Ravenstein Revisited: The Analysis of Migration, Then and Now

Rees, Philip; Lomax, Nik

Mitglied der

Leibniz-Gemeinschaft

gesis
Leibniz-Institut
für Sozialwissenschaften

Ravenstein Revisited: The Analysis of Migration, Then and Now[*]

Philip Rees, Nik Lomax

Abstract: In 1876, 1885 and 1889, Ernst Ravenstein, an Anglo-German geographer, published papers on internal and international migration in Britain, Europe and North America. He generalized his findings as "laws of migration", which have informed subsequent migration research. This paper aims to compare Ravenstein's approach to investigating migration with how researchers have studied the phenomenon more recently. Ravenstein used lifetime migrant tables for counties from the 1871 and 1881 censuses of the British Isles. Data on lifetime migrants are still routinely collected but, because of the indeterminate time interval, they are rarely used to study internal migration. Today, internal migration measures from alternative sources are used to measure internal migration: fixed interval migrant data from censuses and surveys, continuous records of migrations from registers, and "big data" from telecommunications and internet companies.

Ravenstein described and mapped county-level lifetime migration patterns, using the concepts of "absorption" and "dispersion", using migration rates and net balances. Recently, researchers have used lifetime migrant stocks from consecutive censuses to estimate country to country flows for the world. In the last decade, an Australian-led team has built an international database of internal migration flow data and summary measures. Methods were developed to investigate the modifiable areal unit problem (MAUP), in order to design summary internal migration measures comparable across countries. Indicators of internal migration were produced for countries covering 80 percent of the world's population.

Ravenstein observed that most migrants moved only short distances, anticipating the development of "gravity" models of migration. Recent studies calibrated the relationship between migration and distance, using gravity models. For mid-19th century Britain, Ravenstein found the dominant direction of internal migration to be towards the "centres of commerce and industry". Urbanization is still the dominant flow direction in most countries, though, late in the process, suburbanization, counter-urbanization and re-urbanization can occur. Ravenstein focussed on place-specific migration, whereas today researchers describe migration flows using area

---

[*] This article belongs to a special issue on "Internal Migration as a Driver of Regional Population Change in Europe: Updating Ravenstein".

typologies, seeking spatial generality. Ravenstein said little about migrant attributes except that women migrated more than men. In recent decades, the behaviour of migrants by age, sex, education, ethnicity, social class and partnership status have been studied intensively, using microdata from censuses and surveys.

Knowledge about processes influencing internal and international migration has rarely been built into demographic projections. Scenarios that link migration with sub-national or national inequalities and with climate or environmental change are influencing the design of policies to reduce inequalities or slow global warming.

1   Introduction

In the eighth and ninth decades of the 19th Century, Ernst Georg Ravenstein, an Anglo-German cartographer, published three papers that used data from the 1871 and 1881 Censuses to describe the patterns of internal migration across the British Isles (*Ravenstein* 1876, 1885). In a third paper, he extended his analysis to Europe, the United States and Canada, examining "the foreign element" in their populations (*Ravenstein* 1889). These papers continue to influence research into internal and international migration. The aim of this paper is to compare the data, methods and results of Ravenstein's work with research on migration in recent decades, to show both the connections and differences between his ideas and contemporary approaches. This paper is therefore a review of the field of internal migration with excursions into international migration. The raw material used are Ravenstein's papers and a selection of papers and books published in recent decades on internal migration, which connect to themes introduced in the 19th century work.

Ravenstein's most cited paper published in 1885 has the title "Laws of Migration".[1] He presented the paper at the Statistical Society of London (later Royal Statistical Society), followed by a lengthy question and answer session (*Ravenstein* 1885: 228-235). A key outcome of the "Discussion of Mr. Ravenstein's Paper" was that both author and discussants agreed that the "laws" were not immutable rules but rather "empirical generalizations" specific to the time and place of the evidence. Table 1 presents these "Laws of Migration" from the 1885 paper. Tables A1 to A3 in the Appendix report interpretations of the "laws" by later scholars.

---

[1] Google Scholar (9 March 2020) counts 4749 citations for *Ravenstein* (1885), 2159 citations for *Ravenstein* (1889) and 119 citations for *Ravenstein* (1876).

Ravenstein's work on migration has been expertly reviewed and interpreted by subsequent scholars (*Grigg* 1977; *Dorigo/Tobler* 1983 and *Greenwood* 2019) so remarks here will be brief. In his 1876 paper, *Ravenstein* covers much the same ground as his 1885 paper on the British Isles, without the key ingredient that has grabbed the attention of later scholars, his list of seven Laws of Migration (Table 1). *Greenwood* (2019: 271) extracts five key points from *Ravenstein* (1876), which an-

Tab. 1:   Ravenstein's "Laws of Migration" in his 1885 paper

Year-Law#	Text The Laws of Migration, *Ravenstein* (1885)	Pages
1885-1	We have already proved that the great body of our migrants only proceed a short distance, and that there takes place consequently a universal shifting of displacement of the population, which produces "currents of migration" setting in the direction of the great centres of commerce and industry which absorb the migrants. In forming an estimate of this displacement we must take into account the number of natives of each county which furnishes the migrants, as also the population of the towns or districts which absorb them.	p.198
1885-2	It is the natural outcome of this movement of migration, limited in range, but universal throughout the country, that the process of absorption would go on in the following manner: The inhabitants of the country immediately surrounding a town of rapid growth, flock into it; the gaps thus left in the rural population are filled up by migrants from more remote districts, until the attractive force of one of our rapidly growing cities makes its influence felt, step by step, to the most remote corner of the kingdom. Migrants enumerated in a certain centre of absorption will consequently grow less with the distance proportionately to the native population which furnishes them, and a map exhibiting by tints the recruiting process of any town ought clearly to demonstrate this fact. That this is actually the case will be found by referring to maps 3, 4, 8 and 9. These maps show at the same time that facilities of communication may frequently countervail the disadvantages of distance.	p.198-9
1885-3	The process of dispersion is the inverse of that of absorption and exhibits similar features.	p.199
1885-4	Each main current of migration produces a compensating counter-current.	p.199
1885-5	Migrants proceeding long distances generally go by preference to one of the centres of commerce or industry.	p.199
1885-6	The natives of towns are less migratory than those of the rural parts of the country.	p.199
1885-7	Females are more migratory than males.	p.199

Source: *Ravenstein* (1885: 198-199)

ticipate his 1885 Laws of Migration (Table A1). *Ravenstein's* 1885 list (Table 1) mixes elaborate statements of the regularity in which more than one proposition is made (Laws 1885-1 and 1885-2)[2] with pithy summary phrases (Laws 1885-3 to 1885-7). *Grigg* (1977) extends the list of Ravenstein laws to eleven (Table A2) using short sentences or phrases. *Dorigo* and *Tobler* (1983) interpret Ravenstein's statements as constituting precursors for later generalizations about push and pull factors influencing migration flows (Table A3). *Dorigo* and *Tobler* (1983) suggest that one of Ravenstein's Laws (1983-1 in Table A3) anticipates the scale and zonation dimensions of the Modifiable Areal Unit Problem or MAUP (*Openshaw* 1983). Another of Ravenstein's Laws (1983-2 in Table A3), *Dorigo* and *Tobler* propose, anticipates the hypothesis in which previous migrations by "friends and families" channel later migration. This hypothesis was investigated by Swedish geographer *Hägerstrand* (1957), using detailed longitudinal migration records. A third suggestion is that the configuration of the transportation system is likely to play a significant role in channelling migration flows (Ravenstein's Law 1983-3, Table A3). So, we have the seven original Laws proposed by Ravenstein plus four added by Grigg and three by Dorigo and Tobler, fourteen in total. Ideally, this impressive number of empirical generalizations should be tested in any study of new internal migration data. Like most researchers, we select a sub-set to compare the approaches of Ravenstein then and migration scholars now.

We structure the rest of the paper as follows. Section 2 considers the wide set of definitions of migration used in data collection, presenting them in a common graphical framework. Section 3 discusses how "raw" data may be used to estimate "refined" indicators of migration behaviour and how gaps in empirical information can be filled. Section 4 reports on recent research on internal migration patterns, tracking recent work linked to Ravenstein's generalizations. The final section sets out an agenda for future research aimed at improving our knowledge of both internal and international migration.

2   Data: Definitions and sources then and now

To describe and understand migration requires robust definitions of the phenomenon and reliable quantitative and qualitative data based on one or more of the definitions. In this section of the paper, we review definitions of migration and offer a framework that links them.

2.1   What is migration?

For Ravenstein, migration was displacement of population between their place of birth and their place of enumeration at the census because those were the only

---

[2]   This numbering system links to the listing of the laws of migration in papers by date and order in Tables 1, A1, A2 and A3.

comprehensive records available from 19th century censuses. In the contemporary world we need to introduce distinctions between different types of migration based on national or local boundaries. If migration takes place between countries, properly it should be labelled "international migration". If migration takes place within a country it is usually labelled as "internal" or "domestic" migration. Some scholars and some official statistics agencies distinguish between local mobility (migrations within local areas) and inter-area migration. Ravenstein wondered about how urban and rural migration differed but the 1871 and 1881 censuses provided no direct information about migration within cities. We consider this distinction arbitrary because of its dependence on local government boundaries which in many countries are subject to periodic reorganization, destroying temporal comparability. So, the ideal data set should record the geo-coded address of origin and the geo-coded address of destination so that tables of flows for any spatial system or scale can be created.

Several difficulties arise which mean this ideal is rarely achieved. Not everyone agrees with these distinctions. Ravenstein did not distinguish between internal and international migration, treating, for example, the Irish element in the population of England and Wales in a similar way to the foreign element. Papers that discuss international migration will use "migration" without any prior specification that this is only convenient shorthand. The same occurs in papers about internal migration. The second difficulty is that official statistical agencies are reluctant to publish migration information for flows between small areas due to uncertainty or because of the risk of disclosure. Official Statistical offices (such as Australian Bureau of Statistics) often perturb the counts in published census tables and create inconsistency across tables at the same spatial scale and across spatial scales for the same tables. A third difficulty arises because it is difficult to capture geography in a census or survey question. Most statistical agencies rely on a write-in answer, but people's knowledge of places may be deficient. The cost involved in converting a write-in place name answer to a geocoded reference is considerable. So, in 1961, when the General Register Office of England and Wales decided to introduce a question on migration in the year before the census, the tabulation of lifetime migration was restricted to "home" countries with the United Kingdom (UK) and foreign countries, dropping birthplace by county used in censuses from the 1871 to 1951 (*Friedlander/Roshier* 1966).

2.2   Households or individuals?

Human migration is defined as the relocation between dwellings of groups of individuals known as households over time and space. People migrate in small groups (households), which are either a nuclear family unit or a couple or a single person or several unrelated individuals, who share some common aspect of life together such as meals or housework or a set of relationships (husband-wife, child parent, student-fellow student). Household numbers and compositions alter as individuals join to form households or leave households to create or join new ones. Migration is associated with both household formation and dissolution. As activities

shared among members change, so do the precise definitions adopted in census or survey questionnaires. An alternative definition, which is necessary when using administrative records, defines a household as the group of people who live at the same address. However, migration of households is rarely studied, except in the Netherlands (*Van Imhoff/Keilman* 1992), because complex changes take place at the same time in household membership (e.g. separation of partners, death of spouse, birth of a new child). Most research using aggregate data on migration focusses on individuals, whose characteristics are either fixed (e.g. place of birth), stable (e.g. gender) or change systematically (e.g. age). However, there is large and expanding literature that analyses family and household mobility using household survey data. *Mulder* (2018) provides a comprehensive review of the field to which she has contributed many rigorous and insightful studies.

2.3   Migrant attributes

The literature which investigates aggregate migration flows focuses most attention on the age and sex differentials. However, many more attributes are important in migration processes. These include health status (*Wallace/Kulu* 2014), labour force status (*Marois et al.* 2019a), family status (*Mulder* 2018), educational attainment (*Bernard/Bell* 2018), motivations (*Coulter et al.* 2011) and ethnic status (*Darlington-Pollock et al.* 2019). Such attributes have been investigated using census microdata or survey data. Note that migrant attributes tend to be measured at the end of the migration interval and so their impacts are not precisely measured unless longitudinal microdata are available (e.g. from the population registers of Nordic countries or the census-based longitudinal studies of UK devolved administrations).

2.4   Usual residence

The term "usual residence" is the dwelling unit in which people normally reside (sleep, eat and do household work in). Sometimes a detailed definition is provided in the official questionnaire, such as the place where people spend most of their nights. However, working members of a household may spend more time at a distant work location than at the residence they regard as "usual". People who use more than one residence need to choose which to report as their usual residence. Administrative registers ranging from comprehensive population lists through registrations for local or national tax purposes require reporting of a *de jure* (legal) residence though many allow registration of temporary residences.

2.5   A duration criterion for migration or an instantaneity?

To distinguish between usual and temporary residence, statistical offices employ length of residence as a criterion. For example, people who cross international boundaries are asked about how long they will be staying at their destinations. The United Nations specifies that countries use a 12-month duration criterion to distinguish between long-term and short-term migrants, with the exact method of gath-

ering this information left to national statistical offices. This may involve gathering electronic information on all entries and exits, administering a sample survey at entry or exit or deriving the information from registers. The duration criterion is linked to visa regulations governing visits, temporary and permanent migration. When the duration of stay is short, usually up to 3 months but up to 6 months in many countries, the entrant is classified as a visitor. For durations of 3 months/6 months to 12, the entrant is classified as a temporary or short-term immigrant.

2.6 Migration and space

Internal migration measurement is particularly sensitive to the number, size and shapes of regions as migration is recorded as people cross those region boundaries. If a national territory is divided into a few regions, a lower count of migrants or migrations will be recorded than if a larger number of regions, necessarily smaller, is used. The exact shape and borders of regions will also affect the counts of migrants or migrations. This dependence of a migration measures on the spatial systems used in measurement hampers comparison of internal migration across both countries and time periods. *Bell* and *colleagues* (*Bell et al.* 2002, 2015a/b, 2018; *Bell/Muhidin* 2011; *Stillwell et al.* 2014, 2016; *Rees et al.* 2017a; *IMAGE Project* 2020a)[3] have pioneered the development of system wide measures of internal migration, in which the effect of size and zonation of areal units is established and stable summary measures derived for international comparisons of internal migration.

Stillwell et al. (2018) use zone design software that assembles basic areal units into larger zones (*IMAGE Project* 2020b) to investigate the effects of scale and zone configuration on migration indicators and distance decay parameters. The results of this research are that, above a threshold number of regions, many system-wide measures remain reasonably constant. The test data set employed by *Stillwell et al.* (2018) consists of migrant flows between 404 Local Authority Districts in the UK derived from the Special Migration Statistics from the UK's 2011 Census (*Duke-Williams et al.* 2018). Note that the IMAGE software does not "solve" the MAUP. Rather, it allows the user to explore the effect of spatial framework design on their own migration system and select scales and zone systems which can be compared between countries or time periods.

2.7 Full migration histories

It is vital, when using existing migration data or in designing new methods of measurement, to understand the key concepts involved. Here we use a set of graphs (Fig. 1) to identify the most important types of measure.

Ideally, we would like to track the full migration career of individuals. Figure 1A shows an example history for one individual in a time-space diagram. The horizontal axis proceeds from left to right, from time t_b when the individual is born through to

---

[3] The IMAGE project compares **I**nternal **M**igration **A**round the **G**lob**E**.

Fig. 1: Types of migration measure

Fig. 1:   Continuation

1E: Fixed Interval migrants by transition type, from transition demographic accounts

Exist-Survive   Exist-Die   Born-Survive   Born-Die

region i
region j

t_1 t_2 t_1 t_2 t_1 t_2 t_1 t_2

1F: Migrants, conditional on duration of residence

VVVVVVVSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSLLLLLLLLLL

region i

LLLLLLLLLLL

region j

t_1 t_{LV} t_{VS} t_2 t_{SL}

V = Visitor   S= Short-Term Migrant L=Long-Term Migrant
t_1 = start of interval t_2 = end of interval t_{VS} = time of Visitor to Short-Term transition
t_{SL} = time of Short-Term to Long-Term transition

1G: Migrations of different types (from registers)

Return migration   Repeat Migration   All Migrations

region i
region j

t_1 t_2 t_1 t_2 t_1 t_2

Source: Authors' design

time t_d when they die. Vertical timelines, t_1 and t_2 bracket a period of interest. The vertical axis divides space into two regions, i and j, between which the person can migrate. The open arrows, which straddle the border between regions i and j, represent a migration. The person experiences one migration from region i to region j in the t_b to t_1 interval, a return migration from region j to region i in time interval t_1 to t_2 and a further migration to region j in time interval t_2 to t_d. The individual makes three migrations over their lifetime, the second of which is return migration to region of birth and the third a return migration to a region with a previous spell of residence. A full set of migration histories from which migration could be measured would consist of all lifelines that entered, stayed in or left the system of interest. To measure the intensity of migration we would divide the number of events of migration by the person time of exposure in the regions of interest. Only the highest quality population registers, in the Nordic countries, provide the real-world equivalent of Figure 1A. In longitudinal studies based on censuses or surveys such as the Longitudinal Study of England and Wales (*CeLSIUS* 2019) there are missing elements. Reconstruction of migration histories through retrospective questions in surveys suffer from survivor bias as respondents who have died or emigrated cannot be interviewed.

2.8 Lifetime migration

Lifetime migration data analysed by *Ravenstein* (1876, 1885), compare only place of birth with place of current residence (Fig. 1B). In the 1871 and 1881 Censuses, a question was asked about a person's birthplace, using county as the geographical unit. Tables were constructed for the populations of England and Wales, Scotland and Northern Ireland by county of residence and county of birth. There is no information about the exact timing of the migration or about how many migrations occurred between birth and date of enumeration. In the decade 1871-1880, male life expectancy was estimated as 41.4 years and female as 44.6 years (*ONS* 2015). So, it is likely that most migration took place between 1831 and 1871. Similar tables of lifetime migration at county scale were produced in censuses from 1891 to 1951, except for 1931 (*Friedlander/Roshier* 1966).

In 1961, lifetime migration tables were restricted to UK home countries and a fixed one-year interval question substituted ("where were you living one year ago?"). In the sample Census of 1966 and the full Census of 1971, two migration questions were asked using one-year and five-year intervals. From 1981 to 2011, only the one-year question was asked.

2.9 Migrants by last residence

Common questions included in censuses are: "where was your last residence" (open interval) or "where was your last residence in the previous X years" (fixed interval), because this is the easiest question for respondents to answer. Figure 1C shows the time-space graph for the open interval case where the timing is uncertain. Figure 1D presents the fixed interval case where the location at the start of the fixed interval

is uncertain. Uncertainty is indicated by using a pecked timeline or lifeline for the migrant. Most countries that collect last migration data (with timing undefined) do also collect duration of residence. These data can be used to approximate fixed interval measures, but only if the duration is reported in finely graded intervals, and if the space to which the duration refers is clearly specified. These conditions are rarely satisfied. This type of question provides important descriptive information but like the lifetime migration question picks out only one migration without precise information about when the migration occurred. *Rees* (1985) showed that it was difficult to use the results as an input for estimating or forecasting the population. However, *Schmertmann* (1999) proposes three methods for estimating multistate transition hazard rates from last migration data. The methods are compared using a micro-simulation model with invented test data. The author then applies the methods to estimate fixed period transition flows from last migrant data for a Brazilian three state system (Parana, Sao Paolo, Rest of Brazil). He demonstrates that the back-projection method produces plausible estimates.

2.10 Fixed interval migrants

The left-hand graph in Figure 1D shows the lifeline of a typical fixed interval migrant, based on a retrospective census or survey question. Data on fixed interval migrants are generated by asking the question "where were you living X years ago?" (where X is typically, 1, 5 or 10 years) or "where were you living at the time of the last census?" (as in the French census, where censuses have been held at irregular rather than regular intervals). It is difficult to compare migration volumes and intensities of questions with different time intervals because of return and repeat migration (*Rees* 1977). Papers from the IMAGE project (*Bell et al.* 2015b; *Stillwell et al*. 2016 and *Rees et al.* 2017a) report on both 1-year and 5-year migration results but do not compare them directly. *Rogers et al.* (2003) and *Dyrting* (2018) have developed models for translating between 1-year and 5-year probabilities, so in future harmonization may be possible if the assumptions of the models hold for a country. However, Figure 1D shows three graphs which have fixed start and end states in the period which may not be captured by a census or survey question: the exist-die, born-survive and born-die transitions, all of which might involve an inter-regional migration. These additional life-state to life-state transitions were recognized by *Rees* and *Wilson* (1977) in their book on *Spatial Demographic Analysis* as building blocks of transition demographic accounts. In principle, the born-survive transition should be captured by a census question, either through cross-tabulating the residence region at the census of 0-year olds or 5-year olds against their place of birth, derived from the Births Register but this is rarely done.

2.11 Fixed interval migrants, conditional on duration of stay

Many surveys of international migration (*ONS* 2019) incorporate a condition that to be counted as a long-term migrant that person must reside for 12 months or more, following a recommendation by the United Nations. Durations of stay of 3 months

up to 12 months classify migrants as Short-Term while durations of less than 3 months lead to classification as Visitors (Fig. 1E). *Nowok* and *Willekens* (2011) develop a framework that identifies types of migrants, conditional on how long they stay at their destination after migration. The authors specify a general process of migration drawing statistical theory and, using a microsimulation model of an invented sample of migrants, show how different intensities result in different expected transitions over intervals of different lengths (Fig. 5 and 6 in *Nowok* and *Willekens* 2011). This work should be an important building brick in the development of a tool to handle the Modifiable Temporal Unit Problem (MTUP) discussed later in section 3.5.

2.12 Migrations as boundary crossing events

Figure 1's final panel (1F), focuses on representations of migrations as boundary crossing events. The right most graph illustrates the situation when border crossings are counted. A migrant may make more than one border crossing in a time interval of interest as in the middle graph. This person would have been recognised once as a fixed interval migrant between region i and region j. Such migrants are recognised as "performing" repeat migrations (left hand graph in Fig. 1F). When those migrations are a sequence of region i to j and then region j to i, they are return migrants (right hand graph in Fig. 1F). Later in the paper we show that it is necessary to estimate these two types of migrant and the number of moves they make to convert a migrant count into a migration count.

2.13 Sources of migration data

Bell et al. (2015a) reported on the types of internal migration data gathered by countries across the world. Table 2 identifies the types of retrospective questions used in the 2000 round of censuses. The commonest question is asked about place of birth (lifetime migration), followed by duration of residence, last migration and then the five-year, one-year and other interval questions, upon which most comparative analysis is based. In addition, *Bell et al.* (2015a) record the number of countries reporting migrations (events) from registers, though such data are gathered mainly in European countries.

The categories of Census, Register and Survey cover a range of strategies and coverages of the population. Censuses are normally designed to be complete enumerations of the population and surveys are normally samples, designed to be representative of the national population. Often, difficult-to-handle questions, such as those about migration, are assigned to long-form representative samples to reduce the burden of coding geographic responses. In some cases, the migration questions are left out of the census and included in a separate national survey, as in the US Census of 2010 where no migration question was asked in the main census but included instead in the annual American Community Survey. This strategy is adequate for many kinds of migration measure (total inflows, total outflows, age and gender profiles) for larger administrative units but provides little information on origin-destination flows, because the N categories of interest (e.g. ~3,000 US coun-

Tab. 2: Number of countries collecting internal migration data in the 2000 UN census round by world region

World Region	Observation period						
	1 year	5 year	Other fixed interval	Lifetime	Last migration	Duration of residence	Total countries collecting data
Africa	9	8	8	29	13	17	32
Asia	2	13	8	27	18	24	35
Europe	13	4	12	25	10	12	31
Latin America & the Caribbean	2	17	2	29	12	13	29
Northern America	1	2	0	2	0	0	2
Oceania	2	8	2	10	2	5	13
Total	29	52	32	122	55	71	142

Source: *Bell et al.* (2015a), Table 2, p5.

ties) turn into N^2 flow dyads (3,000 by 3,000 or 9 million data points). The modern US migration researcher would have a hard time replicating Ravenstein's analysis of counties of absorption and dispersion.

Administrative Registers avoid the small sample problem by aiming to cover the whole of a target population such as all usual residents or all taxpayers or all patients in the health care system. So, coverage issues replace the sampling issues of survey or within census samples, substituting the advantages of continuous recording and capacity to generate good time series of migration information. However, registers are often restricted in the range of co-variates available. For example, the National Health Patient Register used to measure UK migration provides just basic demographic information (location, gender, age) for statistical release and imposes very strict safeguards and terms and conditions of use for associated health data. Fully developed registration systems such as that of Finland (*Statistics Finland* 2019) overcome this difficult by linking demographic information from the population register with socio-economic information from the tax and benefit records and employment data from the business register.

For many questions about migration, researchers use national or international household surveys or microdata samples from censuses (*Bell et al.* 2015a). Such sample data are very useful where available, yielding new insights into migration behaviour by migration cohort (*Bernard* 2017) or by migration motivation (*Coulter et al.* 2011) by type of area, but they are too small for the analysis of the place to place migration that Ravenstein studied.

Alternative detailed survey data are collected by commercial polling or marketing firms who also need detailed place-specific data. *Thomas et al.* (2014) explore the potential of a *large commercial survey data*, produced by marketing company (*Axciom* 2019). Under licence the authors had access to unit postcode data for respondent households and their residents. They demonstrate that results for the de-

terminants of migration are close to those from official surveys, and that unweighted and reweighted data produce similar results. *Stillwell* and *Thomas* (2016) use postcode-to-postcode migration data to estimate intra-zonal migration distances. They show that traditional methods of estimating this variable are flawed. Distance decay parameters calibrated for spatial interaction models of migration differ significantly using individual rather than zone data.

There has been huge interest in using *electronic data* generated by users of internet platforms and of smartphone devices to provide statistical information on a comparable basis for member states (*UN* 2019a). To date, there has been relatively little progress. The data belong to private corporations, which make limited data available but retain the right to change the algorithms that generate the information. Migration is not a variable that is directly generated from available internet queries and phone location traces. Corporations are reluctant to release the most detailed data for testing or calibration purposes. Nevertheless, for many countries that lack good official statistics systems, "big data" are a useful source, particularly for establishing the location of most used residence and for monitoring daily journeys (*Manley/Dennett* 2018).

A second example of use of telecommunications data is provided by *Lai et al.* (2019). In a Namibia case study where both census and mobile phone call data records (CDRs) are available, the authors construct a model that predicted migration flows between 13 regions from CDR data and some fixed co-variates. The aim was to develop migration flow data for years after the latest census up to the next. A linear model with co-variates had high goodness of fit ($R^2 = 0.94$). There is no way of fully testing the model until the next census, but it clearly has considerable potential for countries without continuous records of migration. Mobile phone penetration in less developed countries is already high and getting higher.

An important study was carried out by a team of experts for the European Union to investigate the potential of using social media data to fill the gap between the latest official Labour Force Survey data and "now" (*European Commission* 2019). The authors commented as follows: "the first results of the application of the stocks model are experimental, but they are promising" but that "the approach taken to estimate EU Mobility Flows has not yet offered any plausible results" (*European Commission* 2019: 15). The problem, recognised by the authors, was that official statisticians have virtually no control over social media data from Facebook or Twitter. Use of call data records seems more promising.

3   Filling the gaps: Harmonizing and estimating migration

Ravenstein (1876) used lifetime migration from the 1871 Census of Great Britain and Ireland, while his 1885 paper employed similar tables from the 1881 Census. There were two gaps in the data: one spatial and one temporal. The spatial gap was the absence of flow data from between counties in one "home" nation (England and Wales together, Ireland and Scotland) and counties in another (*Ravenstein* 1885: Map 6). Outside the home nation only birthplace by home nation was made avail-

able. The temporal gap was the absence of migration flows from one census to the next. The timing of the transition between birth county and residence county was unknown. Ravenstein did not have the statistical tools to fill these spatial and temporal gaps. Contemporary researchers take the view that once the phenomenon to be studied has been identified, efforts should be made to use the partial data available along with auxiliary data and statistical models to make an estimate of the true flows. We now discuss a sequence of critiques, improvements and new estimates that have generated new estimates of migration, which can inform and improve official statistical practice.

3.1 Harmonizing spatial units used to measure internal migration

Ravenstein was fortunate in being able to study historic counties, whose boundaries did not change between 1871 and 1881. American, Australian and Canadian researchers can use state or province boundaries which are fixed from the date of admission to their respective Union, Commonwealth or Confederation, and French researchers can rely on boundaries for *départéments*, first created in 1790 by the National Constituent Assembly of Revolutionary France. Other European countries have experienced greater volatility in both regional and municipal boundaries. The UK, for example, has experienced frequent changes in the boundaries of its local authorities, which have shrunk in number and grown in population through mergers, in order to achieve economies of scale in administration. Look up tables based on the populations of the smallest geographical units are used to apportion population statistics from the "old" geographic units to "new" geographic units (*Simpson* 2002). This works reasonably well for statistics that apply to one spatial entity. However, for migration flows which refer to two spatial units, the origin and the destination, this method is problematic. To aggregate migration flows from one geography to another, it is essential to aggregate flows between small basic spatial units (BSUs), which ideally nest within the larger units of interest. If there are still overlaps, then the centroids of zones can be used to match BSUs via a point in polygon function (*Bell et al.* 2015a).

3.2 Estimating missing internal migration data

Since the 1966 UK Census (a 10 percent sample), migration origin-to-destination flow data sets have been produced, at first on user demand at cost (1971, 1981) and then as publicly accessible data sets in 1991, 2001 and 2011, free at the point of use, though with restriction to safe settings for the most detailed tables (*Duke-Williams et al.* 2018). These flow data sets provide tables for regions, local authorities, and small areas (e.g. output areas and super-output areas), though the detail of the tables shrinks with spatial scale to avoid disclosure issues. However, the difficulty that Ravenstein experienced of flow tables only being available within UK "home" countries and the lack of harmonised flow data for the years between censuses persists for flow tables derived from the main UK administrative register, the Patient Register of the National Health Service (NHS) (*Lomax* 2013a: Chapters 3 and 4). Flow tables

are available for each mid-year to mid-year interval for each home country but flows between local authority districts (LADs) in one home country and LADs in another are not generated. Figure 2 (based on a similar diagram in *Lomax et al.* 2013b) sets out the status of information in each cell of the matrix of flows. The table has origin and destination LADs sorted by home country. LAD to LAD flows (represented by the green cells) are available within England, Wales and Scotland and between England and Wales. Flows between LADs in Scotland and Northern Ireland and LADs in other home countries are not available (cells shaded in beige). For Northern Ireland LAD to LAD flows are also not available. The grey cells represent marginal totals (total inflows and outflows of various kinds). The question is: "how can the full matrix be estimated?"

The solution is to borrow information from a closely aligned source, in this case the full 2001 or 2011 Census flows matrices and adjust these borrowed flows to agree with the known marginals derived from the home country NHS registers for years 2001-02 to 2010-11. The generic algorithm used is called "iterative proportional fitting" (IPF), applied to each sub-matrix in Figure 2 (*Lomax* 2013a/b; *Lomax/ Norman* 2016). The algorithm has been re-invented and renamed in many social disciplines. *Yule* (1912) used it to standardize cross-tabulations; *Kruithof* (1937) developed a double factor method for use with telephone call data; *Deming* and *Stephan* (1940) employed the method to adjust census tables; *Bishop et al.* (1975) wrote a definitive guide to the method. IPF is known as bi-proportional fitting in spatial interaction modelling (*Wilson* 1971) and employed in migration modelling (*Stillwell* 1978, 1979). When employed in economics, IPF is referred to as the RAS method and used to estimate input-output accounts; when employed to re-weight sample survey results the term is referred to as statistical raking (*Mercer et al.* 2018). The Lomax work achieved consistent year by year estimates of migration flows between 389 local authority districts in the UK using IPF techniques.

American demographer *Andrei Rogers* and colleagues have worked intensively on methods for estimating internal migration when direct data are not available. In the monograph *The Indirect Estimation of Migration* (*Rogers et al.* 2010), the authors summarise achievements of three decades of research on migration estimation, acknowledging the key contribution of *Willekens* (1983) on the use of log-linear models to understand the contributions of origin, destination and interaction factors to migration flow matrices. One of the motivations for writing the book was concern that in 2010 the US Census Bureau would no longer ask a fixed interval question on residence 5 years ago but instead rely on the annual (but sample) American Community Survey to provide information on internal migration with the United States (*US Census Bureau* 2019). Log-linear modelling helps to estimate missing flow data from marginals and proxy interaction data. The Rogers team also worked on modelling the age variation of migration intensity, using the migration by age data assembled in the *Migration and Settlement* study at the International Institute for Applied Systems Analysis (*Rogers/Castro* 1981a; *Rogers/Willekens* 1986), developing a multi-exponential model of the age profile of migrants. The average exponents of the profile across the 17 country inter-regional matrices by age in the study have been widely used to disaggregate or improve migration by age information (e.g. *Sander*

Fig. 2: The estimation of origin-destination migration flows between local authorities in the United Kingdom
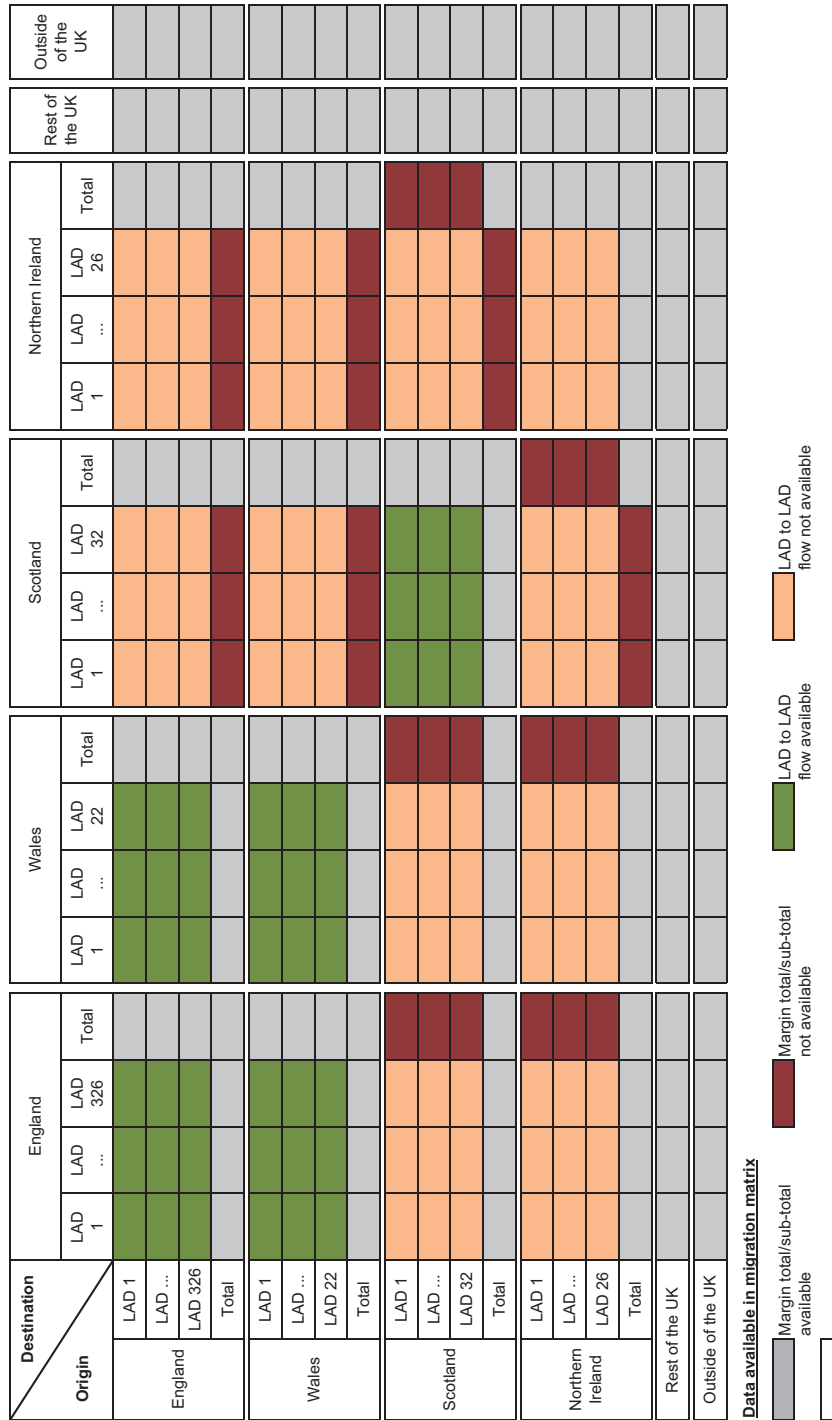
Source: Adapted from *Lomax et al.* (2013b), Fig. 1, p. 268 in the Journal of Population Research, published by Springer.

et al. 2014a). Subsequently, the age profile model has been refined by *Wilson* (2010) and linked to life course event profiles (*Bernard/Bell* 2015). The decomposition of migration by age is hinted at by *Ravenstein* (1876) but no data were available for investigation (*Greenwood* 2019: 271).

3.3   Filling the gaps in international migration data

Despite the prominence of international migration in political discourses in most countries and the existence of national statistics (at least on immigration), there is no system for producing harmonised measures of international migration flows. The best data currently available are the national census tables of migrants by country of residence and country of birth, collected by the Statistics Division of the United Nations and published by the UN's International Organization for Migration (*IOM* 2019). These data are analogous to the 1871 and 1881 census data analysed by Ravenstein, referring to country to country flows rather than county to county flows. The United Nations Population Division assembles data from national statistical offices on net international migration where available and produces estimates as the residual balance after natural change is subtracted from total population change. The IOM data portal (*IOM* 2019) points to the category migrant flows, but these are only available in Europe from EUROSTAT (the Statistical Office of the European Union). Flow estimates were compiled partially from data supplied by national statistical offices of EU member states and partially from estimates compiled by academic researchers working with EUROSTAT (*Raymer et al.* 2011, 2013).

Poulain, Perrin and *Singleton* (2006) carried out a comprehensive survey of sources, definitions and measures of migration available from EU member states. The THESIM Report (Towards Harmonised European Statistics on International Migration) covers EU migration policy and data regulation, administrative systems of data collection, statistics on international migration flows (where available), residence and work permits, asylum applications, removals of persons refused leave to remain together with twenty-five country reports written by national experts using a uniform template. The edited collection by *Raymer* and *Willekens* (2008) builds on the THESIM work by reviewing models available for generating better migration estimates. *Raymer* (2008) proposes a schema for carrying out the estimation by reconciling conflicting origin country and destination country estimates or by borrowing information from other countries where it was not available for a country. In the case of the UK reliable data on flows to/from other EU countries were missing and had to be borrowed from origins or destinations with better statistics.

Raymer and colleagues working together in the MIMOSA (Migration Modelling for Statistical Analysis) project developed a model for estimating international migration flows between member states in the EU and EFTA and flows to and from the Rest of the World. The model combines the procedure for harmonization proposed by *Van der Erf* and *Van der Gaag* (2007), which works from most to least reliable estimates with the log-linear modelling framework set out in *Raymer* (2008), a gravity model for intra-EU and extra-EU flows in the European Union, where flow estimates are missing (*Raymer/Abel* 2009; *Abel* 2010; *Raymer et al.* 2011). The estimates were

further refined in the IMEM (Integrated Modelling of European Migration) project which incorporates expert views about the reliability of flows by country and provides measures of uncertainty/confidence (prediction intervals) around those estimates within a Bayesian framework (*Raymer et al.* 2013).

These estimates of migration flows between EU and EFTA states have been used as inputs in a subsequent demographic forecasting project DEMIFER (Demographic and Migratory Flows Affecting European Regions and Cities) (*Rees et al.* 2010; *Rees et al.* 2012). Scenarios of future migration are developed that are linked to migration and regional development in national and EU policies. The projection model used a multi-level approach (*Kupiszewski/Kupiszewska* 2011), in which migration flows between EU states were explicitly forecast and then allocated to regions within EU states using information on immigration and emigration or population totals by region for EU states. *Dennett* and *Wilson* (2013, 2016) took the MIMOSA/DEMIFER estimates and used them to generate a full matrix of interregional migration across the EU that combines international flows between EU member states and internal interregional migration estimates.

3.4 Estimating period migration data from lifetime migration data

Section 3.2 showed how migration flows can be estimated from partial data, where reliable sub-totals were available together with a full matrix of reliable proxy migration flows to adjust. Section 3.3 discussed ways in which a full flows matrix could be constructed from partial information by using information from destination countries with high quality statistics where origin countries had very poor information. In this section (3.4), we report on methods of estimating period-specific migration flows between counties (internal migration) or countries (international migration) from tables of population classified by current residence and place of birth, the same as the lifetime migrant tables used by *Ravenstein* (1876, 1885). Table 3 sets out the contributions to these estimations.

The first entry identifies Ravenstein's use of lifetime migrant data using information from two censuses. He did not attempt to estimate period migration data but has stimulated subsequent researchers to investigate possible methods of doing so. Important work on the estimation of migrant flows from migrant stocks was carried out by *Friedlander* and *Roshier* (1966). The aim of the analysis was to use lifetime censuses from 1851 to 1951 to create period estimates for inter-census migration between counties in England and Wales. The authors built a model that uses two successive lifetime migrant tables, information on births and deaths in the inter-census intervals and three important assumptions. The first is that persons who die do not make any migrations and die at their place of enumeration at the first census. The second is that only one migration is recognised between censuses and additional (repeat) migrations are ignored. The third assumption is that flows from origin 1 to destination 2 are cancelled out by flows in the opposite direction. Age structures of lifetime migrants available from the 1911 census table are used to estimate age structures at other censuses. Death rates based on decennial life tables for England and Wales are used to compute survivors and non-survivors. All

Tab. 3:   Contributions to the estimation of period migrant flows from lifetime migrant stocks

| Author(s) (Method)	Source Data: Lifetime Migrant Stocks	Number of Territorial Units	Estimated Period Migrant Flows
Ravenstein 1885 (Uses Lifetime Data Only)	UK Census Tables 1881	118 counties	Not attempted
Friedlander/Roshier 1966 (Survival Analysis & Assumptions to generate Net Migrant Flows)	EW Census Tables 1851,1861, 1871, 1881, 1891, 1901, 1911, 1931, 1951	53 counties	Net intercensal migrant bi-lateral flows, 6 ten-year intervals and 2 twenty-year intervals
Abel 2013 (Formulation of problem as IPF estimation of 3 D migrant array, with maximum stayer assumption)	WB Stock Tables 1960, 1970, 1980, 1990, 2000	191 countries	Gross intercensal migrant bi-lateral flows, 4 ten-year intervals
Abel/Sander 2014 (Uses Abel 2013a methods with UN Migrant Stock Table, and Circular Migrant Flows)	UN Stock Tables, 1960, 1970, 1980, 1990, 2000	196 countries	Gross migrant bi-lateral flows, 8 five-year intervals
Dennett 2016 (Uses apportionment probabilities from Global totals; compares estimates against IMEM flow estimates and UN flow estimates used by Kim and Cohen 2010)	WB Stock Tables 1960, 1970, 1980, 1990, 2000	191 countries	Gross migrant bi-lateral flows, 4 ten-year intervals
Abel 2018 (Uses Abel 2013a method with both WB and UN Stock Tables by Gender with successive UN population and vital statistics estimates in comprehensive sensitivity analysis that yields 90 global migration tables)	WB Stock Tables, 1960, 1970, 1980, 1990, 2000, UN Stock Tables, 1980, 1990, 2000, 2010, 2015	WB 196 countries UN 232 countries	Gross bi-lateral migrant flows, for 10-year intervals (1960-1990) and five-year intervals (from 1990)
Azoze/Raftery 2018a, 2018b (Adapts Abel 2013a method using Bayesian methods that relax the maximum stayer assumption. Factor up flows by using ratios from internal migration studies relating 1-year and 5-year migration)	UN Stock Tables: 1990, 1995, 2000, 2005, 2010, 2015	UN 232 countries	Gross migration bi-lateral flows for 5-year intervals, 1990-95 to 2010-15. The method converts estimates of migrants into estimates of migrations

Tab. 3:   Continuation

Author(s) (Method)	Source Data: Lifetime Migrant Stocks	Number of Territorial Units	Estimated Period Migrant Flows
Abel/Cohen 2019 (Test out 6 alternative methods for generating bi-lateral international migrant flows: 2 use stock differencing; 1 uses migration rates, 1 uses open demographic accounting + IPF: 2 use closed demographic accounting + IPF or a pseudo-Bayesian model)	UN Stock Tables: 1990, 1995, 2000, 2005, 2010, 2015 Validation data: 45 countries: Immigration flows Emigration flows WPP2017: Population counts Net migration flows	UN 200 countries	Gross bi-lateral migrant flows, for 5-year intervals (1990-2015). Correlations of flows against validation data

Notes: *Abel/Cohen* (2019) also discuss two contributions (*Beine et al.* 2011 and *Bertoli et al.* 2015) which use the net differencing of country i to country j migrant stocks in two successive censuses. These methods resemble those used by *Friedlander/Roshie*r (1966).

this means that the Friedlander and Roshier model can only estimate minimum net migrant flows. It is not possible to deduce gross flows from net as there are many pairs of gross flow values that would report the same net flow. However, their maps of net migration flows show the dominance of coalfield counties (Glamorgan, Monmouth) and of London as main beneficiaries of net migration in the decades 1861-71, 1871-1881 of Ravenstein's analysis, and 1881-91 to 1901-11. The periods 1911-31 and 1931-51 show a reversal of migration flows from South Wales and the North East as coal mining declined with a persistence of inflows to London and surrounding counties.

Table 3 shows that we had to wait until the 2010s for a methodological breakthrough to be made by *Abel* (2013). He has subsequently refined the estimates in four further papers. Figure 3 sets out the 3-dimensional cube of migration flows invented by *Abel* (2013) that enabled the estimation of period migration from lifetime migrant stocks. Abel skilfully uses simple numerical illustrations to explain the procedures for estimating period migration flows from two consecutive migrant stock tables. Two faces of the cube represent lifetime migrant data at two successive censuses (Matrix A and Matrix B). We wish to estimate the contents of the third face, migrant transitions for a fixed time interval (Matrix C). *Abel's* insight was to realise that if provisional estimates or assumptions of the elements in the three dimensional array were proposed, they could be adjusted to agree with the marginal totals in Matrices A and B. Matrix C elements were simply a sum of array elements over country of birth.

However, it is also useful to set out the general algebraic framework for the estimation, as in Table 4. The notation derives from work on demographic accounting (*Rees/Wilson* 1977; *Rees* 1985 and *Rees/Willekens* 1986). The notation used is set out in Table 5. The variables to be estimated are migrant flows by country of birth

Fig. 3: A framework for estimating period migration from lifetime migration
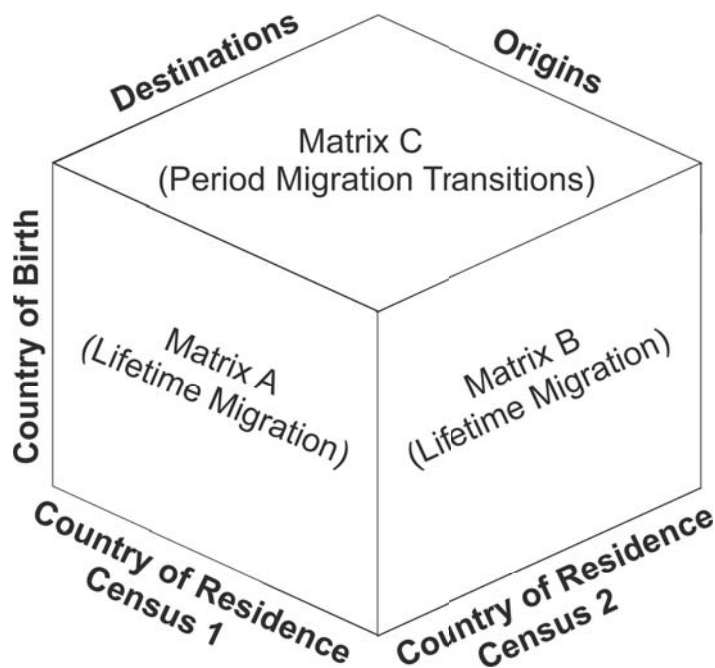
Source: Drawn by the authors, based on a model by *Abel* (2013).

(i), country of origin at the first census (j) and country of destination (k), represented by the variable T_{ijk}^{bes} in Table 4D. How do we arrive at these period estimates of migrant flows by place of birth? We start with information on lifetime population stocks at the first census, P_{ij}^{be} (Matrix A in Fig. 3, Table 4A) and the second census, P_{ik}^{bs} (Matrix B in Fig. 3, Table 4B). These marginal tables can be used to adjust an initial distribution (seed values) of migration flows in the full array X_{ijk}^{bes}. The problem translates into one familiar to users of contingency tables. Iterative proportional fitting methods can be used to estimate adjusted array elements, associated with an underpinning log-linear statistical model (*Deming/Stephen* 1940; *Bishop et al.* 1975; *Willekens* 1983; *Rogers et al.* 2010). Once an array of T_{ijk}^{bes} variables had been generated (Fig. 3), then the required migrant flows matrix (Matrix C in Fig. 3, Table 4E) is obtained as a sum over place of birth of the array variables, that is, $T_{+jk}^{bes} = \sum_i T_{ijk}^{bes}$.

A condition for a successful solution to iterative proportional fitting (IPF) is that the totals of elements in the constraint matrices (Matrix A and Matrix B in Fig. 3) are the same. Otherwise the IPF procedure will never converge. It is necessary to reduce the migrant total for Matrix A by the deaths that occur in the inter-census interval and to reduce the migrant total for Matrix B by the births that occur between censuses, as *Friedlander* and *Roshier* (1966) had earlier incorporated in their methods. In Table 4A the total deaths in a country, D_{+j}^{bd}, are shaded in pink, indicating the counts are available from vital statistics records or UN estimations using, for example, the Demographic and Health Survey. The country of birth of persons dy-

ing is not known, but can be estimated by applying the proportion P_{ij}^{be}/P_{+j}^{be} from the population stocks table to the total deaths over all birthplaces to produce a table of estimated deaths. Births are handled differently, simply by assigning total births to the country of residence at the second census, that is, $B_{ik}^{b+} = B_{k+}^{b+}$. Subtraction of deaths from population by country of residence at time t and subtraction of births from population by country of residence at time u produces two tables of estimated survivors in the third panels of Table 4A and 4B. The totals of these two sub-tables must be made equal, $\sum_i P_{+j}^{be} = \sum_i P_{i+}^{be}$, for example, by constructing an average (*Abel* 2013, 2018; *Abel/Sander* 2014). The two sub-table population elements are adjusted to agree with this common total.

The IPF estimation model requires specification of initial or seed values of all array elements (Table 4C), designated X_{ijk}^{bes}. This process of estimating the origin-destination-birthplace tables involves two steps: (1) making an assumption about the diagonal cells in the array (people who reside in the same country at both start and end of the time interval), and (2) estimating initial values of the off-diagonal elements (people who reside in a country at the end of the period from that at the start). The options for the diagonal include maximization of stayers (as in *Abel* 2013), minimization of stayers or an average of the two. *Abel* (2018) suggests options for the off-diagonal elements include adopting an independence model or introducing an interaction matrix which reflects the difficulty of moving between countries (e.g. spherical distance, airline connections). *Azose* and *Raftery* (2018a/b) use a weighted average of the assumption of maximum stayers and the independence model assumption, developing a method for optimizing the weights. The IPF procedure is carried out conditional on the margin totals and diagonal elements.

A spatial interaction model could be developed for a sub-set of country-to-country migrant flows for which reasonable estimates have been made. Such a model was developed by *Kim* and *Cohen* (2010), fitting their model to a UN data set of international migration flows between industrialized countries. However, the drawback would be that if researchers wanted to fit a global spatial interaction model to estimated flows, there would be an element of "double-counting". Given a choice of the seed values for the migrant flow array, estimates of the full array of migrant flows can be generated and then summed to produce the country to country matrix of flows (Matrix C in Fig. 3, Table 4E).

The Abel estimates generate migrant flows only for persons existing at time t and surviving at time u. Demographic accounting theory (*Rees/Wilson* 1977) demonstrates that for a complete picture of population change and migration it is necessary to consider migrants hidden in the deaths and births statistics. Table 6 expands the matrix in Table 4E by adding transitions for infants born in the period t to u, who survive in the destination country at time u (shaded green), the persons alive in the country of origin at time t who die in the t to u interval (shaded blue) and the infants born in the period who also die before the end of the time interval (shaded brown). The off-diagonal elements in each of these additional sub-matrices in Table 5 represent persons with (at least) one migration in the time interval. The number of such "hidden" migrants is likely to be small (perhaps 1 percent of exist-survive migrants)

Tab. 4: Country to country migrant flows estimated using the Abel (2013, 2018) methods

4A: Lifetime Migrants at Time t

Country of birth	Country of residence at t				
Populations	1	...	j	...	n
1	P_{11}^{be}	...	P_{1j}^{be}	...	P_{1n}^{be}
:	:	...	:	:	:
i	P_{i1}^{be}	...	P_{ij}^{be}	...	P_{in}^{be}
:	:		:		:
n	P_{n1}^{be}	...	P_{nj}^{be}	...	P_{nn}^{be}
Totals	P_{+1}^{be}	...	P_{+j}^{be}	...	P_{+n}^{be}
Deaths	1	...	j	...	n
1	D_{11}^{bd}	...	D_{1j}^{bd}	...	D_{1n}^{bd}
:	:	...	:		:
i	D_{i1}^{bd}	...	D_{ij}^{bd}	...	D_{in}^{bd}
:	:		:		:
n	D_{n1}^{bd}	...	D_{nj}^{bd}		D_{nn}^{bd}
Totals	D_{+1}^{bd}	...	D_{+j}^{bd}	...	D_{+n}^{bd}
Survivors at t	1	...	j	...	n
1	P_{11}^{be}	...	P_{1j}^{be}	...	P_{1n}^{be}
:	:	...	:		:
i	P_{i1}^{be}	...	P_{ij}^{be}	...	P_{in}^{be}
:	:		:		:
n	P_{n1}^{be}	...	P_{nj}^{be}		P_{nn}^{be}
Totals	P_{+1}^{be}	...	P_{+j}^{be}	...	P_{+n}^{be}

4B: Lifetime Migrants at Time u

Country of birth	Country of residence at u				
Populations	1	...	k	...	n
1	P_{11}^{bs}	...	P_{1k}^{bs}	...	P_{1n}^{bs}
:	:	...	:	:	:
i	P_{i1}^{bs}	...	P_{ik}^{bs}	...	P_{in}^{bs}
:	:		:		:
n	P_{n1}^{bs}	...	P_{nj}^{bs}	...	P_{nn}^{bs}
Totals	P_{+1}^{bsu}	...	P_{+k}^{bsu}	...	P_{+n}^{bsu}
Births	1	...	k	...	n
1	B_{11}^{b+}	...	0	...	0
:	:		:		:
i	0	...	B_{ik}^{b+}	...	0
:	:		:		:
n	0	...	0		B_{nn}^{b+}
Totals	B_{1+}^{b+}	...	B_{k+}^{b+}	...	B_{n+}^{b+}
Survivors at u	1	...	k	...	n
1	P_{11}^{bs}	...	P_{1k}^{bs}	...	P_{1n}^{bs}
:	:		:		:
i	P_{i1}^{bs}	...	P_{ik}^{bs}	...	P_{in}^{bs}
:	:		:		:
n	P_{n1}^{bs}	...	P_{nk}^{bs}		P_{nn}^{bs}
Totals	P_{+1}^{bs}	...	P_{i1}^{bs}	...	P_{+n}^{bs}

4C: Lifetime Migration Seed Values

Country of birth i	Country of residence at u				
Country of residence at t	1	...	k	...	n
1	X_{i11}^{bes}	...	X_{i1k}^{bes}	...	X_{i1n}^{bes}
:	:	...	:	:	:
j	X_{ij1}^{bes}	...	X_{ijk}^{bes}	...	X_{ijn}^{bes}
:	:		:		:
n	X_{in1}^{bes}	...	X_{ink}^{bes}	...	X_{inn}^{bes}

4D: Period Migration Transitions

Country of birth i	Country of residence at u					
Country of residence at t	1	...	k	...	n	Totals
1	T_{i11}^{bes}	...	T_{i1k}^{bes}	...	T_{i1n}^{bes}	P_{i1}^{be}
:	:	...	:	:	:	
j	T_{ij1}^{bes}	...	T_{ijk}^{bes}	...	T_{ijn}^{bes}	P_{ij}^{be}
:	:		:		:	
n	T_{in1}^{bes}	...	T_{ink}^{bes}	...	T_{inn}^{bes}	P_{in}^{be}
Totals	P_{i1}^{bs}		P_{ik}^{bs}		P_{in}^{bs}	

4E: Period Migration Transitions Summed over Country of Birth

All countries of birth +	Country of residence at u					
Country of residence at t	1	...	k	...	n	Sums
1	T_{+11}^{bes}	...	T_{+1k}^{bes}	...	T_{+1n}^{bes}	P_{+1+}^{bes}
:	:	...	:	:	:	P_{+j+}^{bes}
j	T_{+j1}^{bes}	...	T_{+jk}^{bes}	...	T_{+jn}^{bes}	P_{+j+}^{bes}
:	:		:		:	
n	T_{+n1}^{bes}	...	T_{+nk}^{bes}	...	T_{+nn}^{bes}	P_{+n+}^{bes}
Sums	P_{++1}^{bes}		P_{++k}^{bes}		P_{++n}^{bes}	T_{+++}^{bes}

Key to colours:

Grey	data input from census tables assembled by World Bank and United Nations
Pink	estimates based on subtracting deaths from the time t population
Green	estimates based on subtracting births from the time u population
Blue	assumptions for initial values of the migrant flows array
Yellow	estimates of final values after running iterative proportional fitting and then summing over country of birth

Source: Based on an interpretation of *Abel* 2013, *Abel/Sander* 2014 and *Abel* 2018.

Tab. 5: Variable and subscript notation for lifetime and period migrants

Variables	Definition
P	Population (stock)
D	Deaths
B	Births
X	Seed (assumed) values of migrant flows before adjustment to migrant stocks
T	Transitions (between state at time t and state at time u)
LM	Long-term migrants (UN definition)
M	Migrations (events of migrating across borders between one area and another
R	Residual balance term in population accounts for migrations
G	Grand total of inputs and outputs in population accounts for migrations
Subscripts	Definitions
t	Time at start of period/interval
u	Time at end of period/interval
t,u	Combination of start and end times, signifying a time period/interval
b	Being born during a time period/interval
e	Existing at the start of a time period/interval
s	Surviving at the end of a time period/interval
d	Dying during a time period/interval
i	Place/zone of birth (e.g. county within country or country within world)
j	Place/zone of residence at time t
k	Place of residence at time u
n	Number of zones in system studies
+	Sum over the subscript replaced

Note: Subscripts can be placed below the variable they are attached to or above.

but it is better to estimate them than include in the closure error. Methods for estimating these migrant flows are described in *Rees* and *Wilson* (1977).

Both *Dennett* (2016) and *Azose* and *Raftery* (2018a) regard the Abel estimates of bi-lateral migrant flows as under-estimates of the "true" flows. Of concern are migrants who return to their origin country during a time interval and migrants who make more than one migration. We have discussed these migration histories earlier (see Fig. 1A). Both authors compared the Abel estimates with the estimates of international migration between EU member states (see earlier for a discussion). The claim of under-estimation derives from a comparison of Abel estimates with migration flow estimates between EU member states (*Raymer et al*. 2011; *Raymer et al.* 2013). However, these international migration estimates are based, largely, on adjusted counts of entries to and exits from national population registers, where commonly no duration of residence criterion is applied. So, these register-based estimates count migration events (Fig. 1F, rightmost graph) rather than (transition)

Tab. 6: Expanded country to country migrant flows adding transitions

		Survive at u					Death t to u					
	Country	1	…	k	…	n	1	…	k	…	n	Totals
Exist at t	1	T_{11}^{es}	…	T_{1k}^{es}	…	T_{1n}^{es}	T_{11}^{ed}	…	T_{1k}^{ed}	…	T_{1n}^{ed}	T_{1+}^{e+}
	⋮	⋮		⋮		⋮	⋮		⋮		⋮	⋮
	j	T_{j1}^{es}	…	T_{jk}^{es}	…	T_{jn}^{es}	T_{j1}^{ed}	…	T_{jk}^{ed}	…	T_{jn}^{ed}	T_{j+}^{e+}
	⋮	⋮		⋮		⋮	⋮		⋮		⋮	⋮
	n	T_{n1}^{es}	…	T_{nk}^{es}	…	T_{nn}^{es}	T_{n1}^{ed}	…	T_{nk}^{ed}	…	T_{nn}^{ed}	T_{n+}^{e+}
Born t to u	1	T_{11}^{bs}	…	T_{1k}^{bs}	…	T_{1n}^{bs}	T_{11}^{bd}	…	T_{1k}^{bd}	…	T_{1n}^{bd}	T_{1+}^{b+}
	⋮	⋮		⋮		⋮	⋮		⋮		⋮	⋮
	J	T_{j1}^{bs}	…	T_{jk}^{bs}	…	T_{jn}^{bs}	T_{j1}^{bd}	…	T_{jk}^{bd}	…	T_{j1}^{bd}	T_{j+}^{b+}
	⋮	⋮		⋮		⋮	⋮		⋮		⋮	⋮
	n	T_{n1}^{bs}	…	T_{nk}^{bs}	…	T_{nn}^{bs}	T_{n1}^{bd}	…	T_{nk}^{bd}	…	T_{nn}^{bd}	T_{n+}^{b+}
	Totals	T_{+1}^{+s}	…	T_{+K}^{+s}	…	T_{+n}^{+s}	T_{+1}^{+d}	…	T_{+k}^{+d}	…	T_{+n}^{+d}	T_{++}^{++}

Colours:

Yellow	Surviving migrant flow estimated from lifetime migration tables at successive censuses/estimates (Table 4)
Green	Estimates of new-born surviving migrant flows
Blue	Estimates of non-surviving migrant flows
Brown	Estimates of new-born, non-surviving migrant flows
Grey	Data from censuses and vital statistics

Notes: Variables: T = transitions. Superscripts: e = exist at time t, s = survive at time t+1, d = die in interval t, t+1, b = born in interval t, t+1. Subscripts: 1 … n = individual countries in world

Source: Based on *Rees/Wilson* 1977.

migrants (Fig. 1D, leftmost graph). In other words, the comparison is between migrants and migrations. Both are "true" flows, conditional on their definitions. The choice of which estimates to use depends on the application they are used in. *Rees* (1985) showed that either type of flow data could be used in a population projection based on a components-of-change model, but that a matching projection model needed to be used. If, for example, you wished to compute the carbon footprint of migrant travel, you would need to use the migration flows rather than the migrant flows.

Azose and *Raftery* (2018a/b) make estimates of the migration events additional to the number of migrants by borrowing results from three pieces of work on internal migration that compared one-year migrant rates with five-year migrant rates in censuses where the migration question had asked respondents where they were living one-year and five-years ago (*Long/Boertlein* 1990; *Rogers et al.* 2003 and *Newbold* 2005). *Azose* and *Raftery* (2018a/b) use an average of 1.5 for the Long-Boertlein Index, the ratio of the measured five-year interval migrants to five times the one-year interval migrants as generated from retrospective questions on previous residence. The method assumes that migration intensity is constant over the five-year interval and that populations are approximately constant. *Azoze* and *Raftery* adopt a one-year retrospective interval equivalent to the 12-month duration (either prospec-

tive or retrospective) recommended for the measurement of long-term international migration by the United Nations. The effect of this procedure is to lift the bilateral flows of Abel by around 200 percent over 5 years (*Abel/Cohen* 2019, Table 7). Note that despite the paper's title, *Azose* and *Raftery* (2018a/b) do not identify the separate contributions of return and transit migrations. Table 7 shows how the Azose and Raftery migration estimates fit into an event-based population accounts table.

Tab. 7: Estimates of country to country migration flows

Migration Flows		Destination country in period t to u					Deaths	
		1	...	k	...	n		Totals
Origin country in period t to u	1	R_1	...	M_{1k}	...	M_{1n}	D_1	P_1^t
	:	:		:		:		
	j	M_{j1}	...	M_{jk}	...	M_{jn}	D_j	P_j^t
	:	:		:		:		
	n	M_{n1}	...	M_{nk}	...	R_n	D_n	P_n^t
Births	+	B_1		B_k		B_n		B_+
	Totals	P_1^u		P_k^u		P_n^u	D_+	G_+

Notes: M = long-term migration using the UN 12-month duration definition. Migrant flow estimates derived from an average of minimum stayer and independence model by multiplication of five-year migrant estimates by the Long-Boertlein ratio (5-year migrant count/5× one-year migrant count) set to 1.5 based on empirical estimates in *Rogers et al.* (2003) and *Newbold* (2005).

M_{jk}	Variables estimated by Azose and Raftery
P_j^t	Variables added to create a demographic account based on migrations

Source: Based on *Azose/Raftery* (2018a/b), *Rees* (1985) and *Rees/Willekens* (1986).

The analyses in Table 3 are not the end of the story. It is necessary to extend the estimation to decompose migration flows by age, as in *Friedlander* and *Roshier* (1966). This would provide a better basis for using international migration flows in a global population projection than assumed the age profile matched the average age profile parameters of the 17 country case studies in *Rogers* and *Castro* (1981a). However, this assumes that country of birth and country of residence tables are available for age. The censuses and surveys containing lifetime migrant data should hold birth date or age at last birthday information, but special tabulations would need to be requested from censuses or might be generated from sample census microdata.

3.5 The "one-year/five-year" problem

The problem of estimating consistent migration probabilities from data with different retrospective intervals had been observed for France by *Courgeau* (1973a), though his solution was to say that good longitudinal migration histories were needed to solve the problem. Courgeau's paper first identified the clear distinction

between migrants (persons making a spatial transition between two points in time) and migrations (the events of changing locations).

In the 1970s and early 1980s, an ambitious study of future population change at regional scale using multi-state methods for 17 countries was carried out at the International Institute for Applied Systems Analysis (IIASA) located in Laxenburg, Austria by *Andrei Rogers* and *Frans Willekens* (*Rogers/Willekens* 1986). As is usual in studies across countries, there were difficulties in harmonizing the data need for input to the software for implementing the projections (*Willekens/Rogers* 1978). The key harmonization problem was to align national migration data with model requirements for migration over five-year intervals, matched with five-year ages. Countries with migration events derived from population registers simply added up five years of migrations. Countries which relied on censuses for migration information could either supply data based on a one-year retrospective migration question or a five-year question. Occasionally, two questions were asked in the national census, as in the UK in 1971 (*Rees* 1979). The existence of two sets of data for the UK enabled *Kitsul* and *Philipov* (1981) to experiment with a high- and low-intensity movers model which gave reasonably good results, bearing in mind that the one-year data (1970-71) were available for just one-year of the five in the five-year data (1966-71).

Having surveyed contemporary research on estimating migrant and migration flows, we turn attention in the next section to contemporary research into a selection of Ravenstein's empirical findings.

4   Recent research on internal migration patterns

This section is not designed to be a comprehensive review of what we know about contemporary internal migration but rather provides illustrations of how contemporary migration scholarship carries out analyses of the internal migration patterns that Ravenstein studied.

4.1   Comparisons of internal migration across countries: case study collections

Using lifetime migration data from censuses in the 1880s and earlier, *Ravenstein* (1889: Maps 1 and 2) describes the regional distribution of foreigners for the countries of Europe by region and the net gains (absorption) and losses (dispersion) in international migration at country scale. Drawing on the 1880 censuses in the United States of America (USA), he maps the distribution of the foreign-born population in 1880 across the states of the USA and the provinces of Canada (Map 3), the gains and losses of state-born, national-born and foreign-born migrants by state and province (Map 4), and examples of migration fields of dispersion (Map 5) and absorption (Map 6). Since the 1880s, the systematic comparison of internal or international migration patterns across countries has not figured prominently, with most work focussing on national case studies, brought together in edited and informative collections (*Nam et al.* 1990; *White* 2016; *Champion et al.* 2018). Com-

parison between countries was difficult because data types, variable definitions and theoretical frameworks differ from chapter to chapter, with the editors making heroic attempts in discussion chapters to overcome the differences in chapter design and achieve a true synthesis of knowledge. *Rogers* and *Willekens* (1986) summarize the results of an international project covering 17 countries in which authors carried out and interpreted the same analysis, a multi-regional population projection incorporating inter-regional migration by age and sex using common software *(Willekens/Rogers* 1978). One common output across the case studies was a table of life expectancy by regions of birth and regions of subsequent residence (e.g. *Rees* 1979: Table 28). The span of life spent outside the region of birth was an index of internal migration intensity. In the late 1990s, a study for the Council of Europe (*Rees/Kupiszewski* 1999) focussed on internal migration within ten member states. Data on internal migration flows supplied by national collaborators at the smallest available scale made possible the elucidation of spatial patterns. These were urban concentration or de-concentration; the relationship of internal migration (or overall population change where internal migration statistics were not available) with population density, the links between internal migration and economic conditions, signalled by unemployment rates, and the relationship of internal migration to gender and the life course. The study identified that many countries were still urbanizing; a few had adopted a counter-urbanization; and other exhibited an intermediate pattern of net flows both down the density gradient from the top and up the density gradient from the bottom.

4.2   Development of a robust methodology for comparing internal migration across countries

A step change in the ability to compare internal migration came with the publication of a paper proposing fifteen measures organized by four measurement dimensions for internal migration: intensity, distance decay, connectivity and effect on population redistribution (*Bell et al*. 2002). The main aim of the paper was to design summary indices of internal migration for a country which could be submitted to international statistical databases (*UN* 2019b/c) and constitute the internal migration component of a Human Migration Database equivalent to the Human Mortality and Human Fertility databases (*HMD* 2019; *HFD* 2019). In the 2010s, a team at the University of Queensland led by Martin Bell have, with funding from the Australian Research Council, been able to implement this agenda across countries of the world (*IMAGE* 2020a) in collaboration with the Asian Demographic Research Institute, Shanghai. *Bell et al.* (2015a) provide an inventory of the data sets used, a mixture of census and register aggregate tables of migration flows, and of census and survey individual microdata sets from which flow data can be computed, held in the an international repository developed by the University of Minnesota (*IPUMS* 2019). Table 8 summarises the significant contribution of Bell and colleagues to our understanding of internal migration across countries that cover 80 of the world's population.

Tab. 8:　Contributions of IMAGE Project Participants to the Analysis of Internal Migration

Topic	Description	Reference
Theory	Defines four structural dimensions and 16 associated indices for describing migration flows	Bell et al. 2002
Data	Reviews sources of internal migration data in countries of the work and describes IMAGE database	Bell et al. 2015a
Software	Describes IMAGE Studio, software for computing indices of migration flows at multiple scales and zoning, to overcome the MAUP problem in migration measurement	Stillwell et al. 2014
Intensity	Develops a method for standardising the measurement of migration that controls for number and size of zones	Bell/Muhidin 2009, 2011; Bell et al. 2015b; Courgeau et al. 2012
Distance	Investigates the effect of distance on migration flows using a spatial interaction model for a set of countries where detailed flow data are available	Stillwell et al. 2016
Impact	Uses net migration and migration effectiveness indices to compare the impact of internal migration on population in two countries	Stillwell et al. 2000, 2001
	Uses IMAGE indicators to compare the impact of internal migration on population redistribution for sets of countries and develops a theory linking migration directions to urbanization-counter-urbanization processes	Rees et al. 2017a
Country comparisons within continents	Uses selected indicators from the IMAGE Studio to compare the structure of migration with continental groups of countries	Latin America: Bernard et al. 2017 Asia: Charles-Edwards et al. 2017a/b, 2019; Bell et al. 2020 Europe: Rowe et al. 2019
Age	Develops methods of harmonizing the age-period-cohort classifications of migration	Bell/Rees 2006;
	Develops simpler models for describing the migration-age relationship Reconceptualises the migration-age relationship as a product of life course transitions	Bernard/Bell 2012; Bernard et al. 2014a,2016; Bernard/Bell 2015 Bernard et al. 2014b
Cohort	Investigates the role of birth cohort in influencing internal migration patterns in European countries and China	Bernard 2017; Bernard/Pelikh 2019; Bernard et al. 2019

Tab. 8:   Continuation

Topic	Description	Reference
Education	Educational differentials in internal migration across the countries of the world	Bernard/Bell 2018
Temporal Change	Extracts information on temporal change for selected countries and assesses the intensity decline hypothesis	Bell et al. 2018

Source: https://imageproject.com.au/framework/

At the core of the IMAGE research was a recognition that the MAUP (*Openshaw* 1983) could potentially invalidate any comparison across countries and that methods were needed to compensate or control for the problem. The MAUP problem has two aspects. The first, the scale problem, is that internal migration measures tend to increase with increasing numbers and decreasing sizes of the zones used. The second, the zoning problem, is that, for any given number of zones, there are multiple ways of creating those zones from smaller basic spatial units. Summary measures may differ drastically across different groupings of the same number of zones (*Openshaw* 1983: 24, Fig. 2). *Stillwell et al.* (2014) developed a software package that, given the input of a matrix of flows between basic spatial units for a country and associated populations, computes aggregations of the flow data for specified numbers of zones and recomputes 15 national summary indices of migration (Table 2 in *Stillwell et al.* 2014). For any aggregation (number of zones), values of the summary indicators are computed for a user-defined number, up to 1,000, of different zoning systems with reports of the median, upper and lower quartile values. It is then possible to evaluate the effect of scale (number of zones) and zonation (number of different constructions of the same number of zones) on the outcome summary indicator.

The scale effect is used to construct summary measures of migration intensity, drawing on an original proposal by *Courgeau* (1973b), further developed by *Bell* and *Muhidin* (2009, 2011) and *Courgeau et al*. (2012). *Bell et al.* (2015b) use the Courgeau-derived methodology to compare migration intensities across the world, linking intensities with development status, with confidence that the intensities were comparable. The authors used the Aggregate Crude Migration Intensity (*ACMI*), designed to be comparable across countries, to reveal the extraordinary variation in migration intensity around the world, from a high of 52 percent per annum in South Korea to a low of only 5 percent in India. Ravenstein did not discuss migration intensity directly in his papers because the period of exposure to the risk of migration in lifetime migrant data was uncertain.

Stillwell et al. (2016) compare the effect of distance on migration volumes using the IMAGE software suite to reveal the scale and distance effects and make possible robust comparisons (see sub-section 4.3). *Rees et al.* (2017) used measures

of the impacts of internal migration to compare the effects of migration on population redistribution, combining the crude migration intensity (CMI) and migration effectiveness index (MEI) measures to create a synthetic index of migration impact (INMI). A key finding was that the MEI was stationary over space down to 20 ASRs (Aggregated Spatial Regions, combinations of BSUs, Basic Spatial Units). This was an unexpected outcome and important because it was key to development of the cross-nationally comparable Index of migration Impact – the INMI. This might be regarded as one of the most important findings of the IMAGE project.

Net migration measures were used to examine re-distribution patterns and a general theory linked to urbanization was proposed. Papers on Latin America, Asia and Europe (*Bernard et al.* 2017; *Charles-Edwards et al.* 2017a/b; *Bell et al.* 2020; *Rowe et al.* 2019) compare on countries within world regions across three of the four dimensions of internal migration identified in *Bell et al.* (2002).

4.3   Distance and migration

The evidence that Ravenstein presented for his short distance generalization (1885-1 in Table 1) and remarks on long-distance migration (1885-5 in Table 1) came from maps of migration rates to or from single counties or through tables of lifetime migrants and so were essentially qualitative observations. In the 20th century, researchers measured distances from migrant origins to destinations and related the size of flows to the distances and other variables. A gravity model of inter-city movement was proposed by *Zipf* (1946), based on Newton's formulation of the attractions of planetary bodies. Since 1946, a very large number of gravity models of migration have been proposed and implemented and the model continues to be used. *Poot et al.* (2016) characterise some recent applications of the gravity model of migration as "the successful comeback of an ageing superstar". The authors make interesting proposals for how the model might be used in other applications, such as population forecasting.

However, the variety of model forms and explanatory variables in the corpus of gravity model case studies make it difficult to compare the impact of distance on migration across countries. The IMAGE project team therefore approached the task of comparing the effect of distance on migration by using a standard model for all countries: the doubly constrained spatial interaction model of *Wilson* (1971) with observed out-migration totals as the origin terms and in-migration totals as the destination terms, a negative power function and distance measured using national grid Cartesian co-ordinates (*Stillwell et al.* 2016). The IMAGE software was used to compute distance decay parameters at different population scales and employing different zoning solutions for a fixed number of zones. Figure 4 presents results for a set of countries reporting fixed interval migration for one-year data. Mean distance decay parameters (beta) are averaged over 1,000 different zoning configurations at each selected zone population size.

The ranges of betas at different zonal population sizes (*Stillwell et al.* 2016, Figure 8) are moderate to small. The graphs of Figure 4 show a feature characteristic of most IMAGE analyses: over ranges where the population sizes are small and

Fig. 4: Mean distance decay parameter by population size, countries with (a) 5-year and (b) 1-year migration data
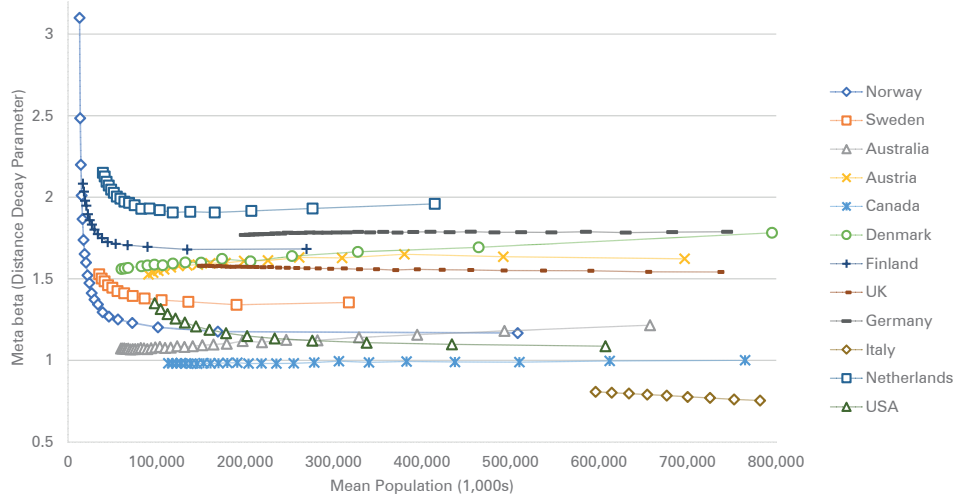
Source: Graph re-drawn based on *Stillwell et al.* (2016): 1667, Figure 6, with permission of the authors.

numbers of zones are large, the distance decay parameter varies substantially by size and moves in different directions in different countries. However, for systems with fewer, larger zones the parameters for each country are relatively stable and can be compared. For example, the mean betas in the population size range 200,000 to 600,000 show few cross overs. We can be sure, for example, that Swiss migrants are more constrained by distance than Australian migrants. For one-year migrants, betas are stable from zone sizes of 100,000 population and larger. Countries exhibit greater variation in betas when migrants are measured over five years than one year. *Stillwell et al.* (2016: 1672) summarise that, of those countries which collect one-year migration data, the frictional effect of distance is lowest in the USA, Canada, Australia and higher in much of Western Europe. For those countries that collect five-year data, those with higher levels of development display lower levels of distance friction.

Most demographic phenomena vary systematically by sex and age. The friction of distance is no exception. Figure 5 plots origin-specific distance decay parameters generated in a study for the UK government by a team led by Stewart Fotheringham and Tony Champion (*ODPM* 2002; *Fotheringham et al.* 2004). Gravity models were fitted to a time series of migration flows (events) between 96 geographic zones in England and Wales over 14 years for 14 sex-age groups. The shaded box shows the upper quartile, median and lower quartile of the distribution of beta values across origin zones. The upper and lower whiskers are located 1.5 times the quartile-median difference, above and below the box. The numbered circles represent outlier zones beyond the whiskers. There is relatively little difference between male and female results by age, because men and women mostly migrate as couples or as

families. There are two exceptions. At ages 16-19 many young people enter higher education, which in England and Wales often involves a migration. The lower median and downshifted box plot for young women probably reflect a preference of either the daughters or their parents that they attend university closer to home. At ages 60 and over, the median for women and the boxplot extent are higher for women. At these ages men and women may no longer be together because of death or divorce. Men die earlier and so are less exposed to the necessity of migration after death of a spouse or partner. At other ages the male and female plots are very similar but vary systematically by age. Younger adult ages experience lower frictions of distance (that is, smaller negative numbers) in a pattern that resembles the plot of migration intensity by age.

4.4   Gravity models of migration

The gravity model has been used in hundreds of applications to understand flows of people, goods, services and investments. Economists have shown a revived interest, as *Poot et al.* (2016) suggest. However, because disciplines still work in silos, *Ramos* (2016), in his review of gravity models, still claims that the recently assembled UN and World Bank databases of lifetime migrants constitute "bilateral data on migratory flows", whereas in section 3.4 we have reviewed the recent attempts to convert what is migrant stock data into estimates of migrant and migration flows over fixed time intervals. There is not space here to treat gravity models (also called spatial interaction models) fully but we briefly describe some important papers. *Sen* and *Smith* (1995) and *Crymble* (2019) provide valuable overviews of gravity models. *Wilson* (1971) developed a family of models based on entropy maximizing methods and *Willekens* (1983) showed how similar log-linear models could be derived from statistical theory. *Flowerdew* and *Murray* (1982), *Flowerdew* and *Lovett* (1988), *Flowerdew* (2010) and *Congdon* (1993) describe Poisson regression models which deal with small numbers or zeroes in origin-destination cells.

    Abel (2010) provides an example of a gravity model used to fill in missing cells in a flow table. He fits a binomial regression model using an expectation-maximization algorithm and co-variates suggested by theory to flows between European countries that are reliably estimated. The model is then used with known co-variates to estimate cells where there are no reliable flow estimates. *Raymer et al.* (2019b) extend this approach to estimating migration flows between ASEAN countries in which there are no reliable estimates at all and gravity model parameters are borrowed from fitting a model to migration flows between European countries instead. The model parameters are used with ASEAN country co-variates. Here we describe some recent work which has advanced our understanding or exposed problems that still need to be solved.

    A key assumption made in gravity models of internal migration is that there are no restrictions on free movements within national spaces whereas between nations restrictions based on immigration policy are ubiquitous. History provides examples of violations of this assumption, ranging from the eviction of crofters in Highland Scotland by landowners leading to mass migration to Scotland's cities and other

Fig. 5: Box plots of origin-specific distance decay parameters by sex and age group generated from a model used to predict internal migration flows between 96 zones (Former Health Service Areas) in England and Wales, 1983-84 to 1997-98

Distance Decay Parameter
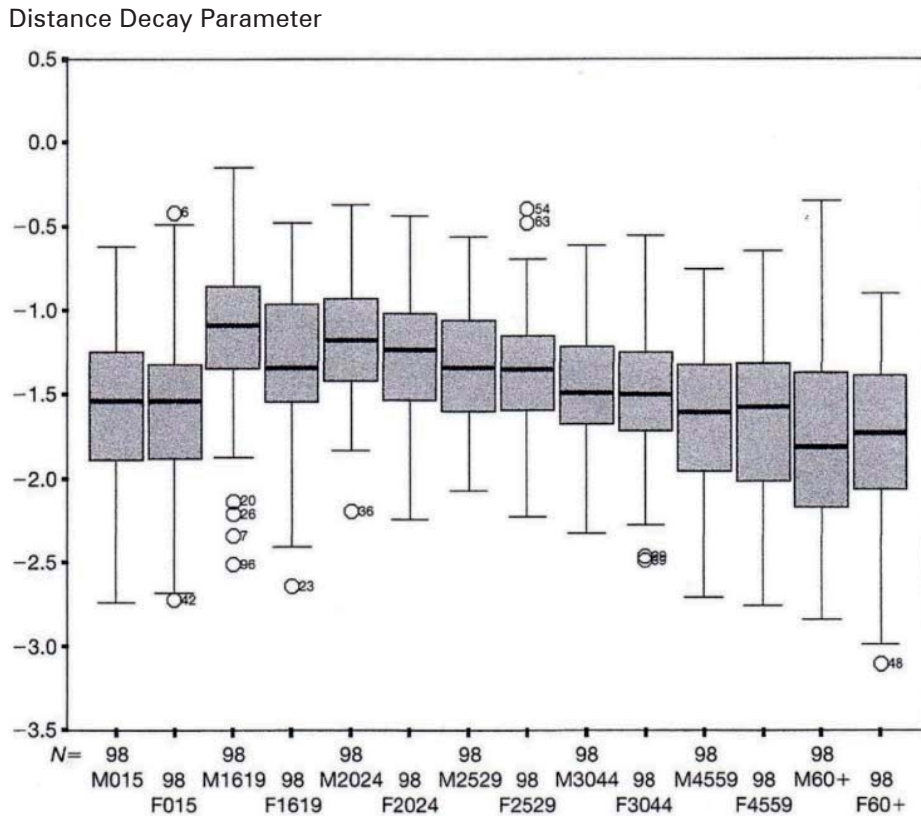
Source: *ODPM* (2002), The Development of a Migration Model, Figure A21.7, p. 301. Crown Copyright.

countries in the Highland Clearances (1750 to 1860) to the example of China's *Hukou* system (1950 to the present), which makes difficult the relocation of rural migrant families to the cities for which they provide vital labour because of restricted access to benefits (*Pradier* 2018). On the other hand, *Fotheringham et al*. (2004) found policy variables at local scale (employment or investment programmes) to have little influence on migration within England and Wales.

Some important decisions need to be made when constructing a gravity model of internal migration flows. First, it is necessary to choose the variable(s) that will represent the impedance between origin and destination. Normally, a simple distance measure is used. However, *Shen* (2016) has shown, in modelling inter-provincial migration in China, that the mis-specification of the spatial interaction effect (predictions of the "friction of distance") leads to absolute mean errors of 32 percent in migration flows compared with 15 percent for each of origin "emis-

siveness" and destination "attractiveness". More attention is needed to investigate alternatives to the minimum distance between origin and destination. Linked questions arise about whether the model should be constructed using intra-zone migration (if available) and how the intra-zone distance might be estimated. *Stillwell* and *Thomas* (2016) show, employing a person level data set recording postcode of origin and destination available from Acxiom Ltd, a marketing technology and services company (*Acxiom* 2019), that conventional geometric methods are inaccurate and that including intra-zonal flows and a more precise intra-zonal distance based on postcode locations improves the goodness of fit of a spatial interaction model substantially. Second, it is necessary to experiment with the mathematical function of impedance which will be used, because predictions of flows are dependent on the decision. *Openshaw* and *Connolly* (1977) provide a methodology for spatial interaction modellers to use. Third, a decision is needed on whether to use a global distance prediction parameter (*Stillwell et al.* 2016) or a set of local parameters either attached to the origin or the destination. A related choice is whether to use a one-step model or a hierarchical model with several steps (*Fotheringham et al.* 2004). Fourth, it is necessary to select or write software that will calibrate the chosen model successfully (*Dennett*/*Wilson* 2016). Can the model be transformed into a linear equation for which standard regression software will fit the coefficients, or should an iterative routine that searches for optimum parameters be used? Fifth, consideration is needed about which variables to include in predicting out-migration from origins and in-migration to destinations. We think Ravenstein would be proud of the sophisticated analyses that his simple "law" about migrants moving short distances has generated.

4.5   Directions of migration flows

Ravenstein's empirical generalisations arose out of his detailed attention to *places*. His papers are filled with list of places (zones) from which lifetime migrants come or to which they go. Today such lists are regarded as the raw material from which we can discover *spatial* generalisations. Places have meaning if you are familiar with the geography of the country being studied, but we try today to attach numerical information to that general knowledge. Ravenstein, being a cartographer by training, drew many maps to illustrate the patterns or directions of flows.

But there is an issue. His maps and ours today are highly selective and convey only a fraction of the information embedded in the data. This can be termed a 3N problem, where N is the number of regions with a country for which you have migration information. To represent the matrix of flows, you need to draw N maps of outflows, N maps of inflows and N maps of the balances. But there are many indexes you can use: the raw numbers, the raw data converted into percentage shares by division by the total inflow or outflow or presented as rates by division of the population of the sending region (outflows) or receiving region (inflows). If V is the number of useful variables for characterising the flows (age, sex, education, ethnicity) then 3N×V maps are needed. In effect, you have created an atlas of evidence for the directions of migration. Could not the maps be reduced by plotting simultane-

ously all the origin to destination flows as lines? Such a plot would be unreadable. *Ravenstein* (1885) does draw some line maps for selected sets of flows but is careful to avoid clashes. *Friedlander* and *Roshier* (1966) plot the largest flows, which reveal the importance of Glamorgan (where most of the South Wales coalfield is located) in the 1890/99 and 1900/09 decades. The French Demographer, Daniel Courgeau, who gave the term "migration field" to the pattern of outflows or inflows around a geographical zone, fails to produce any maps of migration in his monograph on *Les Champs Migratoires en France* (*Courgeau* 1970). For visualization, he relies on graphs of migration against generalised distance bands.

Several techniques can be used to overcome some of these problems. Figure 6 shows how you can plot legible maps for net migration volumes using a proportional circle to represent the absolute number, while using colour to represent direction (gain or loss), for three separate annual intervals during a decade. The maps show continuity of the patterns but with changes in intensity in the third interval when the world was still experiencing the consequences of the global financial crisis of 2008/09. Another technique, promoted by British Geographer Daniel Dorling, is the population cartogram, in which zones are drawn with an area proportional to their population, using hexagonal grids. *Dorling* and *Thomas* (2004) plot an atlas of cartograms based mainly on information from the 2001 Census. Each page offers both the variable on a standard geographic base (a transverse Mercator projection of Great Britain, produced by Ordnance Survey GB) and the hexagonal population cartogram. On pages 61 to 75 of *Dorling* and *Thomas* (2004), lifetime migration from 4 home countries and 11 countries outside the United Kingdom are plotted using the 2001 Census data together with change from the previous census in 1991. These maps are proud successors to those that Ravenstein drew in 1885 and *Friedlander* and *Roshier* produced in 1966. Local knowledge is brought to bear to explain unusual changes. For example, "The English born share of the local population has risen by two percentage points or more in Corby (as Scottish born immigrants to the steel works there have left or died)" (*Dorling/Thomas* 2004: 61).

One very successful graphic for depicting migration flows that is becoming widely used is the circular plot of migration, developed by Nikola Sander, using CIRCOS open source software (*Abel/Sander* 2014; *Sander et al.* 2014a). Figure 7 adapts an example from *Sander et al*. (2018). The plot is for flows between regions in the UK. Design decisions include selecting the number of regions that can be read in the static plot (12) and arranging them in a logical order around the circle. Northern regions are placed at the top, Southern regions at the bottom; regions have neighbours with which they exchange significant flows; only the larger flows are depicted so that number of crossings is minimized. There is potential to design interactive versions of the circular plots, which can be referenced as part of the supplementary material associated with a publication. Failure to take good design decisions can result in what *Sander et al.* (2018) refer to as a "hairball". *Azose* and *Raftery* (2018b) use CIRCOS based plots to demonstrate the differences between their estimates of country to country migration flows and those of *Abel* and *Sander* (2014). Circular plots are used for each of the 15 Asian country cases of internal migration in *Bell et al*. (2020).

Fig. 6: Internal net migration flows for UK local authority districts for mid-year to mid-year intervals 2001/02, 2006/7 and 2010/11
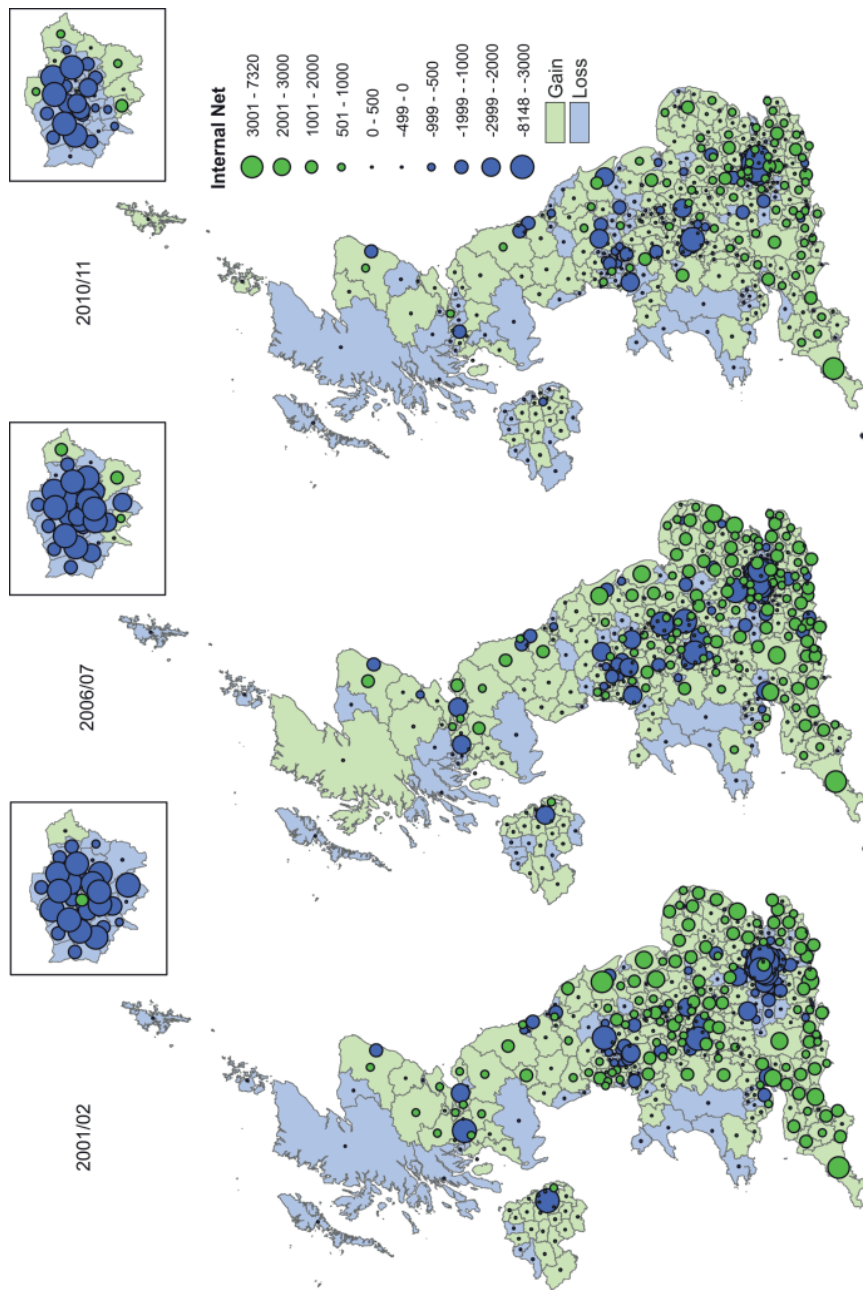
Source: Adapted from *Lomax et al.* (2013b: 282, Fig. 5) in the Journal of Population Research, published by Springer.

Fig. 7:   Circular plots of migration flows between UK regions, 2010-11

Source: Redrawn by the authors based on *Sander et al.* (2018: 205, Fig. 16.1) published by
        Routledge.

4.6   Drivers of migration

It is helpful to consider two kinds of drivers (factors) which influence migration: the context of the places where migrants live and migrate between and the individual characteristics of the migrants which influence their propensity to migrate.

  The most important contextual drivers are the level and changes in economic development, associated with industrialization and urbanization, as Ravenstein recognised. For most of history cities have been the organisers of the space-economy (*Jacobs* 1970). Starting from a low base in the late 18th Century, the share of the world's population which lives in urban places has grown to 58 percent in 2018 (*UN* 2018). In developed countries the urban population makes up between 68 percent and 82 percent of the total, with levels of 50 percent in Asia and 43 percent in Africa. The populations in some cities in eastern Europe, eastern Asia (Japan) and the rust belt of the USA have shrunk in recent decades. The main demographic component contributing to urban population growth has been internal migration supplemented

by international immigration in some developed countries and by natural increase in developing country cities. In selected European countries and in the USA in some recent decades, there have been net flows out of cities (*Rees/Kupiszewski* 1999), which have been termed counter-urbanization. However, often, this counter trend has not persisted. A new wave of urban-ward migration has resumed. The populations in some cities in eastern Europe, eastern Asia (Japan, Korea) and North America have shrunk in recent decades, either as a result of manufacturing decline (e.g. Detroit, St Louis) or because of population ageing (e.g. Nagasaki, Busan). However, cities have also been spreading as new suburbs are added or new towns develop within the commuting field of metropolitan centres. Suburbs and commuting towns are growing through internal migration from centre cities and sometimes through direct immigration from outside the country.

To monitor the migration processes producing urbanization, suburbanization, counter-urbanization and re-urbanization requires careful delineation of functional urban regions (FURs) split into cores, suburbs and peripheries. Many countries have developed geographical definitions of FURs and the European Commission has funded several projects to produce harmonised FUR definitions across EU member states. But FUR definitions differ across countries and most developing countries lack them. FURs also expand their boundaries over time, necessitating decisions about temporal harmonization of boundaries. Different zones within FURs (e.g. core, suburbs, periphery) contain rural settlements; zones outside FURs also contain urban settlements. *Rees et al.* (1996) analyses net internal migration trends using FUR and density frameworks, illustrating the complexities.

To surmount this definitional problem, *Rees* and *Kupiszewski* (1999) used population density by small areas within a country as a means of tracking changing internal migration patterns in ten EU member states, comparing results from the mid-1980s with those from the mid-1990s, straddling the revolution in political systems between 1989 and 1991. Figure 8, re-drawn from the Great Britain case study (*Rees et al.* 1996) illustrates the utility of the density classification. The data used are internal migration flows, classified by broad age for males (the female graphs are very similar), into and out of GB small areas (wards in England and Wales and postal sectors in Scotland) from which net migration rates were computed. Each graph plots the overall GB value for the net internal migration rate in each age band against density classes arranged from least dense (leftmost) to most dense (rightmost).

All age bands except for young adult slope from left to right, with positive net migration into low density areas which falls away to zero at middling densities and becomes negative at the higher densities (the city cores). The declines are steepest for the middle working ages (30-44) of high labour force participation and family formation. The graphs for the childhood ages (1-15) parallel those of 30-44 ages but are lower at the lowest density end, suggesting some differences between family and non-family migration. The downward slopes indicate strong suburbanization and counter-urbanization in the 1990-91 observation period, though this is likely to have been a high water for this process. The slope was less steep for the late working ages, 45 to pensionable age (65 for men and 60 for women in 1991). In the final age band, pensionable age and above, there is net migration towards more dense areas

Fig. 8: Net internal migration gains and losses by density classes for age groups, males, Great Britain, 1990-91
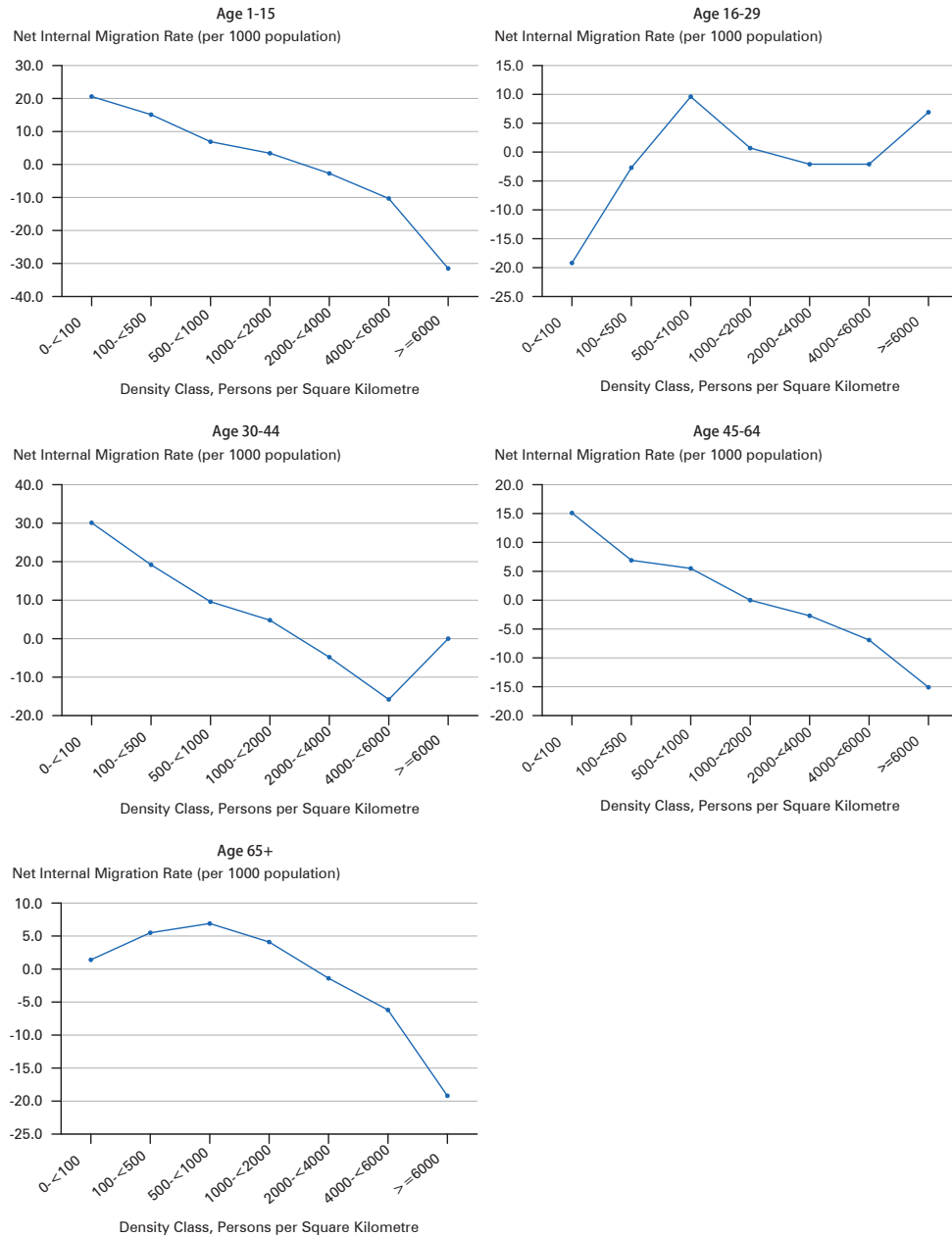
Note: Counts of net internal migrants and population by electoral ward are aggregated to population density classes and rates computed.

Source: Data from the 1991 Census, Special Migration Statistics, Crown Copyright. Indicators computed by the authors. Re-drawn from *Rees et al.* (1996: 72, Fig. 16).

on the left side of the graphs followed by migration away from dense areas. This was interpreted by the authors as a desire among retired migrants not to locate in the most remote areas but in small towns in metropolitan peripheries where health care and other services were more accessible.

Figure 8 represents the relationship between net internal migration and density for one country and time interval. Other researchers had studied this relationship across time in one country (*Courgeau* 1992) and across several countries (*Fielding* 1989). The technique was not to group observation areas into density bands but to plot net migration rates against density for all areas and then fit a linear regression against the logarithm of population density. *Courgeau* (1992) splits the 95 *départements* into rural and urban parts, showing clearly the transition from a strong urbanizing trend in 1954-62 and 1962-68, a flat trend in 1968-75 and a strong counterurbanizing trend in 1975-82.

In the discussion above, we have seen how two individual drivers of migration, age and sex, interact with the context, represented by the population densities of different settlements. The age profiles of internal migration were studied in the Migration and Settlement project at IIASA led by Andrei Rogers. *Rogers et al.* (1978) first proposed a model for describing the variation of migration rate with age as a set of multi-exponential functions linked to the labour force ages, childhood ages, and retirement ages. *Rogers* and *Castro* (1981a) used the model to characterise the internal migration profiles of the inter-regional migration flows in the 17 countries of the project.

Subsequently, additional functions were added for: post-retirement migration by the elderly (*Rogers* 1992) and the double peaks at entry to and exit from higher education (*Wilson* 2010). In a Migration and Settlement case study for the UK, *Rees* (1979) observed that that the retirement peak was confined to flows from big cities to retirement regions and that international migration inflows had a subdued childhood slope and a later peak in the labour force function. *Wilson* (2020) has replaced the rising slope of migration intensity in the age band of retirement with a function that sees an increase followed by a decrease from age onwards. *Rogers* and *Castro* (1981b) analyse migration by age schedules by reason for migration, but data sets providing this additional information are rare in censuses though more common in surveys. The model of the variation of migration by age has been extensively used in other analyses and the schedule parameters have been borrowed to decompose all age migration flows (e.g. in *Sander et al.* 2014b as input to the Wittgenstein world projections of population and human capital).

Bernard, Bell and *Charles-Edwards* have presented an alternative scheme for working with age schedules of migration (*Bernard/Bell* 2012, 2015; *Bernard et al.* 2014a, 2016), which is simpler to implement and more informative. Country profiles are classified by age at peak (years), normalized migration intensity at peak and overall migration intensity. They define three clusters of countries. The first cluster comprises countries with young age peaks (ages 21 or 22), high concentrations at peak ages but low overall intensity. The second cluster has peak ages from 21 to 24 with low concentrations at peak ages and moderate intensities. Countries in a third cluster have later peak ages (26 to 29), moderate to low concentrations at peak ages

and relatively high overall migration intensities (Fig. 5 in *Bernard et al.* 2014a). In an important contribution to migration theory, *Bernard et al.* (2014b: 212) proposed:

> "a framework that links contextual factors to the age structure of migration through proximate determinants that directly affect migration ages: the prevalence, timing, and spread of life-course transitions. We focused on four key transitions that are concentrated at young adult ages and mark the passage to adulthood – exit from education, entry to the labor force, union formation, and childbirth – and sought to link these to the early-to-mid-20s peak commonly found in the age profile of migration. The proposed framework enabled us to identify the determinants of cross-national differences in the age structure of migration and to quantify their relative importance across a global sample of 27 countries".

Other migrant attributes also have important effects. In *World Population and Human Capital in the Twentieth Century* (*Lutz et al.* 2014; for a précis and review see *Rees* 2018), the authors incorporate fertility and mortality rates decomposed by educational attainment categories into a model for projecting the population of world countries. Driven by an independent projection of how educational attainment will change over the current century, improving education raises survival rates but decreases fertility rates in the developing world. The outcome is a set of projected populations significantly lower than those of the United Nations Population Division (*UN* 2015). Education provides skills and knowledge to improve both economic and social well-being. To a certain extent, education also acts as a surrogate for socio-economic status, the occupational and income components of which can change through upward or downward social mobility, whereas the stock of an individual's human capital is more stable, although subject to gradual obsolescence as knowledge grows or to decreases in mental capacity at older ages.

Bernard and *Bell* (2018) showed that, using IPUMS microdata from censuses and surveys, it was possible to measure the differences in propensity to migrate across educational attainment levels in a large set of countries. The authors showed that there was a near universal positive relationship between level of education and the propensity to migrate. Education was one of the few socio-economic variables for which this was the case.

Therefore, it should be possible to introduce education as a driver of internal migration within sub-national population projections. The links between education and migration were largely ignored in *Lutz et al.* (2014). In recent work for the European Commission, *Marois et al.* (2019a/b, 2020) use a large microsimulation model of the European population to connect international migration to the attributes of individuals in the sample microdata in order to assess whether the negative economic consequences of population ageing can be mitigated through higher immigration, higher labour force participation and better integration of new arrivals into the labour force. They find that policies that raise labour force participation rates at older ages, that encourage women to enter the labour market and that promote integration of foreign workers will reduce the increase in the economic dependency rate and mean that immigration levels need not be higher than today.

Internal migration also varies by ethnicity and birthplace. International migrants by race or ethnicity or birthplace have initial distributions that differ from the native population. Subsequently, later generations migrate out of their initial settlement locations to other parts of the country. The consequences for the population composition by race have been monitored in US state projections by *Frey* (2015), for the population composition of local authorities in the UK by *Rees et al.* (2011, 2017b) and for the birthplace composition of Australian state and local populations by *Raymer et al.* (2019a). The USA has experienced a shift in the destinations of Hispanic immigrants from traditional ports of entry to interior cities. In the UK the concentration of ethnic minorities in large cities has been lessened through out-migration. In Australia immigration has been concentrated in the major state capitals but with destinations widening over time. The roles of internal and international migration as contributors to national and local ethnic population change have been assessed for 2001-based projected populations (*Rees et al.* 2013). This analysis extended a decomposition analysis developed for world population regions by *Bongaarts* and *Bulatao* (1999), who claim that, when the world is divided into two regions, developed and developing, migration does not matter. This is a view that does not apply when you drill down to individual countries, sub-national regions and ethnic populations.

4.7   Temporal trends in internal migration

In this review of contemporary analysis of internal and international migration, we have not referred to changes over time. *Abel* and *Sander* (2014) and *Abel* (2018) showed that the view of international migration as an expanding phenomenon was incorrect for the period 1960 to 2015. Yes, the volume of flows showed growth over time in line with world population change, but, when expressed as a rate in relation to population, exhibited fluctuating trends rather than the explosive growth of popular commentators. *Zelinsky* (1971) sets out a profoundly influential set of hypotheses about the "mobility transition". Note that his definition of mobility encompassed human movements including temporary migration, circular migration, daily commuting and visiting, a wider concept than the residential migration reviewed in this paper. He saw "circulation" as replacing the need for migration in later phases of the economic and social development of nation-states (*Zelinsky* 1971, Fig. 2). With hindsight we might extend Zelinsky's definition of circulation to include greater levels/distances of daily commuting, more working from home and telecommunications, including the internet and broad band replacing the need for migration. Other influential theories include the model of urbanization, counter-urbanization and re-urbanization proposed by *Geyer* and *Kontuly* (1993). A theoretical framework linking development to population redistribution through net internal migration was proposed by *Rees et al.* (2017a, Fig. 8) and used in the country studies of Latin America (*Bernard et al.* 2017), Asia (*Charles-Edwards et al.* 2017a/b, 2019; *Bell et al.* 2020) and Europe (*Rowe et al.* 2019). The relationship of internal migration with development was considered in most of the analyses in Table 8, though most are cross-sectional except for *Bell et al.* (2018).

To test these theories and to clarify the directions of recent change requires analysis of long national time series in a robust comparative framework as set out in *Bell et al.* (2002). *Champion* and *Shuttleworth* (2017a) assemble from official statistics on migration based on changes in patient address generated from the National Health Service Register to examine trends in longer and shorter distance migration. *Champion* and *Shuttleworth* (2017b) use the England and Wales Longitudinal Study, which links individuals across census from 1971 to 2011 to look at trends in the distances of migration, by type of person covering age, marital status, country of birth, occupation, and education. They tested the hypothesis, put forward by *Cooke* (2011, 2013), that migration rates were now declining, after demographic composition changes (ageing) had been accounted for. In collaboration, *Champion et al.* (2018) recruited authors from the US, UK, Australia, Japan, Sweden, Germany and Italy to explore trends in internal migration in a set of highly developed countries. *Bell et al.* (2018) contributed a chapter which examined trends in intensity in 66 countries in the IMAGE database for which observations were available for more than one period, focusing on changes in the 2000-2010 decade. Was Cooke's decline trend confirmed? Well, yes and no and maybe (see *Champion et al.* 2018 for the range of case studies with differing outcomes).

5 Summary and discussion

5.1 Since Ravenstein, how far have we come?

In the introduction, we briefly summarised Ravenstein's key insights into internal migration in Britain and in other countries in Europe. His articles also reveal an interest in international migration through his maps of foreign-born immigrants in the British Isles, selected European countries, the USA and Canada. In section 2, we placed the data set used, tables of lifetime migrants by country of residence at the census by country of birth, in a more general conceptual framework of definitions of migrants and migration. In section 3, we tracked further use of this type of migration data to the present decade when solutions have been put forward for estimating fixed period migration from lifetime migration data. We made suggestions for the further refinement of these estimates. In section 4, we showed how some of Ravenstein's empirical and verbal generalizations had been theorized and quantified in 20th and 21st century research, covering the topics of migration and distance, gravity models of migration, migration flow patterns and visualizations, drivers of internal migration and the need for extending comparative internal migration analysis over time. We now set out an agenda for future work on both internal and international migration.

5.2 Agenda for making further progress

The goal in comparing internal migration across countries is to bring the "Cinderella"[4] of population change components (migration) to the demographic ball. Ideally, tables of comparable measures of internal and international migration should appear alongside the tables of indicators for mortality, fertility and net international migration published by the United Nations, the International Organization for Migration and the World Bank. We ask a series of questions, the answers to which may contribute to further progress in estimating, for as many countries as possible, both international and internal migration flows where data are missing or inadequate.

How can we maximise the number of countries and time period for which we have comparable internal migration measures?

The first suggestion focuses on using the rich resources assembled in the IMAGE database so that the measures of migration can be assembled for as many countries as possible and for as many time periods as possible. For example, the models used to produce international migration flow estimates from lifetime migrant data could be applied to improving estimates of internal migration. *Bell et al.* (2015a) find that 109 countries report collection of lifetime migrant data, but little use has yet been made of this resource in IMAGE project analyses. Many countries use census or survey questions on last migration/migrant data which often has an indeterminate time reference. Estimating fixed interval migration from these sources would build up the space-time coverage of comparable internal migration measures. The IMAGE analyses report on 5-year and 1-year fixed interval measures separately. The problem has been investigated many times, but no general solution has been adopted. Further research is needed, building on the literature and the recent model proposed by *Dyrting* (2018). A method for investigating the spatial MAUP has been enabled by the IMAGE studio. An equivalent tool for investigating the temporal MTUP (Modifiable Temporal Unit Problem) needs to synthesize the conditional framework of *Nowok* and *Willekens* (2011) and the migrant heterogeneity models (mover-stayer or high mobility-low mobility) of *Dytring* (2018) and *Schmertmann* (1999). A goal would be to specify a general method for conversion of migration data of different intervals to a common interval for international comparison, that derived the underpinning intensity of migration from the different data sets and then generated the estimation of migration transition rates based on a standard interval.

---

[4] The Cinderella narrative is a global story with Greek and Chinese origins and versions in many different cultures (*Wikipedia* 2019).

How should we improve the estimates of internal and international migration flows?

In section 3, methods for filling gaps in internal and international migration flow data sets were discussed. Innovative measures that make best use of available data collections have been proposed. However, official statistical agencies are slow to take up these methods and use them to provide continuous time series, despite, in some cases, having funded the projects. There are no magic solutions to this problem. It is necessary to engage in an open dialogue with official statisticians and other government departments (including local government) which use migration statistics. Useful venues are conferences attended by official statisticians, other government departments meetings and academics. These conversations are best held outside of formal sessions over coffee or meals. Making solving the problem a shared task is better than an excuse for scoring points.

How can we use knowledge of internal and international migration in policy making?

Most countries and international organizations make use of population projections in formulating medium and long-term plans. Assumptions about internal and international migration are essential ingredients but these are rarely linked to policies. Internal migration assumptions usually involve continuation of inter-regional migration rates for a representative recent period. *Poot et al.* (2016) suggests there is scope to embody gravity models of migration into forecasts, although there would be a challenge in forecasting push factors at the origin and pull factors at the destination, as well as a need for better measures of impedance as *Shen* (2016) has pointed out. The uncertainty around future population numbers can be handled through specifying variant projections based on judgement or through generating probabilistic projections which provide confidence intervals. Scenarios that specifically link international migration to policies (*Rees et al.* 2012; *Abel et al.* 2016; *Cafaro/Dérer* 2019; *Marois et al.* 2020) provide forecasts of future populations conditional on the acceptance of specific policies. Such scenarios need a body of evidence to show that policies will result in a set direction of change in migration. Time series extrapolation, no matter how sophisticated, has a poor track record in forecasting international migration (*Bijak et al.* 2019).

What is the case for a world migration survey?

Willekens et al. (2016: 898-899) have proposed a World Migration Survey (WMS), echoing a proposal by the International Union for the Scientific Study of Population:

> "The WMS should include respondents and partnering institutes in countries and of origin, transit, and destination and, ideally, should be longitudinal."

Such a global migration survey might by implemented through the International Organization for Migration, the UN Population Division and the Organization for Economic Co-operation and Development. The expertise should be harnessed of international population research centres such as the Max Planck Institute for Demographic Research (Rostock), the Wittgenstein Centre for Demography and Global Human Capital (Vienna) and the Economic and Social Research Council Centre for Population Change (Southampton) and global polling organizations such as the Pew Centre or Gallup organization. The WMS will no doubt be besieged by researchers wanting to load their questions into the survey. Procedures should be implemented to require question or design proposals to draw up realistic cost/benefit analyses. Key questions that we suggest should be included are those that would unlock the value of current information so that existing but unsatisfactory data could be converted to fixed interval migration data. *Willekens et al.* (2016: 897) suggest estimation of "flow data" as a target concept without elucidating how this is defined and measured. We hope that the measurement framework set out in section 2 of the paper will help clarify what measures should be included. *Willekens et al.* (2016) recommend "longitudinal data", without specifying exactly what that means: alternatives include regular cross-section surveys, surveys in which individuals are linked over time or full population registers.

What about climate change and migration?

Future research on migration cannot ignore the climate crisis, which threatens the existence of whole island nations and of a share of the world's population that lives in low-lying land vulnerable to sea-level rise. A one degree rise in global mean temperature has already increased storm frequency and the magnitude of droughts and increased flows of persons displaced by climate change. A requirement of future funding of migration research should be to take the climate emergency into as full an account as possible (*Muttarak/Jiang* 2015).

5.3   Concluding remarks

This paper began with a brief survey of the achievements of Anglo-German geographer, Ernst Ravenstein. His papers have influenced much research into migration in the century and a quarter since publication. We have reviewed how scholars now investigate the phenomena he wrote about. He was an international scholar and investigated migration both nationally and internationally. We hope he would be pleased with what the field of migration research has achieved and would not be too daunted by the challenges of understanding migration in the modern world.

References

Abel, Guy J. 2010: Estimation of International Migration Flow Tables in Europe. In: Journal of the Royal Statistical Society A 173: 797-825 [doi: 10.1111/j.1467-985X.2009.00636.x].

Abel, Guy J. 2013: Estimating Global Migration Flow Tables using Place of Birth Data. In: Demographic Research 28,18: 505-546 [doi: 10.4054/DemRes.2013.28.18].

Abel, Guy J.; Sander, Nikola 2014: Quantifying Global International Migration Flows. In: Science 343: 1520-1522 [doi: 10.1126/science.1248676].

Abel, Guy J. et al. 2016. Meeting the Sustainable Development Goals leads to lower world population growth. In: PNAS 113,50: 14294-14299 [doi: 10.1073/pnas.1611386113].

Abel, Guy J. 2018: Estimates of Global Migration Flows by Gender between 1960 and 2015. In: International Migration Review 52,3: 809-852 [doi: 10.1111/imre.12327].

Abel, Guy J.; Cohen, Joel 2019: Bilateral international migration flow estimates for 200 countries. In: Scientific Data 6,82 [doi: 10.1038/s41597-019-0089-3].

Acxiom 2019: Power Exceptional Marketing Experiences with a Unified Data & Technology Foundation [https://www.acxiom.co.uk/, 06.05.2020].

Azose, Jonathan J.; Raftery, Adrian E. 2018a: Estimation of Emigration, Return Migration, and Transit Migration between All Pairs of Countries. In: Proceedings of the National Academy of Sciences 116,1: 116-122 [doi: 10.1073/pnas.1722334116].

Azose, Jonathan J.; Raftery, Adrian E. 2018b: Estimation of Emigration, Return Migration, and Transit Migration between All Pairs of Countries: Supporting Information. In: Proceedings of the National Academy of Sciences [www.pnas.org/cgi/doi/10.1073/pnas.1722334115, 06.05.2020].

Beine, Michel; Docquier, Frédéric; Özden, Çağlar 2011: Diasporas. In: Journal of Development Economics 95,1: 30-41 [doi: 10.1016/j.jdeveco.2009.11.004].

Bell, Martin et al. 2002: Cross-National Comparison of Internal Migration: Issues and Measures. In: Journal of the Royal Statistical Society. Series A: Statistics in Society 165,3: 435-464.

Bell, Martin; Rees, Philip 2006. Comparing migration in Britain and Australia: harmonisation through use of age-time plans. In: Environment and Planning A 38,5: 959-988 [doi: 10.1068/a35245].

Bell, Martin; Muhidin, Salut 2009: Cross-National Comparisons of Internal Migration. In: Human Development Research Paper 2009/30: 1-62 [hdr.undp.org/en/content/cross-national-comparisons-internal-migration, 29.04.2020].

Bell, Martin; Muhidin, Salut 2011: Comparing Internal Migration Between Countries Using Courgeau's k. In: *Stillwell, John; Clarke, Martin* (Eds.): Population Dynamics and Projection Methods. London: Springer: 141-164 [doi: 10.1007/978-90-481-8930-4_7].

Bell, Martin et al. 2015a: Internal Migration Data Around the World: Assessing Contemporary Practice. In: Population, Space and Place 21,1: 1-17 [doi: 10.1002/psp.1848].

Bell, Martin et al. 2015b: Internal Migration and Development: Comparing Migration Intensities Around the World. In: Population and Development Review 41,1: 33-58 [doi: 10.1111/j.1728-4457.2015.00025.x].

Bell, Martin et al. 2018: Global Trends in Internal Migration. In: *Champion, Anthony; Cooke, Thomas; Shuttleworth, Ian* (Eds.): Internal Migration in the Developed World. Are We Becoming Less Mobile? Oxford: Routledge: 76-98 [doi: 10.4324/9781315589282].

Bell, Martin et al. (Eds.) 2020: Internal Migration in the Countries of Asia: A Cross-National Comparison. Springer, in Press.

Bernard, Aude; Bell, Martin 2012: A Comparison of Internal Migration Age Profile Smoothing Methods. In: Working Paper 2012/01. Queensland Centre for Population Research, University of Queensland [https://www.yumpu.com/en/document/view/37450540/a-comparison-of-internal-migration-age-profile-smoothing-methods, 29.04.2020].

Bernard, Aude; Bell, Martin; Charles-Edwards, Elin 2014a: Improved Measures for the Cross-National Comparison of Age Profiles of Internal Migration. In: Population Studies 68,2: 179-195 [doi: 10.1080/00324728.2014.890243].

Bernard, Aude; Bell, Martin; Charles-Edwards, Elin 2014b: Life-Course Transitions and the Age Profile of Internal Migration. In: Population and Development Review 40,2: 213-239 [doi: 10.1111/j.1728-4457.2014.00671.x].

Bernard, Aude; Bell, Martin 2015: Smoothing Internal Migration Age Profiles for Comparative Research. In: Demographic Research 32,33: 915-948 [doi: 10.4054/DemRes.2015.32.33]

Bernard, Aude; Martin Bell; Elin Charles-Edwards 2016: Internal migration age patterns and the transition to adulthood: Australia and Great Britain compared. In: Journal of Population Research 33: 123-146 [doi: 10.1007/s12546-016-9157-0].

Bernard, Aude 2017: Levels and Patterns of Internal Migration in Europe: A Cohort Perspective. In: Population Studies 71,3: 293-311 [doi. 10.1080/00324728.2017.1360932].

Bernard, Aude et al. 2017: Comparing Internal Migration across the Countries of Latin America: A Multidimensional Approach. In: PLoS ONE 12,3 [doi. 10.1371/journal.pone.0173895].

Bernard, Aude; Bell, Martin 2018: Educational Selectivity of Internal Migrants: A Global Assessment. In: Demographic Research 39,29: 835-854 [doi: 10.4054/DemRes.2018.39.29].

Bernard, Aude; Pelikh, Alina 2019: Distinguishing Tempo and Ageing Effects in Migration. In: Demographic Research 40,44: 1291-1322 [doi: 10.4054/demres.2019.40.44].

Bernard, Aude; Bell, Martin; Zhu, Yu 2019: Migration in China: A Cohort Approach to Understanding Past and Future Trends. In: Population, Space and Place 25,6: e2234 [doi: 10.1002/psp.2234].

Bertoli, Simone; Fernández-Huertas Moraga, Jesus 2015: The Size of the Cliff at the Border. In: Regional Science and Urban Economics 51: 1-6 [doi: 10.1016/j.regsciurbeco.2014.12.002].

Bijak, Jakub et al. 2019: Assessing time series models for forecasting international migration: Lessons from the United Kingdom. In: Journal of Forecasting 38,5: 470-487 [DOI: 10.1002/for.2576].

Bishop, Yvonne; Fienberg, Stephen; Holland, Paul 1975: Discrete Multivariate Analysis: Theory and Practice. Cambridge: MIT Press.

Bongaarts, John; Bulatao, Rodolfo 1999: Completing the Demographic Transition. Population and Development Review 25,3: 515-529 [https://www.jstor.org/stable/172345].

Cafaro, Philip; Dérer, Patricia 2019. Policy-based Population Projections for the European Union: A Complementary Approach. In: Comparative Population Studies 44: 171-200 [doi: 10.12765/CPoS-2019-14en].

CeLSIUS 2019: The Office for National Statistics Longitudinal Study (England and Wales). Centre for Longitudinal Study Information and User Support, Institute of Epidemiology and Health Research, University College London, UK [https://www.ucl.ac.uk/epidemiology-health-care/research/epidemiology-and-public-health/research/health-and-social-surveys-research-group/studies-10, 06.05.2020].

Champion, Tony; Shuttleworth, Ian 2017a: Is Longer-Distance Migration Slowing? An Analysis of the Annual Record for England and Wales since the 1970s. In: Population, Space and Place 23,3: 1-14 [doi: 10.1002/psp.2024].

Champion, Tony; Shuttleworth, Ian 2017b: Are People Changing Address Less? An Analysis of Migration within England and Wales, 1971-2011, by Distance of Move. In: Population, Space and Place 23;3: 1-13 [doi: 10.1002/psp.2026].

Champion, Tony; Cooke, Thomas; Shuttleworth, Ian (Eds.) 2018: Internal Migration in the Developed World: Are We Becoming Less Mobile? London: Routledge [doi: 10.4324/9781315589282].

Charles-Edwards et al. 2017a: Internal Migration in the Countries of Asia, and Spatial Impacts. In: ADRI Working Paper Series, ADRI-WP-2017/001. Shanghai: Asian Demographic Research Institute.[http://www.asianmc.org/wp-content/uploads/2017/10/ADRI_Working_Paper_Internal_Migration_in_Asia_Final.pdf].

Charles-Edwards, Elin; Muhidin, Salut; Bell, Martin; Zhu, Yu 2017b: Migration in Asia. In: *White, Michael J.* (Ed.): International Handbook of Migration and Population Distribution. Dordrecht: Springer: 268-284 [doi: 10.1007/978-94-017-7282-2_13].

Charles-Edwards, Elin; Bell, Martin; Bernard, Aude; Zhu, Yu 2019: Internal migration in the countries of Asia: levels, ages and spatial impacts. In: Asian Population Studies 15,2: 150-171 [doi: 10.1080/17441730.2019.1619256].

Congdon, Peter 1993: Approaches to Modelling Overdispersion in the Analysis of Migration. In: *Environment and Planning* A 25: 1481-510.

Cooke, Thomas J. 2011: It is not just the economy: declining migration and the rise of secular rootedness. In: Population, Space and Place 17,3: 193-203 [doi: 10.1002/psp.670].

Cooke, Thomas J. 2013: Internal migration in decline. In: The Professional Geographer 65,4: 664-675 [doi: 10.1080/00330124.2012.724343].

Coulter, Rory; van Ham, Marten; Feijten, Peteke 2011: A longitudinal analysis of moving desires, expectations and actual moving behaviour. In: Environment and Planning A 43,11: 2742-2760 [doi: 10.1068/a44105].

Courgeau, Daniel 1970: Les Champs Migratoires en France. Travaux et Documents, Cahier no58, Institut National d'Etudes Demographiques. Paris: Presses Universitaires de France.

Courgeau, Daniel 1973a: Migrants et migrations. In: Population (French Edition) 28,1: 95-129 [doi: 10.2307/1530972].

Courgeau, Daniel 1973b: Migrations et Découpages du Territoire. In: Population 28,3: 511-537 [doi: 10.2307/1530704].

Courgeau, Daniel 1992: Migration Nette et Densité: la France de 1954 à 1990. In: Population 47,2: 462-467 [doi: 10.2307/1533918].

Courgeau, Daniel; Muhidin, Salut; Bell, Martin 2012: Estimating Changes of Residence for Cross-National Comparison. In: Population (English Edition) 67,4: 631-652 [doi: 10.3917/pope.1204.0631].

Crymble, Adam 2019: Introduction to Gravity Models of Migration and Trade. The Programming Historian [https://programminghistorian.org/en/lessons/gravity-model, 29.04.2020].

Darlington-Pollock, Frances; Lomax, Nik; Norman, Paul 2019: Ethnic Internal Migration: The Importance of Age and Migrant Status. In: The Geographical Journal 185,1: 68-81 [doi: 10.1111/geoj.12286].

Deming, William Edwards; Stephan, Frederick 1940: On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. In: The Annals of Mathematical Statistics 11,4: 427-444 [http://www.jstor.org/stable/2235722].

Dennett, Adam; Wilson, Alan 2013: A Multilevel Spatial Interaction Modelling Framework for Estimating Interregional Migration in Europe. In: Environment and Planning A 45,6: 1491-1507 [doi: 10.1068/a45398].

Dennett, Adam 2016: Estimating an Annual Time Series of Global Migration: An Alternative Methodology for Using Migrant Stock Data. In: *Wilson, Alan* (Ed.): Global Dynamics: Approaches from Complexity Science. Chichester: John Wiley: 125-141 [doi: 10.1002/9781118937464.ch7].

Dennett, Adam; Wilson, Alan 2016: Estimating inter-regional migration in Europe. In: *Wilson, Alan* (Ed.): Global Dynamics: Approaches from Complexity Science. Chichester: John Wiley: 97-124 [doi: 10.1002/9781118937464.ch6].

Dorigo, Guido; Tobler, Waldo 1983: Push and Pull Migration Laws. In: Annals of the Association of American Geographers 73,1: 1-17 [doi: 10.1111/j.1467-8306.1983.tb01392.x].

Dorling, Daniel; Thomas, Bethan 2004: People and Places: A 2001 Census Atlas of the UK. Bristol: The Policy Press.

Duke-Williams, Oliver; Vassilis Routsis; John Stillwell 2018: Census interaction data and their means of access. In: *Stillwell, John* (Ed.): The Routledge Handbook of Census Resources, Methods and Applications. Routledge: London: 110-125.

Dyrting, Sigurd 2018: A Framework for Translating Between One-Year and Five-Year Migration Probabilities [doi: 10.13140/RG.2.2.24261.91362].

European Commission 2019: Measuring Labour Mobility and Migration from Big Data: Exploring the Potential of Social-Media Data for Measuring EU Mobility Flows and Stocks of EU Movers. Luxembourg: Publications Office of the European Union [doi: 10.2767/474282].

Fielding, Antony 1989: Migration and Urbanisation in Western Europe since 1950. In: The Geographical Journal 155,1: 60-69 [doi: 10.2307/635381].

Flowerdew, Robin; Aitkin, Murray 1982: A Method of Fitting the Gravity Model Based on the Poisson Distribution. In: Journal of Regional Science 22,2: 191-202 [doi: 10.1111/j.1467-9787.1982.tb00744.x].

Flowerdew, Robin; Lovett, Andrew 1988: Fitting Constrained Poisson Regression Models to Interurban Migration Flows. In: Geographical Analysis 20,4: 297-307 [doi: 10.1111/j.1538-4632.1988.tb00184.x].

Flowerdew, Robin 2010: Modelling Migration with Poisson Regression. In: *Stillwell, John; Duke-Williams, Oliver; Dennett, Adam* (Eds.): Technologies for Migration and Commuting Analysis: Spatial Interaction Data Applications. IGI, Global: Hershey: 261-279 [doi: 10.4018/978-1-61520-755-8.ch014].

Fotheringham, Stewart et al. 2004: The development of a migration model for England and Wales: overview and modelling out-migration. In: Environment and Planning A 36,9: 1633-1672 [doi: 10.1068/a36136].

Frey, William 2015: Diversity Explosion: How New Racial Demographics Are Remaking America. Washington, DC: Brookings Institution Press.

Friedlander, Dov; Roshier, R. 1966: A Study of Internal Migration in England and Wales: Part I. In: Population Studies 19,3: 239-279 [doi: 10.1080/00324728.1966.10406016].

Geyer, Hermanus S.; Thomas, Kontuly 1993: A Theoretical Foundation for the Concept of Differential Urbanization. In: International Regional Science Review 15,2: 157-177 [doi: 10.1177/016001769301500202].

Greenwood, Michael 2019: The migration legacy of E. G. Ravenstein. In: Migration Studies 7,2: 269-278 [doi: 10.1093/migration/mny043].

Grigg, David B. 1977: Ravenstein and the "Laws of Migration". In: Journal of Historical Geography 3,1: 41-54 [doi: 10.1016/0305-7488(77)90143-8].

Hägerstrand, Torsten 1957: Migration and Area. In: *Hannerberg, David; Hägerstrand, Torsten; Odeving, Bruno* (Eds.): Migration in Sweden: A Symposium. Lund Studies in Geography. Series B, Human Geography N. 13. Lund, Sweden, C.W.K. Gleerup.

HFD 2019: The Human Fertility Database. Max Planck Institute for Demographic Research, Rostock, Germany and Vienna Institute of Demography, Vienna, Austria [https://www.humanfertility.org/cgi-bin/main.php, 29.04.2020].

HMD 2019: The Human Mortality Database. Department of Demography, University of California, Berkeley, USA and Max Planck Institute for Demographic Research, Rostock, Germany [https://mortality.org/, 29.04.2020].

IMAGE Project 2020a: The Image Project: Comparing Internal Migration Around the Globe [https://imageproject.com.au/framework/, 06.05.2020].

IMAGE Project 2020b: IMAGE_Studio_rc. [https://github.com/IMAGE-Project, 06.05.2020].

IOM 2019: Migration Data: The Bigger Picture. International Organization for Migration [https://migrationdataportal.org/?i=stock_abs_&t=2019, 29.04.2020].

IPUMS 2019: Harmonized International Census Data for Social Science and Health Research: Integrated Public Use Microdata Series. Minneapolis, MN, USA: Minnesota Population Center. [https://international.ipums.org/international/, 06.05.2020].

Jacobs, Jane 1970: The Economy of Cities. London: Cape.

Kim, Keuntae; Cohen, Joel E. 2010: Determinants of International Migration Flows to and from Industrialized Countries: A Panel Data Approach Beyond Gravity. In: International Migration Review 44,4: 899-932 [doi: 10.1111/j.1747-7379.2010.00830.x].

Kitsul, Pavel; Philipov, Dimiter 1981: The One-Year-Five-Year Migration Problem. In: *Rogers, Andrei* (Ed.): Advances in Multiregional Demography. Research Report RR-81-006: 1-33. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Kruithof, J. 1937: Telefoonverkeersrekening (Calculation of Telephone Traffic). In: De Ingenieur 52,8: E15-E25 [https://wwwhome.ewi.utwente.nl/~ptdeboer/misc/kruithof-1937-translation.html, 30.04.2020].

Kupiszewski, Marek; Kupiszewska, Dorota 2011: MULTIPOLES: A Revised Multiregional Model for Improved Capture of International Migration. In: *Stillwell, John; Clarke, Martin* (Eds.): Population Dynamics and Projection Methods. Dordrecht: Springer: 41-60 [doi: 10.1007/978-90-481-8930-4_3].

Lai, Shengjie et al. 2019: Exploring the Use of Mobile Phone Data for National Migration Statistics. In: Palgrave Communications 5,34: 1-10 [doi: 10.1057/s41599-019-0242-9].

Lomax, Nikolas 2013a: Internal and cross-border migration in the United Kingdom: harmonising, estimating and analysing a decade of flow data. PhD Thesis. The University of Leeds, School of Geography.

Lomax, Nik et al. 2013b: Subnational migration in the United Kingdom: producing a consistent time series using a combination of available data and estimates. In: Journal of Population Research 30: 265-288 [doi: 10.1007/s12546-013-9115-z].

Lomax, Nik; Norman, Paul 2016: Estimating Population Attribute Values in a Table: "Get Me Started in" Iterative Proportional Fitting. In: The Professional Geographer 68,3: 451-461 [doi: 10.1080/00330124.2015.1099449].

Long John; Boertlein, Celia 1990: Comparing Migration Measures Having Different Intervals. In: Perspectives in Migration Analysis. Current Population Reports, Special Studies Series P-23,166: 1-12. Washington, DC: US Department of Commerce, Bureau of the Census.

Lutz, Wolfgang; Butz, William; KC, Samir (Ed.) 2014a: World Population and Human Capital in the Twenty-First Century. Oxford: Oxford University Press.

Manley, Ed; Dennett, Adam 2018: New Forms of Data for Understanding Urban Activity in Developing Countries. In: Applied Spatial Analysis and Policy 12: 45-70 [doi: 10.1007/s12061-018-9264-8].

Marois, Guillaume; Sabourin, Patrick; Bélanger, Alain 2019a: How reducing differentials in education and labour force participation could lessen workforce decline in the EU28. In: Demographic Research 41,6: 125-160 [doi: 10.4054/DemRes.2019.41.6].

Marois, Guillaume; Sabourin, Patrick; Bélanger, Alain 2019b: Forecasting human capital of EU member countries accounting for sociocultural determinants. In: Journal of Demographic Economics 85,3: 231-269 [doi: 10.1017/dem.2019.4].

Marois, Guillaume; Bélanger, Alain; Lutz, Wolfgang 2020: Does Europe need immigration for demographic reasons? In: PNAS 117,14: 7690-7695 [doi: 10.1073/pnas.1918988117].

Mercer, Arnold; Lau, Arnold; Kennedy, Courtney 2018: For Weighting Online Opt-In Samples, What Matters Most? Washington DC: Pew Research Center.

Mulder, Clara H. 2018: Putting family centre stage: Ties to Non-resident Family, Internal migration, and Immobility. In: Demographic Research 39,43: 1151-1180 [doi: 10.4054/DemRes.2018.39.43].

Muttarak, Raya; Jiang, Leiwen (Eds.) 2015: Special Issue on Demographic Differential Vulnerability to Climate-Related Disasters. In: Vienna Yearbook of Population Research 2015. Vienna: Austrian Academy of Sciences [doi: 10.1553/populationyearbook2015s1].

Nam, Charles; Serow, William; Sly, David (Eds.) 1990: International Handbook on Internal Migration. New York: Greenwood.

Newbold, K. Bruce 2005: Spatial Scale, Return and Onward Migration, and the Long-Boertlein Index of Repeat Migration. In: Papers in Regional Science 84,2: 281-290 [doi: 10.1111/j.1435-5957.2005.00018.x].

Nowok, Beata; Willekens, Frans 2011: A Probabilistic Framework for Harmonisation of Migration Statistics. In: Population, Space and Place 17,5: 521-533 [doi: 10.1002/psp.624].

ODPM 2002: Development of a Migration Model. London: Office of the Deputy Prime Minister [https://www.researchgate.net/publication/259308761_Development_of_a_Migration_Model_Analytical_and_Practical_Enhancements, 06.05.2020].

ONS 2015: How has Life Expectancy Changed over Time? Office for National Statistics [https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/articles/howhaslifeexpectancychangedovertime/2015-09-09, 30.04.2020].

ONS 2019: International Migration — Terms, Definitions and Frequently Asked Questions [https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/methodologies/longterminternationalmigrationfrequentlyaskedquestionsandbackgroundnotes#migration-terms-and-definitions, 30.04.2020].

Openshaw, Stan; Connolly, C. 1977: Empirically Derived Deterrence Functions for Maximum Performance Spatial Interaction Models. In: Environment and Planning A 9,9: 1067-1079 [doi: 10.1068/a091067].

Openshaw, Stan 1983: The Modifiable Areal Unit Problem. Concepts and Techniques in Modern Geography (CATMOG) No.38. Norwich: Geo Books.

Pradier, Pierre 2018: The Chinese Hukou System Explained – What is "Hukou" and How it Works. Shanghai: New Horizons Global Partners Ltd. [https://nhglobalpartners.com/author/pierre/, 06.05.2020].

Poot, Jacques et al. 2016: The gravity model of migration: the successful comeback of an ageing superstar in regional science. In: Investigaciones Regionales/Journal of Regional Research 36: 63-86 [https://core.ac.uk/download/pdf/74338682.pdf, 30.04.2020].

Poulain, Michel; Perrin, Nicolas; Singleton, Ann (Eds.) 2006: THESIM: Towards Harmonised European Statistics on International Migration. Louvain-La-Neuve: Presses Universitaires de Louvain.

Ramos, Raul 2016: Gravity Models: A Tool for Migration Analysis. In: IZA World of Labor 239: 1-10 [doi: 10.15185/izawol.239].

Ravenstein, Ernst 1876: Census of the British Isles, 1871. Birthplaces and Migration. In: The Geographical Magazine July a: 173-177; August b: 202-206; September c: 229-233.

Ravenstein, Ernst 1885: The Laws of Migration. In: Journal of the Statistical Society of London 48,2: 167-235 [doi: 10.2307/2979181].

Ravenstein, Ernst 1889: The Laws of Migration. In: Journal of the Royal Statistical Society 52,2: 241-305 [doi: 10.2307/2979333].

Raymer, James 2008: Obtaining an Overall Picture of Population Movement in the European Union. In: *Raymer, James; Willekens, Frans* (Eds.): International Migration in Europe: Data, Models and Estimates. Chichester: Wiley: 209-234 [doi: 10.1002/9780470985557.ch10].

Raymer, James; Willekens, Frans (Ed.) 2008: International Migration in Europe: Data, Models and Estimates. Chichester: Wiley. [doi: 10.1002/9780470985557].

Raymer, James; Guy, Abel 2009: The MIMOSA Model for Estimating International Migration Flows in the European Union [https://iussp2009.princeton.edu/papers/92501, 30.04.2020].

Raymer, James; De Beer, Joop; Van der Erf, Rob 2011: Putting the Pieces of the Puzzle Together: Age and Sex-Specific Estimates of Migration Amongst Countries in the EU/EFTA, 2002-2007. In: European Journal of Population 27,2: 185-215 [doi: 10.1007/s10680-011-9230-5].

Raymer, James et al. 2013: Integrated Modeling of European Migration. In: Journal of the American Statistical Association 108,503: 801-819 [doi: 10.1080/01621459.2013.789435].

Raymer, James et al. 2019a: Estimating a Consistent and Detailed Time Series of Immigration and Emigration for Sub-State Regions of Australia. In: Applied Spatial Analysis and Policy [doi: 10.1007/s12061-019-09310-w].

Raymer, James; Guan, Qing; Ha, Jasmine Trang 2019b: Overcoming Data Limitations to Obtain Migration Flows for ASEAN Countries. In: Asian and Pacific Migration Journal 28,4: 385-414 [doi: 10.1177/0117196819892344].

Rees, Philip 1977: The Measurement of Migration from Census and Other Sources. In: Environment and Planning A 9,3: 247-272 [doi: 10.1068/a090247].

Rees, Philip; Wilson, Alan 1977: Spatial Population Analysis. London: Edward Arnold.

Rees, Philip 1979: Migration and Settlement: 1. United Kingdom. In: Research Report RR-79-3. Laxenburg, Austria: International Institute for Applied Systems Analysis. [http://pure.iiasa.ac.at/id/eprint/1034/1/RR-79-003.pdf, 30.04.2020].

Rees, Philip 1985: Does It Really Matter Which Migration Data You Use in a Population Model? In: *White, Paul; Van der Knaap, Gert* (Eds.): Contemporary Studies of Migration. Norwich: Geobooks: 55-77.

Rees, Philip; Willekens, Frans 1986: Data and accounts. In: *Rogers, Andrei, Willekens, Frans* (Eds.): Migration and Settlement: A Multiregional Comparative Study. Dordrecht: Reidel: 19-38.

Rees, Philip; Durham, Helen; Kupiszewski, Marek 1996: Internal Migration and Regional Population Dynamics in Europe: United Kingdom Case Study. In: Working Paper 96/20. United Kingdom: School of Geography, University of Leeds [http://www.geog.leeds.ac.uk/fileadmin/documents/research/csap/working_papers_new/1996-20.pdf, 30.04.2020].

Rees, Philip; Kupiszewski, Marek 1999: Internal Migration and Regional Population Dynamics in Europe: A Synthesis. In: Population Studies No.32. Strasbourg: Council of Europe Publishing.

Rees, Philip et al. 2010: DEMIFER: Demographic and Migratory Flows Affecting European Regions and Cities. Applied Research Project 2013/1/3. Annex D6 Report on Scenarios and a Database of Scenario Drivers, Annex D7 Regional Population Dynamics: A Report Assessing the Effects of Demographic Developments on Regional Competitiveness and Cohesion. Luxembourg: The ESPON 2013 Programme. [https://www.espon.eu/demifer, 06.05.2020].

Rees, Philip et al. 2011: A local analysis of ethnic group population trends and projections for the UK. In: Journal of Population Research 28,2-3: 149-184 [doi: 10.1007/s12546-011-9047-4]

Rees, Philip et al. 2012: European Regional Populations: Current Trends, Future Pathways, and Policy Options. In: European Journal of Population 28,4: 385-416 [doi: 10.1007/s10680-012-9268-z].

Rees, Philip; Wohland, Pia; Norman, Paul 2013: The Demographic Drivers of Future Ethnic Group Populations for UK Local Areas 2001-2051. In: The Geographical Journal 179,1: 44-60 [doi: 10.1111/j.1475-4959.2012.00471.x].

Rees Philip et al. 2017a: The Impact of Internal Migration on Population Redistribution: An International Comparison. In: Population, Space and Place 23,6 [doi: 10.1002/psp.2036].

Rees, Philip et al. 2017b. Population Projections by Ethnicity: Challenges and a Solution for the United Kingdom. In: *Swanson, David A.* (Ed.): The Frontiers of Applied Demography. Switzerland: Springer: 383-408 [doi: 10.1007/978-3-319-43329-5_18].

Rees, Philip 2018: Education and Demography: A Review of World Population and Human Capital in the 21st Century (A Demographic Debate Contribution). In: Vienna Yearbook of Population Research 16: 1-17 [doi: 10.1553/populationyearbook2018s037].

Rogers, Andrei; Raquillet, Richard; Castro, Luis 1978: Model migration schedules and their applications. In: Environment and Planning A 10,5: 475-502 [doi: 10.1068/a100475].

Rogers, Andrei; Castro, Luis 1981a: Model Migration Schedules. In: Research Report 81-30. Laxenburg, Austria: International Institute for Applied Systems Analysis [https://www.researchgate.net/publication/301286423_Model_MIgration_Schedules, 30.04.2020].

Rogers, Andrei; Castro, Luis 1981b: Age Patterns of Migration: Cause-Specific Profiles. In: Rogers, Andrei (Ed.): Advances in Multiregional Demography. In: Research Report RR-81-6: 125-159. Laxenburg, Austria: International Institute for Applied Systems Analysis. [http://pure.iiasa.ac.at/id/eprint/1556/, 30.04.2020].

Rogers, Andrei; Willekens, Frans (Ed.) 1986: Migration and Settlement: A Multiregional Comparative Study. Dordrecht: Reidel.

Rogers, Andrei 1990: Requiem for the Net Migrant. In: Geographical Analysis 22,4: 283-300 [doi: 10.1111/j.1538-4632.1990.tb00212.x].

Rogers, Andrei; James Raymer; Newbold, K. Bruce 2003: Reconciling and Translating Migration Data Collected over time Intervals of Differing Widths. In: The Annals of Regional Science 37,4: 581-601 [doi: 10.1007/s00168-003-0128-y].

Rogers, Andrei (Ed.) 1992: Elderly Migration and Population Redistribution. London: Belhaven.

Rogers, Andrei; Little, Jani; Raymer, James 2010: The Indirect Estimation of Migration: Methods for Dealing with Irregular, Inadequate, and Missing Data. Dordrecht: Springer [doi: 10.1007/978-90-481-8915-1].

Rogerson, Peter A. 1990: Migration Analysis Using Data with Time Intervals of Differing Widths. In: Papers of the Regional Science Association 68,1: 97-106 [doi: 10.1007/BF01933910].

Rowe, Francisco et al. 2019: Impact of Internal Migration on Population Redistribution in Europe: Urbanisation, Counterurbanisation or Spatial Equilibrium? In: Comparative Population Studies 44: 201-234 [doi: 10.12765/CPoS-2019-18en].

Sander, Nikola; Abel, Guy; Riosmena, Fernando 2014a: The Future of International Migration. In: Lutz, Wolfgang, Butz, William; KC, Samir (Eds): World Population and Human Capital in the Twenty-First Century. Oxford: Oxford University Press: 333-396.

Sander, Nikola; Abel, Guy; Bauer, Ramon; Schmidt, Johannes 2014b: Visualising migration flow data with circular plots. Working Paper No.2. Vienna: Vienna Institute of Demography [http://download.gsb.bund.de/BIB/global_flow/VID%20WP%20Visualising%20Migration%20Flow%20Data%20with%20Circular%20Plots.pdf, 04.05.2020].

Sander, Nikola; Stillwell, John: Lomax, Nikolas 2018: Circular Migration Plots. In: Stillwell, John (Ed.): The Routledge Handbook of Census Resources, Methods and Applications: Unlocking the 2011 Census. London: Routledge: 203-209.

Schmertmann, Carl 1999: Estimating Multistate Transition Hazards from Last-Move Data. In: Journal American Statistical Association 94,445: 53-63.

Sen, Ashish; Smith, Tony 1995: Gravity Models of Spatial Interaction Behavior. Springer. [doi: 10.1007/978-3-642-79880-1].

Shen, Jianfa 2016: Error analysis of Regional Migration Modeling. In: Annals of the Association of American Geographers 106,6: 1253-1267 [doi: 10.1080/24694452.2016.1197767].

Simpson, Ludi 2002: Geography Conversion Tables: A Framework for Conversion of Data between Geographical Units. In: International Journal of Population Geography 8,1: 69-82 [doi: 10.1002/ijpg.235].

Solzhenitsyn, Aleksandr; 1991: The Gulag Archipelago, 1918-1956: An Experiment in Literary Investigation. English Translation by Ericson, Edward. London: Harvill.

Statistics Finland 2019: Tilastokeskus. Statistics Finland PxWeb Databases [http://pxnet2.stat.fi/PXWeb/pxweb/en/StatFin/, 06.05.2020].

Stillwell, John 1978: Interzonal Migration: Some Historical Tests of Spatial-Interaction Models. In: Environment and Planning A 10,10: 1187-1200 [doi: 10.1068/a101187].

Stillwell, John 1979: Migration in England and Wales: A Study of Inter-Area Patterns, Accounts, Models and Projections. PhD Thesis. Leeds, UK: University of Leeds.

Stillwell, Johnet al. 2000: Net migration and migration effectiveness: a comparison between Australia and the United Kingdom, 1976-96. Part 1: total migration patterns. In: Journal of Population Research 17,1: 17-38 [doi: 10.1007/BF03029446].

Stillwell, John et al. 2001: Net migration and migration effectiveness: a comparison between Australia and the United Kingdom, 1976-96. Part 2: Age-related migration patterns. In: Journal of Population Research 18,1: 19-38 [doi: 10.1007/BF03031953].

Stillwell, John et al. 2014: The IMAGE Studio: A tool for internal migration analysis and modelling. In: Applied Spatial Analysis and Policy 7,1: 5-23 [doi: 10.1007/s12061-014-9104-4].

Stillwell, John; Thomas, Michael 2016: How far do internal migrants really move? Demonstrating a new method for the estimation of intra-zonal distance. In: Regional Studies, Regional Science 3,1: 28-47 [doi: 10.1080/21681376.2015.1109473].

Stillwell John et al. 2016: Internal Migration Around the World: Comparing Distance Travelled and its Frictional Effect. In: Environment and Planning A 48,8: 1657-1675 [doi: 10.1177/0308518X16643963].

Stillwell, John; Daras, Kostantinos; Bell, Martin 2018: Spatial Aggregation Methods for Investigating the MAUP Effects in Migration Analysis. In: Applied Spatial Analysis and Policy 11: 693-711 [doi: 10.1007/s12061-018-9274-6].

Thomas, Michael; Stillwell, John; Gould, Myles 2014: Exploring and Validating a Commercial Lifestyle Survey for its use in the Analysis of Population Migration. In: Applied Spatial Analysis and Policy 7: 71-95 [doi: 10.1007/s12061-013-9096-5].

UN 2015: World Population Prospects: The 2015 Revision. Key Findings and Advance Tables. New York: United Nations, Department of Economic and Social Affairs, Population Division [https://population.un.org/wpp/Publications/Files/Key_Findings_WPP_2015.pdf, 04.05.2020].

UN 2018: 2018 Revision of World Urbanization Prospects. New York: United Nations, Department of Economic and Social Affairs [https://esa.un.org/unpd/wup/, 06.05.2020].

UN 2019a: Big Data: UN Global Working Group. United Nations, Department of Economic and Social Statistics, Statistics Division [https://unstats.un.org/bigdata/, 06.05.2020].

UN 2019b: Demographic and Social Statistics: International Migration. United Nations, Department of Economic and Social Statistics, Statistics Division [https://unstats.un.org/unsd/demographic-social/sconcerns/migration/index.cshtml, 06.05.2020].

UN 2019c: Demographic and Social Statistics: Demographic Yearbook System. New York: United Nations, Department of Economic and Social Affairs, Statistics Division [https://unstats.un.org/unsd/demographic-social/products/dyb/index.cshtml, 06.05.2020].

US Census Bureau 2019: Census Flows Mapper. [https://flowsmapper.geo.census.gov/map.html, 06.05.2020].

Van der Erf, Rob: van der Gaag, Nicole 2007: An iterative procedure to revise available data in the double entry migration matrix for 2002, 2003 and 2004. Discussion Paper. The Hague: Netherlands Interdisciplinary Demographic Institute [http://mimosa.gedap.be/Documents/Erf_2007.pdf, 06.05.2020].

Van Imhoff, Evert; Keilman, Nico 1992: LIPRO 2.0: An Application of a Dynamic Demographic Projection Model to Household Structure in the Netherlands. NIDI-CBGS Publication 23. Amsterdam: Swets and Zeitlinger.

Wallace, Matthew; Kulu, Hill 2014: Migration and Health in England and Scotland: A Study of Migrant Selectivity and Salmon Bias. In: Population, Space and Place 20,8: 694-708 [doi: 10.1002/psp.1804].

White, Michael (Ed.) 2016: International Handbook of Migration and Population Distribution. Dordrecht: Springer [doi: 10.1007/978-94-017-7282-2].

Wikipedia 2019: Cinderella [https://en.wikipedia.org/wiki/Cinderella].

Willekens, Frans; Rogers, Andrei 1978: Spatial Population Analysis: Methods and Computer Programs. In: Research Report RR-78-18. Laxenburg, Austria: International Institute for Applied Systems Analysis [pure.iiasa.ac.at/id/eprint/825/1/RR-78-018.pdf, 04.05.2020].

Willekens, Frans 1983: Log-linear Modelling of Spatial Interaction. In: Papers of the Regional Science Association 52: 187-205 [doi: 10.1007/BF01944102].

Willekens, Frans et al. 2016: International migration under the Microscope. In: Science 352,6288: 897-899 [doi: 10.1126/science.aaf6545].

Wilson, Alan G. 1971: A family of spatial interaction models, and associated developments. In: Environment and Planning A 3,1: 1-32 [doi: 10.1068/a030001].

Wilson, Tom 2010: Model Migration Schedules Incorporating Student Migration Peaks. In: Demographic Research 23,8: 191-222 [doi: 10.4054/demres.2010.23.8].

Wilson, Tom 2020: Modelling Age Patterns of Internal Migration at the Highest Ages. Draft paper. Melbourne School of Population and Global Health, University of Melbourne [https://www.researchgate.net/requests/r69779130, 06.05.2020].

Yule, G Udny 1912: On the Methods of Measuring Association Between Two Attributes. In: Journal of the Royal Statistical Society 75,6: 579-652 [doi: 10.2307/2340126].

Zelinsky, Wilbur 1971: The Hypothesis of the Mobility Transition. In: Geographical Review 61,2: 219-249.

Zipf, George Kingsley 1946: The P1 P2/D Hypothesis: On the Intercity Movement of Persons. In: American Sociological Review 11,6: 677-686 [doi: 10.2307/2087063].

Prof. Dr. Philip Rees (✉), Dr. Nik Lomax. School of Geography, University of Leeds. Leeds, United Kingdom. E-mail: p.h.rees@leeds.ac.uk, N.M.Lomax@leeds.ac.uk
URL: https://environment.leeds.ac.uk/geography/staff/1094/professor-philip-rees
　　　https://environment.leeds.ac.uk/geography/staff/1064/dr-nik-lomax

Appendix: Summaries of Ravenstein's Findings by Later Authors

Tab. A1: Key points from *Ravenstein's* 1876 papers as interpreted by *Greenwood*

Year-#	Text	Pages
1876-1	Centres of industry and commerce tend to grow more rapidly than agricultural areas, with large towns growing more rapidly than surrounding rural areas. The observed differences in population growth patterns are due to migration.	p.271
1876-2	People migrate to towns from nearby rural areas, and migrants from more distant places fill the gaps left in the rural areas. Thus, Ravenstein sees migration occurring in stages, and he believes that with respect to growing towns the nearby potential labour supply is quickly depleted.	p.271
1876-3	Migration decreases with distance, an observation based on London and to a lesser extent on Glasgow, but other places as well.	p.271
1876-4	Migration is not determined by distance alone, but rather also is shaped by 'facility of access' and 'local circumstances', which at least in part apparently means the quality of roads and the transportation system. Thus, at least in some limited sense, the public sector may facilitate migration.	p.271
1876-5	Intervening opportunities tend to absorb migrants to any given urban destination.	p.271

Source: *Ravenstein* (1876), *Greenwood* (2019).

Tab. A2: The summary of Ravenstein's Laws by *Grigg* (1977)

Year-#	Text	Pages
1977-1	The majority of migrants go only a short distance.	p.44
1977-2	Migration proceeds step by step.	p.47
1977-3	Migrants proceeding long distances generally go by preference to one of the great centres of commerce and industry.	p.48
1977-4	Every migratory current has a counter-current.	p.48
1977-5	The natives of towns are less migratory than those of rural districts.	p.48
1977-6	Females are more migratory than males within the kingdom of their birth, but males move more frequently abroad.	p.49
1977-7	Most migrants are adults; families rarely migrated.	p.49
1977-8	Towns grow more by migration than natural increase.	p.50
1977-9	Migration increases as industries develop and the means of transport improves.	p.52
1977-10	The major direction of migration is from the rural areas to the towns.	p.52
1977-11	The main causes of migration are economic.	p.53

Source: *Grigg* (1977).

Tab. A3: Key points made about Ravenstein's work by *Dorigo* and *Tobler* (1983)

Year-Law#	Text Push and Pull Migration Laws, *Dorigo/Tobler* (1983), quoted in *Greenwood* (2019)	Pages
1983-1	'Counties having an extended boundary in proportion to their area, naturally offer greater facilities for an inflow … than others with a restricted boundary' (*Ravenstein* 1885: 175). Boundaries, or the geographic size of the spatial unit over which migration is measured, are clearly of great importance, but this observation was not noted by others.	p.5
1983-2	(Migration streams) 'sweep along with them many of the natives of the counties through which they pass (and) deposit, in their progress, many of the migrants, which have joined them at their origin' (*Ravenstein* 1885: 191). With some imagination, we might view this statement of Ravenstein as an early indication of the family and friends hypothesis regarding migration. However, Ravenstein does not indicate any thoughts about what forces underlie the influence of family and friends-forces such as information provided by earlier migrants, their social ties, or the miniature private social benefits due to the presence of prior migrants from their origin who may provide food, shelter, and information to the newly arrived migrants.	p.10
1983-3	'Migratory currents flow along certain well defined geographical channels' (*Ravenstein* 1889: 284). This point is related to that immediately above, but is broader (*Dorigo/Tobler*, p. 12). Features of the transportation system are likely to play a critical role in this regard. A good example is found in US historical migration from the South to the North. Railroad and highway systems, as well as the Mississippi river, generally ran more or less directly south to north with a chain of mountains (the Appalachians) separating the two south-to-north paths. Consequently, northern seaboard cities received migrants disproportionately from Florida, Georgia, the Carolinas, and Virginia, whereas Midwestern cities like Chicago, Detroit, Cleveland and St. Louis received migrants from Alabama, Mississippi, Louisiana, Texas, and Tennessee.	p.12

Source: *Dorigo* and *Tobler* (1983).