Estimates of Global Bilateral Migration Flows by Gender Between 1960 and 2010.

Guy J. Abel*
Wittgenstein Centre (IIASA, VID/ÖAW, WU)
Vienna Institute of Demography/Austrian Academy of Sciences

Abstract

Measures of international migration flows are often limited in both availability and comparability. This paper aims to address these issues at a global level using an indirect method to estimate country to country migration flows from more readily available bilateral stock data. Estimates are obtained over five and ten-year periods between 1960 and 2010 by gender, providing a comprehensive picture of past migration patterns. The estimated total amount of global international migrant flows is shown to generally increase over the 50 year time frame. The intensity of migration flows over five and ten-year periods fluctuate at around 0.6 and 1.25 percent of the global population respectively, with a noticeable spike during the 1990-95 period. Gender imbalances in the estimated flows between selected regions were found to exist, such as recent movements into oil rich Gulf States from South Asia. The sensitivity of flow estimates to alternative input stock and demographic data as well as changes in political geography are explored.

1 Introduction

Global migration is a complex system influenced by a mix of social, economic, 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 informs policy makers, the media and academics about the level and direction of population movements. Migration data provides a means for users to asses the determinants of which people move, why they move and where people move between.

Moves in populations can be quantified using either migrant stock or migration flow data. Unlike a static stock measure, flow data are dynamic, summarising movements over defined period and consequently allow for 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 Nation have served as the sole comprehensive source of global migration flow data. However, as with any net measure, they are susceptible to distorting and disguising the underlying patterns (Rogers, 1990) and hence are of limited explanatory use. More detailed measures, such as the

*Correspondence Address: Vienna Institute of Demography, Wohllebengasse 12-14, 6th floor, 1040 Vienna, Austria. Email: guy.abel@oeaw.ac.at

immigration and emigration counts into each country, or the 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; Kupiszewska and Nowok, 2008; Nowok, Kupiszewska, and Poulain, 2006). In the European context, where flow data are more plentiful, methodologies to harmonise existing data have been developed (Abel, 2010; Beer et al., 2010; Raymer, 2007; Raymer et al., 2013; Wiśniowski et al., 2013). Each are severely limited in their application to a global setting where missing data becomes 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, Haas, and Miller (2013) have been based on a patchwork of net migration flow data, changes in bilateral stocks over time and available, unharmonised flow data from predominately rich Western countries. A growing literature of work has developed based on bilateral migrant stock data (Beine, Docquier, and Özden, 2011; Czaika and Haas, 2014) to explain changes in contemporary migration patterns. However, in countries where there are significant return migration or mortality among foreign populations the analysis of stock data can provide a misleading portrait of migration systems (Massey et al., 1999, p.200). Further, stock data based on place of birth are an aggregate of moves over individuals lifetimes causing a number of problems when used to measure contemporary migration. For example, patterns of intermediary moves between individuals birth and the date of data collection are not captured. Native born stocks are likely to lower in younger countries, where the sum of years lived, and hence the cumulative time available to people to emigrate, are lower. Recent studies of global migration patterns such as Hawelka et al. (2014); State, Weber, and Zagheni (2013); Zagheni and Weber (2012) or Zagheni et al. (2014) have focused on short term mobility measures derived from data sources based on individuals geo-located of internet activities such as twitter messages or logins to email services. As they authors note, the users for these studies are based on may not be fully representative of the whole global population.

Indirect methods have recently been used to estimate global bilateral migration flows using changes in published bilateral migrant stock data. Abel (2013) used global bilateral stock tables from the World Bank to derive global bilateral flow estimates between 1960 and 2000 over four ten-year periods via a proposed flows from stocks methodology. The methodology was altered slightly, and then applied by Abel and Sander (2014) 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 alteration in the methodology allowed the difference of the estimated immigration and emigration flow totals to match the net migration estimates of United Nations Population Division (2011).

In this paper, the methodology of Abel (2013) and Abel and Sander (2014) is extended once more and applied to estimate five and ten-year migrant flows separately by gender between 1960 and 2010. The application of the revised methodology allows for an updated view of international migration over a far longer time period. Estimates by gender quantify the differences in male and female global migration flow patterns which are known to have a distinct variations in the origin and destinations by gender (Donato et al., 2006; Zlotnik, 1995) and have only previously been measured using net migration statistics. Estimates over both five and ten-year periods enable for contrasts between possible different global migration transitions rates to be identified.

Estimates of flows will also be 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 birth and death numbers over a given interval. The culmination of the country to country flows estimates varying by different gender, time periods, intervals, stock and demographic data, provides a combined set of 108 estimated migrant flow tables, far exceeding those of the previously discussed flow estimation studies.

In the next section the methodology to indirectly estimate origin-destination flow tables from changes in bilateral stock data is outlined. In Section 3 an overview on the various migrant stock and demographic data, required as inputs for the estimation methodology is provided. In Section 4 the results from the estimated flow tables are shown at different levels of analysis. The sensitivity of the estimates to alternative demographic input data and changes in political geography are then discussed before the final summary and conclusion.

2 Methodology Background

Available bilateral migration data can be categorised as either a stock measures, that represent a static number of foreign populations by some characteristic such as their place of birth or a flow measures, that represent the dynamic movements of populations between origin and destinations. In comparison with flow data, the static nature of stock data leads to far fewer issues in its measurement and collection. As a result migrant stock data tends to be more widely available across longer time periods than migrant flow data. Groups at both the United Nations and the World Bank have collated together estimates of past stock data to build global bilateral tables of migrant stocks.

Changes in bilateral migrant stock sizes over time, defined by the place of birth of individuals, can be the result of 1) 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 data on both bilateral migrant stocks at the start and end of period are available it is possible to indirectly derive the number of bilateral migrant flows by viewing each population stock as part of 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 1 illustrated using a schematic of a simple demographic account framework based on dummy 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 start and end of interval. They are grouped together by the country of birth. For example, for those in born in country A are shown in the top left; 100 are native born citizens, living in country A at time t. A further two sets of 10 people born in A are living abroad in countries B and C, whilst 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, whilst the stock living in country B and D has increased. Note, the total population of those born in A but residing in any country does not change over the

		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	15	0	100

Table 1: Estimated origin-destination flow table based on the changes in the bilateral migrant stock data illustrated in Figure 1.

time period as there are no births or deaths, and birthplace is fixed characteristic that cannot alter over time.

There are many thousands of possible combinations of moves that can take place over the time period to match the changes in these migrant stock. However, at a minimum at least 20 migrants must leave A and arrive in B, and a further 10 must leave A and arrive in D. The minimum amount of migrant transitions for all birth place populations in a global system can be indirectly estimated using an iterative proportional algorithm, details of which are given in Appendix. The results of the applied indirect estimation method for the global system of four countries are shown by the arrows in Figure 1. The estimates of the flows in Figure 1 can be used to derive a traditional origin-destination migrant flow table in Table 1 by summing over places of birth. For example, the 25 moves from D to B in Table 1 are comprised of 10 from those born in A, 15 from those born in C and 10 from those born in D (each shown in Figure 1).

The estimated flow in both Figure 1 and Table 1 are based on a number of migrant transition over the period. Migration can alternative be measured as the number of movements during a period between given origin and destinations. A movement definition of a migration flow captures multiple changes in location over a defined period including intermediate moves. Although the number of movements will be at least as high and the number of translations, there is no simple mathematical solution to estimate one from the other.

The demographic framework in Figure 1 can be extended to account for these demographic changes from both births and deaths, which are likely to have large impacts on the changes in bilateral migrant stocks data. 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 1 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 birth place itself is a defining characteristic of bilateral migrant stock data, the number of newborns can be subtracted only from the native born populations at time $t+1$[1]. The reduction accounts for potential increase in migrant stocks from time t which might otherwise

[1]Note, 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, which is assumed to not be occur.

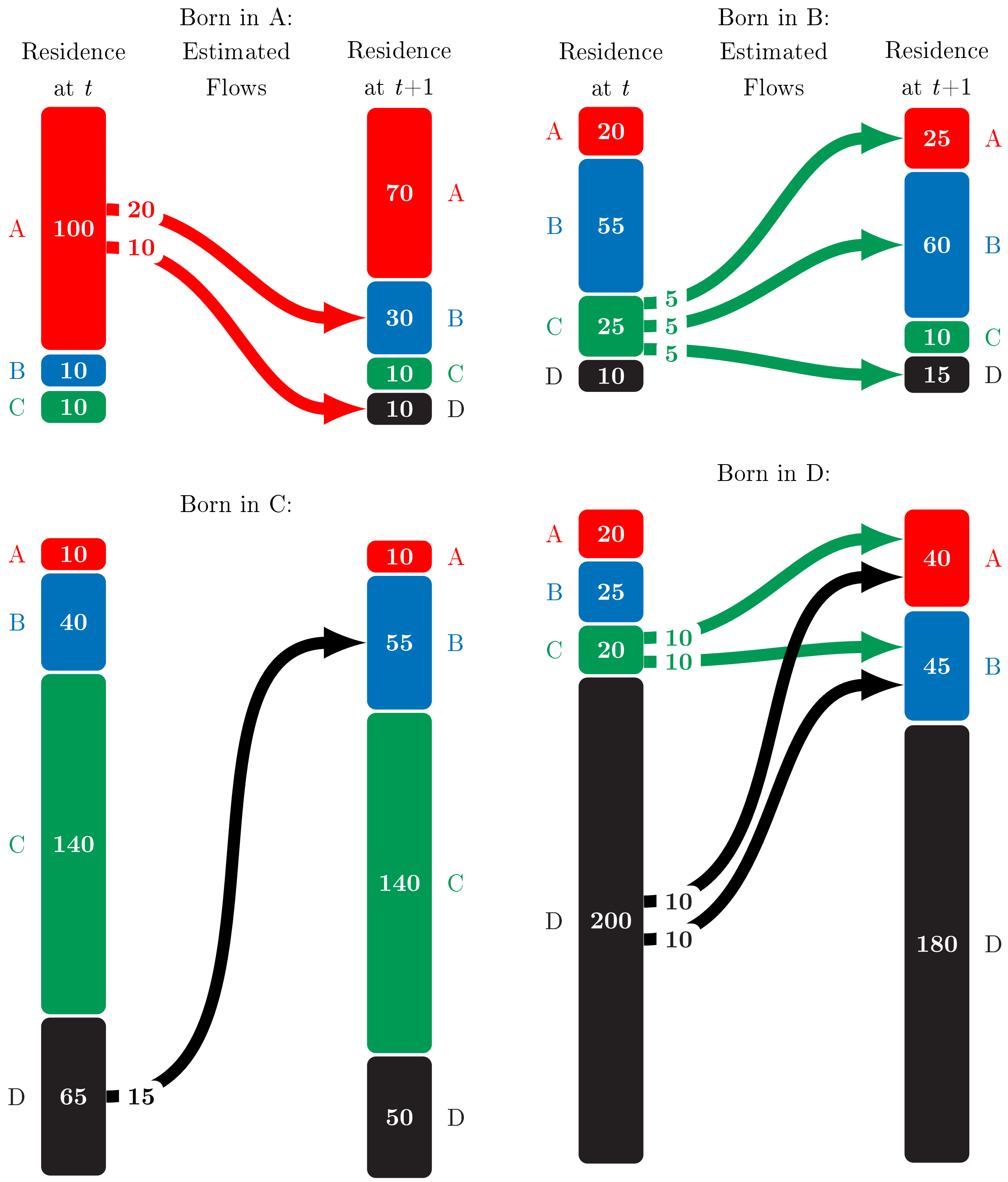


Figure 1: Schematic of a demographic accounting framework to link changes in bilateral migrant stock data via estimated migrant flows. Note, for each birthplace there are no births or deaths during the time interval, hence the equal heights in the total populations. The estimated flow sizes displayed in the arrows are the minimum number of migrant transitions required to match the known migrant stock data given in each block.

result in an increase in the estimate of return migrants to their birthplace. More details of the demographic accounting framework and adjustments for births and deaths are given in the Appendix.

3 Input Data

The estimation of international bilateral migration flow tables requires two sets of input data. First, bilateral stock tables are required at the start and end of a given period. Currently, both the United Nations (UN) and World Bank provide bilateral stock estimates. Each set of estimates cover differing periods and use different estimation methods to derive missing data. Second, demographic data on the number of births, deaths and population. Births and death information are required to alter stock data for natural change. Population data is needed to derive estimates of the native born population size 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) provide foreign born migration stock tables at the start of each of decade, from 1960 to 2000, for 226 countries[2]. Data are primarily based on place of birth responses to Census questions or details collected from population registers. Where no data was available, alternative stock measures such as citizenship or ethnicity are used. For countries where no stock measures were available, missing values are imputed using various propensity and interpolation methods, typically dependent on foreign born distributions from available countries in the region.

The United Nations Population Division (2013a) provide foreign born migrant stock tables at the start of each of the last three decades (1990, 2000, and 2010) covering 230 countries[3]. As with the World Bank estimates, data are primarily based on place of birth responses to Census questions and register informations. Adjustments to estimates are 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 are made based on the change in the overall populations size to align all estimates at the same time point. For countries or areas without any data sources, 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 the place of birth (Other North and Other South). These counts originate from either regional aggregations or non-standard areas used by national statistical agencies to enumerate foreign born stocks which the UN are then unable to redistribute into its 230 countries. For the vast majority of countries the counts of unknowns comprised less than five percent of the total foreign born population.

Demographic data on births, deaths and population totals are available from the World Population Prospects (WPP) of the United Nations Population Division (2011, 2013b). Every two years the UN release update versions of the WPP incorporating updated 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

[2]Data available from `http://data.worldbank.org/data-catalog/global-bilateral-migration-database`
[3]Data available from `http://esa.un.org/unmigration/TIMSO2013/migrantstocks2013.htm`

a gender disaggregation. 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 two most recent versions of WPP are used, WPP2010 and WPP2012 in order to determine what, if any, effect updated demographic data has on bilateral migration flow estimates.

4 Results

Using the flow from stock estimation method and different combinations of input data discussed in the previous two sections, 108 different country to country bilateral migration flow tables were produced. Of these, 36 were of estimated migrant flow transitions over six ten-year periods between 1960 and 2010 where the 1990-2000 flow was calculated twice; once from the World Bank data and once from the UN data. In each period multiple flow tables were calculated based on alternative stock data for each gender; male, female and both, and alternative demographic data; WPP2010 and WPP2012. The remaining 72 migration flow tables were based on 12 five-year periods over the interval, where the 1990-95 and 1995-2000 flows were calculated from both the UN and World Bank stock data, with alternative input stock and demographic data.

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

The culmination of the country to country flows estimates vary by different gender, time periods, intervals, stock and demographic data provides a combined data set with over four 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 summarised at the regional level. Full estimates of country to country flows are provided in the supplementary materials or from contacting the author. Note, throughout the remainder of this article, when referring to an estimated flow, the estimate have 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 estimates of births and deaths. 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 inherit assumptions in the methodology used to estimate the flow.

4.1 Global Level Summary Statistics

In Figure 2 summary statistics for estimated global migration flows over time are displayed using the *ggplot2* package (Wickham, 2009) in R (R Development Core Team, 2012). The symbol type of each point corresponds to the stock data source used as input data when estimating the flow table. Only estimated based on WPP2010 demographic data are shown.

The estimated sum of the number of migrants for each of the 54 flow tables that used WPP2010 input data are shown on the left hand side. An upward trend in the global level of

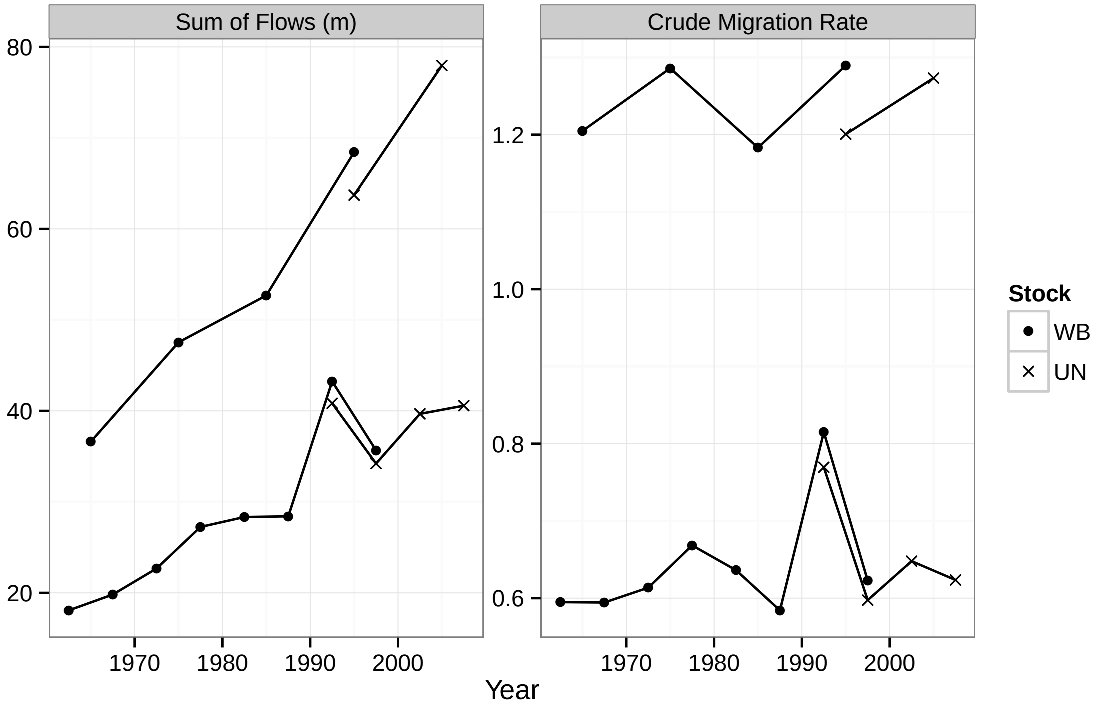


Figure 2: Total estimated country to country bilateral flows and crude global migration rate. The stock data source used to estimate the flow is indicated by the point type. Note, flows estimated over a ten-year period are represented by a single point at the mid-decade. Flows estimated over five-year periods are represented at the mid points of the first and second half of the decade.

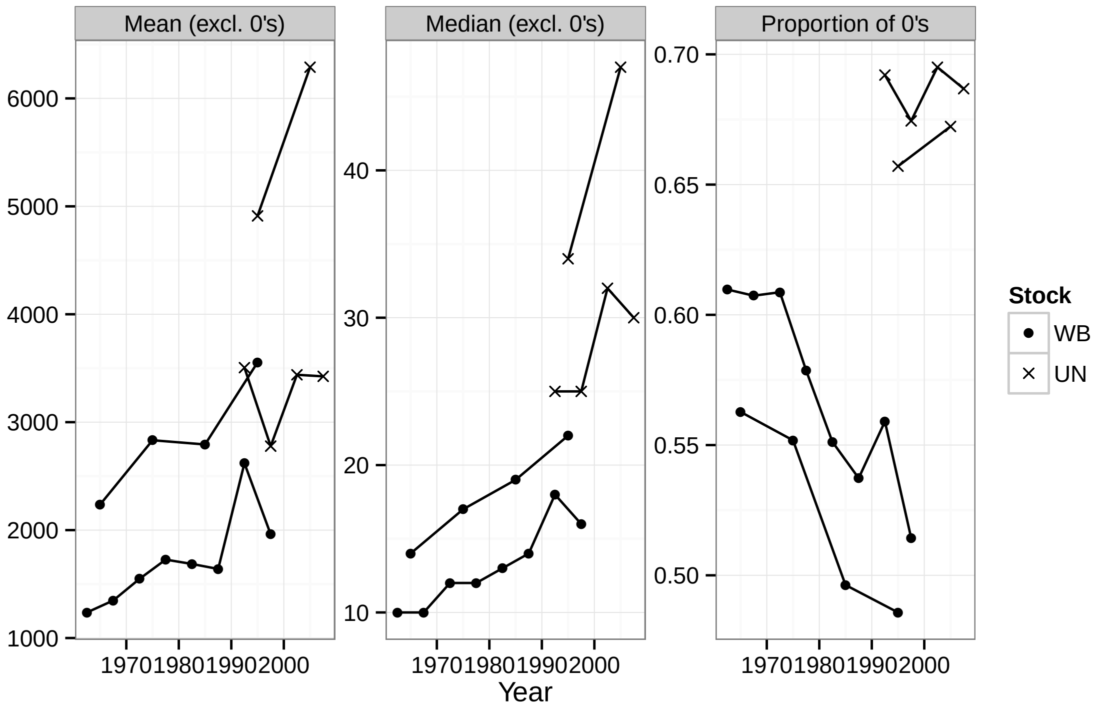


Figure 3: Further summary statistics for estimated country to country bilateral flows. The stock data source used to estimate the flow is indicated by the point type. Note, flows estimated over a ten-year period are represented by a single point at the mid-decade. Flows estimated over five-year periods are represented at the mid points of the first and second half of the decade.

migrants over time is apparent. The upper lines are based on the total flows over ten-year period, plotted at the mid-decade point on the x-axis. In the 1990-2000 period, when an estimate of flows from both the World Bank and UN are available, the total flows from the World Bank stock data is 68.43 million people, 4.72 million higher than the estimates from the UN data. The lower lines are based on flows over five-year periods, plotted at the mid-point of the relative period. A sharp rise in the total amount of migrants during the 1990-95 period is evident, 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 from countries that were experiencing the effects of armed conflicts during the period, such as Kuwait, Rwanda, Afghanistan and Liberia. These movements are not fully captured in the ten-year interval estimates for 1990-2000, where for example crisis migrants might have returned to their original place of residence by end of the period.

The right hand side of Figure 2 illustrates the percentage of the global population that were estimated to migrate during the relevant interval derived by dividing the sums on the left hand side by the WPP2010 global population at the beginning of the corresponding time interval. The percentage remains relatively constant, at around 1.25 percent for migrant transitions over a ten-year interval. The estimates based on the five-year interval also remain fairly constant at around 0.6 percent, with the exception of the 1990-95 period.

Figure 3 illustrates further summary statistics for the estimated bilateral tables. On the left hand side is a plot of the mean of non-zero estimated flows in each period. The mean flow size follow a broad upward trend over time. Non zero 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 non-zero estimated flows are not constant across time, as illustrated on the right hand side of Figure 2. Zero 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 end of 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 around 67 percent of bilateral foreign born stocks are zero in the 1960 data. This percentage falls to 55 percent by 2000. In the UN data stock data approximately 75 percent of stock estimates are zero throughout the data period. Second, the number of countries included in the estimated flow tables differ according to the stock data. For flow estimates based on the World Bank stock data, origin-destination flows are obtained for 194 countries where both demographic data and stock data are available. In comparison, estimates based on the UN stock data are possible for 195 countries. Of these, 193 were common to both sets of estimates[4]. Estimates based on the World Bank stocks included an additional country; the Netherlands Antilles, whilst estimates based on the UN stocks included two additional countries; the Channel Islands and Western Sahara.

The estimated median of the non-zero flows are shown in the middle panel of Figure 3. These broadly follow a similar pattern as the mean, although at much lower levels indicating a large skew in the distributions of estimated global bilateral flows towards smaller counts

4.2 Bilateral Patterns

In order to illustrate the pattern of estimated bilateral relationships, a set of six circular migration plots are shown in Figure 4. Plots were created in R using the *circlize* package (Gu et al., 2014)[5]. The direction of the flow is indicated by the arrow head. 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 movements are in millions. Each plot is based on flows over a ten-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, 5.20 million; Europe, 4.95 million). Many movements within the first of these regions were at time not international movess, such as Russia to Ukraine (0.88 million) or Russia to Kazakhstan (0.79 million). The total estimates transitions flows during 1970-1980 increased globally from the previous period, as illustrated in Figure 2. Whilst such an increase is difficult to view from comparing circular migration plots (a) and (b) in Figure 4, changes in the bilateral relationships between the selected regions are clearly visible. Most noticeable is a large increases in the share of global migrants moving within Southern Asia. During 1970-1890, 4.38 million movements were estimated from Bangladesh to India and another 1.81 million from India to

[4]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 was derived from the aggregation of the individual country information provided in each WPP.

[5]Full details of their implementation for migration data can be found in (Sander et al., 2014).

Pakistan most likely driven by the Indo-Pakistani War of 1971.

Changes in regional migration patterns over time are more easily detected in Figure 5, which provides plots of estimated immigration and emigration totals by UN Population Division demographic regions[6]. As noted in the methodology section, at the country level, the estimated net migration, obtained from differencing the immigration and emigration values, matches those implied by the demographic data. In the first two time periods in Southern Asia there is a sharp rise in immigration and emigration where as 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 4 (a) and (b). Estimated flows into and within Europe during 1970-80 decreased on the decade before. Sizeable moves into West Asia from countries such as Egypt (0.45 million), India (0.16 million) and Pakistan (0.13 million) to Saudi Arabia began to develop. Moves within Africa also increased, including large flows out of Ethiopia (0.95 million to Sudan) and into the Ivory Coast (0.40 million from Burkina Faso).

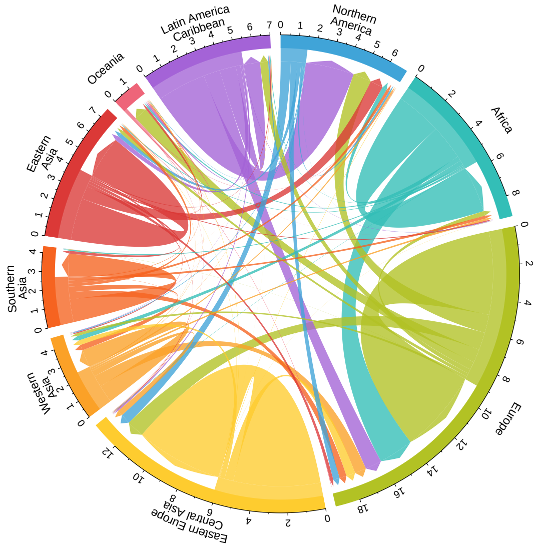
Estimated flows during 1980-1990 increased in most regions in comparison to previous periods. Most noticeable is the further rise in movements from Latin America and the Caribbean to North America in Figure 4 (c) in comparison to (a) and (b). The largest flow during the period was estimated from Mexico (3.17 million). The number of movements within Eastern Europe also increased, including 1.12 million from Ukraine to Russia.

In Figure 4 (d) and (e) are circular migration flow plots based on estimates during 1990-2000 period using different stock data sources. In (d) estimates based on the World Bank stock data. It shows increasing immigration to North America (also shown in Figure 5) from a wider variety of origin regions, including Eastern and Southern Asia. Moves into Europe, especially from other European countries increased, as does immigration into West Asia. The plot of the estimates during the same period but based on the UN stock data is shown in Figure 4 (e). Many of the same estimated bilateral flow patterns are similar to those based on the World Bank data in (d). However, some distinct difference in the size of movements are apparent from the immigration and emigration summary plots in Figure 5. 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. In all but the last of these regions, flow estimates from the World Bank stock data result in higher levels. The difference in the estimates are driven by larger (or smaller) changes in the foreign born stock values provided by the World Bank in 1990 and 2000 in comparison to those of the UN stock data. For example, the largest estimated flow into Europe based on the UN stock data is from Kazakhstan to Germany, to match an increase in Kazakh born residents in Germany (10.2 thousand in 1990 to 490 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.

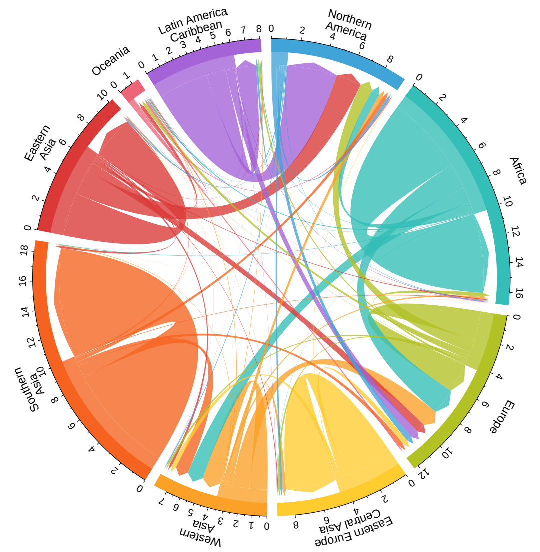
The circular migration flow plots related to the final ten-year time period between 2000 and 2010 is shown in Figure 4 (f). Based upon UN stock data, there are further increases of immigration flows into North America from Asia and into Europe from Asia, Africa, North and Latin America. Some of the largest increases of estimated flows into Europe are into Southern European countries, as shown in Figure 5, the largest being 0.55 million from Morocco to Spain. There are also sizeable increases in the estimated flows from South America countries such as

[6] With the exception of Polynesia, Melanesia and Micronesia which are aggregated to a Pacific Island region.

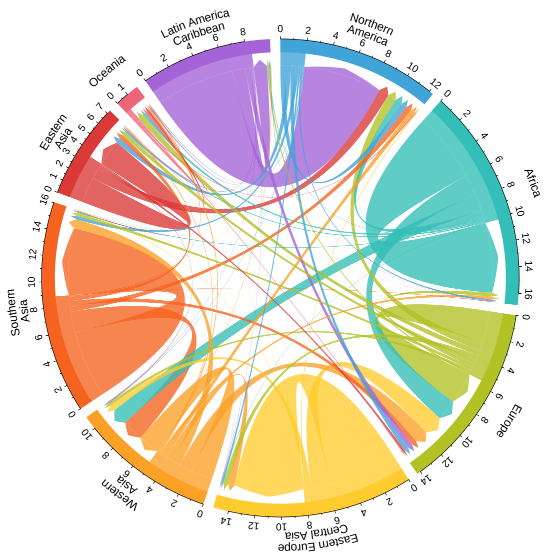
Figure 4: Estimated 10 year migrant flows over time aggregated by selected regions.



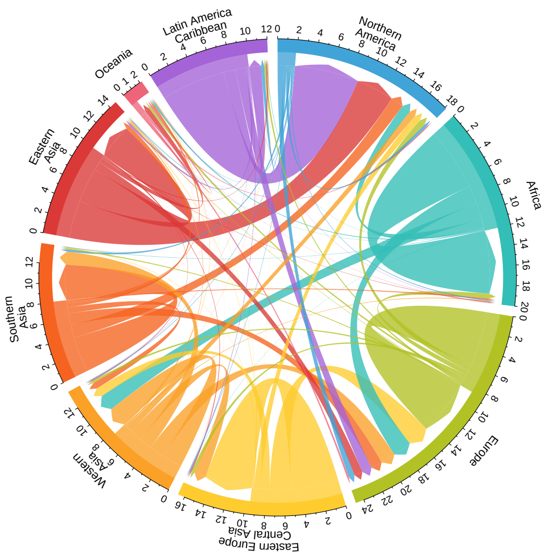
(a) Total 1960-1970. Based on WB stock data



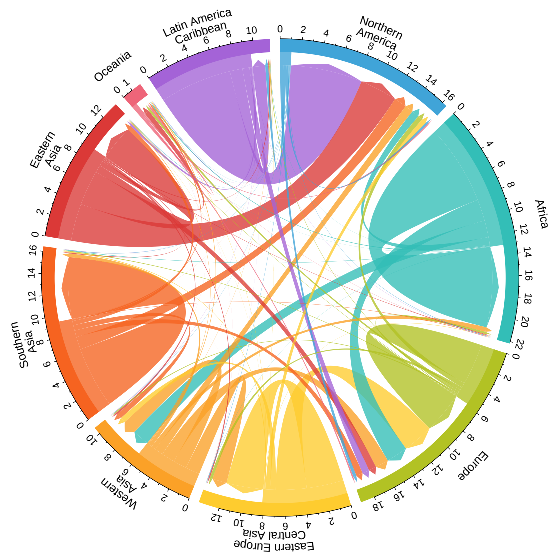
(b) Total 1970-1980 Based on WB stock data



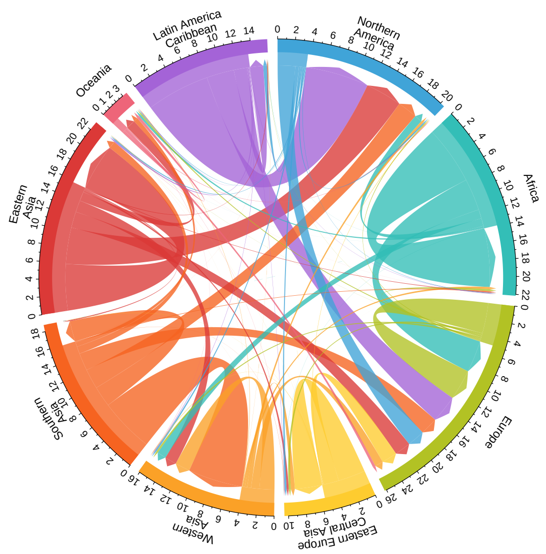
(c) Total 1980-1990 Based on WB stock data



(d) Total 1990-2000 Based on WB stock data



(e) Total 1990-2000 Based on UN stock data



(f) Total 2000-2010 Based on UN stock data

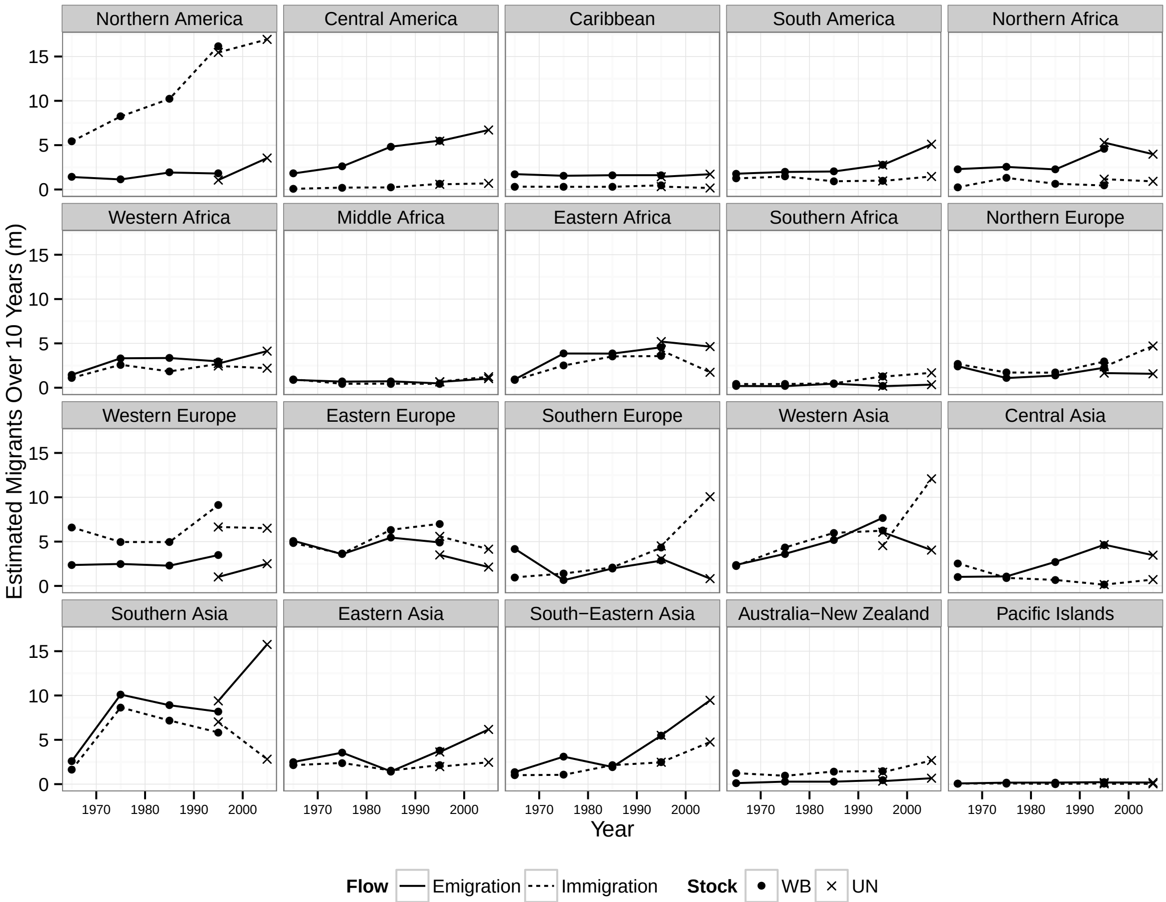


Figure 5: Total estimated immigration and emigration flows over a 10 year periods. Estimates based on aggregations of country to country bilateral flows, both World Bank and UN stock data (WB or UN) and WPP2010 demographic data.

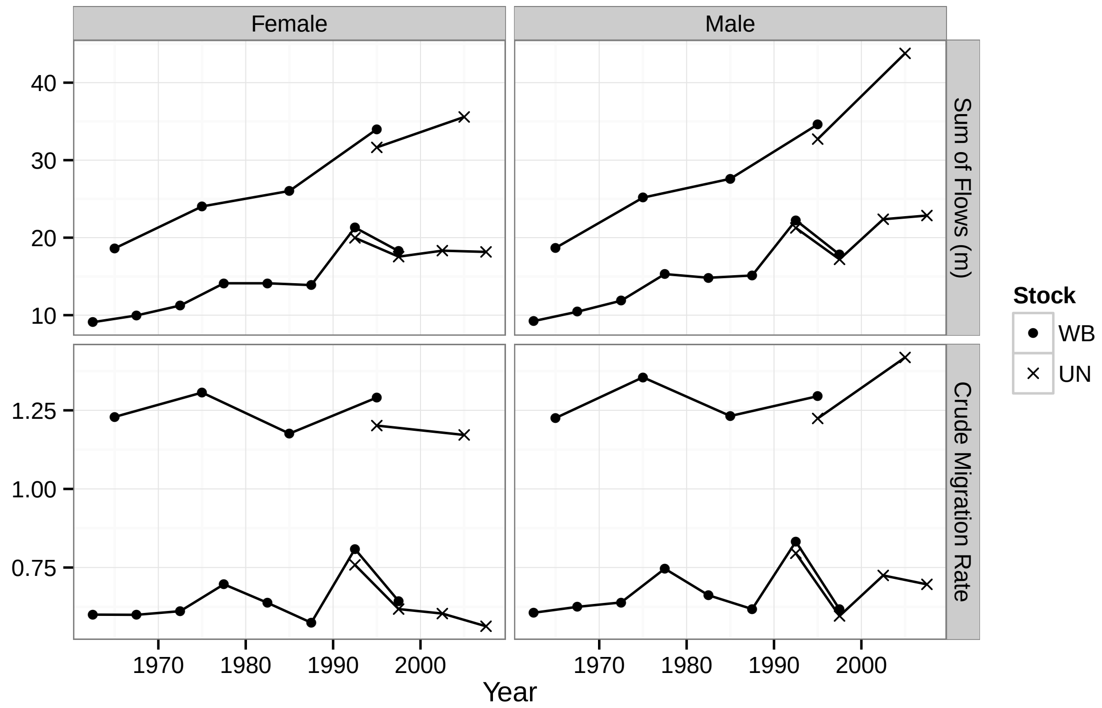


Figure 6: Total global migration flows and crude rate for estimated country to country bilateral flows by gender. The stock data source used to estimate the flow is indicated by the point type. Note, flows estimated over a ten-year period are represented by a single point at the mid-decade. Flows estimated over five-year periods are represented at the mid points of the first and second half of the decade.

Bolivia and Colombia into Southern Europe. Immigration into West Asia further increases, as does movements within South-Eastern Asia, including an estimated 1.16 million people moving from Myanmar to Thailand over the ten-year period.

4.3 Flows by Gender

Female and male total flows and crude migration rates are shown in Figure 6. The estimated patterns of both statistics follow similar paths as the that estimated from the total stocks in the Figure 2. There is little difference between the sum of male and female flows in all but the most recent time periods. In the last decade (2000-10), male flows increased faster than the females. For example, the total male flows over a ten-year interval, based on UN stock data and WPP2010 demographic data, rose from 32.71 million during 1990-2000 to 43.77 million during 2000-10, whilst the equivalent female flow total only increased from 31.64 million to 35.57 million.

Selected circular migration flow plots for both estimated males (left) and females (right) are shown for two time periods in Figure 7. Estimates are based on gender-specific stock and demographic data. In each of the time periods both male and females migration patterns are broadly 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 ten-year migration flows for each regions over all time periods shown in Figure 8. Viewed together, both Figure 7 and 8 allow for a more detailed discussion of male and female estimated flows.

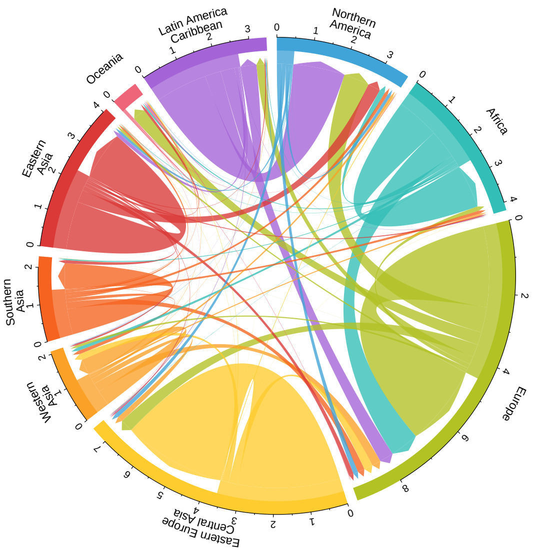
During the 1960-70 period, as shown in the top two circular migration plots in Figure 7, estimated male migration within Africa and into Europe are distinctly higher than the female estimated flows. The origin or higher males flows can be more easily identified in the time series plots of Figure 8. For the case of Northern Africa, the source of the imbalance are large flows into Europe, in particular France, to match the changes in its stock data. For example, 0.52 million males and 0.19 million females are estimated to move from Algeria to France. In the case of Southern Africa, the large imbalance is predominately due to greater increases of the male stocks of people born in Lesotho, Swaziland and Namibia in South Africa, creating larger male flows. Also evident is the high share of male migrants out of North America in the first two decades. Large male flows were estimated to originate from the USA, going to Canada, Poland, Russia and Great Britain. In each case, a large share of the moves are of non-US born making return moves back to the country of birth or onwards moves to a third country. For example the largest country to country migration flow out of North America during the period is of 163 thousand from USA to Canada. Of this estimated flow, 67 thousand were estimated to be Canadian born, moving back to Canada to match decreasing male Canadian stocks in the USA and increasing numbers of Canadian males in Canada (after controlling for the number of births). Another estimated 35 thousand moving from the USA to Canada are American born, moving to match increasing males stock of US born in Canada. Further amounts of European born (Italy, Serbia and Montenegro and Czech Republic) were estimated to move from USA to Canada, to match increases in their resident population stocks in Canada and decreasing stocks in the USA over the period. Similar changes in the female stock data are not present, and hence only small flows, such as from the USA to Canada or elsewhere are generated.

Higher shares of females are estimated to emigrate out of Eastern Europe as well as Central, South-Eastern and East Asia during 1960-1970. In each of these cases, the largest flows tend to be to other countries in the same region. In subsequent time periods, the share of estimated male to female migrant flows tend not to follow any clear and consistent patterns. Flows that were distinctly male dominated in 1960-70, such as emigration from Southern Africa or North America, become more evenly split by 2000-10. Other flows fluctuate either side of 50:50 male-female share, such as immigration and emigration out of Eastern Asia. However, there are some noticeable exceptions. First, in South-Eastern Asia particularity large male immigration and emigration flows are estimates during 1970-80. Second, in Western Asia, in particular oil rich Gulf States, large male immigration flows are estimated in 2000-10. As shown in the circular migration plots of Figure 7 (c), these flows are predominately from Southern Asia, South-Eastern Asia, other countries is in West Asia and Africa, where similar strong bilateral links are not present in the female plot of (d).

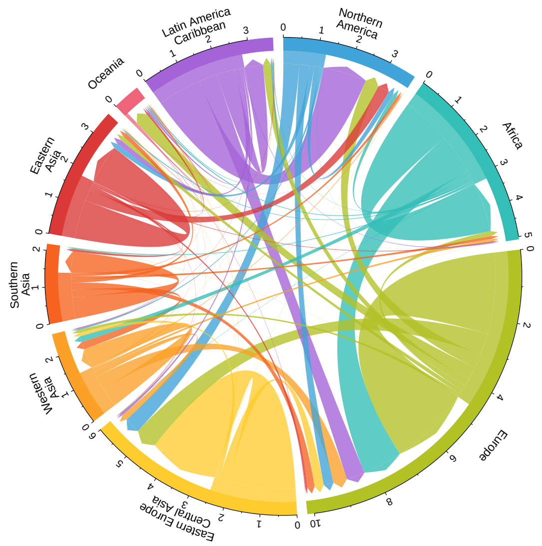
5 Sensitivity Analysis

Estimates of migrant flows from stocks can potentiality be sensitive to the input data used. As discussed in the previous section, during the 1990-2000 period where two sets of stock data are available, flow estimates may not necessarily be the same. Further, as discussed in Abel

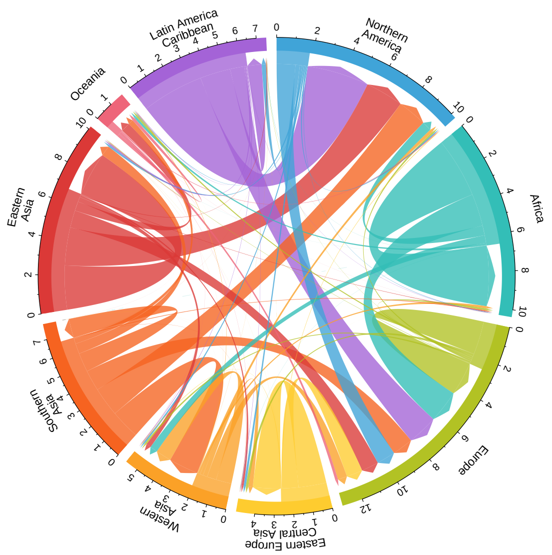
Figure 7: Estimated 10 year migrant flows by gender for selected time periods.



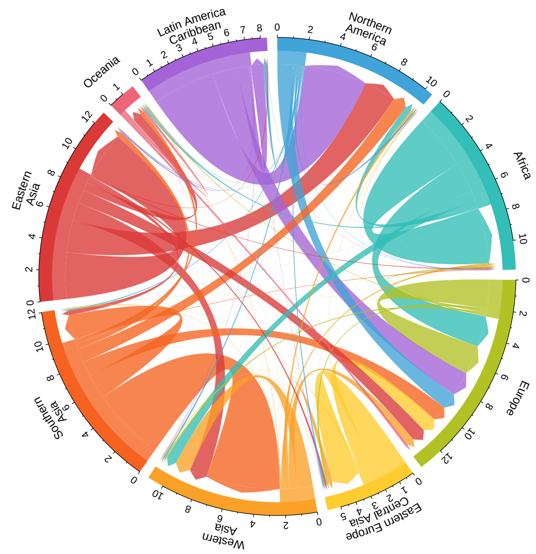
(a) Males, 1960-1970. Based on WB stock data



(b) Females, 1960-1970. Based on WB stock data



(c) Males, 2000-2010. Based on UN stock data



(d) Females, 2000-2010. Based on UN stock data

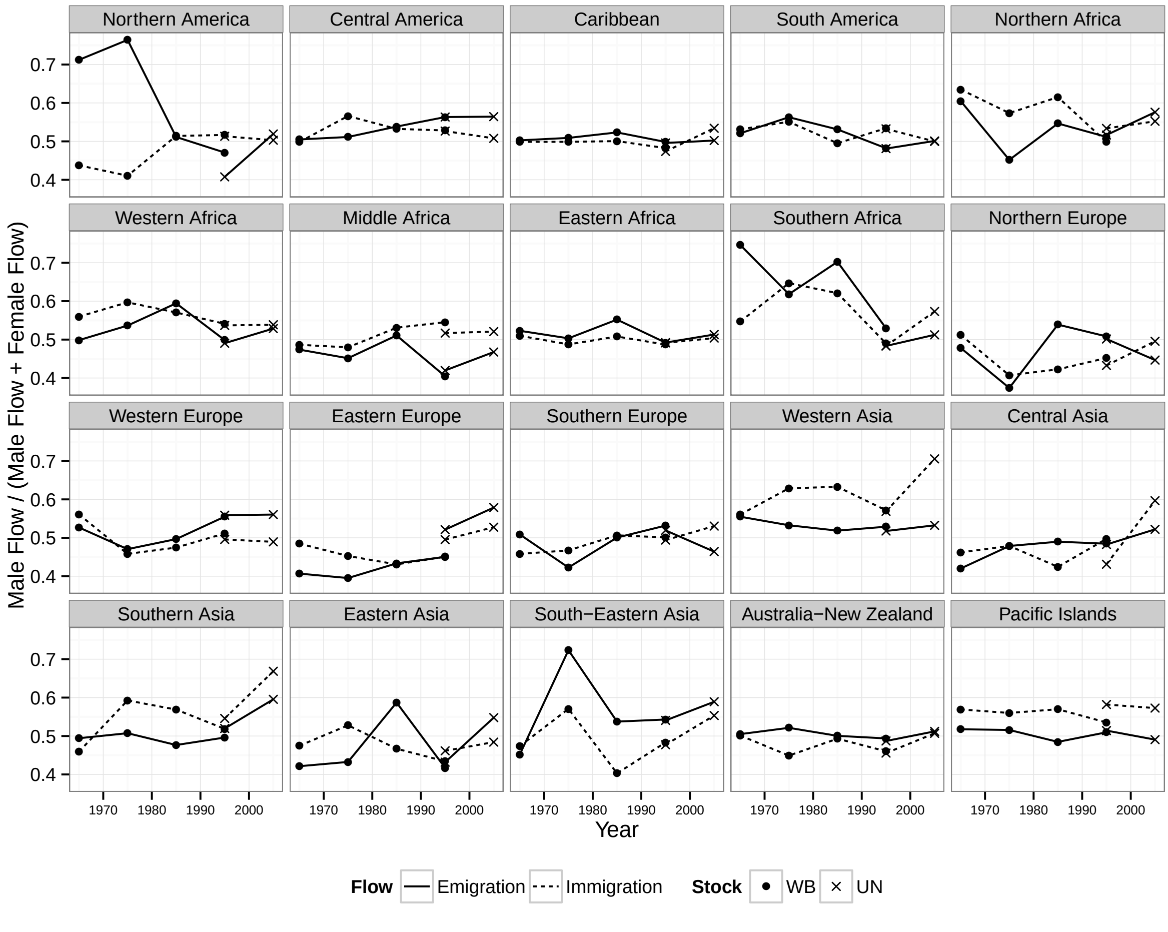


Figure 8: Male percentage of total estimated immigration and emigration flows over a 10 year periods. Estimates based on aggregations of country to country bilateral flows, both World Bank and UN stock data (WB or UN) and WPP2010 demographic data.

(2013) a handful of unexpected flow estimates result from peculiarities in the input stock data. The same unexpected flows area also found in the estimates presented in this paper 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 movement in the reported stocks creates a large estimated outflow of Chinese in the 1960's. These emigrants are estimated to move to countries where there are increases in the number of Chinese born, including but not exclusively, China. In turn, during the 1970's there is a large estimated inflow back into Hong Kong of Chinese born, to meet the sudden increase in their migrant stock.

In the remainder of this section a further analysis of the sensitivity of estimates to alternative demographic data, the other source of input data required to estimate flows from stocks, is analysed. This is followed by a comparison between the flows presented in the previous section with those adjusted for the changes in past political geography are drawn.

5.1 Demographic Data Source

The UN Population Division updates demographic estimates for all countries every two years. The results presented so far have all been based on the WPP2010 version. Total migration flow estimates and crude global migration rate based on the WPP2012 are shown by the dashed line in Figure 9. The total flows from the WPP2010 data are given by the solid line and match those in Figure 2. The updated demographic data has a noticeable effect on increasing the total estimated flows of both the ten-year and five-year interval estimates during the last decade. For example, the estimate of all flows for the ten-year interval 2000-10 is 82.85 million, 4.89 million higher than estimates using WPP2010 version.

Higher migration from the WPP2012 data are partly due to an increased number of countries used. For estimates based on World Bank stock data, flows for four additional countries; Antigua and Barbuda, Kiribati, Seychelles and Taiwan, can be estimated, as complete demographic data is available in the WPP2012[7]. For one country, Netherlands Antilles, estimates based on the World Bank stock data were no longer possible, as data for its population, births and deaths are not provided in WPP2012 (but was provided in WPP2012). For estimates based on differences in UN stock data, flows for four additional countries; Antigua and Barbuda, Kiribati, Seychelles and Curacao, can be obtained, as complete demographic data is available in the WPP2012.

In earlier periods, the effect of alternative demographic data has 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 10 plots differences of both immigration and emigration by region.

The largest difference between the estimates from alternative demographic data sources appear in the emigration estimates from Southern Asia during 2000-10. During this period, there are over four million more estimated emigrants based on the WPP2012 input data than from the WPP2010 data. This difference results from changes in the estimates of the number

[7]Note, although World Bank provided stock data for each of these countries, estimates based on WPP2010 data were not possible due to the lack of demographic data in WPP2010.

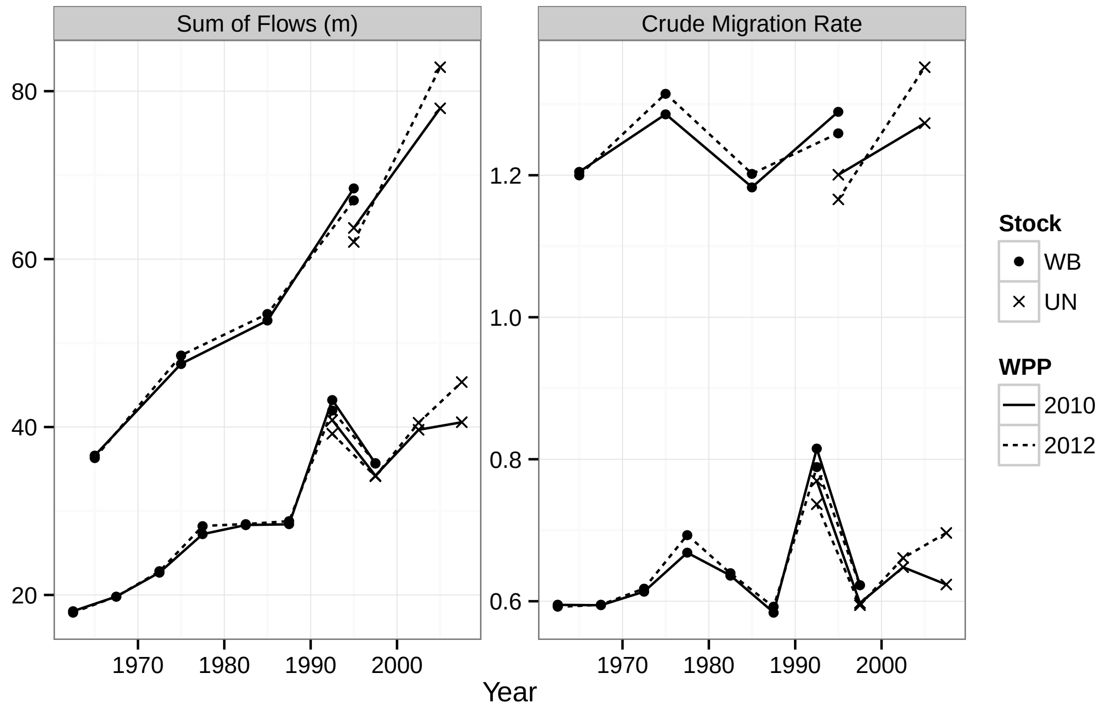


Figure 9: Total estimated migration flows and crude migration rate over a five and ten-year periods from alternative demographic data sources. Estimates based on aggregations of country to country bilateral flows and both World Bank and UN stock data (WB or UN).

of birth, deaths and the population at the start and end of the decade. The 2000 and 2010 population totals were revised down by 12.35 and 22.74 million respectively. The number of births and deaths over the period were also revised down by 12.00 and 5.77 million. When the updated demographic data of the WPP2012 were used to estimate the flows, the immigration into Southern Asia increases from 2.80 million (from WPP2010) to 3.04 million. Emigration from South Asia increases from the 15.76 million estimated using WPP2010 data to 20.21 million from WPP2012, to match the lower population totals in the WPP20102 data. The patterns of bilateral migrant flows out of Southern Asia, shown in Figure 11, appears to be similar where the migrant stock data used has not changed. However, the size of major flows to destinations in Western Asia, North America and Europe have increased.

One further large change during the 2000-10 period from alternative demographic data was for flows into Russia. Russia positive net migration over the ten-year period from WPP2010 data was 2.7 million whilst in WPP2012 the value increased to 3.89 million. This revision is passed to input data via a higher 2010 population (revised up by 0.66 million) and a lower number of deaths (revised down by 0.5 million). Consequently, larger flows of Russian born from 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 USA (356 thousand, up from 56 thousand for estimates based on WPP2010 data) and Germany (175 thousand, up from 6.5 thousand).

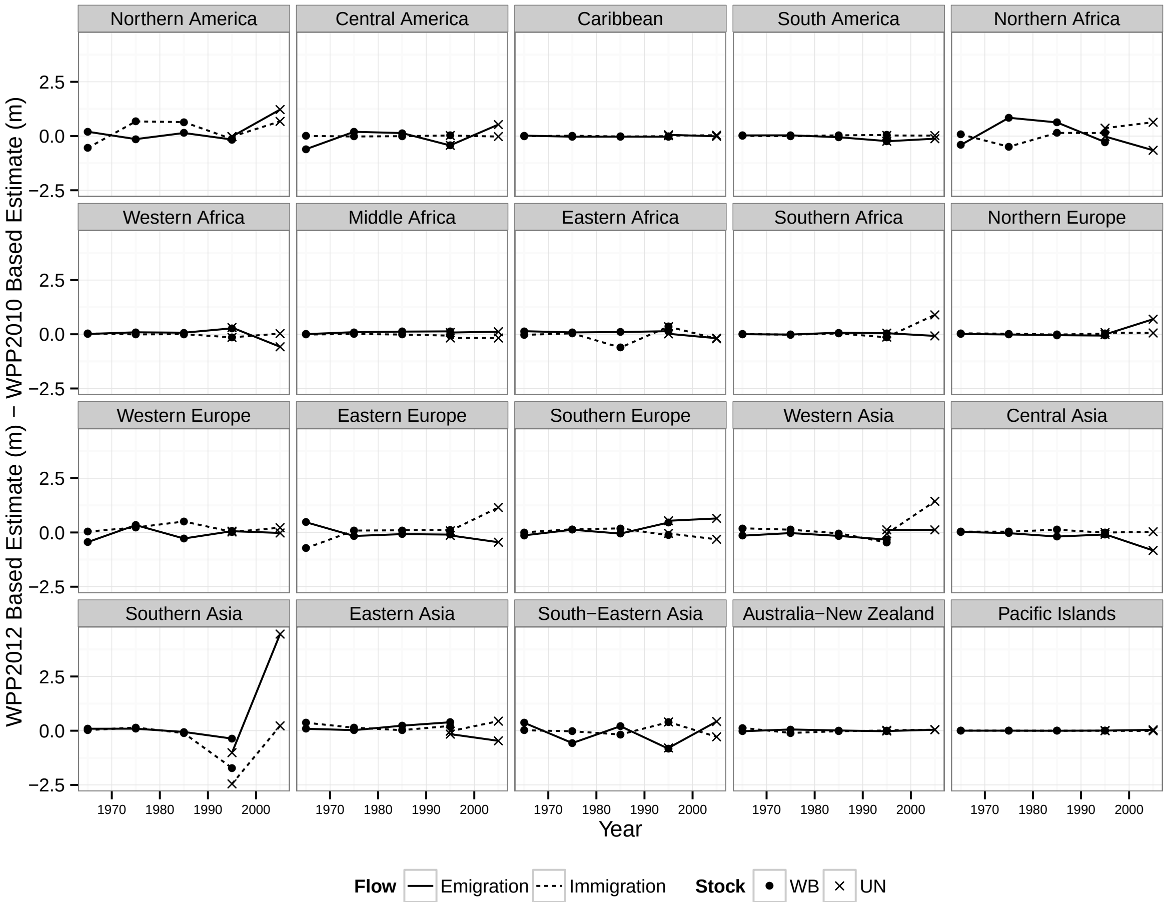
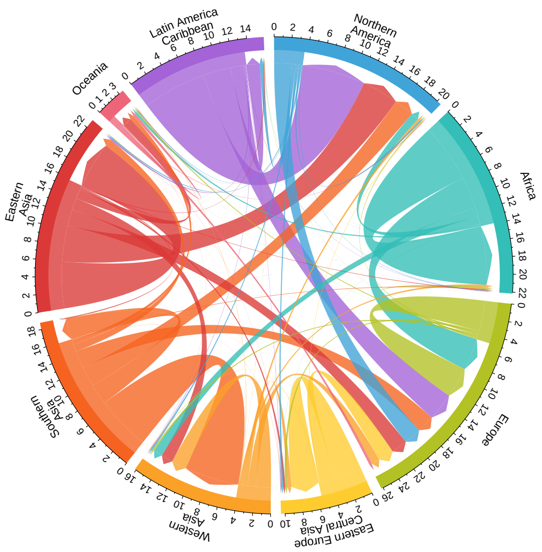
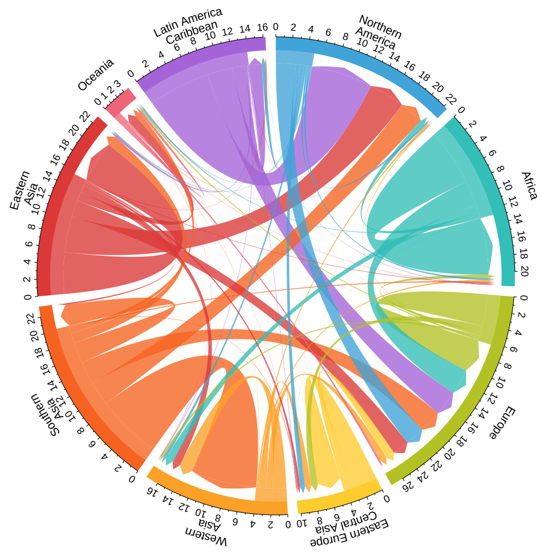


Figure 10: Differences in estimated migration flows over a 10 year period from alternative demographic data sources. Estimates based on aggregations of country to country bilateral flows and both World Bank and UN stock data (WB or UN).

Figure 11: Estimated 10 year global migrant flows over time.



(a) Total 2000-2010. Based on UN stock data and WPP2010 demographc data.



(b) Total 2000-2010. Based on UN stock data and WPP2012 demographc data.

In some regions, such as North America, Northern Africa, Western Europe or South-Eastern Asia, the choice of demographic data leads to different immigration and emigration estimates, depending on the period at hand. As shown in Figure 10 these differences tend to be less than a million either way. For other regions, such as Western Africa, Northern Europe, Western Asia the demographic data used, only has 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 has very little effect on the gross number of immigrants and emigrants estimated.

At the country level further contradictions between the input demographic and stock data were found due to unexpected estimated flows that they produced. For example, during 2005-10 the demographic data a imply net migration for Poland of +55 thousand (WPP2010) or -70 thousand (WPP2012). These differences contradict the large increases in the UN stock data of Polish born in major destinations counties over the same period, such as the UK and Germany. As the estimation methodology is crude global demographic account, the increases in Polish stocks in the UK and Germany are matched with estimated flows from reported decreases in Polish born populations in the stock data, mainly in France, the US and Canada. Only small amounts of flows from Poland are estimated when the WPP2012 data is used, as the methodology is constrained to the implied net migration (-70 thousand) in the demographic data.

5.2 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 result 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 been fully fledged separate nation states. For example, past bilateral migrant stock information are 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 movements. Secondly, the set of countries based on recent stock data from the UN only provides information on new countries in the 2010 bilateral tables. As a result, separate estimates into, out and between both Serbia and Montenegro and Sudan and South Sudan can not be obtained as differences in the changes in stock patterns can not be utilised[8][9].

In order to analysis the effect of the first of these features; changes in political geography, estimates of flows which at the time would be 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 both a smaller and changing set of countries 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, Czechoslovakia into the Czech Republic and Slovakia. Flow estimates between an 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. Further 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 12 the total flows are plotted using a broken line for estimates adjusted for changes in political geography and the solid line for the original estimates with a fixed set of countries throughout the period. In comparison to the total flows based on the fixed set of countries, the adjusted estimates previous to 1990 are lower. Estimates of both the five and ten-year flows during the earlier periods no longer show a steady increase. 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 12 during this period falls more sharply than estimates based on a fixed number of countries.

6 Summary and Discussion

Global international migration is an ever changing demographic 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 migration flows provide a basis to allow for a better understanding of contemporary movement patterns during a given period, where no existing source of flow data exists.

In this paper, global bilateral flow tables were estimated by gender from 1960 through to 2010. Results were predominantly presented by regions (results for all countries are available in the supplementary materials). The total estimated international migration flows over time are shown to generally increase. The percentage of global population estimated to migrate over five or ten-year period remained fairly steady at 0.6 and 1.25 percent of the global population

[8]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

[9]Demographic data, not provided for the unified areas were obtained by combing data from the separate countries

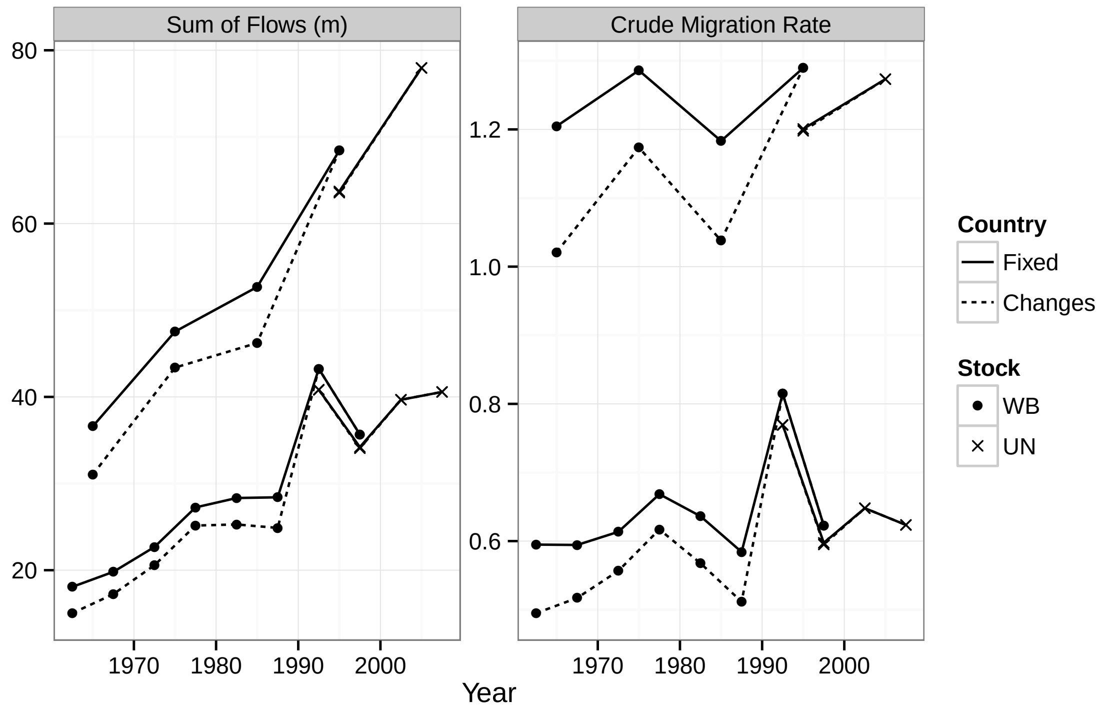


Figure 12: Total estimated intonational migrant flows by length of migration period in each decade. Estimates based on aggregations of country to country bilateral flows, both World Bank and UN stock data (WB or UN) and WPP2010 demographic data. Country definitions are either fixed according the union of stock and demographic data or changing to reflect altering political geographies.

respectively, with a noticeable spike during the 1990-95 period. Some regions were estimated to have continuously increasing numbers of migrants arriving (North America and West Asia) or leaving (Central or South America), whilst 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 movements were estimated from South to West Asia and from Asia to North America, most likely related to economic changes. Large moves were also estimated in selected time periods within Africa or Eastern Europe during times of armed conflicts or political change. Differences in the size of gender of imbalances were found to exist between selected regions, such as recent movements into oil rich Gulf States from South Asia.

A large proportion of the estimated flows were zero and the overall distribution of non-zero flows were skewed to small counts. These characteristics are driven by the stock data and the inherent assumptions required in flows from stock estimation methodology. In both the World Bank and UN there are large amounts of zero bilateral migrant stocks throughout, which results in a zero flow estimates. The estimated flows are not approximations to a true flow, but instead are the minimum amount of migrants required to match the demographic data and changes in migrant stocks.

The estimates of bilateral migration flows presented provide a number of new insights to the global migration flow literature. First, the methodology of Abel (2013) 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. Whilst the patterns are drivers of migration flows are discussed in far less detail in this paper, many of the observed trends of 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 non-existent population. Further, they are sensitive to changes in both immigration or emigration patterns and hence their time series are often volatile. Fourth, whilst ten-year estimates based on the World Bank data are comparable to those of Abel (2013), the estimates in this paper 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).

The choice of input data used were found to bare influence on the estimated flows. Stock data was 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 UN and World Bank to provide complete and comparable bilateral stock estimates. Demographic data was found to bare some 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 time scales. Estimates over five-year periods were found to detect large movements, such as those induced by armed conflicts or political changes that were not as clearly identifiable for estimates over ten-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 allows for the first time a comprehensive system to compare global demographic data for inconsistencies, check for errors and match available data with conceptual models or known migrant flows. The estimates presented in this paper uncovered a number of unexpected results which can be sourced to errors or inconsistencies in the stock or demographic data used as inputs. For example, large flows of Chinese born were estimated to go in and then out of Hong Kong during 1960's and 1970's. 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. 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 deficiencies was handled by an extension to the methodology outlined in the Appendix. In both of the latter two cases the unusual flow estimates are derived from underestimates of the native born populations, which are themselves 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 demographic estimates of population size are relatively more simple to produce. Further, in the case of Russia it is likely that the next revision of stock estimates by the UN will utilise the latest population size data to estimate foreign born counts, as has been the case with past versions. Future World Bank stock estimates might also employ improved estimation procedures, as used for more recent estimates of bilateral stock tables by gender and skill level (Artuç et al., 2012). 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.

International migration depends on a complex mix of social, economic, political and demographic factors. 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 paper, 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 international global migration trends.

References

Abel, G. J. (2010). "Estimation of international migration flow tables in Europe". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 173.4, pp. 797–825.

Abel, G. J. (2012). *migest: Useful R Code for the Estimation of Migration.*

Abel, G. J. (2013). "Estimating global migration flow tables using place of birth data". In: *Demographic Research* 28.March, pp. 505–546.

Abel, G. J. and N. Sander (2014). "Quantifying global international migration flows." In: *Science* 343.6178, pp. 1520–2.

Artuç, E. et al. (2012). "A global assessment of human capital mobility: the role of non-OECD destinations".

Beer, J. et al. (2010). “Overcoming the Problems of Inconsistent International Migration data: A New Method Applied to Flows in Europe”. In: *European Journal of Population/Revue européenne de Démographie* 26.4, pp. 459–481.

Beine, M., F. Docquier, and c. Özden (2011). “Diasporas”. In: *Journal of Development Economics* 95.1, pp. 30–41.

Castles, S., H. de Haas, and M. J. Miller (2013). *The Age of Migration: International Population Movements in the Modern World*. 5th ed. Vol. 17. Palgrave Macmillan.

Czaika, M. and H. de Haas (2014). “The Globalization of Migration: Has the World Become More Migratory?” In: *International Migration Review* Summer, n/a–n/a.

Donato, K. M. et al. (2006). “A Glass Half Full? Gender in Migration Studies”. In: *International Migration Review* 40.1, pp. 3–26.

Gu, Z. et al. (2014). “circlize implements and enhances circular visualization in R.” In: *Bioinformatics*, pp. 1–2.

Hawelka, B. et al. (2014). “Geo-located Twitter as proxy for global mobility patterns”. In: *Cartography and Geographic Information Science* 41.3, pp. 260–271. arXiv: 1311.0680.

Kelly, J. J. (1987). “Improving the Comparability of International Migration Statistics: Contributions by the Conference of European Statisticians from 1971 to Date”. In: *International Migration Review* 21.4, pp. 1017–1037.

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*. Ed. by F Willekens and J Raymer. London, England: Wiley. Chap. 3, pp. 41–73.

Lee, R. D. (2011). “The outlook for population growth.” In: *Science* 333.6042, pp. 569–73.

Martin, P. and J. Widgren (2002). “International migration: Facing the challenge”. In: *Population Bulletin* 57.1.

Massey, D. S. et al. (1999). *Worlds in Motion: Understanding International Migration at the End of the Millennium*. International Studies in Demography. Oxford, United Kingdom: Oxford University Press.

Mayer, T. and S. Zignago (2011). “Notes on CEPII’s Distances Measures: The GeoDist Database”. In: *CEPII Working Paper* 2011-25.

National Research Council (2000). *Beyond six billion: Forecasting the world’s population*. Ed. by J Bongaarts and R. Bulatao. Washington D.C., USA: National Academy Press. Chap. 3. P. 258.

Nowok, B., D. Kupiszewska, and M. Poulain (2006). “Statistics on International Migration Flows”. In: *Towards the Harmonisation of European Statistics on International Migration (THESIM)*. Ed. by M Poulain, N Perrin, and A Singleton. d. Louvain-La-Neuve, Belguim: UCL–Presses Universitaires de Louvain. Chap. 8, pp. 203–233.

Özden, c. et al. (2011). “Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960–2000”. In: *World Bank Economic Review* 25.1, pp. 12–56.

R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Raymer, J. (2007). “The estimation of international migration flows: a general technique focused on the origin–destination association structure”. In: *Environment and Planning A* 39.4, pp. 985–995.

Raymer, J. et al. (2013). “Integrated Modeling of European Migration”. In: *Journal of the American Statistical Association* 108.503, pp. 801–819.

Rees, P. H. (1980). “Multistate demographic accounts: measurement and estimation procedures”. In: *Environment and Planning A* 12.5, pp. 499–531.

Rogers, A. (1990). “Requiem for the Net Migrant”. In: *Geographical Analysis* 22.4, pp. 283–300.

Sander, N. et al. (2014). “Visualising Migration Flow Data with Circular Plots”. In: *Vienna Institute of Demography: Working Papers* 02.

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*. New York, New York, USA: ACM Press, p. 265.

United Nations Population Division (2011). *World Population Prospects: The 2010 Revision. CD-ROM Edition*. New York, USA.

United Nations Population Division (2013a). *Trends in International Migrant Stock: Migrants by Destination and Origin (United Nations database, POP/DB/MIG/Stock/Rev.2013)*. New York, USA.

United Nations Population Division (2013b). *World Population Prospects: The 2012 Revision. DVD Edition*. New York, USA.

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer New York.

Wiśniowski, A. et al. (2013). “Utilising Expert Opinion to Improve the Measurement of International Migration in Europe”. In: *Journal of Official Statistics* 29.4, pp. 583–607.

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*, pp. 1–10.

Zagheni, E. et al. (2014). “Inferring International and Internal Migration Patterns from Twitter Data”. In: *Proceedings of the companion publication of the 23rd international conference on World wide web companion*. Seoul, Korea, pp. 439–444.

Zlotnik, H (1999). “Trends of international migration since 1965: what existing data reveal.” In: *International Migration* 37.1, pp. 21–61.

Zlotnik, H. (1995). “The South-to-North migration of women.” In: *International Migration Review* 29.1, pp. 229–254.

Appendix

A Representing Bilateral Migrant Stock Data in Flow Tables

Bilateral migrant stock data can be represented in contingency tables. For example, the bilateral migrant stock data at time t and $t+1$ given in Figure 1 can be displayed in two contingency tables as in Table 2. In each table, the rows represent a categorization of the population, which in Table 2 is place of birth. 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 its place of birth and place of residence at a specified time. Values in diagonal cells represent the number native born. 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

		Place of Residence (t)							*Place of Residence* ($t+1$)				
		A	B	C	D	Sum			A	B	C	D	Sum
Birthplace	A	100	10	10	0	120	*Birthplace*	A	70	30	10	10	120
	B	20	55	25	10	110		B	25	60	10	15	110
	C	10	40	140	65	255		C	10	55	140	50	255
	D	20	25	20	200	265		D	40	45	0	180	265
	Sum	150	130	195	275	750		Sum	145	190	160	255	750

Table 2: Contingency table of bilateral migrant stock data given in Figure 1.

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 some of other measure, such as citizenship (not considered in this paper), the column totals may no longer represent a total population, but a count over the number of citizens or nationals. In such cases, the column totals can potentially be greater than the population as persons 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 example data used in Table 2, where there are no births or deaths during the period between, the row totals are the same, as country of birth is fixed characteristic. In this case, where births and deaths have been accounted for, Abel (2013) 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 Table 3.

The (originally) unknown cells within the migrant flow table represent birthplace specific non-movers on the diagonal and birthplace specific migrant transitions on the non-diagonal. Abel (2013) outlined two assumptions to estimate these unknown cells. The first is based on a maximising assumption to fix the diagonal terms to their highest possible value, conditional on the known marginal stock counts. These are illustrated by the italic typeface in Table 3. Flow counts within the birthplace migrant flow table are then estimated by assuming a log-linear model adapted to account for the known margins and diagonal elements over the multiple tables. The parameters can be obtained using an iterative proportional fitting type algorithm fully detailed in Abel (2013) and available in the *migest* R package (Abel, 2012). Given the parameter estimates, the imputed values for the non-diagonal elements shown in each of the birth place specific flow tables can be estimated. These estimates match those shown in Figure 1 and can be summed over birthplaces to provide the origin-destination flow table shown in Table 1.

Underlying the estimation of the flows is log-linear model which includes an offset term. This factor allows for a single set of auxiliary information to be used in the iterative proportional algorithm to augment the estimated flows and provide more realistic estimates. In the example of Table 3 the auxiliary variable is set to one for all bilateral combinations. In previous applications to real bilateral migrant stock data, and as used in the calculations discussed in this article, a distance function based data provided by the Centre d'Etudes Prospective et d'Informations Internationales (Mayer and Zignago, 2011) was used. Given distance measures (d_{ij}) between all

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**

Table 3: Bilateral stock data of Table 3 arranged as birthplace specific origin destination migrant flow tables.

capital cities at origin country i and destination j, the offset term is calculated as $m_{ij} = d_{ij}^{-1}$ to weight in favour of moves to closer countries.

B Controlling for Natural Population Change

The application of the flow from stock methodology outlined in the section above relies upon an equal row totals (the population of people born in each area) at the beginning and end of the period. In the case where there are births and deaths over the period, row totals are unlikely to be equal, as is the case in the dummy example data of Table 4, where new data are provided in time period $t + 1$

In order to account for births and deaths over the period, Abel and Sander (2014) showed that the stock data can be adjusted through a multi-step correction process, resulting in altered migrant stock tables with equal row totals. In the first step of the process, deductions to the bilateral stock data are made to account for natural population change. The number of deaths during the time interval, supplied from basic demographic data, are subtracted from the available bilateral stocks data at time t. As the disaggregation of the number of deaths in each place of residence by birthplace is typically not known, estimates can be obtained by assuming equal mortality rates for all migrant stocks and the native population. The count of the number of deaths can then be proportionally spread to each sub-population[10]. This step is illustrated on

[10] In the case of Abel and Sander (2014), additional information on the median age of native born and foreign migrant stocks (as a whole) in each country between 1990 and 2010 were used to proportion less deaths to the

		Place of Residence (t)							Place of Residence ($t+1$)				
		A	B	C	D	Sum			A	B	C	D	Sum
Birthplace	A	100	10	10	0	120	Birthplace	A	75	25	5	0	105
	B	20	55	25	10	110		B	20	45	30	15	110
	C	10	40	140	65	255		C	30	40	150	60	280
	D	20	25	20	200	265		D	30	30	20	230	310
	Sum	150	130	195	275	750		Sum	155	140	205	305	805

Table 4: Dummy example of place of birth data with births and deaths over the interval

the left hand side of Step 1 in Table 5. The known total number of deaths, given in bold type face in the final row, is proportionally split according to the population stock sizes in time t.

To control for increases in native born population totals from newborns, the number of births between t and $t+1$ is subtracted from the reported stock data at time $t+1$. As with deaths, data are typically only available on the total number of births during an interval, and not their subsequent location at time $t+1$. Assuming there is no migration of newborns, the total number of births can be subtracted from the native born populations. This is illustrated on the right hand side of Step 1 in Table 5, where the known total number of births is given in bold type face. These newborn population totals are assumed to reside in the same place of residence at time $t+1$ for countries A, C and D.

In country C an alteration is made to previous demographic accounting methodologies presented in Abel (2013) and Abel and Sander (2014). When initially applying previous methodologies to the range of stock and demographic data used in this paper occasional negative estimated migration flows would result. These cases occurred where place of birth stock data imply a change in native born population which directly conflicts with changes in demographic data unless there is mass emigration of all newborns. For example, for the State of Palestine the native born population size, calculated as the remainder of the total population from UN demographic data and the foreign born population from the World Bank migrant stock data is estimated to be 0.54 million in 1970. The number of births over the ten-year period prior is given by United Nations Population Division (2011) as 0.59 million. If the number of births are subtracted from the native born population, a negative total results in the bilateral stock post an adjustment for natural population change. Further on in the estimation process, negative flows are estimated to match the negative native born stock total. In order avoid this case, the number of births are distributed proportionally to all migrant stocks at time $t+1$ cases where there appears to be a direct conflict between the migrant stock and demographic data. Note, in all the 108 estimated migrant flow tables discussed in this article, this alternative assumption was required for only one country (the State of Palestine) when calculating ten-year flows in a few selected decades.

An adjusted stock tables, where both the death and birth estimates of the previous step are subtracted cell-wise from the original data in Table 4, are shown in Step 2 of Table 5. The

relatively younger group. This approach is not used to calculate the results in this paper as no equivalent data on the median age of migrants exist for the World Bank stock data.

Step 1: Control for Natural Changes

Place of Death

Birthplace	A	B	C	D
A	20.0	0.8	2.1	0.0
B	4.0	4.2	5.1	0.4
C	2.0	3.1	28.7	0.4
D	4.0	1.9	4.1	7.3
Sum	**30**	**10**	**40**	**10**

Place of Residence $(t+1)$

Birthplace	A	B	C	D
A	10.0	8.9	0.0	0.0
B	0.0	16.1	0.0	0.0
C	0.0	14.3	25.0	0.0
D	0.0	10.7	0.0	60.0
Sum	**10**	**50**	**25**	**60**

Step 2: Estimated Altered Stocks

Place of Residence (t)

Birthplace	A	B	C	D	Sum
A	80.0	9.2	7.9	0.0	97.2
B	16.0	50.8	19.9	9.6	96.3
C	8.0	36.9	111.3	62.6	218.8
D	16.0	23.1	15.9	192.7	247.7
Sum	120.0	120.0	155.0	275.0	660.0

Place of Residence $(t+1)$

Birthplace	A	B	C	D	Sum
A	65.0	16.1	5.0	0.0	85.7
B	20.0	28.9	30.0	15.0	90.2
C	30.0	25.7	125.0	60.0	240.2
D	30.0	19.3	20.0	170.0	238.9
Sum	145.0	90.0	180.0	245.0	655.0

Step 3: Re-estimated Altered Stocks

Place of Residence (t)

Birthplace	A	B	C	D	Sum
A	76.7	8.2	6.8	0.0	91.6
B	17.0	49.8	18.8	9.5	95.1
C	9.2	39.4	114.4	66.7	229.8
D	17.0	22.6	15.0	188.8	243.5
Sum	120.0	120.0	155.0	265.0	660.0

Place of Residence $(t+1)$

Birthplace	A	B	C	D	Sum
A	68.6	17.4	5.6	0.0	91.6
B	19.5	29.0	31.4	15.2	95.1
C	27.3	24.1	121.9	56.5	229.8
D	29.6	19.5	21.1	173.3	243.5
Sum	145.0	90.0	180.0	245.0	660.0

Table 5: Multi-step demographic account framework using stock data from Table 4

overall totals of the altered stock tables are now equal, as the difference in the original data between the two periods ($805 - 750 = 55$) is fully accounted for by the natural increase from births and deaths ($145 - 90 = 55$). However, as the new altered stock tables do not have equal row totals, further adjustments are required in order to estimate migrant flow. These differences are likely to represent the sum of differences in migrant stock data collection procedures of each region. As in Abel and Sander (2014), a simple iterative proportional fitting scheme is applied to scale each stock table to 1) maintain their column totals in Step 2 and 2) fix the row totals to an average of the adjusted row totals obtained in Step 2, and 3) maintain the same interaction structure with in the re-estimated stock totals as in those calculated in Step 2. The resulting re-estimated altered stock are shown in Step 3 of Table 5.

As the re-adjusted estimates shown in Stage 3 of Table 5 have the same row totals in t and $t+1$, the updated stocks can be considered as a set of marginal totals for a set of birthplace specific flow tables as given in by the boldface type in the top panel of Table 6. The same assumptions as in Table 3 can be used to set the diagonal cell values for the non-movers in italics, allowing the iterative proportional fitting algorithm of Abel (2013) to estimate the unknown flows in the non-diagonals. Aggregating over all birthplaces and removing those with same origin and destination in the diagonal elements gives a traditional origin-destination flow table of migrant transitions during the time period t to $t+1$ shown in the bottom panel of Table 6. Estimates are not directly comparable with previous flow shown in Table 1 and Table 3 as they are formed from a different set of migrant stock data in $t+1$.

Estimates of Origin Destination Place of Birth Flow Tables:

Birthplace=A

		Destination				
		A	B	C	D	Sum
Origin	A	*68.6*	8.1	0.0	0.0	**76.7**
	B	0.0	*8.2*	0.0	0.0	**8.2**
	C	0.0	1.1	*5.6*	0.0	**6.8**
	D	0.0	0.0	0.0	*0.0*	**0.0**
	Sum	**68.6**	**17.4**	**5.6**	**0.0**	**91.6**

Birthplace=B

		Destination				
		A	B	C	D	Sum
Origin	A	*17.0*	0.0	0.0	0.0	**17.0**
	B	2.5	*29.0*	12.6	5.7	**49.8**
	C	0.0	0.0	*18.8*	0.0	**18.8**
	D	0.0	0.0	0.0	*9.5*	**9.5**
	Sum	**19.5**	**29.0**	**31.4**	**15.2**	**95.1**

Birthplace=C

		Destination				
		A	B	C	D	Sum
Origin	A	*9.2*	0.0	0.0	0.0	**9.2**
	B	10.8	*24.1*	4.5	0.0	**39.4**
	C	0.0	0.0	*114.4*	0.0	**114.4**
	D	7.2	0.0	3.0	*56.5*	**66.7**
	Sum	**27.3**	**24.1**	**121.9**	**56.5**	**229.8**

Birthplace=D

		Destination				
		A	B	C	D	Sum
Origin	A	*17.0*	0.0	0.0	0.0	**17.0**
	B	2.1	*19.5*	1.0	0.0	**22.6**
	C	0.0	0.0	*15.0*	0.0	**15.0**
	D	10.5	0.0	5.1	*173.3*	**188.9**
	Sum	**29.6**	**19.5**	**21.1**	**173.3**	**243.5**

Estimates of Total Origin Destination Flow Table:

		Destination				
		A	B	C	D	Sum
Origin	A		8.1	0.0	0.0	15.7
	B	15.5		18.1	5.7	53.8
	C	0.0	1.1		0.0	3.9
	D	17.7	0.0	8.0		32.7
	Sum	33.2	9.2	26.1	5.7	74.2

Table 6: Bilateral stock data of Table 4, controlling for natural change as in Table 5, arranged as birthplace specific origin destination migrant flow tables with resulting flow estimates.