, or its "equilibrium" variant $h(t)$ obtained by assuming the initial distribution to be the stationary one.[4]

$$\breve{h}(t) := - \sum_{j,k} f_j(0) w_{jk}(t) \ln w_{jk}(t) \quad h(t) := - \sum_{j,k} \pi_j w_{jk}(t) \ln w_{jk}(t) \quad (51)$$

---

[4]$h(t)$ is the *entropy rate* of the dynamical process governed by $W(t)$ (see e.g. Cover and Thomas, 1991).
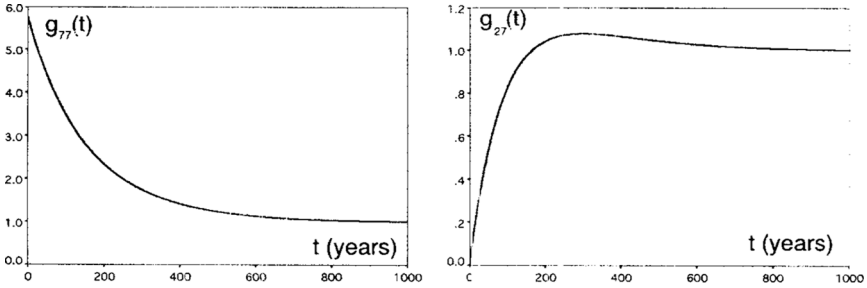
FIGURE 2 Typical temporal behavior of diagonal (left) and off-diagonal (right) accessibility components $g_{jk}(t)$, for the Swiss migrations data, extrapolated from data observed during $T = 5$ years only.

Consider also $k(t)$, the relative entropy between the actual and the asymptotic distribution (Theil (1972) considers the variant $K(\pi||f(t))$):

$$k(t) := K(f(t)||\pi) = \sum_k f_k(t) \ln\frac{f_k(t)}{\pi_k} \tag{52}$$

The typical behavior of $h(t)$ and $k(t)$ is illustrated in Figure 3, and follow from theorem 9.

$\breve{h}(t)$ and $h(t)$ constitute mobility indexes, while $k(t)$ is an immobility index. Their temporal behavior illustrates the trivial fact that the longer a system is observed, the more mobile it appears. However, the *rates* at which the short- and long-term limits are attained provide natural candidates for time-invariant mobility indices.
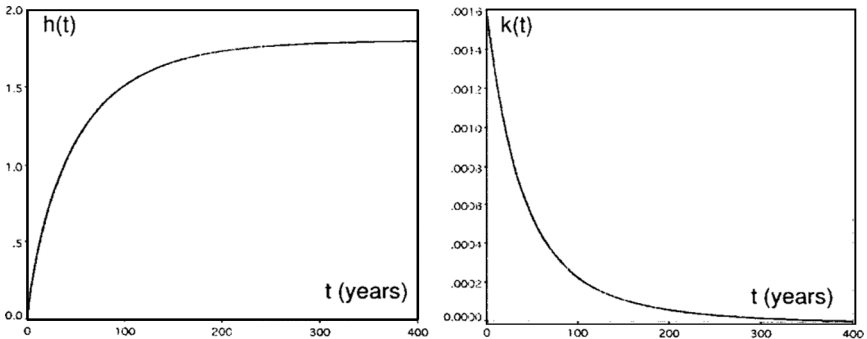

FIGURE 3 Increase of the conditional uncertainty (left) and decrease of the relative entropy (right), whose asymptotic behaviour are controlled by $|\lambda_2|$. By contrast, the small-time behavior of $h(t)$ is governed by τ_{Av}.

Theorem 9 (proof in appendix).

(a) $\breve{h}(t)$ *and* $h(t)$ *are non-negative and increasing, with small-time behavior* $t \to 0$

$$\breve{h}(t) = -\frac{t}{\breve{\tau}_{Av}} \ln \frac{t}{\breve{\tau}_{Av}} + 0(t^2 \ln t) \quad h(t) = -\frac{t}{\tau_{Av}} \ln \frac{t}{\tau_{Av}} + 0(t^2 \ln t)$$

(*where* $\frac{1}{\breve{\tau}_{Av}} := \sum\limits_j \frac{f_j(0)}{\tau_j}$), *and large-time behavior* $t \to \infty$

$$\breve{h}(t) = -\sum_j \pi_j \ln \pi_j - A\lambda_2^{\frac{2t}{T}} + o(\lambda_2^{\frac{2t}{T}}) \quad h(t) = -\sum_j \pi_j \ln \pi_j - \lambda_2^{\frac{2t}{T}} + o(\lambda_2^{\frac{2t}{T}})$$

where $A := \sum\limits_j f_j(0)x_{j2}^2 \geq 0$.

(b) $k(t)$ *non-negative and decreasing, with small- and large-time behavior*

$$k(t) = \sum_j f_j(0) \ln \frac{f_j(0)}{\pi_j} - Bt + 0(t^2) \quad k(t) = C\lambda_2^{\frac{2t}{T}} + o(\lambda_2^{\frac{2t}{T}})$$

where $B := -\sum\limits_{jk} r_{jk} \ln \frac{f_k(0)}{\pi_k} \geq 0$ *and* $C := \left(\sum\limits_j f_j(0)x_{j2} \right)^2 \geq 0$.

(c) $\sum\limits_j u_j(f_j(T) - f_j(0)) \geq \frac{1}{2} \sum\limits_j f_j(T) \ln \frac{f_j(T)}{f_j(0)} \geq 0$.

From Theorem 9 emerge two natural candidates for immobility indices:

- The harmonic average sojourn time τ_{Av} (27) as well as its variants $\breve{\tau}_{Av}$ (28) or $\breve{\tau}_{Av}$, controlling the short-term dynamics: the larger τ_{Av}, the larger the tendency for an individual to remain in the same state. The quantity τ_{Av} does not seem to have been proposed so far in the mobility literature in contrast to the uniform sojourn time $\hat{\tau}_{uniform}$ (34), whose dynamical significance is, however, questionable.
- The second eigenvalue $\lambda_2 < 1$, controlling the asymptotic dynamics (Theil, 1972; Shorrocks, 1978; Sommers and Conlisk, 1979). One finds $\lambda_2 = .97$ for the Swiss migrations, $\lambda_2 = .65$ for the Danish opinion shifts (and $\lambda_2 = .47$ for the British mobility). The larger λ_2, the longer the relaxation time towards equilibrium.

Exact and approximate sojourn times are upper-bounded by the second eigenvalue as

Theorem 10 (proof in appendix)**.**

$$\tau_j \le \frac{T}{(1 - \pi_j)(- \ln \lambda_2)} \quad \hat{\tau}_j \le \frac{T}{(1 - \pi_j)(1 - \lambda_2)} \qquad (53)$$

In particular, one gets $\tau_j = 0$ under perfect mobility $\lambda_2 \to 0$. Inequalities (53) become equalities in the special case $\lambda_2 = \cdots = \lambda_m \equiv \lambda$, yielding (by (38) and (39)) the so-called *stayers-movers* model $w_{jk} = \lambda \delta_{jk} + (1 - \lambda)\pi_k$ (Blumen et al., 1955) originally conceived as resulting from the mixture of an immobile population and a perfectly mobile population.

Under aggregation, the harmonic sojourn time $\tilde{\tau}_{Av}$ and second eigenvalues $\tilde{\lambda}_2$ transform oppositely, namely

$$\tilde{\tau}_{Av} \ge \tau_{Av} \quad \tilde{\lambda}_2 \le \lambda_2$$

where the first inequality follows from (37), and the second one from a classical result on eigenvalues interlacing (Haemers, 1995). By contrast, under equilibrium, (37) shows that $\tau_{Av}^A \ge \tau_{Av}$. Thus the aggregated uncertainty $\tilde{h}(t)$ increases more rapidly at the origin and reaches more quickly equilibrium than $h(t)$ does.

6. CONCLUSION

The use of embeddable Markov models for modelling social mobility have, with very good reasons, been criticized as early as the 1970s (see, e.g., Morisson, 1967; McGinnis 1968; Brown, 1970 or McFarland, 1970 and references therein). Typically, the dynamics of different populations should be distinguished, and disaggregations initiated in the stayers-movers model should be pursued. Also, the probability of remaining in a state should increase with the duration of prior residence in that state, which would require the use of semi-Markov models.

These sophisticated considerations and other (multilevel analysis, inter-dependence of individual choices, non-linear modelling) contrast the relative crudeness of the conclusions and of the further analysis potential of most empirical reports, whose indices turn out difficult to compare on different time scales or aggregation levels. The nature of the obstacle is arguably purely formal, that is, lies at the endogenous level and not at the exogenous level of real interest for the social scientist. We hope the present work can modestly contribute to alleviate this difficulty, to clarify the logical

relationships between the gravity and the Markov approaches and to encourage further endogenous studies in more advanced modelling.

REFERENCES

Anderson, E. B. (1990) *The Statistical Analysis of Categorical Data*, Berlin: Springer.

Bavaud, F. (1998). Models for spatial weights: a systematic look. *Geographical Analysis*, *30*, 153–171.

Bavaud, F. (2002). The quasi-symmetric side of gravity modelling. *Environment and Planning A*, *34*, 61–79.

Berger, J. & Snell, J. L. (1957). On the concept of equal exchange. *Behavioral Science*, *2*, 111–118.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete Multivariate Analysis*, Cambridge, MA: The MIT Press.

Blumen, I., Kogan, M., & McCarthy, P. J. (1955). *The Industrial Mobility of Labor as a Probability Process*, Ithaca, NY: Cornell University Press.

Boudon, R. (1973). *Mathematical Structures of Social Mobility*, Amsterdam: Elsevier.

Brown, L. A. (1970). On the use of Markov chains in movement research. *Economic Geography*, *46*, 393–403.

Caussinus, H. (1965). Contribution statistique à l' analyse des tableaux de corrélation. *Annales de la Faculté des Sciences de l' Université de Toulouse*, *29*, 77–182.

Çinlar, E. (1969). Markov renewal theory. *Adv. Appl. Prob.*, *1*, 123–187.

Courgeau, D. (1980). *Analyse quantitative des migrations humaines*, Paris: Masson.

Cover, T. M. & Thomas, J. A. (1991). *Elements of Information Theory*, New York: Wiley.

Fotheringham, A. S. & O'Kelly, M. E. (1989). *Spatial Interaction Models: Formulations and Applications*, Dordrecht: Kluwer.

Geweke, J., Marshall, R. C., & Zarkin, G. A. (1989). Mobility indices in continuous time Markov chains. *Econometrica*, *54*, 1407–1423.

Glass, D. V. (ed.) (1954). *Social Mobility in Britain*, London: Routledge.

Haemers, W. H. (1995). Interlacing eigenvalues and graphs. *Linear Algebra and Its Applications*, *227–228*, 593–616.

Israel, R. B., Rosenthal, J. S., & Wei, J. Z. (2001). Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings. *Mathematical Finance*, *11*, 245–265.

Karlin, S. & Taylor, H. M. (1975). *A First Course in Stochastic Processes*, New York: Academic Press.

Kijima, M. (1997). *Markov Processes for Stochastic Modeling*, London: Chapman and Hall.

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1988). *Multivariate Analysis*, New York: Academic Press.

McGinnis, R. (1968). A stochastic model of social mobility. *American Sociological Review*, *33*, 712–722.

McFarland, D. D. (1970). Intragenerational social mobility as Markov process: including a time–stationary Markovian model that explains observed declines in mobility rates over time. *American Sociological Review*, *35*, 463–476.

Morrison, P. A. (1967). Duration of residence and prospective migration: the evaluation of a stochastic model. *Demography*, *4*, 553–561.

Prais, S. J. (1955). Measuring social mobility. *J. R. Statistical Soc.*, *A118*, 56–66.

Sen, A. & Smith, T. E. (1995). *Gravity Models of Spatial Interaction Behavior*, Berlin: Springer.

Reilly, W. J. (1931). *The Law of Retail Gravitation*, New York: The Knickerbrocker Press.

Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica*, *46*, 1013–24.

Singer, B. & Spilerman, S. (1976). The representation of social processes by Markov models. *American Journal of Sociology*, *82*, 1–54.

Sommers, P. M. & Conlisk, J. (1979). Eigenvalue immobility measure for Markov chains. *Journal of Mathematical Sociology*, *6*, 253–276.

Theil, H. (1972). *Statistical Decomposition Analysis*, Amsterdam: North-Holland.

Weidlich, W. & Haag, G. (eds.) (1988). *Interregional Migration*, Berlin: Springer-Verlag.

Wilson, A. G. (1971). A family of spatial interaction models, and associated developments. *Environment and Planning*, *3*, 1–32.

Willekens, F. (1983). Log-linear modeling of spatial interaction. *Papers of the Regional Science Association*, *52*, 87–205.

APPENDIX

Proof of Theorem 1

Suppose n_{jk} to be quasi-symmetric, i.e. of the form $n_{jk} = c_{jk}b_k$ with $c_{jk} = c_{kj}$ (this is equivalent to the original definition $n_{jk} = \alpha_j \beta_k \gamma_{jk}$ with $\gamma_{jk} = \gamma_{kj}$).

Defining $u_j := \frac{1}{2}\ln b_j$ yields

$$n_{jk} = \sqrt{n_{jk}n_{kj}}\sqrt{\frac{n_{jk}}{n_{kj}}} = \sqrt{n_{jk}n_{kj}}\sqrt{\frac{b_k}{b_j}} = \sqrt{n_{jk}n_{kj}}\exp(u_k - u_j) \qquad (54)$$

Conversely, suppose n_{jk} to be of the form (54) (where u_j is defined up to an additive constant). Then $\frac{1}{2}\sum_l a_l \ln n_{lj}/n_{jl} = \sum_l a_l(u_j - u_l) = u_j - \sum_l a_l u_l$. Fixing the constant so that $\sum_l a_l u_l = 0$ yields (10).

Proof of Theorem 4

Under an equi-expansive aggregation, flows (4) transform as

$$n_{JK} = n\left(\sum_{j\in J}\sum_{k\in K} f_j^m f_k^m g_{jk}\right)\exp(c_K - c_J)$$

which is quasi-symmetric. Extracting back the parameters as indicated in Theorem 1 proves the assertion.

Proof of Theorem 5

π being the stationary distribution of the embeddable transition matrix $W(T)$, it is easy to show that the stationary distribution of $W\left(\frac{a}{b}T\right)$ (for any positive integers a and b) must be also π; by continuity,

π is the stationary distribution of $W(t) = \exp(tR)$ for any $t > 0$. Expanding the exponential, one gets the result $\sum_j \pi_j r_{jk} = 0$, that is, from (23):

$$\sum_{j|j \neq k} \pi_j \frac{\kappa_{jk}}{\tau_j} - \pi_k \frac{1}{\tau_k} = 0 \quad \text{i.e.} \quad \sum_j \frac{\pi_j}{\tau_j} \kappa_{jk} = \frac{\pi_k}{\tau_k}$$

thus proving that the normalized profile σ is the stationary distribution of κ. Also

$$\sum_j \hat{\sigma}_j \hat{\kappa}_{jk} \overset{(a)}{=} \tau_{\text{Av}} \sum_{j|j \neq k} \frac{\pi_j}{\hat{\tau}_j} \frac{n_{jk}}{n_{j\bullet} - n_{jj}} \overset{(b)}{=} \frac{\tau_{\text{Av}}}{T} \sum_{j|j \neq k} \pi_j \frac{n_{jk}}{n_{j\bullet}} \overset{(c)}{=} \frac{\tau_{\text{Av}}}{T} \pi_k \left(1 - \frac{n_{kk}}{n_{k\bullet}}\right)$$

$$= \tau_{\text{Av}} \frac{\pi_k}{\hat{\tau}_k} = \hat{\sigma}_k$$

where (a) and (b) follow from (24) and (c) from the fact that π is the stationary distribution of $W(T)$.

Proof of Theorem 6

Theorem 6 is well known (see e.g. Kijima (1997)) and its proof is sketched for completeness. By reversibility, $f_{jk}(T) := w_{jk}(T)\sqrt{\pi_j/\pi_k}$ is symmetric, and can hence be spectrally decomposed as $F = U\Lambda U'$ where $U = (u_{j\alpha})$ is orthogonal and contains the eigenvectors, and Λ is diagonal and contains the real eigenvalues λ_α in decreasing order. In components, $f_{jk} = \sum_{\alpha=1}^m \lambda_\alpha u_{j\alpha} u_{k\alpha}$ with $\sum_j u_{j\alpha} u_{j\beta} = \delta_{\alpha\beta}$ and $\sum_\alpha u_{j\alpha} u_{k\alpha} = \delta_{jk}$. By the Perron-Froebenius theorem, $\lambda_1 = 1$, $u_{j1} \equiv \sqrt{\pi_j}$ and $\lambda_m \geq -1$. Regularity excludes the case $\lambda_2 = 1$ corresponding to a reducible Markov chain. Embeddability implies that the eigenvalues corresponding to $F(t)$ (for any $t > 0$) are $\lambda(t) = \lambda_\alpha^{t/T}$ (with identical eigenvectors), which entails $\lambda_\alpha \geq 0$. Defining the so-called *eigencoordinates* $x_{j\alpha} := u_{j\alpha}/\sqrt{\pi_j}$ achieves the proof.

Proof of Theorem 7

Proceeding similarly as above, one shows that $\pi_j w_{jk}(T) = \pi_k w_{kj}(T)$ (for all $j \neq k$, the case $j = k$ being trivial) iff $\pi_j w_{jk}(t) = \pi_k w_{kj}(t)$ for any $t > 0$, that is iff $\pi_j r_{jk} = \pi_k r_{kj}$, that is iff $\pi_j/\tau_j \kappa_{jk} = \pi_k/\tau_k \kappa_{kj}$, that is iff κ is reversible. On the other hand, $\hat{\sigma}_j \hat{\kappa}_{jk} = \hat{\sigma}_k \hat{\kappa}_{kj}$ iff $\pi_j/\hat{\tau}_j n_{jk}/n_{j\bullet} - n_{jj} = \pi_k/\hat{\tau}_k \, n_{kj}/n_{k\bullet} - n_{kk}$ iff $\pi_j n_{jk}/n_{j\bullet} = \pi_k n_{kj}/n_{k\bullet}$ iff $W(T)$ is reversible.

Proof of Theorem 9

(a) Non-negativity and monotonicty follow from Jensen's inequality; they constitute standard results in Information Theory (see e.g. Cover

and Thomas 1991). Large-time behavior directly follows from (38). To show the small-time behavior of $h(t)$, one uses $w_{jk}(t) = \delta_{jk} + t r_{jk} + 0(t^2)$, and finds $h(0) = 0$, $h'(0) = \infty$ and $\lim_{t \to \infty} t\, h''(t) = -1/\tau_{\mathrm{Av}}$. Define $\varepsilon(t) := -t/\tau_{\mathrm{Av}} \ln t/\tau_{\mathrm{Av}}$, obeying $\varepsilon(0) = 0$, $\varepsilon'(0) = \infty$ and $\varepsilon''(t) = -1/(t\tau_{\mathrm{Av}})$. Double application of l'Hospital rule yields

$$\lim_{t \to 0} \frac{h(t)}{\varepsilon(t)} = \lim_{t \to 0} \frac{h'(t)}{\varepsilon'(t)} = \lim_{t \to 0} \frac{h''(t)}{\varepsilon''(t)} = -\tau_{\mathrm{Av}} \lim_{t \to 0} t\, h''(t) = 1$$

Small-time behavior of $\check{h}(t)$ is shown analogously.

(b) Let $f_k(t) := \sum_j f_j(0) w_{jk}(t)$ and $g_k(t) := \sum_j g_j(0) w_{jk}(t)$ be two distributions following the same Markov dynamics. Then their relative entropy $\sum_k f_k(t) \ln f_k(t)/g_k(t)$ is nonincreasing in t, for all $t \geq 0$ (see, e.g., Cover and Thomas, 1991 for a proof). Setting $g_j(0) := \pi_j$ proves the assertion.

(c) As $k(t)$ is decreasing, $\sum_j f_j(T) \ln\big(f_j(T)/f_j(0) f_j(0)/\pi_j\big) \leq \sum_j f_j(0) \ln f_j(0)/\pi_j$. Using (16) proves the assertion.

Proof of Theorem 10

Equations 23, 38, 24, 39 and 41 yield

$$\frac{1}{\tau_j} = -r_{jj} = -\frac{\pi_j}{T} \sum_{1 < \lambda_\alpha < 0} (\ln \lambda_\alpha) x_{j\alpha}^2 \geq -\ln \lambda_2 \frac{\pi_j}{T} \sum_{\alpha \geq 2} x_{j\alpha}^2$$

$$= -\ln \lambda_2 \frac{\pi_j}{T} \left(\frac{1}{\pi_j} - 1\right) \frac{1}{\hat{\tau}_j}$$

$$= \frac{1}{T}(1 - w_{jj}) = \frac{1}{T}\left(1 - \pi_j - \pi_j \sum_{\alpha \geq 2} \lambda_\alpha x_{j\alpha}^2\right) \geq \frac{1 - \pi_j}{T}(1 - \lambda_2)$$