www.sciencemag.org/content/343/6178/1520/suppl/DC1

Supplementary Material for

Quantifying Global International Migration Flows

Guy J. Abel and Nikola Sander*

*Corresponding author. E-mail: nikola.sander@oeaw.ac.at

Published 28 March 2014, *Science* **343**, 1520 (2014)
DOI: 10.1126/science.1248676

This PDF file includes:

Materials and Methods

Tables S1 to S5

Full References List

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/content/343/6178/1520/suppl/DC1)

Databases S1 and S2 as separate Excel files

Materials and Methods

The methodology for estimating bilateral migration flows from migrant stock data while maintaining known net migration flows is outlined in this section. The methodology is illustrated using a set of simple hypothetical data in three steps. First, the link between bilateral migrant stock tables and bilateral migration flow tables is introduced. Second, the methodology to estimate flows from stocks, introduced in *(13)* is briefly reviewed. Third, an extension to this method is proposed to allow the net migration flows implied by successive population, birth, and death data to be maintained in the estimated flow tables.

Representing Bilateral Migrant Stock Data in Flow Tables

Bilateral migrant stocks data are commonly represented in square tables, as presented for different time periods in the top panel of Table S1. Rows represent a categorization of the population, such as place of birth[1]. The columns in bilateral migrant stock typically represent the place of residence. Values in non-diagonal cells represent the size of a migrant stock by its place of birth (or some other measure) in a given place of residence at a specified time. Values in diagonal cells represent the number native born. As these do not measure a form of mobility they are sometimes not shown.

When the diagonal cells in a bilateral migrant stock table are included, the column totals represent the total population in the region, so long as the rows represent a set of mutually exclusive categories, such as place of birth[2]. In the hypothetical place of birth stock data in Table S1 there are no births or deaths. This results in three important features when comparing

[1]Migrant stocks are occasionally categorized by national statistics institutes using alternative measures such as citizenship or ethnicity

[2]When rows represent some of other measure, such as citizenship, the column totals may no longer represent a total population, but a count over the number of citizens or nationals. This total can potentially be greater than the population when persons with dual citizenship or dual nationalities are counted twice.

the stock tables. First, the row totals in each time period remain the same, as the number people born in each region cannot increase or decrease. Second, differences in cells must implicitly be driven solely by migration flows. These movements occur by individuals changing their place of residence (moving across columns), while their birthplace (row) characteristic remains fixed as represented by the different shadings in Table S1. Third, when stock data are categorized according to the place of birth characteristic and there are zero births and deaths, the change in the column (total population) sums represent the net migration flow for each country between the two time points. For example, in region A there is a net migration flow of $1160 - 1155 = 5$ people is implied from the two migrant stock tables.

We can alternatively view the top panel of Table S1 as a set of four birthplace specific migration flow tables where the marginal totals are known, shown in the bottom panel of Table S1. These are formed by considering each row of the two consecutive stock tables as a set of separate margins of a migration flow table. Place of residence totals at time t from the stock data now become origin margin (row) totals for each birthplace specific population. Similarly, place of residence totals at time $t + 1$ from the stock data now become destination margin (column) totals for each birthplace specific population. As the row totals from the stock tables are equal, the row and column margins in each of the birthplace specific migration flow tables in Table S1 are also equal.

Estimating Flows From Stocks

Within each birthplace specific table in the bottom panel of Table S1, missing non-diagonal cells must represent the counts of migrants whose location at time t is different to that at the $t + 1$. These are commonly known in the migration data literature as migrant transition flows (see *(2)* for a full exposition of migration flow data measures). Diagonal cells represent people who have the same location at each time point, known as "stayers". In order to estimate the missing migrant transition flows, we must first make an assumption for the number of stayers.

We set the diagonal cells to their maximum possible value, given the known row and column totals. This assumption allows the remaining missing (non-diagonal) cells to represent the minimum number of migrant transition flows. Alternative assumptions, where diagonal cells are set to values below their maximum would result in more migrant transition flows. We make the maximizing assumption for three reasons. First, people are far more likely to stay than make an international move. Second, there is no available country specific information on the intensities of circular and return migration which could be utilized to set diagonal values below their maximum. Third, further experimentation has shown that even a slight reduction of the diagonal values below their maximum has a substantial impact on the number of estimated moves. Regardless of the diagonal assumption the implied net migration from the demographic accounting, discussed in the next subsection, remains.

Log-linear models of *(20)* are often used as a framework to estimate missing bilateral migration flows in internal migration studies, where data on the marginal totals are known, for a number of reasons *(21,22)*. First, in order to estimate parameters in log-linear models, only the marginal (sufficient) statistics relating to the parameter in question are required. For example, in order to estimate the origin (row) parameters in a log-linear model from data in a complete migration flow table, only knowledge on the row totals are required. Hence, if the data within the flow table is missing, the parameter can still be identified. Second, the fitted values from a log-linear model are constrained in their estimation to match the corresponding observed marginal totals. For example, if a log-linear model is fitted with origin and destination (row and column) parameters, estimated fitted values will have row and column sums equal to the observed row and column sums. Third, the estimation of parameters in log-linear models require an assumption that data follow a Poisson distribution. Consequently, fitted values are maximum likelihood estimates, and hence possess a number of desirable asymptotic properties such as consistency, asymptotic normality and asymptotic robustness (see *(23)* for discussion of maximum likelihood estimates in relation to migration models).

A log-linear model for the number of migrants in transition n_{ijk} from origin i to destination j born in k during the respective time interval, as in the migrant flow tables in the bottom panel of Table S1, can be represented as:

$$\log y_{ijk} = \log \alpha_i + \log \beta_j + \log \lambda_k + \log \gamma_{ik} + \log \kappa_{jk} + \log \delta_{ijk} I(i = j) + \log m_{ij}, \qquad (1)$$

where y_{ijk} is the expected number of migrant transitions from origin i to destination j of people born in k, during the respective time interval and $i, j, k = 1, 2, \ldots, R$, for R origins, destinations, and birthplaces. The α_i, β_j and λ_k parameters represent background factors that relate to the characteristics of the origins, destinations and birthplaces respectively. The γ_{ik} and κ_{jk} parameter sets represent the factors specific to each origin-birthplace and destination-birthplace specific combinations respectively. The $I(\cdot)$ is the indicator function,

$$I(i = j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases},$$

and the corresponding δ_{ijk} parameter set represents the factors specific to each set of stayers. The m_{ij} variable represents some auxiliary information on migration flows. This is typically additional data related to migration between the same origins and destinations that might inform the estimation of y_{ijk}. We use the inverse of the distance between each region, $m_{ij} = d_{ij}^{-1}$. Other alternatives have been suggested in the estimation of internal migration flows, see *(22)* for a brief review.

Using the sufficient statistics shown in Table S1 and values for m_{ij}, we can obtain estimates of the parameters in the model (1) above using a conditional maximization iterative scheme[3] outlined in *(13)*. The maximum likelihood estimates of y_{ijk}, the expected number of migrant transitions can then be derived. These values are shown in the top panel of Table S2 where all

[3]Equivalent to an iterative proportional fitting scheme *(12)*

non-diagonal elements of m_{ij} are set to unity ($m_{ij} = d_{ij}^{-1} = 1$). Summing over all birthplaces and deleting stayers in the diagonal elements gives a traditional flow table of migrant transitions from origin i to destination j during the time period t to $t + 1$ shown in the bottom panel of Table S2. Note, that the net migration flows discussed in the original migrant stock table are still present in the estimated flow table. For example in region A the total immigration flow (column total) is 55, while the total emigration flow (row total) is 50, resulting in the net migration flow of 5 people. As discussed previously, these nets are implied in the original stock data via the difference in the populations (the column sums of Table S1) accounting for natural change from births and deaths (both of which are set to zero). Also note that the total net migration in this simplified global migrant flow table is zero.

Net Migration Flows Implied from Population, Birth, and Death Data

In reality, natural changes from births and deaths in the population occur, causing differences in the row totals of the stock data over time. In *(13)* these changes were controlled for through a number of demographic accounting procedures to adjust stock totals to have equal row totals[4]. Once these controls were made, the row totals of the migrant stock tables were equal allowing a representation of the stock tables as birthplace specific flows tables with known margins as in Table S1. As shown in the previous section, when stock data are represented in birthplace specific flow tables, missing flows in the non-diagonal cells can be estimated by assuming the log-linear model (1). However, in *(13)* the adjustments to control for birth and deaths led to the net migration flows that did not equal those implied by the demographic data. What follows is a new approach to control for changes in stock totals that maintain the net migration flow total implied in the original migrant stock data. This is carried out in a three-step procedure,

[4]Additional adjustments were also made for migrants moving to or from regions outside those under consideration.

illustrated using a new set of hypothetical data for $t+1$, displayed in Table S3. In this new data, on the right hand side, differences in row totals from births and deaths now exist.

First, alterations to the stock tables to account for sources of natural population change must be made. In order to avoid estimating migrant transitions to meet decreases in stock totals from mortality, the number of deaths in the time interval t to $t+1$ is subtracted from the reported stock data at time t. While a decomposition of the numbers of deaths by birthplace is typically missing from official statistics, the total number of deaths in each place of residence is known. To estimate this breakdown, and hence adjust each native and foreign born population stocks, the total number of deaths is proportionally allocated out to each population stock. This is illustrated on the left hand side of Step 1 in Table S4, where the total number of deaths, given in bold type face in the final sum row, is known. These totals are proportionally split according to the reported population stocks in time t, to provide estimates of the number of deaths by each birthplace. This allocation could be further refined given information on the age structure of native and foreign born populations. If migrant stocks are relatively young, the number of deaths in these groups would be expected to be relatively low and could be adjusted accordingly.

In order to avoid estimating migration flows to meet increases in native born totals from newborns, the number of births between t and $t+1$ is subtracted from the reported stock data at time $t+1$. As with deaths, we tend to only have information on the total number of births, where ideally more detail on the place of residence of newborns at time $t+1$ is desired. In order to adjust stock totals for natural increases, births are assumed to only affect the native born stocks, assuming there is no migration of newborns. This is illustrated on the right hand side of Step 1 in Table S4, where the total number of births, given in bold type face in the final row sum, is initially known. These totals of newborns are allocated to reside in their birthplace at time $t+1$.

A new set of adjusted stock tables that account for natural population change are shown in

Step 2 of Table S4, where both the death and birth estimates of the previous step are subtracted cell-wise from the original data in Table S3. The new altered stock tables still do not have equal row totals. However, they do have equal table totals, as the difference in the stock totals ($3060 - 3000 = 60$) between the two periods is fully accounted for by the natural increase from births and deaths ($220 - 160 = 60$). If the estimates (and assumptions) about the changes to population stocks from natural causes are true, the remaining differences between the row totals in the altered stock tables are likely to represent the sum of differences in migrant stock data collection procedures of each region[5]. In order to adjust for these differences, we make one further alteration to the stock totals. Using a simple iterative proportional fitting scheme we adjust each stock table to 1) maintain their column totals in Step 2 and 2) fix the row totals to the average of those calculated in Step 2, and 3) maintain the same interaction structure with in the re-estimated stock totals as in those calculated in Step 2[6]. The new set of altered stock totals that adjust for difference in stock totals beyond natural change are shown in Step 3 of Table S4. These now have matching row totals, required to estimate flows using the methodology outlined in the previous subsection.

The re-adjusted estimates shown in Stage 3 of Table S4 can be considered as a set of $R = 4$ birthplace specific flow tables, shown in the margins in the top panel of Table S5, just as the data in Table S1 were. Using these marginal data and the converged parameter estimates in the log linear model (1) we can obtain the maximum likelihood estimates of y_{ijk}, the expected number of migrant transitions, controlling for natural population changes. These values are

[5] In *(13)* an alternative assumption was made, whereby an additional calculation of the minimum amount of migrant transitions to or from external regions beyond A to D was derived. These estimated further movements were then deduced from the altered stock tables in Step 2 to provide a new set of re-estimated stock tables with matching row totals in each time period

[6] The re-estimated results are derived using the `ipf2` routine in the `migest` R package *(24)* which fits independent log-linear models with offset for a two-way table, given row and column totals, and the adjusted values in Step 2 as the offset term.

shown in the cells of the tables in the top panel of Table S5[7] Summing over all birthplaces and deleting stayers in the diagonal elements gives a traditional flow table of migrant transitions from origin i to destination j during the time period t to $t+1$ shown in the bottom panel of Table S5. Estimates are not directly comparable with previous flow tables as they are formed from a different set of migrant stock data in $t+1$.

Application

Place of birth data published by the U.N. *(11)* provide foreign born migrant stock tables at 1 July at the start of each of the last three decades (1990, 2000, and 2010) covering 232 countries. The data are primarily based on place of birth responses to Census questions, details collected from population registers and refugee statistics. In order to create a complete data for the same mid-year date the U.N. undertook a number of extrapolations to available data and imputations for missing data. For full details the reader is referred to *(11)*. Of the 232 countries for which stock data were available, 196 also had the demographic data from the World Population Prospects of the U.N. *(25)*[8] throughout the time period, as required for estimating flow methodology outlined in the previous section. None of the dropped countries had populations in 2010 in excess of 100,000 people. Diagonal elements in each stock table, of the native-born population totals in each place of residence j, ($P_j^{k=j}$), are not provided in the U.N. stock data. These were derived as a remainder ($P_j^{k=j} = P_j^{+} - \sum_{k \neq j} P_j^{k}$) using annual population totals from the U.N. (32) (P_j^{+}), and the column sums of the foreign born populations in each place of residence ($\sum_{k \neq j} P_j^{k}$). This procedure constrained the column totals of the stock tables to meet those of the reported populations at the start of each decade.

[7]The iterative procedure to estimate parameters and y_{ijk}, controlling for flows to and from outside regions is undertaken using the `ffs` routine in the `migest` R package *(24)* and setting the argument `method = "stocks"`. By default, in the `ffs` routine all elements of m_{ij} are set to unity ($m_{ij} = 1$) and the diagonal element are set to their maximum possible values given the known margins.

[8]Data available from `http://esa.un.org/unpd/wpp/`.

In order to estimate five year migrant flow tables we estimated the mid-decadal (1995 and 2005) stock table following a similar procedure used by the U.N. to align census and survey data at the beginning of each decade. First, we interpolated the proportions of each foreign born stock in the bilateral flow table to its mid-decadal value. We then multiplied up the proportion to using the population total of the appropriate year. Demographic data on the number of births and deaths in each country, required in the multi-step estimation shown in Table S4, were also taken from *(25)*. In addition, we used data on the age structure of foreign born populations in each country from the U.N. *(11)* to weight the distribution of the number deaths down in each column of the death by birthplace table according to the population size and the mean age of the relevant (native born or foreign born). Auxiliary data for use in the offset term of the estimation procedure were taken from the Centre d'Etudes Prospective et d'Informations Internationales data base on geographic distance *(26)*, which provides a distance measure between all capital cities. The offset term was calculated as $m_{ij} = d_{ij}^{-1}$. The multi-step accounting method was undertaken to adjust reported stock totals for births and deaths while maintaining the implied net migration flow totals from the demographic data. The conditional maximization routine was then run to calculate the five-year migrant flow tables using the stock tables at the beginning and end of each period[9].

[9]Both of these processes were undertaken within the `ffs` routine in the `migest` R package *(24)*.

Table S 1: Dummy Example of Place of Birth Migrant Stock Data

Place of Birth Data in Stock Tables:

Place of Residence (t)

Birthplace	A	B	C	D	Sum
A	1000	100	10	0	1110
B	55	555	50	5	665
C	80	40	800	40	960
D	20	25	20	200	265
Sum	1155	720	880	245	3000

Place of Residence ($t+1$)

Birthplace	A	B	C	D	Sum
A	950	100	60	0	1110
B	80	505	75	5	665
C	90	30	800	40	960
D	40	45	0	180	265
Sum	1160	680	935	225	3000

Place of Birth Data in Flow Tables:

Birthplace=A

Origin \ *Destination*	A	B	C	D	Sum
A	950				1000
B		100			100
C			10		10
D				0	0
Sum	950	100	60	0	1110

Birthplace=B

Origin \ *Destination*	A	B	C	D	Sum
A	55				55
B		505			555
C			50		50
D				5	5
Sum	80	505	75	5	665

Birthplace=C

Origin \ *Destination*	A	B	C	D	Sum
A	80				80
B		30			40
C			800		800
D				40	40
Sum	90	30	800	40	960

Birthplace=D

Origin \ *Destination*	A	B	C	D	Sum
A	20				20
B		25			25
C			0		20
D				180	200
Sum	40	40	0	180	265

Table S 2: Estimates of Migrant Transition Flow Tables Based on Stock Data in Table S1, with Known Diagonals

Estimates of Origin Destination Place of Birth Flow Tables:

Birthplace=A

		Destination				
		A	B	C	D	Sum
Origin	A	950	0	50	0	1000
	B	0	100	0	0	100
	C	0	0	10	0	10
	D	0	0	0	0	0
	Sum	950	100	60	0	1110

Birthplace=B

		Destination				
		A	B	C	D	Sum
Origin	A	55	0	0	0	55
	B	25	505	25	0	555
	C	0	0	50	0	50
	D	0	0	0	5	5
	Sum	80	505	75	5	665

Birthplace=C

		Destination				
		A	B	C	D	Sum
Origin	A	80	0	0	0	80
	B	10	30	0	0	40
	C	0	0	800	0	800
	D	0	0	0	40	40
	Sum	90	30	800	40	960

Birthplace=D

		Destination				
		A	B	C	D	Sum
Origin	A	20	0	0	0	20
	B	0	25	0	0	25
	C	10	10	0	0	20
	D	10	10	0	180	200
	Sum	40	45	0	180	265

Estimates of Total Origin Destination Place of Birth Flow Tables:

		Destination				
		A	B	C	D	Sum
Origin	A		0	50	0	50
	B	35		25	0	60
	C	10	10		0	20
	D	10	10	0		20
	Sum	55	20	75	0	150

Table S 3: Dummy Example of Place of Birth Data

		Place of Residence (t)							*Place of Residence* $(t+1)$				
		A	B	C	D	Sum			A	B	C	D	Sum
Birthplace	A	1000	100	10	0	1110	*Birthplace*	A	1060	60	10	10	1140
	B	55	555	50	5	665		B	45	540	40	0	625
	C	80	40	800	40	960		C	70	75	770	70	985
	D	20	25	20	200	265		D	30	30	20	230	310
	Sum	1155	720	880	245	3000		Sum	1205	705	840	310	3060

Table S 4: Multi-Step Correction to Stock Data

Step 1: Control for Natural Changes

Birthplace	*Place of Death* (t) A	B	C	D
A	60.6	4.2	0.6	0
B	3.3	23.1	2.8	0.2
C	4.9	1.7	45.5	1.6
D	1.2	1.0	1.1	8.2
Sum	**70**	**30**	**50**	**10**

Birthplace	*Place of Residence* $(t+1)$ A	B	C	D	Sum
A	80	0	0	0	**80**
B	0	20	0	0	**20**
C	0	0	60	0	**60**
D	0	0	0	60	**60**

Step 2: Estimated Altered Stocks

Birthplace	*Place of Residence* (t) A	B	C	D	Sum
A	939.4	95.8	9.4	0.0	1044.7
B	51.7	531.9	47.2	4.8	635.5
C	75.2	38.3	754.5	38.4	906.4
D	18.8	24.0	18.9	191.8	253.4
Sum	1085	690	830	235	2840

Birthplace	*Place of Residence* $(t+1)$ A	B	C	D	Sum
A	980	60	10	10	1060
B	45	520	40	0	605
C	70	75	710	70	925
D	30	30	20	170	250
Sum	1125	685	780	250	2840

Step 3: Re-estimated Altered Stocks

Birthplace	*Place of Residence* (t) A	B	C	D	Sum
A	942.0	101.0	9.4	0.0	1052.3
B	48.4	523.5	43.7	4.6	620.2
C	76.3	40.9	758.7	39.7	915.7
D	18.3	24.5	18.2	190.7	251.7
Sum	1085	690	830	235	2840

Birthplace	*Place of Residence* $(t+1)$ A	B	C	D	Sum
A	976.1	56.5	9.9	9.8	1052.3
B	48.4	528.8	43.0	0.0	620.2
C	69.8	70.6	706.6	68.7	915.7
D	30.7	29.0	20.5	171.5	251.7
Sum	1125	685	780	250	2840

Table S 5: Estimates of Migrant Transition Flow Tables Based on Stock Data Derived in Table S4, with Known Diagonals

Estimates of Origin Destination Place of Birth Flow Tables:

Birthplace=A

		Destination				
		A	B	C	D	Sum
Origin	A	942.0	0.0	0.0	0.0	942.0
	B	34.1	56.5	0.6	9.8	101.0
	C	0.0	0.0	9.4	0.0	9.4
	D	0.0	0.0	0.0	0.0	0.0
	Sum	976.1	56.5	9.9	9.8	1052.3

Birthplace=B

		Destination				
		A	B	C	D	Sum
Origin	A	48.4	0.0	0.0	0.0	48.4
	B	0.0	523.5	0.0	0.0	523.5
	C	0.0	0.7	43.0	0.0	43.7
	D	0.0	4.6	0.0	0.0	4.6
	Sum	48.4	528.8	43.0	0.0	620.2

Birthplace=C

		Destination				
		A	B	C	D	Sum
Origin	A	69.8	3.3	0.0	3.2	76.3
	B	0.0	40.9	0.0	0.0	40.9
	C	0.0	26.4	706.6	25.7	758.7
	D	0.0	0.0	0.0	39.7	39.7
	Sum	69.8	70.6	706.6	68.7	915.7

Birthplace=D

		Destination				
		A	B	C	D	Sum
Origin	A	18.3	0.0	0.0	0.0	18.3
	B	0.0	24.5	0.0	0.0	24.5
	C	0.0	0.0	18.2	0.0	18.2
	D	12.4	4.5	2.3	171.5	190.7
	Sum	30.7	29.0	20.5	171.5	251.7

Estimates of Total Origin Destination Flow Table:

		Destination				
		A	B	C	D	Sum
Origin	A		3.3	0	3.2	6.6
	B	34.1		0.6	9.8	44.5
	C	0	27.1		25.7	52.8
	D	12.4	9.1	2.3		23.8
	Sum	46.6	39.5	2.8	38.8	127.7

Additional Data Table S1 (Flow estimates by region 2005.xls)

Bilateral flow estimates by region, 2005-10

This is a 15*15 matrix stored as an excel file. Rows correspond to origins, columns to destinations.

Additional Data S2 (Flow estimates by country 1990-2010.xls)

Bilateral flow estimates by country, 1990-95 to 2005-10

This is a 196*196 matrix stored as an excel file. Rows correspond to origins, columns to destinations. Countries a indicated by their iso-3 code.

References and Notes

1. B. Nowok, D. Kupiszewska, M. Poulain, in *THESIM: Towards Harmonised European Statistics on International Migration*, M. Poulain, N. Perrin, A. Singleton, Eds. (Presses universitaires de Louvain, Louvain-la-Neuve, 2006), pp. 203–231.

2. P. Rees, F. Willekens, in *Migration and Settlement: A Multiregional Comparative Study*, F. Willekens, A. Rogers, Eds. (Reidel, Dordrecht, Netherlands, 1986), pp. 19–58.

3. H. Zlotnik, The concept of international migration as reflected in data collection systems. *Int. Migr. Rev.* **21**, 925–946 (1987). doi:10.2307/2546498 Medline

4. S. Castles, M. J. Miller, *The Age of Migration: International Population Movements in the Modern World* (Macmillan, London, 2009).

5. D. S. Massey, R. M. Zenteno, The dynamics of mass migration. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5328–5335 (1999). doi:10.1073/pnas.96.9.5328 Medline

6. M. Czaika, H. de Haas, *The Globalisation of Migration* (IMI Working Papers, WP-682013, 2013).

7. G. J. Abel, Estimation of international migration flow tables in Europe. *J. R. Stat. Soc. A* **173**, 797–825 (2010). doi:10.1111/j.1467-985X.2009.00636.x

8. J. de Beer, J. Raymer, R. van der Erf, L. van Wissen, Overcoming the problems of inconsistent international migration data: A new method applied to flows in Europe. *Eur. J. Popul.* **26**, 459–481 (2010). doi:10.1007/s10680-010-9220-z Medline

9. J. Raymer, A. Wiśniowski, J. J. Forster, P. W. F. Smith, J. Bijak, Integrated modeling of European migration. *J. Am. Stat. Assoc.* **108**, 801–819 (2013). doi:10.1080/01621459.2013.789435

10. C. Özden, C. R. Parsons, M. Schiff, T. L. Walmsley, Where on earth is everybody? The evolution of global bilateral migration 1960–2000. *World Bank Econ. Rev.* **25**, 12–56 (2011). doi:10.1093/wber/lhr024

11. UNPD, *Trends in International Migrant Stock: Migrants by Destination and Origin, The 2013 Revision* (United Nations, Department of Economic and Social Affairs, Population Division, New York, 2013).

12. W. Deming, F. Stephan, On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.* **11**, 427–444 (1940). doi:10.1214/aoms/1177731829

13. G. J. Abel, Estimating global migration flow tables using place of birth data. *Demogr. Res.* **28**, 505–546 (2013). doi:10.4054/DemRes.2013.28.18

14. M. Bell, E. Charles-Edwards, *Cross-National Comparisons of internal Migration: An Update of Global Patterns and Trends* (United Nations, Department of Economic and Social Affairs, Population Division, New York, 2013).

15. P. H. Rees, The measurement of migration, from census data and other sources. *Environ. Plan. A* **9**, 247–272 (1977). doi:10.1068/a090247

16. M. Bell, M. Blake, P. Boyle, O. Duke-Williams, P. Rees, J. Stillwell, G. Hugo, Cross-national comparison of internal migration: Issues and measures. *J. R. Stat. Soc. A* **165**, 435–464 (2002). doi:10.1111/1467-985X.00247

17. M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, M. A. Marra, Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009). doi:10.1101/gr.092759.109 Medline

18. S. Henning, B. Hovy, Data sets on international migration. *Int. Migr. Rev.* **45**, 980–985 (2011). doi:10.1111/j.1747-7379.2011.00874_2.x

19. J. S. Passel, R. Suro, Rise, *Peak, and Decline: Trends in US Immigration 1992–2004* (Pew Hispanic Center, Washington, DC, 2005).

20. M. W. Birch, Maximum likelihood in Three-Way tables. *J. R. Stat. Soc., B* **25**, 220–233 (1963).

21. J. Raymer, G. J. Abel, P. W. Smith, Combining census and registration data to estimate detailed elderly migration flows in England and Wales. *J. R. Stat. Soc. A* **170**, 891–908 (2007). doi:10.1111/j.1467-985X.2007.00490.x

22. F. Willekens, Modeling approaches to the indirect estimation of migration flows: From entropy to EM. *Math. Popul. Stud.* **7**, 239–278, 308 (1999). doi:10.1080/08898489909525459 Medline

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

24. G. J. Abel, "migest: Useful R code for the Estimation of Migration" (Retrieved 25 April, 2013); available at http://cran.r-project.org/web/packages/migest.

25. UNPD, World Population Prospects: the 2010 Revision (United Nations Population Division, New York, 2011).

26. T. Mayer, S. Zignago, "Notes on CEPII's Distances Measures: The GeoDist Database" (CEPII document No. 2011-25, retrieved 1 August 2013, from www.cepii.fr/anglaisgraph/workpap/pdf/2011/wp2011-25.pdf.