IMR SAGE

Estimates of Global Bilateral Migration Flows by Gender between 1960 and 2015[1]

Guy J. Abel
Asian Demographic Research Institute, Shanghai University
Wittgenstein Centre (IIASA, VID/ÖAW, WU) International Institute for Applied Systems Analysis

An indirect estimation method is used to derive country to country migration flows from changes in global bilateral stock data. Estimates are obtained over five- and 10-year periods between 1960 and 2015 by gender, providing a comprehensive picture of past migration patterns. The estimated total of global international migrant flows generally increases over the 55-year time frame. The global rate of migration over five- and 10-year periods fluctuate at around 0.65 and 1.25 percent of the population, respectively. The sensitivity of estimates to alternative input stock and demographic data are explored.

INTRODUCTION

Global migration flows form a complex system influenced by a mix of socioeconomic, political, and demographic factors. In many developed countries, international migration is an important driver of demographic growth, often accounting for over half of the population change Lee (2011). Comparable international migration data inform policymakers, the media, and academics about the level and direction of population movements and allows hypotheses on the determinants and patterns of people's moves to be tested. This study presents new estimates of comparable bilateral migration between all countries from 1960 to 2015 based on range of available data sources.

[1]This work was supported by National Science Foundation of China Research Fund for International Young Scientists. The author is grateful for suggestions provided by three anonymous reviewers as well as Tom King and Anne Goujon.

DOI: 10.1111/imre.12327

Moves in populations can be quantified using either migrant stock or migration flow data. Unlike a static stock measure, flow data are dynamic, summarizing migration over a defined period and consequently leading to a better understanding of past patterns and the prediction of future trends. Until recently, net migration flow estimates produced every two years by the United Nations have served as the sole source of global migration flow data. However, as with any net measure, they are susceptible to distorting or disguising the underlying patterns (Rogers 1990) and hence are of limited explanatory use. More detailed measures, such as the immigration and emigration counts, or country-to-country bilateral flows are far better equipped to explain and predict global migration trends. Currently, only a minority of countries collect detailed flow data. When comparing available flow data, major problems exist stemming from the use of different definitions and measures employed by national statistic institutes and the availability of data over different time horizons (Kelly 1987; Nowok, Kupiszewska, and Poulain 2006; Kupiszewska and Nowok 2008). In the European context, where flow data are more abundant, methodologies to harmonize existing data have been developed (Raymer 2007; Abel 2010; de Beer et al. 2010; Raymer et al. 2013; Wiśniowski et al. 2013). Each is severely limited in its application to a global setting where missing data become a major issue. Hence, in order to obtain an understanding of global migration patterns, indirect methods must be used to estimate international flows using alternative data sources.

Previous studies of global migration patterns such as those of Zlotnik (1999), National Research Council (2000) Martin and Widgren (2002) or Castles, de Haas, and Miller (2013) have been based on a patchwork of net migration measures, changes in bilateral stocks over time, and available, unharmonized flow data from predominately rich Western countries. A growing literature on the analysis of bilateral migrant stock data (Beine, Docquier, and Özden 2011; Docquier et al. 2012; Czaika and De Haas 2014) to explain changes in contemporary migration patterns has recently developed. However, as stock data only record the place of birth and current residence, they can easily misrepresent contemporary migration patterns. This is particularly true in countries where there are significant return migration or mortality among foreign populations (Massey et al. 1999, 200). Furthermore, recent moves by migrants already living outside their country of birth are also not covered using stock measures. These drawbacks can potentially result in countries with longer migration histories becoming overrepresented in comparison with those with younger populations, where the cumulative time available to people to emigrate is lower. Other studies of global migration patterns such as Zagheni and Weber (2012), State, Weber, and Zagheni (2013), Hawelka et al. (2014) or Zagheni et al. (2014) have focused on short-term mobility measures derived from data sources based on individuals geo-location of Internet activities such as Twitter messages or logins to email services. As the authors noted, their data may not be fully representative of the global migration patterns and are not always publicly available.

Indirect methods have recently been used to estimate global bilateral migration flows based upon changes in published bilateral migrant stock data. Abel (2013a) used global bilateral stock tables from the World Bank to derive global bilateral flow estimates between 1960 and 2000 over four 10-year periods via a proposed flows from stocks methodology. Abel and Sander (2014) extended the method to estimate bilateral migration flows over four five-year periods between 1990 and 2010, based on the changes in global bilateral stocks of the United Nations. The extension allowed the difference of estimated immigration and emigration flow totals to match the net migration estimates of United Nations Population Division (2011).

In this study, these previous estimation methods are further developed and applied to a much larger range of input data, allowing for a number of new insights from the resulting global bilateral flow estimates that can be grouped into three areas. First, the new estimates cover each gender, quantifying for the first time differences in male and female global migration flow patterns. Piper (2005) and Zlotnik (2003) noted an overall rise in the share of female in migrant stocks, rising from 46.6 to 48.8 percent of the global migrant stock between 1960 and 2000. Using available flow and stock data, Donato and Gabaccia (2015) discussed the persistence of a convergence to gender-balanced migration over a much longer period of time, facilitated in recent decades by an increased restriction and management of moves and a gendered shift in demand for migrant labor. Distinct gender variations in the migration patterns are known to exist from stock data and localized studies (Zlotnik 1995; Curran and Saguy 2001; Donato et al. 2006). This is often related to the social factors that influence migrating women's and men's roles, access to resources, facilities, and services which have been the focus of research often based on arrivals to a single country–see for example the compilations of Piper (2013) or Truong et al. (2014). The role of gender differentials in international moves for domestic workers is often highlighted. The International Labour Office estimated that there are between 53 and 100 million domestic workers worldwide (accounting for hidden and unregistered people). Approximately 83 percent of these workers are women or girls and many are migrant workers (International Labour Office 2013). When considering education levels, Docquier, Lowell, and Marfouk (2009) and Docquier et al. (2012) found evidence to suggest skilled women exhibit greater propensities to make international moves during recent decades than skilled men.

Second, the methodology of Abel (2013a) and Abel and Sander (2014) is extended to account for contradictions between demographic and stock data which previously limited the estimation of bilateral flows between some countries in certain periods. The revised method is applied to estimate five- and 10-year migrant flows by gender between 1960 and 2015, to provide an updated view of international migration over a longer period of time. Estimates over both five- and 10-year periods enable differences in global migration transition rates to be identified.

Third, estimates of migrant flows in this study are based on a variety of migrant stock and demographic data to study their sensitivity to alternative bilateral stocks (of the United Nations and World Bank) and revised estimates in the number of births and deaths over a given interval. The culmination of bilateral flow estimates varying by different gender, time periods, intervals, stock, and demographic data provides a combined set of 273 estimated migrant flow tables, far exceeding those in the previously discussed flows from stocks estimation studies (four tables in each).

In the next section, the methodology to estimate origin–destination flow tables from changes in bilateral stock data will be outlined. This includes an overview of the various migrant stock and demographic data required as inputs for the estimation methodology and some details on the assumptions and limitations underlying the estimates. In section "Results," results from the estimated flow tables are shown at different levels of analyses. The sensitivity of the methodology to alternative input data and changes in political geography are then discussed followed by a comparison of the estimates to reported data from national statistics institutes. Finally, the results are summarized and discussed in reference to current work on global migration data. The Appendix S1 provides a detailed review of the flows from stocks methodology as well as some further sensitivity analyses.

METHODOLOGICAL BACKGROUND

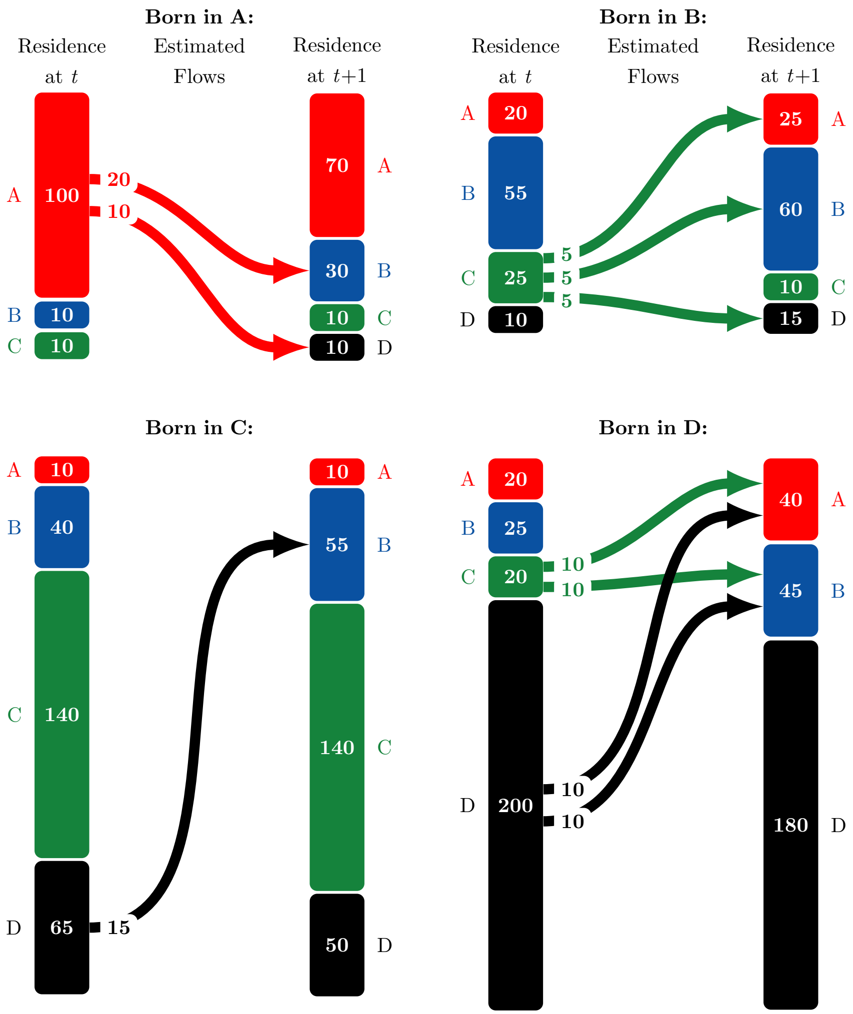
Available bilateral migration data can be categorized as either a stock measure, which represents a static number of a foreign population defined by a characteristic such as their place of birth, or a flow measure, which represents the dynamic movements of populations between origin and destinations. As a stock data measure is at a single point of time, there are far fewer limitations in its measurement and collection than there are for flow measures. Hence, migrant stock data are available across a wider range of countries and over longer time periods than migrant flow data. Groups at both the United Nations and the World Bank have collated together stock data from national statistical institutes to build global bilateral migrant stock tables for different time points. In this section, an outline is given for the methodology to indirectly estimate bilateral migrant flows from sequential bilateral migrant stock tables while accounting for demographic changes over the period. This is followed by some additional discussions on representing migration data in contingency tables form. Then, the use of log-linear models, which forms the statistical heart of the flow from stock methodology, is highlighted. Background information on the input data required for the application of the methodology to estimate global bilateral migration flows are then discussed. Finally, details on some possible alternative assumptions for the estimation methodology and the estimates limitations are addressed.

Flows from Stocks Method

Changes over time in bilateral migrant stock sizes, defined by the place of birth of individuals, can be the result of (1) an increase in the size of native-born populations from births, (2) reductions in the size of both foreign- and native-born populations from deaths, and (3) migrant flows that can either increase or decrease migrant stock sizes. When the data of bilateral migrant stocks at the beginning and end of periods are available, it is possible to indirectly derive the number of bilateral migrant flows (having accounted for births and deaths) by viewing each population stock as part of a demographic accounting system.

Consider the hypothetical case where there are no births and deaths over a given time interval. Changes in bilateral stocks in each location must be solely due to migrant transitions. Figure I illustrates this case using a

Figure I. **Schematic of a Demographic Accounting Framework to Link Changes in Bilateral Migrant Stock Data Via Estimated Migrant Flows**



Notes: For each birthplace there are no births or deaths during the time interval; hence, the total birthplace populations are the same at time t and t + 1, represented by the equal heights in each set of stacked blocks. The estimated flow sizes displayed in the arrows are the minimum number of migrant transitions required to match changes in the known bilateral migrant stock data given in each block.

schematic of a simple demographic account framework based on a dummy data, for example, data at time t and $t + 1$ and a global migration system consisting of four countries. Blocks represent the size of bilateral migrant stocks at the beginning and the end of the interval. They are grouped together by the country of birth. For example, those born in country A are shown in the top left; 100 are native-born citizens, living in country A at time *t*. Two sets of 10 people born in A are living abroad in countries B and C, while none live in country D. At time $t + 1$, the distribution of those born in country A alters. The native-born population has dropped by 30, while the stock living in countries B and D has increased. Note, the total population of those born in A (120) residing in any country does not change over the period of time as there are no births or deaths, and the birthplace is a fixed characteristic that cannot alter over time.

There are many possible combinations of migration events that can take place over the period to match the changes in the bilateral migrant stocks. However, a minimum of 20 migrants must leave A and arrive in B, and a further 10 must leave A and arrive in D. The *minimum* number of migrant transitions for all birthplace populations in a global system can be indirectly estimated using a log-linear model, details of which are given in the next subsection.

The results of the indirect estimation method for the global system of four countries are shown by the arrows in Figure I. The estimated flows are based on a number of migrant transition over the period. Alternatively, migration can be measured by the number of migrant movements during a period between a given origin and destination. A movement definition of a migration flow captures multiple changes in location over a defined period including intermediate moves to a third country or return to the original place of residence after a short stay elsewhere. Although the number of movements will be at least as high and the number of transitions, there is no simple mathematical solution to estimate one from the other.

The demographic framework in Figure I can be extended to account for demographic changes from both births and deaths, which are likely to have large impacts on the changes in bilateral migrant stock data over a five- or 10-year period of time. In the case of deaths over a given time period, the migrant stocks can be adjusted by subtracting the estimated number of deaths in each population block at time *t* in Figure I before any flows are calculated. The reduction accounts for potential drops in migrant stocks at time $t + 1$ which might otherwise result in higher estimates of the number of outward migrants. A similar procedure can also be performed to account for changes in stocks from births. As the birthplace is a defining characteristic of bilateral migrant stock data, the number of newborns can be subtracted only from the native-born populations at time $t + 1$.[2] The reduction accounts for a potential increase in migrant stocks from time t which might otherwise result in an increase in the estimate of migrant that returned to their birthplace. More details of the demographic accounting framework and adjustments for births and deaths are given in the Appendix S1.

Contingency Table Representation

The bilateral migrant stock data at time t and $t + 1$ given in Figure I can be displayed in two contingency tables as shown in the top panel of Table 1. In each, rows represent a categorization of the population, which for the stock data used in this study is birthplace. The columns in bilateral migrant stock represent the place of residence. The values in non-diagonal cells represent the size of a migrant stock cross-classified by place of birth and place of residence at a specified time. Values in diagonal cells represent the native-born population size. They are sometimes not shown in migration tables as they do not measure a form of mobility. 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. When rows represent another measure, such as citizenship or nationality (not considered in this study), the column totals can potentially be greater than the population because people with dual citizenship or nationalities can be counted twice.

The row totals in each contingency table represent the count of people born in a given location across all places of residence. For the dummy data used in the top panel of Table 1, the row totals are the same, as the country of birth is a fixed characteristic of all individuals and there are no births or deaths during the period. In this case, where births and deaths have been accounted for, Abel (2013a) showed that bilateral migrant stock data can be re-represented as birthplace-specific origin–destination migration flow tables with known margins as shown in the bold typeface in the middle panel of Table 1.

[2]Note that if a newborn has a mother that is living outside her country of birth, the newborn itself will belong to the native-born population at the end of the time period unless they migrate before the end of the time period (a transition which is assumed to not occur).

TABLE 1
Deriving Bilateral Flow Tables Estimates (Bottom) from Sequential Bilateral Stock Data (Top) Via Birthplace-Specific Flow Tables (Middle)

Bilateral stock data:

		Place of residence (t)				
		A	B	C	D	Sum
Birthplace	A	100	10	10	0	120
	B	20	55	25	10	110
	C	10	40	140	65	255
	D	20	25	20	200	265
	Sum	150	130	195	275	750

		Place of residence ($t + 1$)				
		A	B	C	D	Sum
Birthplace	A	70	30	10	10	120
	B	25	60	10	15	110
	C	10	55	140	50	255
	D	40	45	0	180	265
	Sum	145	190	160	255	750

Estimates of origin–destination–place of birth flow tables:

Birthplace = A

		Destination				
		A	B	C	D	Sum
Origin	A	*70*	20	0	10	**100**
	B	0	*10*	0	0	**10**
	C	0	0	*10*	0	**10**
	D	0	0	0	*0*	**0**
	Sum	**70**	**30**	**10**	**10**	**120**

Birthplace = B

		Destination				
		A	B	C	D	Sum
Origin	A	*20*	0	0	0	**20**
	B	0	*55*	0	0	**55**
	C	5	5	*10*	5	**25**
	D	0	0	0	*10*	**10**
	Sum	**25**	**60**	**10**	**15**	**110**

Birthplace = C

		Destination				
		A	B	C	D	Sum
Origin	A	*10*	0	0	0	**10**
	B	0	*40*	0	0	**40**
	C	0	0	*140*	0	**140**
	D	0	15	0	*50*	**65**
	Sum	**10**	**55**	**140**	**50**	**255**

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 origin–destination flow table:

		Destination				
		A	B	C	D	Sum
Origin	A		20	0	10	30
	B	0		0	0	0
	C	15	15		5	35
	D	10	25	0		35
	Sum	25	60	0	15	100

Note: The stock data form the margins, in bold font, of the birthplace-specific tables. The assumed number of stayers is given in italic font.

In each birthplace-specific flow tables, row totals in the middle panel of Table 1 are comprised of the corresponding stock data at time t while the column totals are obtained from the stock data at time $t + 1$. The remaining cells of each table are unknown. The diagonal cells differ in character from the non-diagonal cells. The diagonal cells represent populations which have the same country of residence at the beginning and end of the time period and hence are counts on the number of "stayers." Abel (2013a) outlined a maximizing assumption to fix the diagonal terms to their highest possible value, conditional on the known marginal stock counts. These are provided in the italics typeface in Table 1. The remaining diagonal cells represent the flow counts within the birthplace migrant flow table. These can be indirectly estimated using a log-linear model outlined in the next subsection. These estimated flows can then be used to obtain a traditional origin–destination migrant flow table in the bottom panel of Table 1 by summing over places of birth. For example, the 25 migrating from D to B are an aggregation of 10 D to B transitions from those born in A, 15 from those born in C, and 10 from those born in D (each shown in Figure I and the middle panel of Table 1).

Log-Linear Models

Log-linear models are a form of Poisson regression model, commonly used to explain or predict count variables that are cross-classified by one or more categorical explanatory variables. In this study, they are used to predict missing cells in migration flow tables that match known marginal totals. Log-linear models are suitable for this purpose as (1) the parameters in the models can be estimated without knowing the cell totals and (2) the fitted and observed values in a log-linear models have the same marginal totals when the corresponding categorical variable is used.

In order to predict migration flows given the marginal and diagonal values in the middle panel of Table 1, a log-linear model is carefully specified to have parameters corresponding to their available sufficient statistics;

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

where $I(\cdot)$ is the indicator function,

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

and y_{ijk} is the expected number of migrant transitions from origin i to destination j of people born in birthplace 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 are used to ensure the overall origin-, destination-, and birthplace-imputed totals match the observed values. The γ_{ik} and κ_{jk} parameter sets ensure specific origin–birthplace and destination–birthplace-imputed margin totals match each observed margin from the stock data. The δ_{ijk} parameter set acts similarly for the imputed diagonal cells, to provide a saturated model fit ensuring they match the assumed number of stayers given in italics in Table 1.

Each of the parameter values can be obtained using only the known marginal sums and the diagonal cells (the sufficient statistics) from an iterative proportional fitting (IPF) type algorithm fully detailed in Abel (2013a, 16) and available in the *migest* R package Abel (2013b). For more details on IPF algorithms in relation to estimating bilateral migration flow, please refer to Willekens (1999), Raymer, Abel, and Smith (2007), or Abel (2013a). For a more general statistical background, see Agresti (2003). Given the parameter estimates, the imputed values for the non-diagonal elements shown in each of the birthplace-specific flow tables can be estimated. These estimates match those shown in Figure I and the middle panel of Table 1.

The log-linear model in (1) includes an offset term m_{ij}. It allows for a single set of auxiliary information to be used in the IPF algorithm to augment the estimated flows and provide more realistic estimates. In the example of Table 1, the auxiliary variable is set to one for all bilateral combinations. For the flow estimates derived in the Results section of this study, distances between capital cities were calculated using longitude and latitude information from United Nations Population Division (2014) and the *geosphere* R package (Hijmans 2015). Given distance measures (d_{ij}) between all capital cities at origin country i and destination j, the offset term is calculated as $m_{ij} = d_{ij}^{-1}$ to weight in favor of migrant flows to closer countries.

Input Data

The estimation of international bilateral migration flow tables requires two inputs. First, bilateral stock tables are required at the beginning and the end of a given period. Currently, both the United Nations and World Bank provide sets of bilateral stock data that include more than one time period. Additional bilateral migrant stock data, such as estimates by Parsons et al. (2007), Ratha and Shaw (2007), Jean-Christophe, Spielvogel, and Widmaier (2010), or Artuç et al. (2015) are not used, as they are restricted in their global coverage or provide stock data at only one or two time points, limiting the number of periods for indirect estimates of flows to be derived. Second, demographic data on the number of births, deaths, and population are also required to estimate bilateral flows. Birth and death information is needed to alter stock data for natural change over the period of time for which flow estimates are being derived. Population data are needed to obtain the size of the native-born population, typically not given in bilateral stock tables but required to estimate flows using the method outlined in the previous section. Background details for each of these input data sources are discussed in the remainder of this section.

The World Bank (Özden et al. 2011) provided migration stock tables at the start of each decade, from 1960 to 2000, for 226 countries primarily based on place of birth responses to census questions or details collected from population registers.[3] Where no data were available, alternative stock measures such as citizenship or ethnicity were used. For countries where no stock measures were available, missing values were imputed using, typically based on interpolations from patterns of foreign-born distributions in countries within the same region.

The United Nations Population Division (2015b) provided a sequence of foreign-born migrant stock tables 5 years apart, beginning in 1990 until 2015 covering 232 countries.[4] Previous versions by the United Nations Population Division (2012, 2013a) provided stocks only at the start of each of the last three decades (1990, 2000, and 2010). As with the World Bank estimates, stock data were primarily based on place of birth responses to census questions and from population registers. Adjustments to estimates were made to include available refugee statistics. As data on foreign-born stocks might be collected in census years that are not at the start of the decade, extrapolations were made based on the change in the overall population size to align all estimates at the same mid-year time point. For countries or areas without any data source, a similar country or group of countries were used to estimate missing bilateral stocks. Unlike the World Bank stocks, the UN estimates have categories for foreign-born populations with an unknown place of birth (Other North and Other South). These

[3]Data available from http://data.worldbank.org/data-catalog/global-bilateral-migration-database (accessed May 2017)

[4]Data available from http://www.un.org/en/development/desa/population/migration/data/ (accessed May 2017)

counts originated from either regional aggregations or nonstandard areas used by national statistical agencies to enumerate foreign-born stock data that the United Nations were unable to redistribute into each country. For the vast majority of countries, the counts of unknowns comprised less than five percent of the total foreign-born population.

In this study, all three versions of the UN stock data (from now on referred to as UN2012, UN2013, and UN2015) are used, alongside the data of Özden et al. (2011) (referred to as WB2011). Estimates based on the different input stock data will allow the sensitivity of the flow estimates to alternate stock data sets to be studied, and an indirect comparison of the stock data sets themselves.

Demographic data on births, deaths, and population totals are available from the World Population Prospects (WPP) of the United Nations Population Division (2011, 2013b, 2015c). Every two to three years, the United Nations releases an updated version of the WPP incorporating revised estimates of past demographic statistics for all countries. Data on the total population and number of deaths are typically given by gender in each WPP. Data on the number of births are usually given without a disaggregation by gender of the child. However, estimates of the number of births by gender can be derived using supplementary data on the sex ratio of birth also contained in each WPP. In this study, the three most recent versions of WPP are used, WPP2010, WPP2012, and WPP2015, in order to determine what effect, if any, updated demographic data have on bilateral migration flow estimates.

Limitations and Alternative Assumptions

Two possible assumptions in the flows from stock methodology could be potentially altered and result in different bilateral migration flow estimates that still matches the marginal totals from migrant stock data. First, an alternative set of auxiliary information could be used for the offset term. Second, the number of stayers in the flows-from-stocks methodology outlined above is assumed to be the maximum possible values implied by the corresponding stock data. As a result, the estimated flows are the minimum number of transitions required to match the changes in migrant stocks. In this subsection, each of these is explored in turn.

Alternative Offset. Within the log-linear model framework, the estimation of missing migration flows uses auxiliary data in the offset term whose parameter is fixed to one. The remaining parameters in the log-linear model of (1) ensure imputed values match the known marginal and diagonal elements in the origin–destination–birthplace arrays. To investigate the sensitivity in estimation methodology from auxiliary data, an alternative distance matrix based on sets of a Pythagorean triple (3,4,5) is used to estimate a new set of flows. This distance matrix is shown in the left-hand panel of Table S1, in the Supporting information.

The estimated flows, given on the right-hand side of Table S1, can be directly compared to those in Table 1 which were based on a distance matrix where all values are set to one and the same migrant stock data given in top panel of Table 1. Both the estimated immigration and emigration totals in the row and column margins are the same, as the stock and demographic data (where births and deaths over the interval are zero) are unchanged. Only a small difference occurs in the origin–destination flow estimates, where, for example, slightly larger number of migrants are estimated to leave from D and arrive in A (9.8 as opposed to 10) and a corresponding reduction in the migrants transitioning from D to B. Similar small differences were also found in the global migration flows when contrasting estimates based on the inverse distances between capital cities (presented in the next section) and estimates based on a distance matrix where all values were set to one.

Alternative Stayers' Assumption. In order to estimate the unknown number of migrations, an assumption is required on the number of each bilateral migrant stock who does not move. The population of stayers is represented by the diagonal elements of birthplace-specific migration flow tables such as in the middle panel of Table 1. There exist a limited amount of possible values for which the diagonal cells can take. For example, at a minimum, the number of people born in A who start the period residing in A and stay in A (the top-left cell of the top-left array in the middle panel of Table 1) is 50. Any lower, and the estimated flows, conditional on the margins would need to become negative.

The alternative extreme assumption for the diagonal cells is to minimize (rather than maximize) the number of stayers. In Table S2, all stayers in the diagonal cells are set to their minimum possible values (shown in italic font) conditional on the marginal stock information (shown in bold font) which is the same as in the middle panel of Table 1. The resulting flows are shown in the non-diagonal cells. In the lower panel, estimates of the origin–destination flow table are shown having summed over all the birthplace-specific tables. The total estimated flow (555) from the minimum diagonal assumption is far greater than the total flows from the maximum diagonal assumption (100) shown in the bottom panel of Table 1.

In order to obtain a clearer picture of the relationship of assumption on the number of stayers and the overall flow total, a range of diagonal values between the maximum and minimum assumptions can be set and the resulting origin–destination flow table derived. The plot in Figure S1 shows this relationship. On the horizontal axis is the total number of stayers, ranging between the minimum of 195 and the maximum of 650 (shown in Tables S2 and 1, respectively). On the vertical axis is the estimated total flow given the corresponding number of assumed stayers. Unsurprisingly, the total estimated flow falls as the number of stayers increases.

While it is clear from migration literature that setting the diagonal to a minimum number of stayers is implausible (implying that there is cost to staying in all countries, rather than moving), it is unclear just how far from the maximum diagonal we could reasonably set the number of stayers. The dashed lines in Figure S1 represent the total estimated flows from the independence model (where the cost of staying is equal to the cost of moving) when there are no constraints on the diagonal values. It provides an indication that any reasonable assumption on the diagonal must be on the right-hand side of this line, when the number of stayers exceeds 346.

In reality, when working in an applied setting, there are a couple of logical arguments to make a maximizing assumption on the number of stayers. First, any other level of stayers would require some empirical information on the probability staying in their same country of residence during each time period and for each bilateral migrant stock. Such data are not available for the past periods and are unlikely to be forthcoming in the near future. Second, international migration is known to be a rare event. The number of stayers is likely to be much closer to the maximum assumed values than those from the independence model as (1) there are many barriers to international migration, and (2) estimated flows in this study are measures of migrant transitions over relative long periods (five and 10 years). During these periods, migrant transitions become an even rarer event, as short-term moves for returning migrants during the period are not be considered as a migration transition.

RESULTS

In order to better to understand the past patterns of global migration flows and asses the role of various components of the flows from stock estimation methodology, flow tables were estimated using all available combinations of demographic and stock data for each gender and in each period. This estimation procedure was undertaken in two rounds.

In the first round, flows over 10-year periods were estimated. The 1960–1970, 1970–1980, and 1980–1990 flow tables were calculated nine times each, based on alternative combinations of gender (male, female, and both), demographic data (WPP2010, WPP2012, and WPP2015), and stock data (WB2011). During 1990–2000, 36 flow tables were calculated, based on alternative gender, demographic data (both with the same three demographic data sets as in the previous periods), and stock data (WB2011, UN2012, UN2013, and UN2015). In the last 10-year period, 2000–2010, 27 flow tables were calculated, based on the alternative gender, demographic data (varying as in the previous periods), and stock data (UN2012, UN2013, and UN2015). This resulted in 90 estimated flow tables in total.

In the second round, flows over five-year periods between 1960 and 2010 were estimated. These were based on the same combination of period-specific gender, demographic data, and stock data when estimating the 10-year flows, providing 180 estimated flow tables. Three further flow tables were also estimated for the 2010–2015 period based on each gender combination (male, female, and both) for the WPP2015 demographic data and the most recent UN migrant stock data. Previous versions of demographic or stock data did not include information for 2015.

In order to estimate the five-year migrant flow tables, all but the latest UN stock data, estimates of the mid-decade stock tables were required. In each decade, these were imputed through a procedure similar to that previously used by the UN to align census and survey data at the beginning of each decade. This process consists of first interpolating the proportions of each bilateral foreign-born population in the stock table to its mid-decade value. The proportions are then multiplied by the available mid-decade population total of the appropriate year to provide complete bilateral stock estimates.

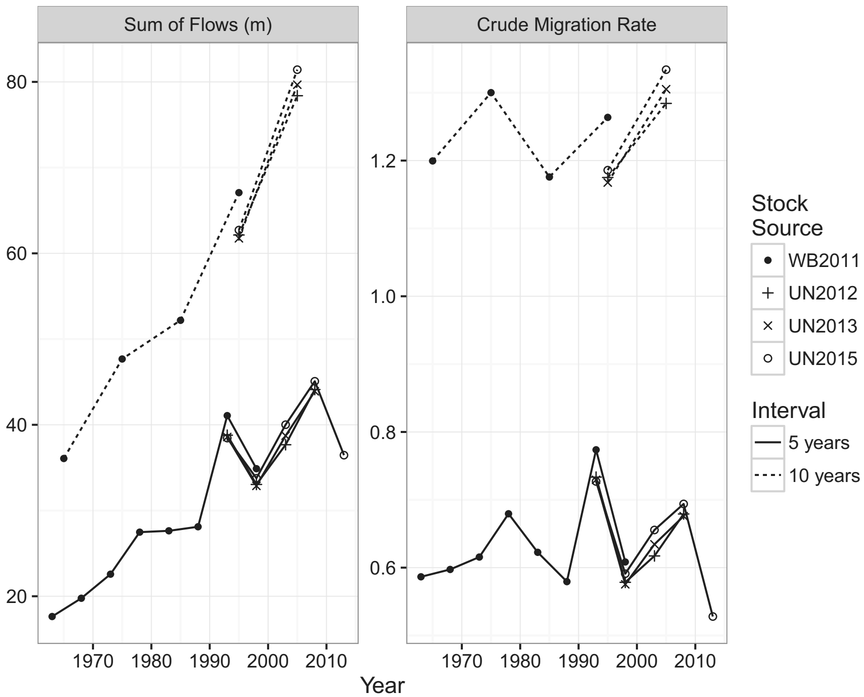
The culmination of the country-to-country flow estimates varying by different gender, time period, interval length, and stock and demographic data, provided a combined data set with over 10 million entries. The results in this section are first discussed with regard to summary statistics of the flow tables. Then, the bilateral patterns as well as immigration and emigration trends are summarized at the regional level. Full estimates of the country-to-country flows are available in the Appendix S2. Note, throughout the remainder of this article, when referring to an estimated flow, the estimate has the properties outlined in the methodology section, namely a minimum number of migrant transitions required to match the changes in the given stock data, controlling for births and deaths in each country over the period. The true migrant transition flow may well be higher, and an estimate itself is subjected to errors propagated from varying degrees of inaccuracy in the stock or demographic data as well as the inherent assumptions in the methodology used to estimate the flow, some of which are discussed at length in the next section.

Global-Level Summary Statistics

In Figure II, summary statistics for estimated global migration flows over time are displayed using the *ggplot2* package (Wickham 2009) in R (R Development Core Team 2016). The symbol type of each point corresponds to the stock data source used as input data when estimating the flow table.

The estimated sum of the number of migrants for each of the 31 flow tables that used WPP2015 input data is shown on the left-hand side. An upward trend in the global level of migrants over time is apparent. The upper lines are based on the total flows over 10-year period, plotted at the mid-decade point on the horizontal axis. In the 1990–2000 period, when an estimate of flows from both the World Bank and United Nations is available, the total flow from the World Bank stock data is 67.08 million people, 4.36 million higher than the estimate from the UN2015 data. Estimates from the UN2012 and UN2013 during this period are within a million migrants of the UN2015-based estimate. The range of the estimates is wider for the flows during the 2000–10, with a high of 81.42 million based on the UN2015 data and a low of 78.39 million from the UN2012 data. The lower lines represent totals from flows over five-year periods, plotted at the midpoint of the corresponding period on the horizontal axis. A sharp rise in the total amount of migrants during the 1990–1995 period was evidently driven by a number of factors including increased moves between countries of the former USSR around the fall of the Iron Curtain. Large flows are also estimated

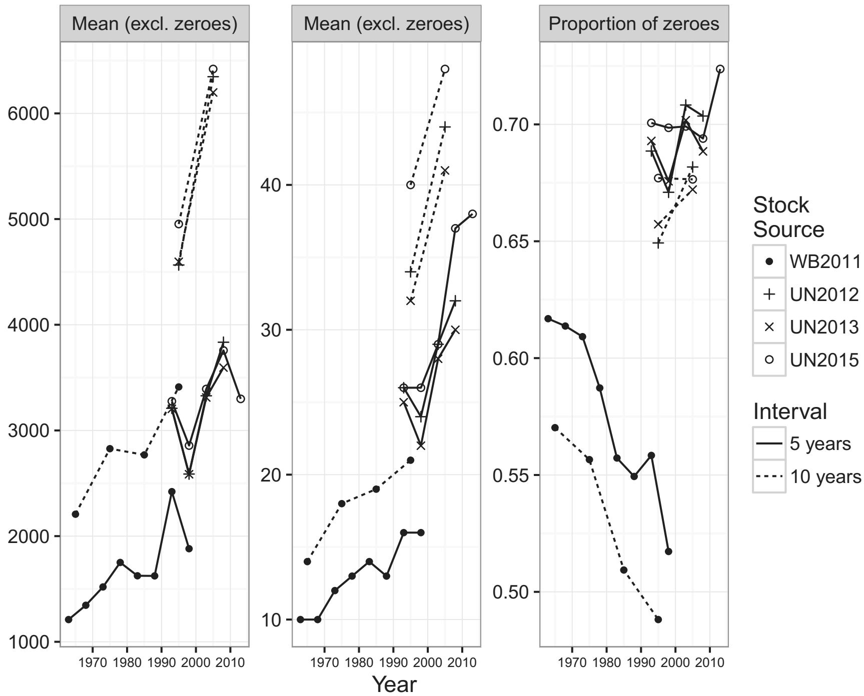
Figure II. **Total Estimated Country-to-Country Bilateral Flows and Crude Global Migration Rate Varying by Stock Data Source and Interval. All Based on WPP2015 Demographic Data. On the Horizontal Axis, Points are Plotted at the Midpoint of Their Corresponding Interval**



from countries that were experiencing armed conflicts during the period, such as Kuwait, Rwanda, Afghanistan, and Liberia. These flows are not fully captured in the 10-year interval estimates for 1990–2000, where, for example, crisis migrants might have returned to their original place of residence by the end of the period. During the most recent period, 2010–2015, large labor-related flows from Latin America to North America, parts of Asia to the Gulf States, and flows into Europe from both Asia and Latin America all fell, contributing to a decrease in the estimated number of global migrant flows.

The right-hand side of Figure II illustrates the percentage of the population that were estimated to migrate during the relevant interval derived by dividing the sums on the left-hand side by the populations in each origin at the beginning of the corresponding time interval. The

Figure III. **Summary Statistics for Estimated Country-to-Country Bilateral Flows Varying by Stock Data Source and Interval. All Based on WPP2015 Demographic Data. On the Horizontal Axis, Points are Plotted at the Midpoint of Their Corresponding Interval**



percentage remains relatively constant, at around 1.25 percent for migrant transitions over a 10-year interval. The estimates based on five-year intervals also remain fairly constant at around 0.65 percent, except during the 1990–1995 period.

Figure III illustrates further summary statistics for the estimated bilateral tables. On the left-hand side is a plot of the mean of nonzero estimated flows in each period. The mean flow size follows a broad upward trend over time. Nonzero flows based on UN stocks are higher on average than the flows derived from World Bank stocks during the 1990s. This difference occurs for a couple of reasons. First, the number of nonzero estimated flows are not constant across time, as illustrated in the plot on the right-hand side of Figure II. Zero flow estimates are directly related to the number of zeros in the stock data. If a foreign-born stock in a particular country is zero at both the beginning and the end of the period, the resulting estimate of flows will also be zero, as there is no change in the foreign-born stock over the time period. In the World Bank stock data, 60 percent of bilateral foreign-born stocks are zero in 1960. This percentage falls to 45 percent by 2000. The number of zero flow estimates from the World Bank stocks follows a similar decline. In the older versions of the UN data stock data, approximately 70 percent of stock estimates are zero throughout the data period. For the latest UN2015 stock data, the number zeros are slightly higher, around 75 percent at each time point. The flow estimates from the UN also contain similar levels of zeros.

The second cause of differences in the mean flow is due to the variation in the number of countries included in the estimated tables. Origin–destination flow estimates based on the World Bank stock data are obtained for 196 countries where both demographic data (WPP2015 in this case) and stock data are available. In comparison, estimates based on the UN2012 stock data are possible for 197 countries. Of these, 195 were common to all sets of estimates.[5] Estimates based on the World Bank stocks included an additional country, Taiwan, while estimates based on the UN2012 stocks included two additional countries, the Channel Islands and Western Sahara. Estimates based on UN2013 and UN2015 stock data cover 198 countries, the same 197 as the UN2012 plus Curacao. In the 2010–2015 period, estimates based on the UN2015 data include 200 countries, as separate estimates for bilateral flows to and from Montenegro, Serbia, Sudan, and South Sudan at the start and end of the period are available. In previous periods, only data for the previously unified countries were available.

The estimated median of the nonzero flows is shown in the middle panel of Figure III. This broadly follows a similar pattern as the mean, although at much lower levels indicating a large skew in the distributions of estimated global bilateral flows toward smaller counts.

Bilateral Patterns

In order to illustrate the pattern of estimated bilateral relationships, a set of six circular migration plots are shown in Figure IV. Plots were created

[5]Bilateral stocks were available for the aggregation of Serbia and Montenegro and Sudan and South Sudan in both the World Bank and UN data. The corresponding demographic data were derived from the aggregation of the individual country information provided in each WPP.

in R using the *circlize* package (Gu et al. 2014). The direction of the flow is indicated by the arrowhead. The size of the flow is determined by the width of the arrow at its base. Numbers on the outer section axis, used to read the size of migration flows, are in millions. Each plot is based on flows over a 10-year period, aggregated to selected regional levels.

The first four plots (a–d) are flow estimates based on World Bank stock data. In the first period, the largest estimated flows occur within the defined regions (Eastern Europe and Central Asia, 5.45 million; Europe, 4.77 million). Many migration flows within the first of these regions were not international moves at the time, such as Russia to Ukraine (0.99 million) or Russia to Kazakhstan (0.87 million). Then, total estimated flows during 1970–1980 increased globally from the period before, as discussed in the previous subsection. Although this increase in size is difficult to view from comparing circular migration plots in Figure IV (a) and (b), changes in the share of global migration flows between selected regions can be easily detected. Most noticeable is a large increase in the share of global migrants moving within Southern Asia. During 1970–1980, 4.37 million migrant flows were estimated from Bangladesh to India and another 1.76 million from India to Pakistan most likely driven by the Indo-Pakistani War of 1971.

Changes in the sizes of regional migration flows over time are more easily viewed in Figure V, which provides plots of estimated immigration and emigration totals by United Nations Population Division demographic regions.[6] At the country level, the estimated net migration, obtained from differencing the immigration and emigration values, matches those from the demographic data, due to the method used to control for births and deaths over the period (discussed in the Appendix S1). In the first two time periods in Southern Asia, there is a sharp rise in immigration and emigration whereas the net migration, the gap between the immigration and emigration lines, during the same period is almost constant. Further changes in the global bilateral flows are apparent from comparing Figure IV (a) and (b). Estimated flows into and within Europe during 1970–1980 decreased from the decade before. Sizeable migration into West Asia from countries such as Egypt (0.39 million) and India (0.16 million) to Saudi Arabia began to develop. Migration within Africa also increased, including large flows out of

[6]Except for Polynesia, Melanesia, and Micronesia which are aggregated to a Pacific Island region.

Figure IV. Estimated 10-Year Bilateral Migrant Flows Over Time Aggregated to Selected Regions. All Based on WPP2015 Demographic Data

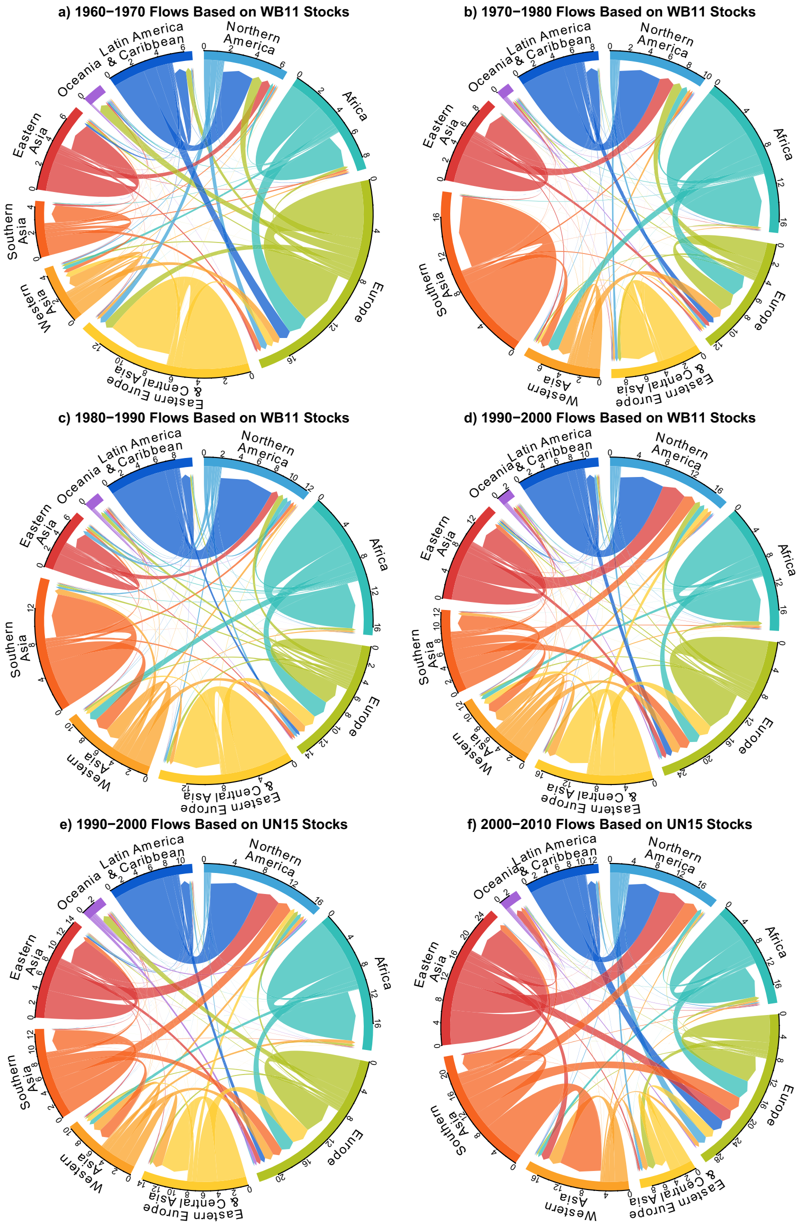
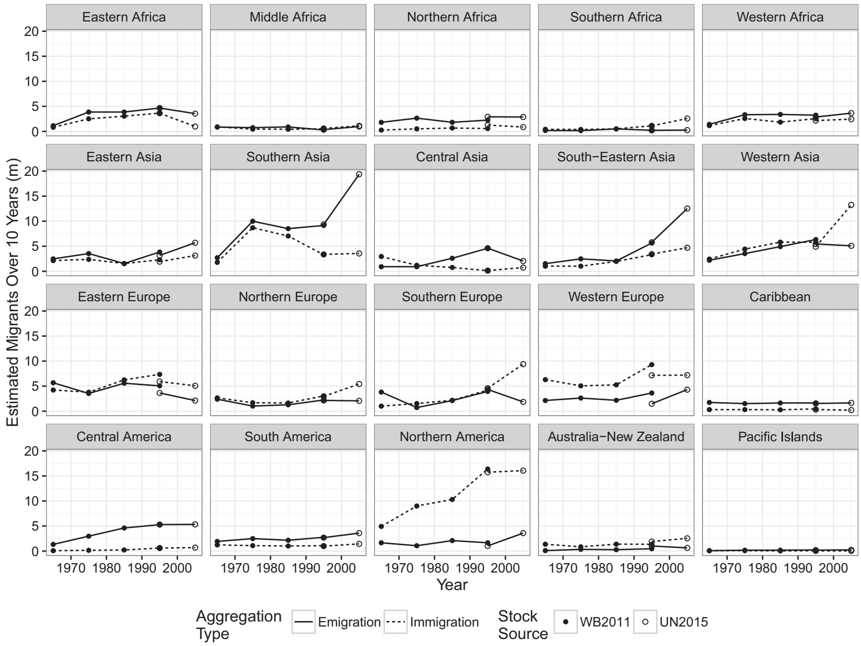


Figure V. Total Estimated Immigration and Emigration Flows Over 10-Year Periods. All Based on WPP2015 Demographic Data. On the Horizontal Axis, Points are Plotted at the Midpoint of Their Corresponding Interval



Ethiopia (0.95 million to Somalia) and Burkina Faso (0.41 million to Ivory Coast).

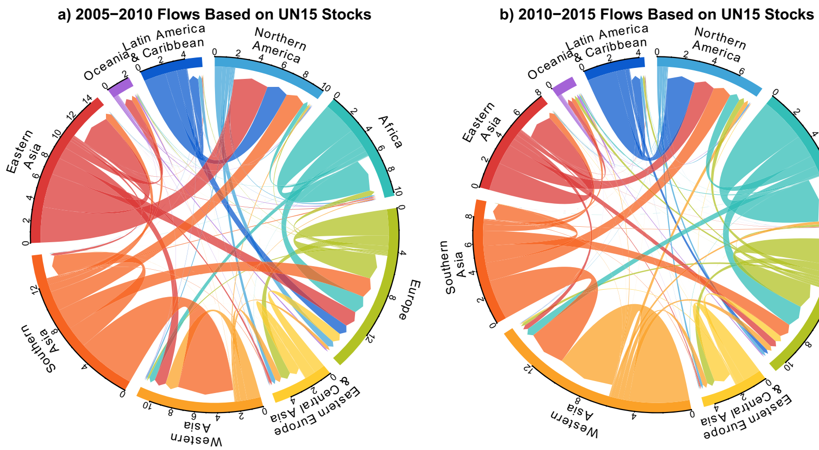
Estimated flows during 1980–1990 increased in most regions in comparison with previous periods. Most noticeable is the further rise in flows from Latin America and the Caribbean to North America in Figure IV (c) in comparison with (a) and (b). The largest flow during the period was estimated from Mexico (3.09 million). The number of migrant transitions within Eastern Europe also increased, including 1.03 million from Ukraine to Russia.

In Figure IV (d) and (e) are circular migration flow plots based on estimates during 1990–2000 period using different stock data sources. In plot (d), estimates based on the World Bank stock data are shown. The level of immigration in North America (also shown in Figure V) is estimated to increase from a wider variety of origins, including Eastern and Southern Asia. Migrant flows into Europe, especially from other European countries increased, as do immigration into West Asia. The plot of the estimates during the same period, but based on the UN2015 stock data, is shown in Figure IV (e). Many of the same estimated bilateral flow patterns are similar, as a share of the global migration system, to those based on the World Bank data in (d). However, some distinct differences in the size of flows are apparent from the immigration and emigration summary plots in Figure V. In Western and Eastern Europe and Western and Southern Asia, there are some large disparities in the level of the total immigration and emigration flows based on different input stock data. In all but the last of these regions, estimates from the World Bank stock data result in higher flow levels. The differences in the estimates are driven by larger changes in the foreign-born stock values provided by the World Bank in 1990 and 2000 in comparison with those of the UN stock data. For example, the largest estimated flow into Europe based on the UN2015 stock data is from Kazakhstan to Germany, to match an increase in Kazakh-born residents in Germany (10.2 thousand in 1990 to 487 thousand in 2000). In comparison, the same foreign-born stock in the World Bank data increased from 18.9 to only 21.4 thousand over the same period, resulting in a much smaller estimated flow.

The circular migration flow plots related to the final 10-year time period between 2000 and 2010 is shown in Figure IV (f). Based on UN2015 stock data, there are further increases in immigration flows into North America from Asia and into Europe from Asia, Africa, and North and Latin America. Some of the largest increases of estimated flows into Europe are into Southern European countries, as shown in Figure V, the largest being 0.62 million from Morocco to Spain. There are also sizeable increases in the estimated flows from South American countries such as Bolivia and Colombia into Southern Europe. Immigration into West Asia further increases, as do migrant flows within South-Eastern Asia, including an estimated 1.42 million people moving from Myanmar to Thailand over the 10-year period.

In Figure VI, circular migration flow plots for the two most recent five-year periods are given. As shown in Figure II, estimated migration flows dropped considerably from 45.08 million during 2005–10 to 36.46 million during 2010–2015. The origin–destination patterns also underwent some considerable change. For example, large flows within Western Asia appear in 2010–2015 based on migration out of Syria to Turkey (1.51 million) and Lebanon (1.22 million). In contrast, flows into Europe from Latin America and Eastern Asia fell sharply, from 1.06 and

Figure VI. **Estimated Five-Year Bilateral Migrant Flows in Recent Time Periods, Aggregated to Selected Regions. All Based on WPP2015 Demographic Data**

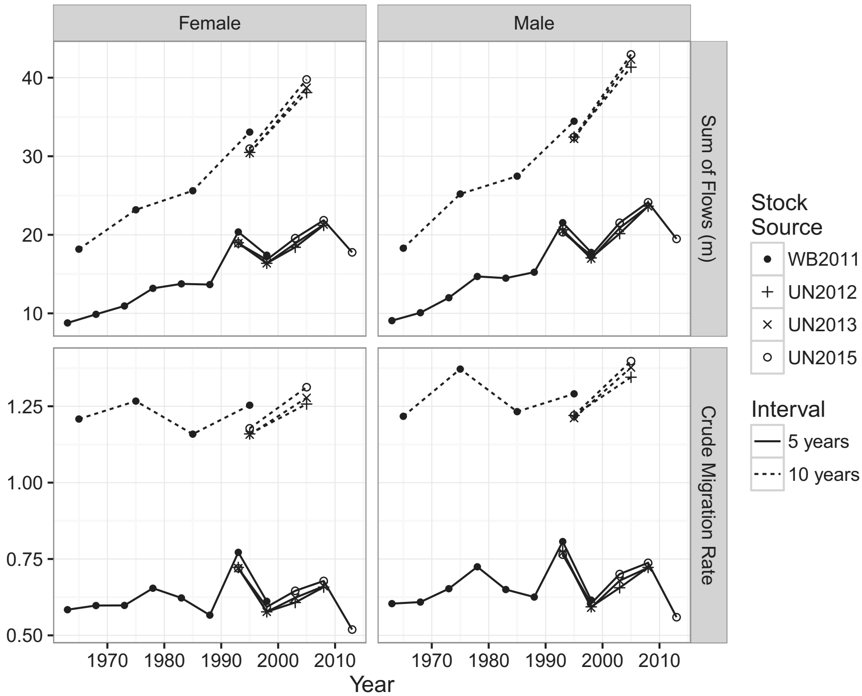


1.63 million to 0.30 and 0.57 million, respectively, driven by reduced flows into Southern European countries such as Spain. Similar drops were also estimated into Northern America, where migration flows from Eastern Asia fell from 3.40 to 1.59 million. Migration from South Asia to Western Asia also decreased, where, for example, the estimated migrant flows from India to the United Arab Emirates fell from 1.38 million during 2005–2010 to 0.45 million during 2010–2015. Widely reported flows of refugees into Europe during the summer of 2015 do not appear in the estimates. This is due to the UN2015 stock data, which (1) is based at mid-year 2015 and (2) does not show any large increase in populations born in conflict-afflicted countries. For example, the UN stock has 51,330 Syrian-born migrants in Germany in 2010, with only a small increase to 53,099 by 2015.

Flows by Gender

Female and male total flows and crude migration rates are shown in Figure VII. The patterns of both statistics follow similar paths as the estimates based on both genders combined, discussed in the previous subsections. The sums of male flows are slightly larger in most time periods. During 2000–2010, male flows increased faster than the females,

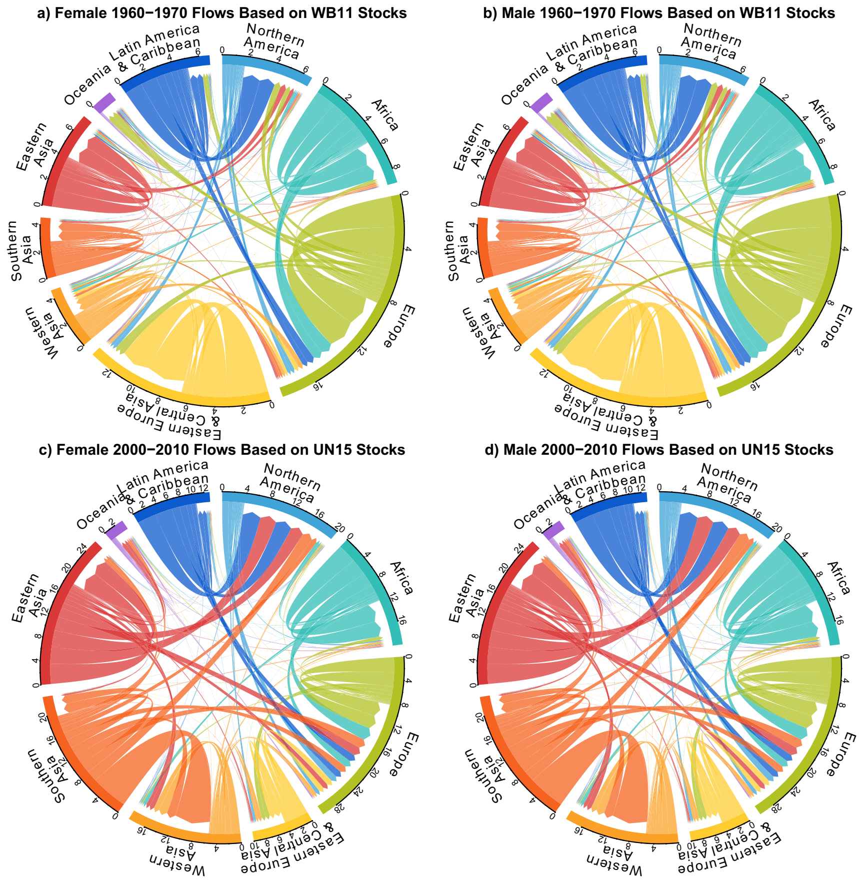
Figure VII. **Total Global Migration Flows and Crude Rate for Estimated Country-to-Country Bilateral Flows by Gender. All Based on WPP2015 Demographic Data. On the Horizontal Axis, Points are Plotted at the Midpoint of Their Corresponding Interval**



reaching their peak of 42.96 million compared to a female total of 39.79 million (based on UN2015 stock data and WPP2015 demographic data). The disjoint between the World Bank and UN stocks that was apparent for the total flows is also evident in the gender-specific flows. Note, the combined total of male and female flows does not match the total based on the estimated migration flows from the non-gender-disaggregated stock data presented above (e.g., 81.42 million for 2000–2010 based on UN2015 stock data and WPP2015 demographic data) as the estimation methodology is run separately for each of the total, male, and female sets of migrant stock tables.

Selected circular migration flow plots for both estimated males (left) and females (right) are shown for two time periods in Figure VIII. Estimates are based on gender-specific stock and demographic data. In each of the time periods, both male and female migration patterns are broadly

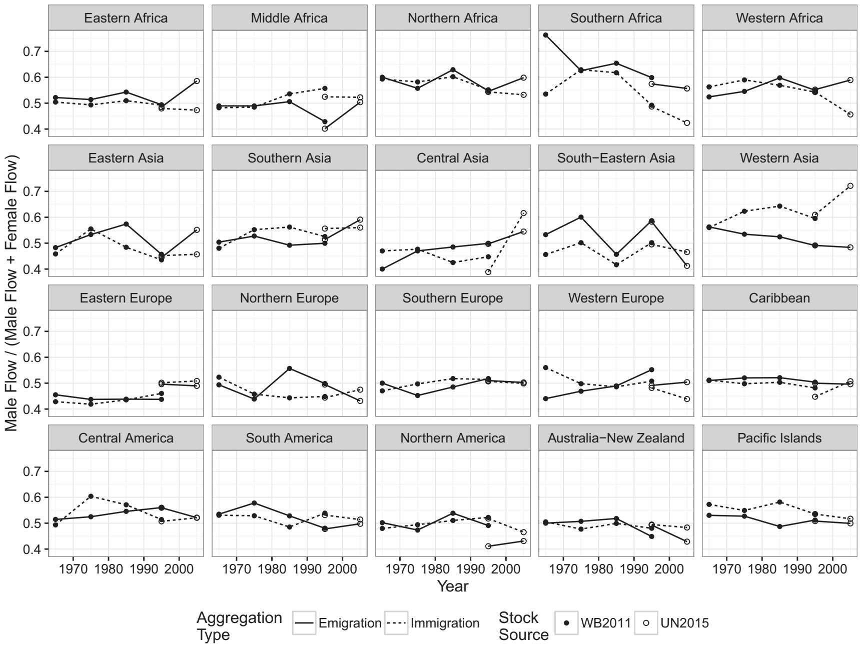
Figure VIII. **Estimated 10-Year Bilateral Migrant Flows Over Time Aggregated to Selected Regions. All Based on WPP2015 Demographic Data. On the Horizontal Axis, Points are Plotted at the Midpoint of Their Corresponding Interval.**



similar. However, in particular periods and regions, some distinct differences occur. The changes are more clearly illustrated using a plot of the proportion of male to female estimated 10-year migration flows for each region over all time periods shown in Figure IX.

Male-dominated flows (where the immigration or emigration lines are above 0.5) occur almost entirely throughout the period for migration in and out of Northern, Southern, and Western Africa, as well as for flows into Western Asia. With the exception of migrant flows into South-

Figure IX. **Male Percentage of Total Estimated Immigration and Emigration Flows Over 10-Year Periods. All Based on WPP2015 Demographic Data. On the Horizontal Axis, Points are Plotted at the Midpoint of Their Corresponding Interval**



Eastern Asia, female-dominated flows during the entire period are less common. In most regions, the shares of the estimated male to female migrant flows do not to follow any clear and consistent patterns. Two particular data points stand out when considering the data in Figure IX. First, in Southern Africa, estimated flows during the 1960–1970 period are overwhelming male. This is predominately due to greater increases of the male stocks of people born in Lesotho, Swaziland, and Namibia residing in South Africa, creating larger estimates of male flows, where similar changes in female stocks do not occur. Second, in particular oil-rich Gulf States, large male immigration flows are estimated in 2000–2010. As shown in the circular migration plots of Figure VIII (c), these flows are predominately from Southern Asia, South-Eastern Asia, and other countries in West Asia and Africa, where similar strong bilateral links are not present in the female plot of (d).

SENSITIVITY ANALYSIS

Estimates of migrant flows from stock data can potentiality be sensitive to the input data used in the methodology. For example, in Abel (2013a) a handful of unexpected flow estimates result from peculiarities in the input stock data were highlighted. The same unexpected flows are also found in the estimates presented in this study using the updated methodology. For example, in 1960 there were a reported 1.5 million Chinese-born in Hong Kong. This stock drops to 16,823 in 1970 and rises back up to almost 1.9 million in 1980. This dramatic change in the reported stocks creates a large estimated outflow of Chinese in the 1960s. These emigrants are estimated to migrate to countries where there are increases in the number of Chinese-born, including but not exclusively, China. In turn, during the 1970s, there is a large estimated inflow back into Hong Kong of Chinese-born, to meet the sudden increase in their migrant stock.

Unexpected estimates such as these have potential advantages and disadvantages. On the one hand, the flow estimate methodology combines demographic and stock data in an enhanced global demographic accounting system. The unusual flow estimates are easier to detect than separate comparisons of stock data at the beginning and end of the period or with net migration estimates from demographic data. Once identified, the input data that generate the unusual flow estimate might be of use to data producers when considering future revisions to their stock or demographic estimates. On the other hand, potential users of the flows should be aware that the estimates are not always close to the true flows as they are constrained to match imperfect demographic and bilateral stock data. In this section, a further analysis of the sensitivity of estimates to alternative stock and demographic data is studied. This is followed by a comparison between the flows presented in the previous section with those adjusted for changes in political geography.

Bilateral Stock Data Sources

As shown in Figures II and III, there are some differences between the summary statistics estimated from the previous version of the UN stock data. Figure IV shows in more detail the differences in bilateral migration flow patterns estimated from the WB2011 and UN2015 stock data during the overlapping 1990–2000 period. In Figure S2, regional bilateral estimated flows from each of the four bilateral stock data tables available at the beginning and end of the 1990–2000 are plotted alongside each other. There appears to be only minor differences in the bilateral patterns from UN migrant stock data (b–d). However, distinct differences occur when comparing estimates from the World Bank (a) to those from the United Nations. These stem from a number of causes including an earlier release data by the World Bank, different methods to impute missing bilateral stock data and a slight difference in the set of countries covered in the World Bank and UN data sets.

Demographic Data Sources

The UN Population Division updates demographic estimates for all countries every 2–3 years. The results presented so far have all been based on the WPP2015 version. Total migration flow estimates and crude global migration rate based on the WPP2010 and WPP2012 are shown by the dashed lines in Figure S3.

The total flows from the WPP2015 data are represented by the solid line and match those in Figure II. The updated demographic data have a noticeable effect on the total estimated flows of both the 10-year and five-year interval estimates during the last decade. For example, the estimate of all flows for the 10-year interval 2000–2010 is 78.30 based on WPP2010 and 82.96 based on WPP2012, compared to 81.42 million in the WPP2015 version. Differences in the totals are partly due to a non-constant number of countries used. In general, the more recent WPP data have allowed for more countries to be included. For example, the 2000–2010 estimates based on UN2015 stock data and WPP2015 involve 198 countries, whereas the WPP2010 version includes only 194 countries. In earlier periods, the effect of alternative demographic data had little impact on the total estimated flows. This is not too surprising. Revisions to demographic data tend to be larger in more recent periods as more up-to-date estimates are obtained from census and surveys. In order to detect the regions where the data revisions have the largest impact on estimated migration flows, Figure S4 plots differences of both immigration and emigration by region.

Some of the largest differences between the estimates from alternative demographic data sources appear in the estimated migrants flows from Southern Asia during 2000–2010 and to the region during 1990–2000. These were due to revisions in the demographic data, predominately the population total in 2000, which was revised down by 8.27 in WPP2015 compared to WPP2010. In the later period, 2000–2010, this alteration is matched in the flow estimation, in both some increase in immigration from 2.85 million (estimated from WPP2010 data) to 3.55 million and a larger climb in the emigration, from the 15.80 million estimated using WPP2010 data to 19.37 million from WPP2015.

In some regions, such as Northern Africa, Northern America, or Southern Europe, the choice of demographic data leads to different immigration and emigration estimates, depending on the period at hand. As shown in Figure S4, these differences tend to be less than a million either way. For other regions, such as Western Africa, Northern Europe, Western Asia, or South-Eastern Asia, the demographic data used only have an effect on estimates during the later time periods. In other regions, such as the Caribbean, Middle Africa, or Australia–New Zealand, the demographic data have very little effect on the gross number of immigrants and emigrants estimated.

At the country level, estimates can also be sensitive to alternative demographic data. One of the most prominent examples is the flows into and out of Russia. In the WPP2010 data, Russia had a positive net migration of 2.7 million over the 10-year period while in WPP2015 the value increased to 3.89 million. This revision is included in the flow estimation procedure through the demographic input data via a higher 2010 population (revised up by 0.2 million) and a lower number of deaths (revised down by 0.51 million). Consequently, larger flows of Russian-born abroad are estimated to return to match the greater native-born population in Russia. The biggest of the estimated flows come from countries with high Russian-born populations predominately in other Eastern European and Central Asian nations, as well as the United States (301 thousand, up from 56 thousand for flow estimates based on WPP2010 data) and Germany (531 thousand, up from 58 thousand).

Contradictions between the input demographic and stock data were discovered due to the unexpected estimated flows they produced. In the remainder of this section, a couple of these are highlighted. First, during 2005–2010 the demographic data imply net migration for Poland of +55 thousand (WPP2010) or −70 thousand (WPP2012). These differences somewhat contradict the large increases in the UN stock data of Polish-born in major destinations countries over the same period, such as the United Kingdom and Germany (increases of over 300 thousand in each country). As the estimation methodology is a crude global demographic account, the increases in Polish stocks in the United Kingdom and Germany are matched with estimated flows from reported decreases in Polish-born populations in the stock data, mainly in France, the United States, and Canada. Only small amounts of flows from Poland to the United Kingdom or Germany are estimated when the WPP2012 data are used, as the methodology is constrained by the population, birth, and death data to allow only 70 thousand migrants to leave Poland over the period. Second, in the United Arab Emirates, the total male population given in WPP2010 is 5.22 million whereas in the UN2013 and UN2015 stock data the male foreign-born populations are 5.46 million, 0.24 million higher than the total population. With these combinations of demographic and stock data, the flow estimation procedure produces negative flows. When using the more recent versions of demographic data in combination with UN2013 and UN2015 stock data, the total male population is higher than the reported male foreign-born population, leading to plausible flow estimates.

Changes in Political Geography

The estimates presented thus far are based on the availability of information from both migrant stocks and the demographic data. The results are flows over sets of countries with two noticeable features. Firstly, both historical migrant stocks and demographic data are provided for countries which at given periods of time might not necessarily be fully fledged separate nation states. For example, past bilateral migrant stock information is provided by the World Bank for what were at the time republics of the USSR. This results in estimates of international migrant flows into, out of, and between Soviet Republics which at the time could be considered as internal migrant flows. Secondly, the set of countries used in the UN stock data only provides information on new countries in 2010. As a result, separate estimates, in, out, and between both Serbia and Montenegro and Sudan and South Sudan could not be obtained as there were no foreign-born stock data for these new countries in previous decades.[7]

In order to analyze the effect of the first of these features, changes in political geography, estimates of flows which at the time would be

[7]Estimates for flows in Abel and Sander (2014) incorrectly treat UN stock data for Serbia, Montenegro, Sudan, and South Sudan in 1990 and 2000 as separate countries.

considered internal migration can be set to zero. Then, flows into and out of the old set of unified countries can be aggregated, resulting in a new set of bilateral flow estimates between a set of countries that varies over time.

This procedure was implemented for estimates before 1990 for the split of the former USSR into 15 countries, as well as Yugoslavia into Bosnia and Herzegovina, Croatia, Serbia and Montenegro, Slovenia and Macedonia, and Czechoslovakia into the Czech Republic and Slovakia. Flow estimates between a unified Eritrea and Ethiopia as well as Namibia and South Africa before 1990 were also set to zero. Estimates before 2000 were adjusted to combine Timor-Leste with Indonesia. Other potential adjustments, such as Bangladesh and Pakistan before 1970 or former European colonies with their ruling governments, are not implemented, as the resulting estimates would imply an internal migration between non-contiguous areas.

In Figure S5, the total flows for estimates adjusted for changes in political geography are plotted using a broken line and the original estimates with a fixed set of countries throughout the period are plotted using the solid line. In comparison with the total flows based on the fixed set of countries, the adjusted estimates previous to 1990 are lower. Estimates of both the five- and 10-year flows during the earlier periods are less smooth. Instead, global migrant flow numbers remain somewhat level during the late 1970s up until the late 1980s. Consequently, the percentage of estimated migrants, shown in the bottom panel of Figure S5 during this period, falls more sharply than estimates based on a fixed number of countries.

VALIDATION

As there is no existing data set on past bilateral migration flow between all countries, any comprehensive validation of the estimates presented in this study is difficult. Further, the estimated net migration for each country matches those from the United Nations as (1) the input demographic data on population, births, and deaths from the UN sum to the UN net migration data and (2) the changes in migrant stock data and native-born population from births and deaths during each period are accounted for in the flow estimation procedure to ensure that their residual net migration is not altered. Nevertheless, for a few countries, past data on immigration and emigration flows exist. In Figure S6, the proportion of immigration flows by continent as reported by each destination country is plotted against the estimated five-year flows based on changes in migrant stock data and the WPP2015 demographic data. As the reported immigration data are provided for each year, proportions based on a five-year average are taken. Shading represents the time period of the comparison, where a deeper shade represents more recent data points. A diagonal line is plotted for each country to indicate where there is a perfect agreement of the proportions in the reported data and estimates calculated in the previous section.

The immigration flow data on the vertical axis are taken from the United Nations Population Division (2015a) which is based on data collected by national statistical offices. Unlike other estimates of migration flows such as Raymer et al. (2013), it covers non-European countries and has a relatively long history. However, as noted earlier, there are a number of challenges with collections of data taken from individual nations. For example, a wide variety of definitions are used which precludes direct comparisons on the level of flows, hence the use of proportions. In each country, the duration of stay is either permanent, one year or less, such as six, three, or one month. Furthermore, in some countries, there are large counts of immigrant flows with unknown origins but are included in a total flow value. For the data in Figure S6, the origin in most countries is based on the previous residence of migrants, but in some cases the country of birth or citizenship (i.e., a stock measure) is used by the reporting countries. Despite these important differences, there is some broad agreement between the origin proportions into each country, where many points are close the diagonal lines in Figure S6. There is an overall correlation of 0.77 between the proportions from the reported immigration data and the estimates. A few notable exceptions can be seen, such as the proportion of flows from Europe to Czech Republic, Finland, Germany, Iceland, Italy, Romania, and Latvia during various periods, where the estimates based on the changes in stocks are smaller than those reported by the individual countries.

In Figure S7, the proportion of emigration flows by continent, as reported by each origin country, is plotted against the estimated five-year flows based on changes in migrant stock data and the WPP2015 demographic data. There are fewer countries reporting emigration data; hence, there are fewer comparisons to be made. However, in most cases there is an agreement between the proportions, with an overall correlation of 0.91 between the proportions from the reported emigration data and the estimates. Notable exceptions include the lower proportions in emigration flows estimated from changes in stocks to Europe and Asia from the United Kingdom.

SUMMARY AND DISCUSSION

Global international migration is an ever-changing process. Migrant stock data, based predominately on a single transition from the place of birth to the place of residence, only manages to capture part of the dynamic nature of international migration. Indirect estimates of flows provide a more robust basis to understand contemporary migration patterns during a given period, where no comprehensive source of global migration flow data exists.

In this study, global bilateral flow tables were estimated by gender from 1960 through to 2015. Results were predominantly presented by region (the estimates for all countries are available in the Supporting informations). The total estimated international migration flows over time are shown to generally increase. The percentage of the global population estimated to migrate over five- or 10-year period remained fairly steady at 0.65 and 1.25 percent of the global population, respectively, with a noticeable spike during the 1990–1995 period. Some regions were estimated to have continuously increasing numbers of migrants arriving (North America and West Asia) or leaving (Central or South America), while others showed fluctuating patterns over the time period. The patterns of bilateral flows also varied across regions and time. For example, a growing numbers of migration flows were estimated from South to West Asia and from Asia to North America, most likely related to economic changes. Large migrant transitions were also estimated in selected time periods within Africa or Eastern Europe during times of armed conflicts or political change. These results concur with the findings from Czaika and De Haas (2014). In their analysis of the World Bank stock data, they found global international migration had not accelerated in recent decades but rather shifted in directions linked to major geopolitical and economic shifts.

During 2010–2015, the estimated global migration flow fell from previous five-year periods. At first glance, this finding is different from those in recent reports by the World Bank (2016) and United Nations Population Division (2016), which claim an all-time high level of international migration at the beginning of 2016. However, their findings are based solely on peaks in migrant stock data, the culmination of migration flows over individual's life times. Without flow data, it is not easy to detect how much of this record high is due to increases in the global population level (also at a record high) rather than a rise in migration. As illustrated in this study, flow estimates based on changes in the stock data between 2010 and 2015 indicate a decreasing number of migrants transitions out of Latin America and Asia toward Southern Europe, North America, and Gulf States compared to previous five-year periods. These decreases are not fully compensated for through increased estimated flows elsewhere, such as out of Syria to neighboring countries. The flow estimates during the most recent five-year period use recently published UN stock data, which is based on the mid-year populations. The stock data do not intend to reflect some of the large movements of refugees into Europe during the summer of 2015, after the end of June. Future revisions in the stock data might include some increases in foreign-born populations; however, the bulk of recent flows is likely to be captured in the next five-year period if they remain outside their country of birth by 2020.

Overall, the trends in global migration flow for each gender followed similar paths, with slightly higher male's shares throughout. In recent periods such as 2000–2010, the male flows were relatively larger than female flows during the same period. This imbalance at the global-level flows is predominantly due to large increases in male migrant stocks of South Asians in oil-rich Gulf States. It partly counters the common narrative regarding the feminisation of global migration which was based on older migrant stock data. However, beyond the overall total, the bilateral patterns of estimates in this study show that the gender composition of migration to or from some regions varies, where in some regions female flows are more dominant. Similar findings are discussed in Kofman (2000) based on a range of stock and flow data over an extended time period. The differences in the estimates from this study could be further explored in future work using regression modeling techniques based on the estimates presented in this study and potentially related explanatory variables. Information on the growth and size of specific migration types such skilled labor, construction, or domestic work would be particularly helpful in this exercise.

There are some important characteristics of the migration estimates that should be noted by potential users. There are many ways to measure migration. As highlighted in the methodology section, the estimated flows in this study are based on the absolute minimum number of migrant transitions required to match the demographic data and changes in migrant stocks. Consequently, a large proportion of the estimated flows were zero and the overall distribution of nonzero flows were skewed toward small counts driven by the large numbers of zero bilateral migrant stocks in both the World Bank and UN stock data. There is no obvious empirical measure to help inform the estimation method on both the number of non-migrants in each period and migrant stock to allow for a higher measure of migrant transitions to be estimated (beyond the minimum). Alternative migration measures, such as the number of movements, cannot be simply derived from migrant transition estimates. Circular or return migration over a short period is not being captured, such as away from a person's country of birth for a couple of years to then return home, unless the individual is captured in the foreign-born stock data when abroad. The use distance measures for the offset term was found to have only a very minor impact on the estimated results.

The estimates of bilateral migration flows presented provide a number of new insights to the global migration flow literature. First, the methodology of Abel (2013a) and Abel and Sander (2014) is extended and applied to cover a wider time period and to estimate migrant flows separately by gender. Second, the presented bilateral flows provide an updated view of international migration to those outlined by Zlotnik (1999) and National Research Council (2000), who used a patchwork of migration data. While the patterns and drivers of migration flows are discussed in far less detail in this study, many of the observed trends in previous studies are represented in the estimates. Third, estimates of bilateral flows provide a far greater depth to the understanding of international migration than can be obtained from net migration measures. As Rogers (1990) details, net migration statistics are fundamentally flawed as they are based on a nonexistent population. Furthermore, they are sensitive to changes in both immigration and emigration patterns, and hence, their time series are often volatile. Fourth, while 10-year estimates based on the World Bank data are comparable to those of Abel (2013a), the estimates in this study use an extension of the methodology developed in Abel and Sander (2014). This extension results in different estimated bilateral flow tables, with the net migration matching those of the United Nations Population Division (2011, 2013b, 2015c). Fifth, in countries and time periods where reported migration flow data exist, comparisons are made with the estimates. There is a broad correspondence between the proportion of migration to and from each continent in the selected countries where data for a validation exercise exist.

The choice of input data used was found to have an influence on the estimated flows. Stock data were found to have an impact on the average estimated flow size and the number of zero flows. These are caused by the different methods used by the United Nations and World Bank to provide complete and comparable bilateral stock estimates. Demographic data were found to have a strong influence on the scale of migration flows particularly for estimates during more recent periods where data revisions to demographic data were greater. Flows were estimated over both five- and ten-year periods in order to quantify migrant transitions over a variety of timescales. Estimates over five-year periods were found to detect large migration flows, such as those induced by armed conflicts or political changes that were not as clearly identifiable for estimates over 10-year periods.

Estimating flows from changes in the stocks and controlling of births, deaths, and population sizes forms a crude global account of demographic data. This account allows a comprehensive system to compare global demographic data for inconsistencies, check for errors, and match available data with conceptual models or known migrant flows. Unusual data points in either the bilateral stock or demographic data are be carried forward to the flow estimates. As a result, a number of unexpected estimates were uncovered. For example, large flows of Chinese-born were estimated to go in and then out of Hong Kong during the 1960s and 1970s. Revisions in demographic data resulted in large increase in the stock of the native-born population in Russia, which was accounted in the methodology through increased return flow of Russian-born migrants elsewhere. Initial estimates for flows into and out of the State of Palestine were implausible due to the size of incompatibilities between the demographic and stock data. This deficiency was handled by an extension to the methodology outlined in the Appendix S1. In both of the latter two cases, the unusual flow estimates are derived from underestimates of the native-born populations, which are in turn derived from the difference in the total and foreign-born populations. It is most likely that the source of the underestimation lies with counts of foreign-born rather than the total population, as the demographic estimates of population size are relatively simpler to produce. Future World Bank stock estimates over extended periods might also employ improved estimation procedures, as used for more recent estimates of bilateral stock tables by gender and skill level (Artuç et al. 2015). In other cases, as with the estimates of recent Polish migration flows, there appears to be a contradiction between the stock and demographic data, caused by a lower-than-expected net migration estimate. At a bare minimum, the methodology in this study serves as useful tool to detect unusual data points as it links both demographic and stock data in a global demographic account. The unusual estimated flows could be utilized in future research to detect, revise, and improve future migrant stock or demographic data where contradicting reported values currently exist.

Comparable international migration flow data are needed to better understand the role of these factors and effectively govern. To this end, bilateral estimates, such as those presented in this study, provide a more comprehensive insight into past migration patterns, by gender and over different period lengths, than previously available. It is hoped that they can serve future migration scholars to better explain and predict global international migration trends.

REFERENCES

Abel, G. J.
2010 "Estimation of International Migration Flow Tables in Europe." *Journal of the Royal Statistical Society. Series A: Statistics in Society* 173(4):797–825.
———
2013a "Estimating Global Migration Flow Tables Using Place of Birth Data." *Demographic Research* 28:505–46.
———
2013b *migest: Methods for the Indirect Estimation of Bilateral Migration.*
———, and N. Sander
2014 "Quantifying Global International Migration Flows." *Science* 343(6178):1520–2.
Agresti, A.
2003 *Categorical Data Analysis*, 2nd ed. Wiley Series in Probability and Statistics, 734. Hoboken, NJ: Wiley-Interscience.
Artuç, E., F. Docquier, Ç. Özden, and C. R. Parsons.
2015 "A Global Assessment of Human Capital Mobility: The Role of Non-OECD Destinations." *World Development* 65:6–26.
de Beer, J., J. Raymer, R. Van der Erf, and L. van Wissen.
2010 "Surmonter les problèmes d'incohérences des données sur les migrations internationales: une nouvelle méthode appliquée aux flux en Europe." *European Journal of Population* 26(4):459–81.
Beine, M., F. Docquier, and Ç. Özden
2011 "Diasporas." *Journal of Development Economics* 95(1):30–41.
Castles, S., H. de Haas, and M. J. Miller
2013 *The Age of Migration — Castler & Miller*, 5th ed., Vol. 17. New York: Palgrave Macmillan.
Curran, S. R., and A. C. Saguy
2001 "Migration and Cultural Change: A Role for Gender and Social Networks?" *Journal of International Women's Studies* 2(3):54–77.

Czaika, M., and H. De Haas
2014 "The Globalization of Migration: Has the World Become More Migratory?" *International Migration Review* 48(2):283–323.
Docquier, F., B. L. Lowell, and A. Marfouk
2009 "A Gendered Assessment of Highly Skilled Emigration." *Population and Development Review* 35(2):297–321.
———, A. Marfouk, S. Salomone, and K. Sekkat
2012 "Are Skilled Women More Migratory Than Skilled Men?" *World Development* 40 (2):251–65.
Donato, K. M., D. Gabaccia, J. Holdaway, M. Manalansan IV, and P. R. Pessar
2006 "A Glass Half Full? Gender in Migration Studies." *International Migration Review* 40(1):3–26.
———, and ———
2015 *Gender and International Migration: From the Slavery Era to the Global Age*. New York, NY: Russell Sage Foundation.
Gu, Z., L. Gu, R. Eils, M. Schlesner, and B. Brors
2014 "Circlize Implements and Enhances Circular Visualization in R." *Bioinformatics* 30 (19):2811–12.
Hawelka, B., I. Sitko, E. Beinat, S. Sobolevsky, P. Kazakopoulos, and C. Ratti
2014 "Geo-Located Twitter as Proxy for Global Mobility Patterns." *Cartography and Geographic Information Science* 41(3):260–71.
Hijmans, R. J.
2015 *Geosphere: Spherical Trigonometry.*
International Labour Office
2013 *Domestic Workers Across the World: Global and Regional Statistics and the Extent of Legal Protection*. January. Geneva: International Labour Organization.
Jean-Christophe, D., G. Spielvogel, and S. Widmaier
2010 *International Migrants in Developed, Emerging and Developing Countries: An Extended Profile*. OECD Social, Employment and Migration papers No. 114. OECD Social, Employment and Migration Working Papers. Paris: OECD.
Kelly, J. J.
1987 "Improving the Comparability of International Migration Statistics: Contributions by the Conference of European Statisticians From 1971 to Date." *International Migration Review* 21(4):1017–37.
Kupiszewska, D., and B. Nowok
2008 "Comparability of Statistics on International Migration Flows in the European Union." In *International Migration in Europe: Data, Models and Estimates*, Chap. 3, edited by F. Willekens, and J. Raymer, 41–73. London: Wiley.
Lee, R. D.
2011 "The Outlook for Population Growth." *Science* 333(6042):569–73.
Martin, P, and J. Widgren
2002 "*International Migration: Facing the Challenge*," Population Bulletin 57(1). Washington, DC: Population Reference Bureau.
Massey, D. S., J. Arango, G. J. Hugo, A. Kouaouci, A. Pellegrino, and J. E. Taylor
1999 *Worlds in Motion: Understanding International Migration at the End of the Millennium*. Oxford: Oxford University Press.
National Research Council
2000 "International Migration." In *Beyond Six Billion: Forecasting the World's Population*, Chap. 3, edited by J. Bongaarts, and R. Bulatao, 258. Washington, DC: National Academy Press.

Nowok, B., D. Kupiszewska, and M. Poulain
2006 "Statistics on International Migration Flows." In *Towards the Harmonisation of European Statistics on International Migration (THESIM)*, Chap. 8, edited by M. Poulain, N. Perrin, and A. Singleton, 203–33. d. Louvain-La-Neuve, Belguim: UCL Presses Universitaires de Louvain.
Özden, Ç., C. R. Parsons, M. Schiff, and T. L. Walmsley
2011 "Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960–2000." *World Bank Economic Review* 25(1):12–56.
Parsons, C. R., R. Skeldon, T. L. Walmsley, and L. A. Winters
2007 *Quantifying International Migration: A Database Of Bilateral Migrant Stocks*. Policy Research Working papers No. 4165. Washington, DC: World Bank.
Piper, N.
2005 *Gender and Migration*. Tech. rep. September, p. 55. Geneva, Switzerland: Policy Analysis and Research Programme of the Global Commission on International Migration.
———
2013 In *New Perspectives on Gender and Migration: Livelihood, Rights and Entitlements*, edited by N. Piper. New York: Routledge.
R Development Core Team
2016 *R: A Language and Environment for Statistical Computing*. https://cran.r-project.org/doc/FAQ/R-FAQ.html#Citing-R
Ratha, D., and W. Shaw
2007 *South-South Migration and Remittances*. Working paper 102, 1–68. Washington DC: World Bank.
Raymer, J.
2007 "The Estimation of International Migration Flows: A General Technique Focused on the Origin Destination Association Structure." *Environment and Planning A* 39 (4):985–95.
———, G. J. Abel, and P. W. Smith
2007 "Combining Census and Registration Data to Estimate Detailed Elderly Migration Flows in England and Wales." *Journal of the Royal Statistical Society. Series A: Statistics in Society* 170(4):891–908.
———, A. Wiśniowski, J. J. Forster, P. W. F. Smith, and J. Bijak
2013 "Integrated Modeling of European Migration." *Journal of the American Statistical Association* 108(503):801–19.
Rogers, A.
1990 "Requiem for the Net Migrant." *Geographical Analysis* 22(4):283–300.
State, B., I. Weber, and E. Zagheni
2013 "Studying Inter-National Mobility Through IP Geolocation." In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining - WSDM '13.* 265 New York, NY: ACM Press. https://doi.org/dl.acm.org/citation.cfm?doid=2433396.2433432
Truong, T.-D., D. Gasper, J. Handmaker, and S. I. Bergh (Eds)
2014 *Migration, Gender and Social Justice*. Berlin, Heidelberg: Springer Berlin Heidelberg.
United Nations Population Division
2011 *World Population Prospects: The 2010 Revision. CD-ROM Edition*. New York, NY: United Nations, Department of Economic and Social Affairs.

———
2012 *Trends in International Migrant Stock: Migrants by Destination and Origin (United Nations Database, POP/DB/MIG/Stock/Rev.2012).* New York, NY: United Nations, Department of Economic and Social Affairs.

———
2013a *Trends in International Migrant Stock: Migrants by Destination and Origin (United Nations Database, POP/DB/MIG/Stock/Rev.2013).* New York, NY: United Nations, Department of Economic and Social Affairs.

———
2013b *World Population Prospects: The 2012 Revision. DVD Edition.* New York, NY: United Nations, Department of Economic and Social Affairs.

———
2014 *World Urbanization Prospects: The 2014 Revision, Highlights.* New York, NY: United Nations, Department of Economic and Social Affairs.

———
2015a *International Migration Flows to and From Selected Countries: The 2015 Revision.* New York, NY: United Nations, Department of Economic and Social Affairs.

———
2015b *Trends in International Migrant Stock: The 2015 Revision.* New York, NY: United Nations, Department of Economic and Social Affairs.

———
2015c *World Population Prospects: The 2015 Revision. DVD Edition.* New York, NY: United Nations, Department of Economic and Social Affairs.

———
2016 *International Migration Report 2015 Highlights.* New York, NY: United Nations, Department of Economic and Social Affairs.

Wickham, H.
2009 *ggplot2: Elegant Graphics for Data Analysis.* New York, NY: Springer New York.

Willekens, F.
1999 "Modeling Approaches to the Indirect Estimation of Migration Flows: From Entropy to EM." *Mathematical Population Studies* 7(3):239–78, 308.

Wiśniowski, A., J. Bijak, S. Shristiansen, J. J. Forster, N. Keilman, J. Raymer, and P.W.F. Smith.
2013 "Utilising Expert Opinion to Improve the Measurement of International Migration in Europe." *Journal of Official Statistics* 29(4):583–607.

World Bank
2016 *Migration and Remittances Factbook 2016.* Washington, DC: World Bank.

Zagheni, E., and I. Weber
2012 "You Are Where You E-mail: Using E-mail Data to Estimate International Migration Rates." In *Proceedings of the 3rd Annual ACM Web Science Conference.* https://doi.org/dl.acm.org/citation.cfm?doid=2380718.2380764.

———, V. R. K. Garimella, I. Weber, and B. State
2014 "Inferring International and Internal Migration Patterns From Twitter Data." In *Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion*, 439–44. New York, NY: ACM Press.

Zlotnik, H.
1995 "The South-to-North Migration of Women." *International Migration Review* 29 (1):229–54.

———
1999 "Trends of International Migration Since 1965: What Existing Data Reveal." *International Migration* 37(1):21–61.

———
2003 *The Global Dimensions of Female Migration.* Washington, DC: Migration Policy Institute.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site:

Appendix S1. Extended methodology to control for natural population change.

Appendix S2. CSV file of the estimates of country to country bilateral migration gender in all time periods and readme text file with descriptions of the data frame format.

Table S1. Distance matrix and corresponding flow estimates based on the sequential stock table given in Table 1.

Table S2. Bilateral stock data of Table 1 arranged as birthplace-specific origin–destination migrant flow tables with minimal stayers assumption and the corresponding estimated origin–destination flow table.

Table S3. Example of place of birth data with births and deaths over the interval.

Table S4. Bilateral stock data, controlled for natural change, arranged as birthplace-specific origin–destination migrant flow tables with resulting flow estimates.

Table S5. Multi-step demographic account framework using stock data from Table S3.

Figure S1. Total estimated flows from stock data in Table S3 under various assumptions for the number of stayers.

Figure S2. Estimated 10-year bilateral migrant flows during 2000–10 based on alternative stock and demographic data.

Figure S3. Total estimated country-to-country bilateral flows and crude global migration rate varying by demographic data source and interval.

Figure S4. Regional differences in estimated migration flows over 10-year periods from alternative demographic data sources.

Figure S5. Total estimated international migrant flows under changing geographic boundaries.

Figure S6. Comparisons of estimated flows and reported flows on the proportions of migrants from each continent.

Figure S7. Comparisons of estimated flows and reported flows on the proportions of migrants to each continent.