VIENNA INSTITUTE OF DEMOGRAPHY

Working Papers

Guy J. Abel

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

Vienna Institute of Demography

OAW 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 2015 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 0.65 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. Global migration during 2010-15 fell in comparison previous periods. The sensitivity of flow estimates to alternative input stock and demographic data as well as changes in political geography are explored. Estimates are validated through comparisons with existing reported migration flows statistics.

Keywords

International migration, migration estimation, migration flows, migration stocks, global migration, bilateral migration.

Authors

Guy J. Abel, Professor, Asian Demographic Research Institute, Shanghai University, China and Research Scholar, Wittgenstein Centre (IIASA, VID/ÖAW, WU), Vienna Institute of Demography, Austria. Email: guy.abel@oeaw.ac.at

Acknowledgements

Estimates of Global Bilateral Migration Flows by Gender Between 1960 and 2015

Guy J. Abel

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 and allows hypotheses on the determinants and patterns of peoples moves to be tested.

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 Nations 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 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; 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 measures, changes in bilateral stocks over time and available, unharmonised flow data from predominately rich Western countries. A growing literature on the analysis of bilateral migrant stock data (Beine, Doquier, and Özden, 2011; Czaika and Haas, 2014; Docquier et al., 2012) 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, p.200). Further, 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 to those with younger populations, where the cumulative time available to people to emigrate is lower. Recent 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-located of internet activities such as twitter messages or logins to email services. As the authors note, their data may not be fully representative of the global migration patterns and is not always publicly available.

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, a number innovations on previous global bilateral flows are made. First, bilateral flow estimates are produced for each gender, quantifying for the first time, differences in male and female global migration flow patterns. Previously, both Piper (2005) and Zlotnik (2003) note 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. Distinct gender variations in the migration patterns are known to exist from the stock data and localised studies (Donato et al., 2006; Zlotnik, 1995). 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). In particular, the role of gender differentials in international moves for domestic workers is often highlighted. The International Labour Office estimated 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 (2013) and Abel and Sander (2014) is extended to account for contradictions between demographic and stock data. The revised method is applied to estimate five and ten-year migrant flows separately by gender between 1960 and 2015, to provide an updated view of international migration over a far longer time period. Estimates over both five and ten-year periods enable for contrasts between possible different global migration transitions rates to be identified.

Third, estimates of migrant flows in this paper 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 births and deaths 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 273 estimated migrant flow tables, far exceeding those in the previously discussed flows from stocks 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 methodology to alternative demographic input data and changes in political geography are discussed followed by a comparison of the estimates to reported data from national statistics institutes. Finally, the results are summarised and discussed in reference to current work on global migration data. The appendix provides a detailed review of the flows from stocks methodology outlined in Section 3 as well as some further sensitivity analyses.

2   Methodological Background

Available bilateral migration data can be categorised as either a stock measure, that represents a static number of a foreign population defined by a characteristic such as their place of birth or a flow measure, that represents 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 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 a general outline on how bilateral migrant flows data can be indirectly estimated from sequential bilateral migrant stock tables whilst accounting for demographic changes over the period. This is followed by some additional discussion on log-linear models which forms the statistical heart of the flow from stock methodology.

2.1   Flows from Stocks Method

Changes in bilateral migrant stock sizes over time, 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 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 illustrates this case using a schematic of a simple demographic account framework

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

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

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 residing in any country does not change over the 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 log-linear model, details of which are given in the next subsection.

The results of the applied indirect estimation method for the global system of four countries are shown by the arrows in Figure 1. These estimates 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 may alternative be measured as the number of migrant 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 transitions, there is no simple mathematical solution to estimate one from the other.

The demographic framework in Figure 1 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 stocks data over a sizeable time period (such as five or ten years). 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
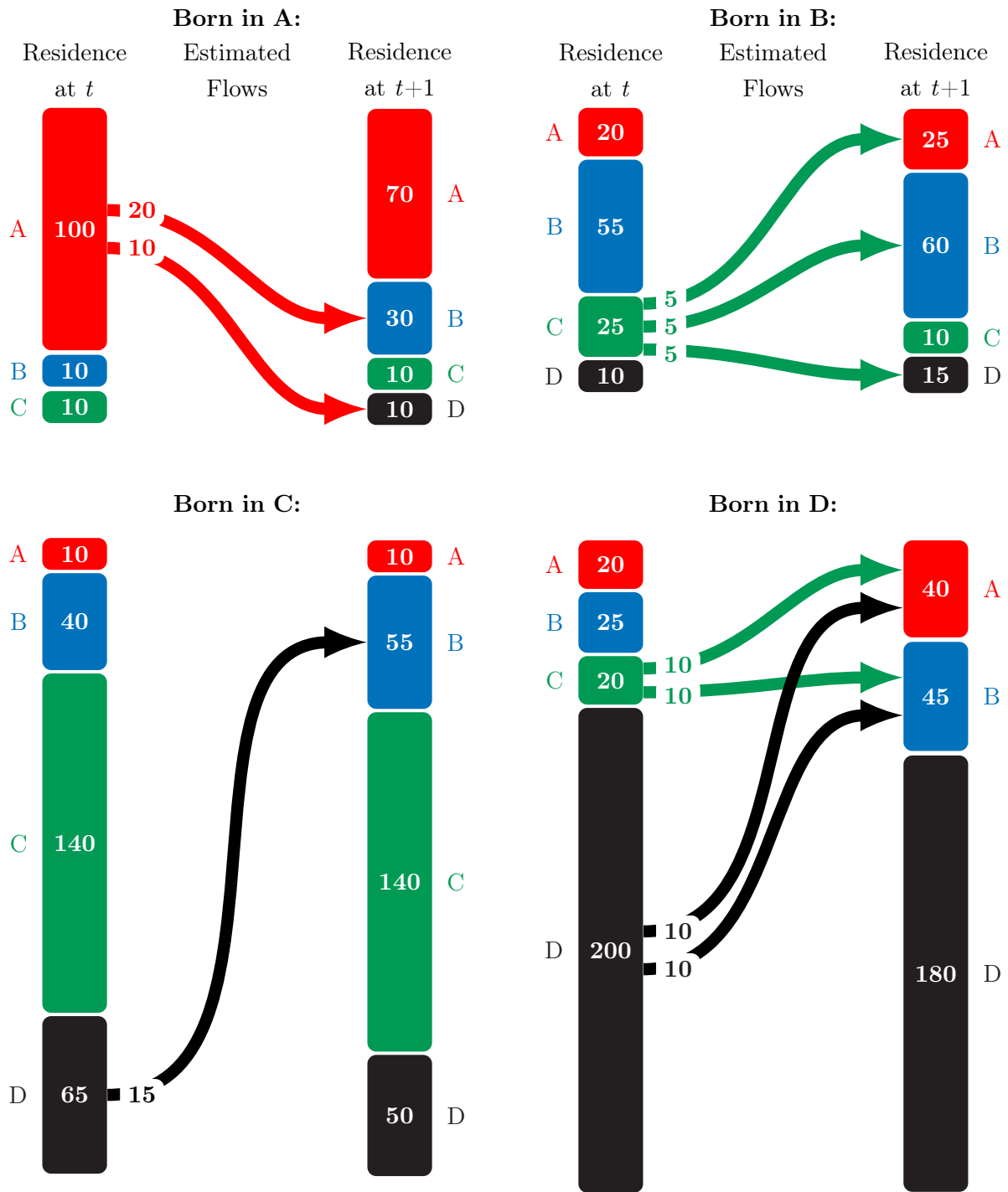
Figure 1: Schematic of a demographic accounting framework to link changes in bilateral migrant stock data via estimated migrant flows. Note, for each birthplace there are no births or deaths during the time interval. Thus, the total birthplace populations are the same at time t and $t+1$, represented 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.

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 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.

2.2   Log-linear Models

Log-linear models are a form of Poisson regression model, where the explanatory variables are all categorical. They can be used to predict missing cells in migration flow tables that match known marginal totals as 1) 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 the Appendix of this paper, details are given on how migrant stocks, such as those in Figure 1, can be represented as the marginal totals of a three-way array of origin-destination flow tables. A log-linear model can be fitted to the array, with categorical explanatory variables for the origin, destination, birthplace and some of their interactions corresponding to the known marginal totals. Parameters are estimated for the log-linear model with an Iterative Proportional Fitting (IPF) algorithm. Missing migrant flows, with values summing to the known marginal totals, are predicted from the log-linear model using the converged parameter values.

Two further extensions can be made to the log-linear model to help estimate migration flows from marginal totals derived from migrant stock data. First, as with any Poisson regression model, an offset term whose parameter estimate is fixed to unity, can be included to provide auxiliary data to aid the estimation of missing flows without altering constraints on the known marginal sums. In the estimation of migration flows distance measures are typically used. The effect of auxiliary data when estimating flows from migrant stock tables is relatively minor due to the marginal constraints imposed in the methodology. This is studied at more length in the Appendix, by comparing of estimates of migration flows from changes in stocks using a log-linear model with and without an offset term.

A second possible extension is to include further parameters in the log-linear model to account for diagonal cells in a migration flow table. These cells represent populations which have the same country of residence at the start and end of the time period and hence by definition are counts on the number of stayers. Additional parameters for these diagonal cells allow the log-linear model to have the same fitted diagonal counts as those observed. Consequently, imputations for the non-diagonal cells can be provided using the model equation. These imputations will sum to match the constrained margin totals whilst accounted for the number of stayers in the diagonal.

In previous applications of the log-linear model for estimating flow from stocks, the number of stayers have been assumed to be the maximum possible values implied by the corresponding

---

[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 (a transition which is assumed to not occur).

stock data. As a result, the estimated flows are the minimum number of transitions required to match the changes in migrant stocks, as illustrated in Figure 1. In the Appendix a relaxation of the maximum stayers assumption is explored. The total number of migrant flows are shown to linearly decrease as the number of stayers are increased towards their maximum values. However, it is unclear what reduction, if any, in the number of stayers away from the maximum is optimal without detailed knowledge of the international migration propensities in each country and time period. As no strong empirical evidence is available, the maximum number of stayers assumptions is kept for subsequent estimates in the remainder of this paper.

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 sets of bilateral stock data that include more than one time period. Additional bilateral migrant stock data does exist, such as estimates by Artuc et al. (2015), Dumont, Spielvogel, and Widmaier (2010), Ratha and Shaw (2007) or Parsons et al. (2007) but are not used, as they either are restricted in their global coverage or provide stocks only at a 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. Births and death information is needed to alter stock data for natural change over the time period for which flow estimates are being derived. Population data is 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) provide foreign born migration stock tables at the start of each 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 (2015b) provides a sequence of foreign born migrant stock tables five years apart, beginning in 1990 until 2015 covering 232 countries[3]. 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 are primarily based on place of birth responses to census questions and from population registers. 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

---

[2]Data available from `http://data.worldbank.org/data-catalog/global-bilateral-migration-database`
[3]Data available from `http://www.un.org/en/development/desa/population/migration/data/`

group of countries are 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 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 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 alternative 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 UN release an updated versions 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 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 three most recent versions of WPP are used, WPP2010, WPP2012 and WPP2015, in order to determine what effect, if any, updated demographic data has on bilateral migration flow estimates.

4   Results

In order to understand the role of varying components of the flows from stock estimation methodology as well as better understand past patterns of global migration flows, 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 ten-year periods were estimated. The 1960-70, 1970-80 and 1980-90 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 options as in the previous periods) and stock data (WB2011, UN2012, UN2013 and UN2015). In the last ten-year period; 2000-10, 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 the 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 and stock data when estimating the ten-year flows, providing 180 estimated flow tables. A further three flow tables were also estimated for the 2010-15 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 five-year migrant flow tables, for 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 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 flows estimates vary by different gender, time period, interval length, stock and demographic data, provided a combined data set with over ten 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 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.

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, 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 are 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 ten-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 UN are 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 between 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 mid-point of the corresponding period on the horizontal axis. 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 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. During the most
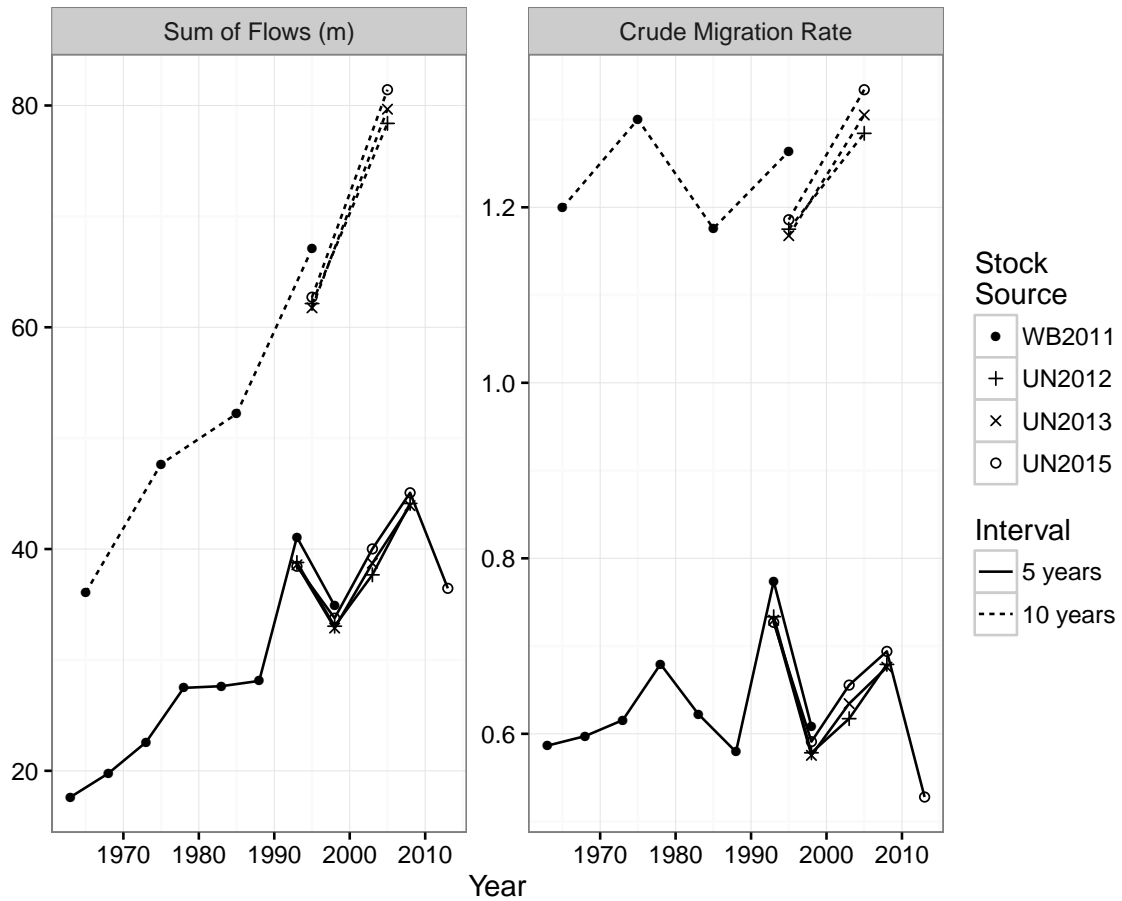
Figure 2: Total estimated country to country bilateral flows and crude global migration rate varying by stock data source used and interval of flow estimate. Only estimates based on WPP2015 demographic data are shown. On the horizontal axis, points are plotted at the midpoint of their corresponding interval.
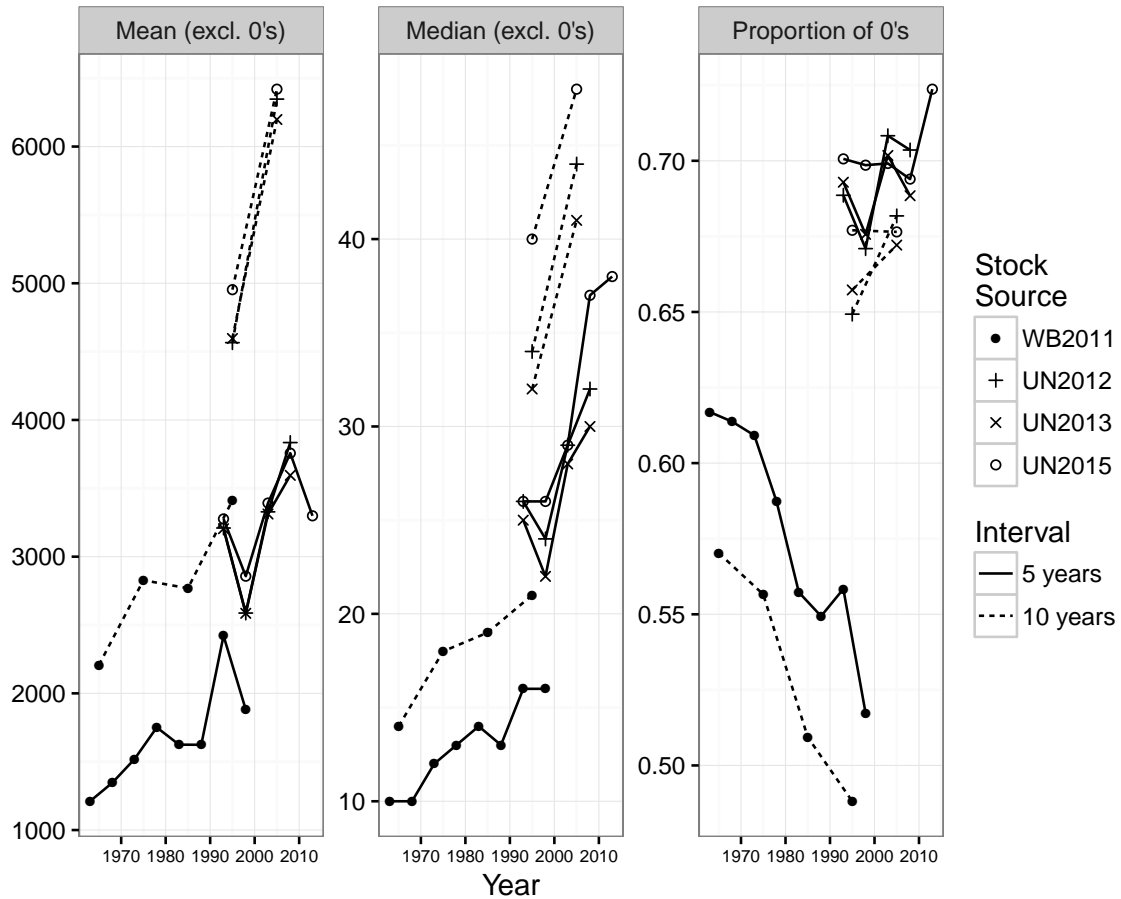
Figure 3: Further summary statistics for estimated country to country bilateral flows varying by stock data source used and interval of flow estimate. Only estimates based on WPP2015 demographic data are shown. On the horizontal axis, points are plotted at the mid-point of their corresponding interval.

recent period, 2010-15, large labour related flows from Latin America to North America, parts of Asia to the Gulf States and moves 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 2 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 WPP2015 populations in each origin 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 five-year intervals also remain fairly constant at around 0.65 percent, except during 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 1990's. This difference occurs for a couple of reasons. First, the number of non-zero estimated flows are not constant across time, as illustrated in the plot on the right hand side of Figure 2. 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 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 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 follow a similar decline in Figure 2. 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 is slightly higher, around 75 percent in 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[4]. Estimates based on the World Bank stocks included an additional country; Taiwan, whilst 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 2010-15 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 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). 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, 5.45 million; Europe, 4.77 million). Many movements 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 1970-80 increased globally from the previous period, as illustrated in Figure 2. Although this increase in size is difficult to view from comparing circular migration plots in Figure 4 (a) and (b), changes in the share of global moves between selected regions can be easily detected. Most noticeable is a large increase in the share of global migrants moving within Southern Asia. During 1970-80, 4.37 million movements

---

[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.

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 regional migration flows over time are more easily viewed in Figure 5, which provides plots of estimated immigration and emigration totals by UN Population Division demographic regions[5]. 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 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 4 (a) and (b). Estimated flows into and within Europe during 1970-80 decreased from the decade before. Sizeable moves into West Asia from countries such as Egypt (0.39 million) and India (0.16 million) to Saudi Arabia began to develop. Moves within Africa also increased, including large flows out of Ethiopia (0.95 million to Somalia) and Burkina Faso (0.41 million to Ivory Coast).

Estimated flows during 1980-90 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 (3.09 million). The number of movements within Eastern Europe also increased, including 1.03 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 plot (d), estimates based on the World Bank stock data are shown. The level of immigration in North America (also shown in Figure 5) is estimated to increase from a wider variety of origins, 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 UN2015 stock data is shown in Figure 4 (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 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 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 ten-year time period between 2000 and 2010 is shown in Figure 4 (f). Based upon UN2015 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

---

[5]Except for Polynesia, Melanesia and Micronesia which are aggregated to a Pacific Island region.
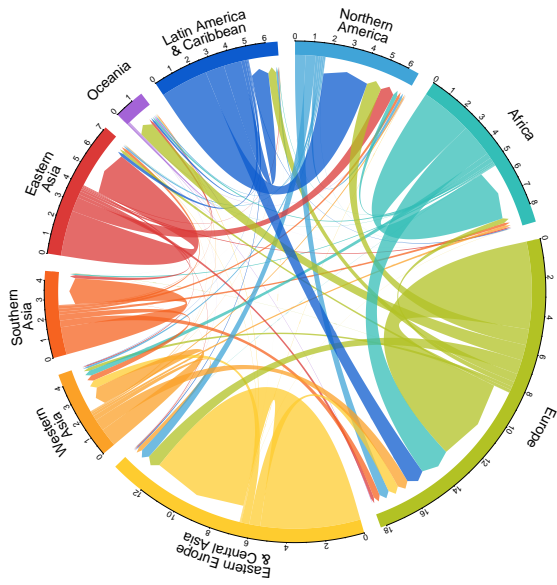
(a) 1960-70 based on WB2011 stock data
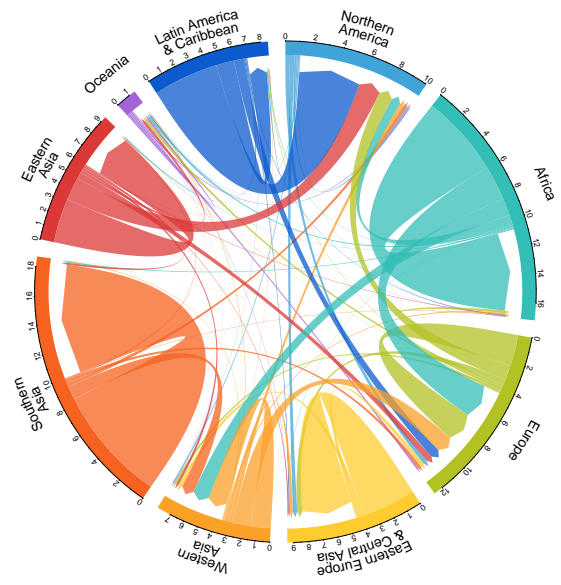
(b) Total 1970-80 based on WB2011 stock data
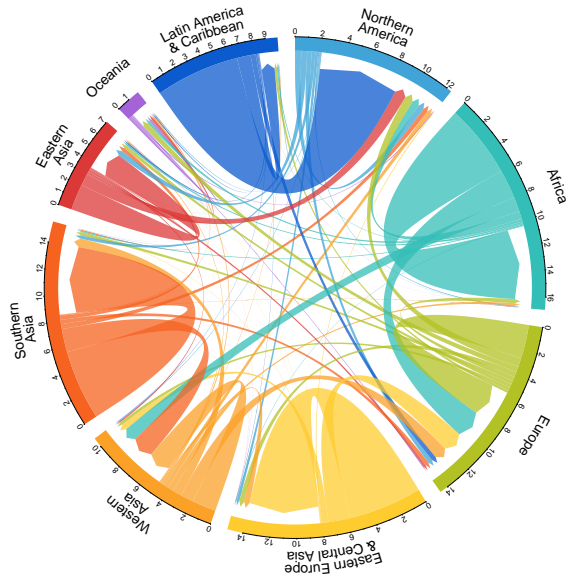
(c) 1980-90 based on WB2011 stock data
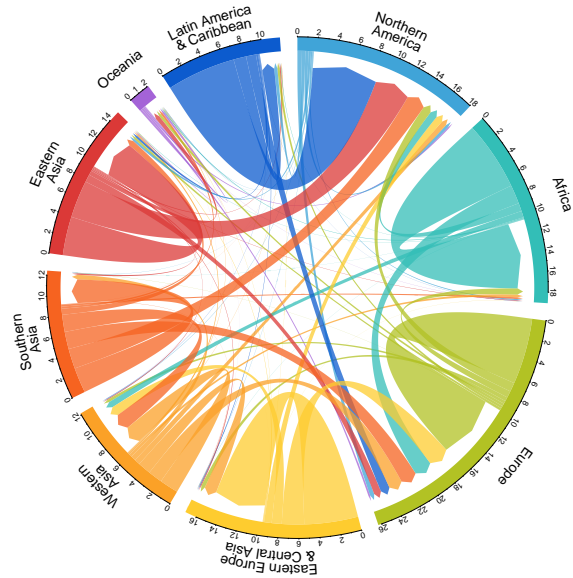
(d) 1990-2000 based on WB2011 stock data
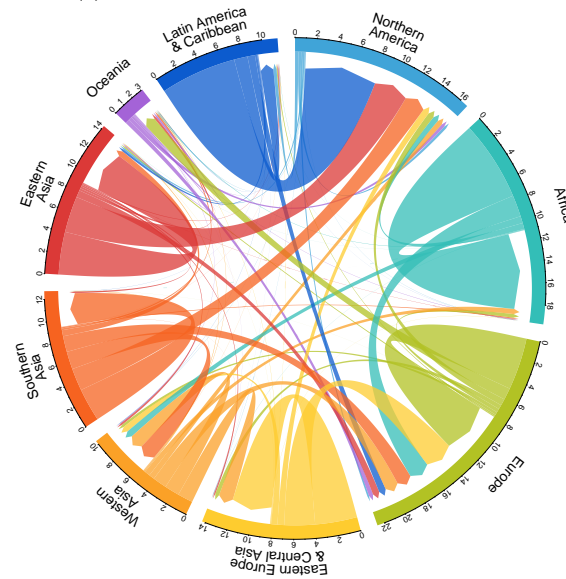
(e) 1990-2000 based on UN2015 stock data
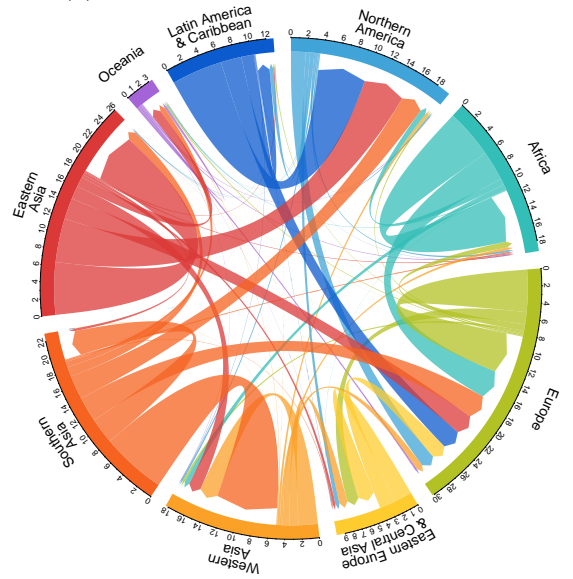
(f) 2000-10 based on UN2015 stock data

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

15
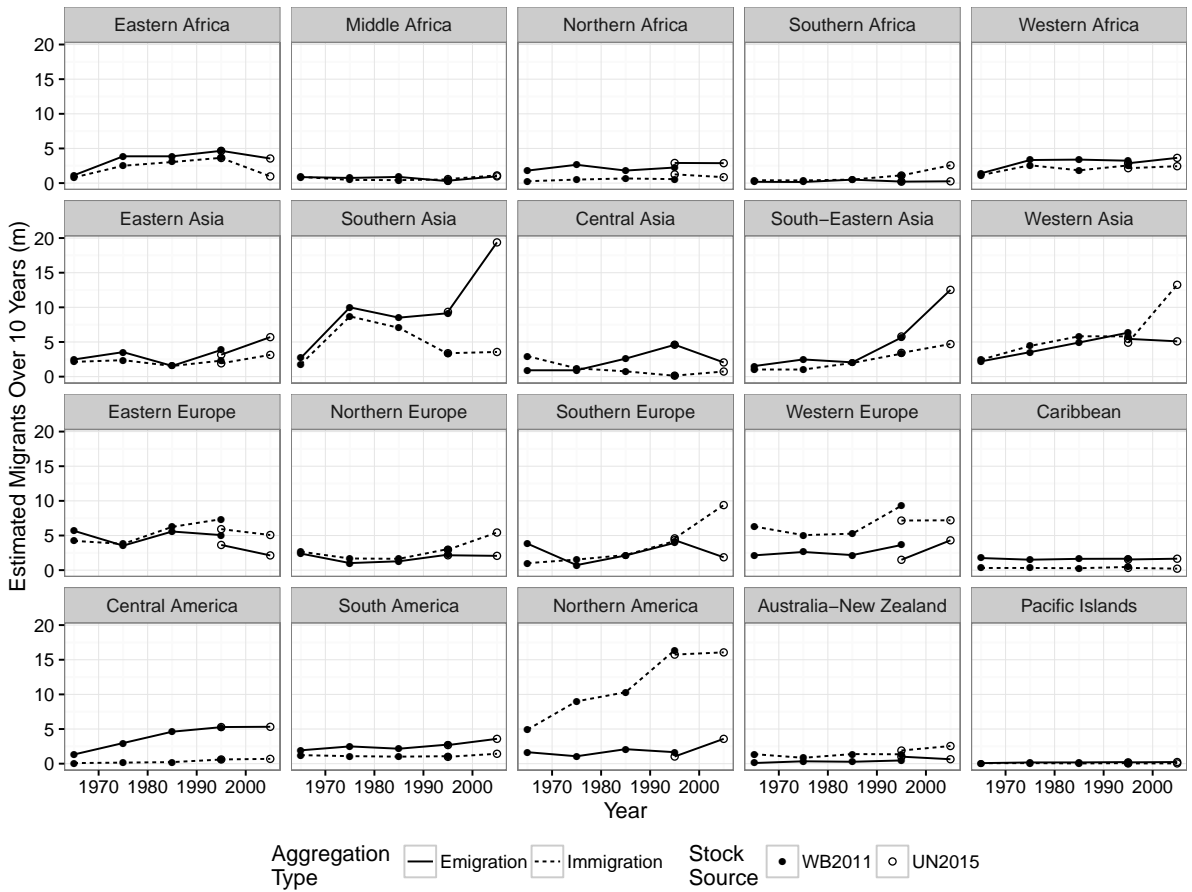
Figure 5: Total estimated immigration and emigration flows over a 10 year periods. Estimates based on aggregations of country to country bilateral flows, both World Bank and UN stock data (WB or UN) and WPP2015 demographic data.

(a) 2005-2010 based on UN2015 stock data   (b) 2010-2015 based on UN2015 stock data
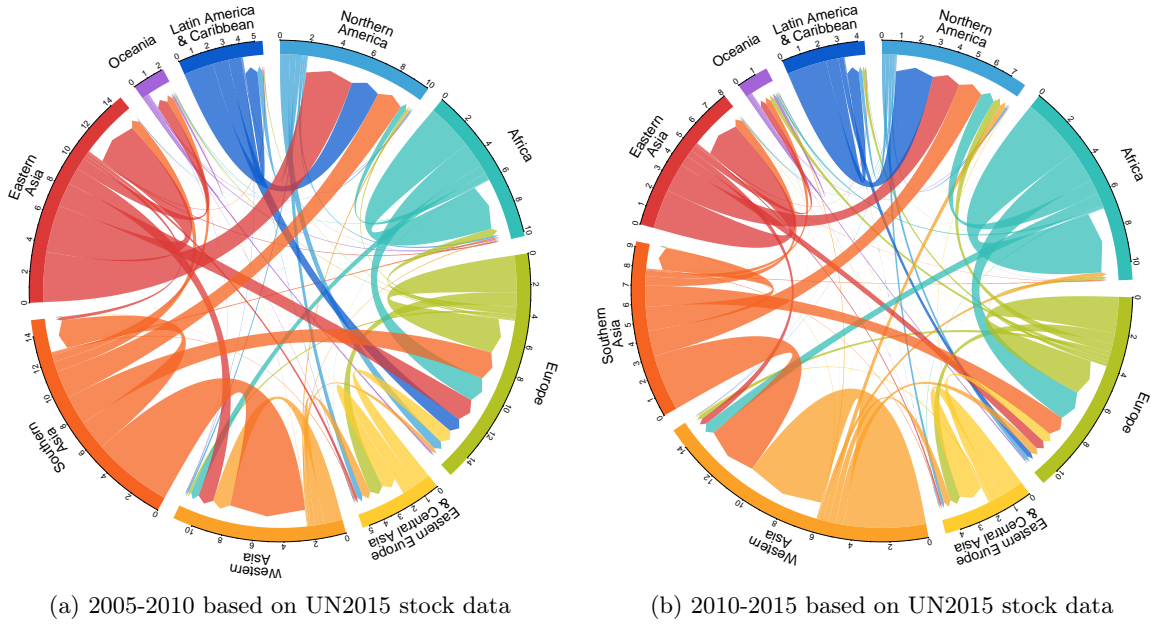
Figure 6: Estimated 5 year migrant flows in recent periods aggregated by selected regions. Both based on WPP2015 demographic data.

European countries, as shown in Figure 5, 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 movements within South-Eastern Asia, including an estimated 1.42 million people moving from Myanmar to Thailand over the ten-year period.

In Figure 6, circular migration flow plots for the two most recent 5-year periods are given. As shown in Figure 2, estimated migration flows dropped considerably from 45.08 million during 2005-10 to 36.46 million during 2010-15. The origin-destination patterns also underwent some considerable change. For example, large flows within Western Asia appear in 2010-15 based on movements out 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 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 moves from Eastern Asia fell from 3.40 to 1.59 million. Moves from South Asia to Western Asia also decreased, where for example the estimated number of migrants from India to the United Arab Emirates fell from 1.38 million during 2005-10 to 0.45 million during 2010-15.

4.3 Flows by Gender

Female and male total flows and crude migration rates are shown in Figure 7. The patterns of both statistics follow similar paths as those in Figure 2 for the estimates based on the differences between the total stock tables. There is sum of male flows are slightly larger in most time periods. During 2000-10, male flows increased faster than the females, 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
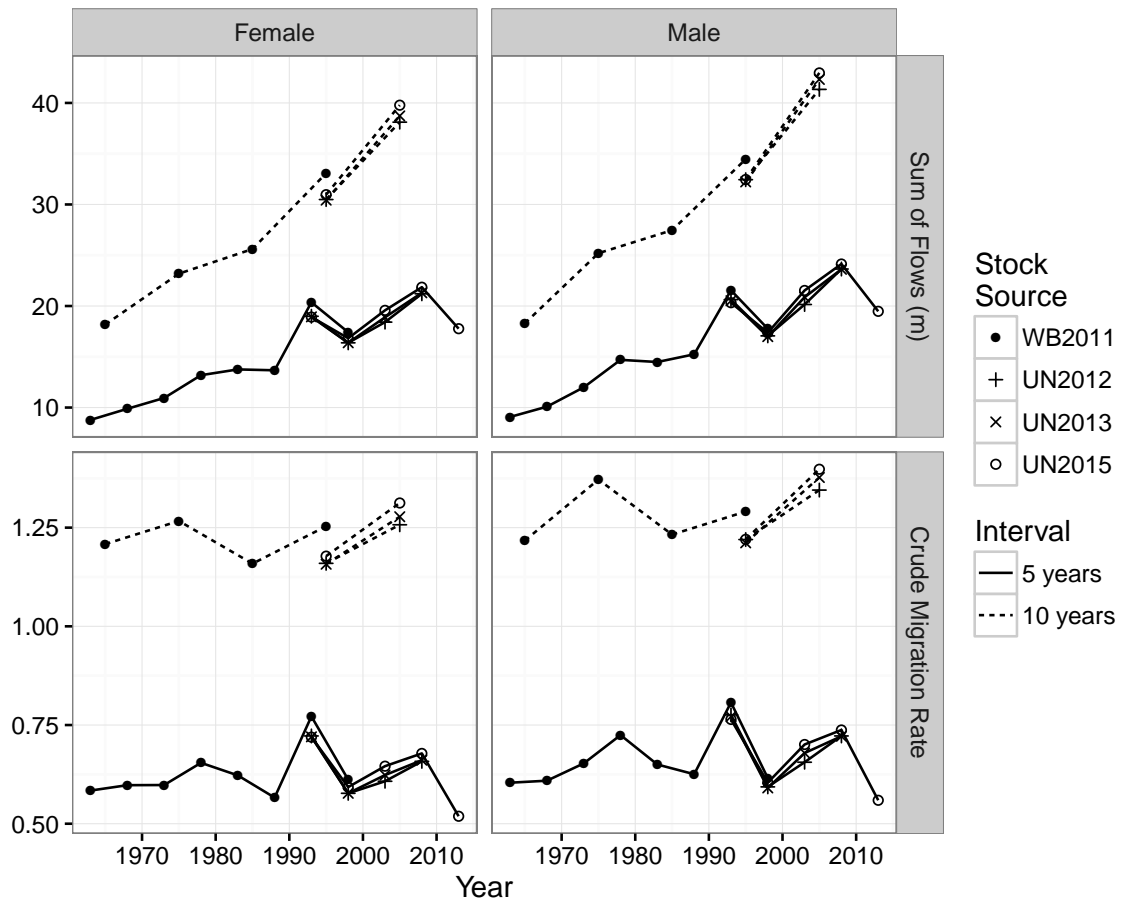
Figure 7: Total global migration flows and crude rate for estimated country to country bilateral flows by gender. Based on WPP2015 demographic data.

apparent for the total flows is also evident in the gender specific flows.

Selected circular migration flow plots for both estimated males (left) and females (right) are shown for two time periods in Figure 8. Estimates are based on gender-specific stock and demographic data. In each of the time periods both male and female 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 9.

Male dominated flows (where the immigration or emigration lines are above 0.5) occur almost entirely throughout the period for moves in and out of Northern, Southern and Western Africa, as well as for moves into Western Asia. Except for moves into South-Eastern Asia, female dominated flows during the entire period are less common. In most regions the share of estimated male to female migrant flows do not to follow any clear and consistent patterns. Two particular data points stand out when considering Figure 9 overall. First, in Southern Africa, estimated flows during the 1960-70 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 2000-10. As shown in the circular migration plots of Figure 8 (c), these flows are predominately
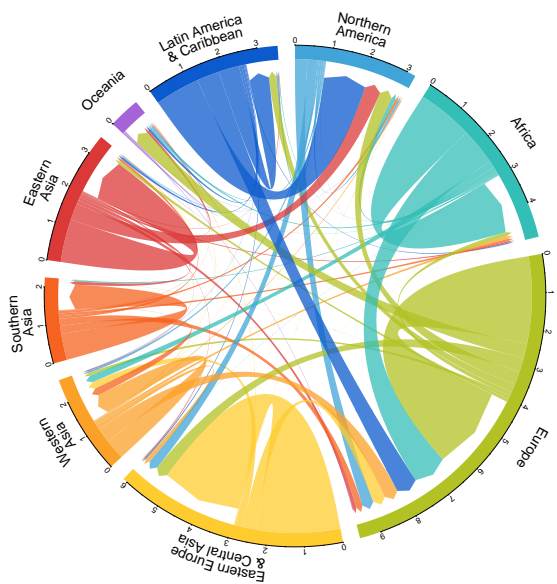
(a) Males, 1960-70. Based on WB2011 stock data
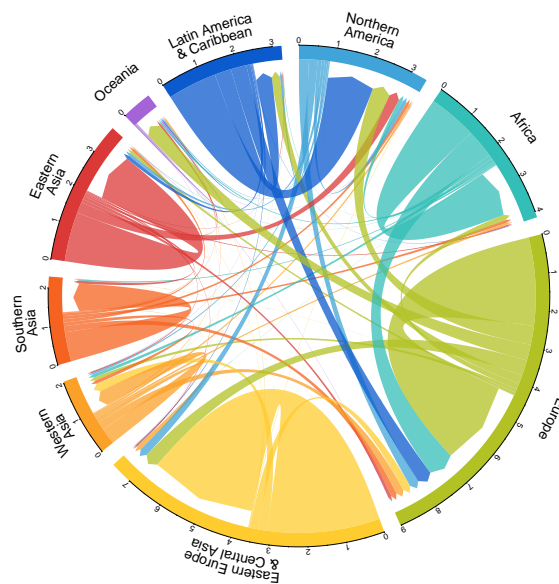
(b) Females, 1960-70. Based on WB stock data
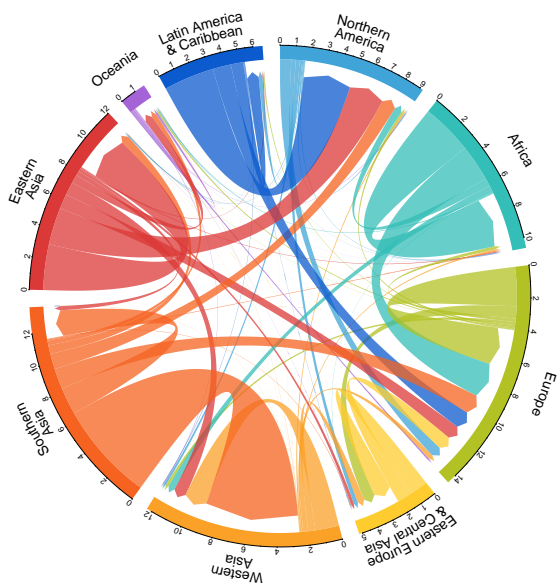
(c) Males, 2000-10. Based on UN2015 stock data
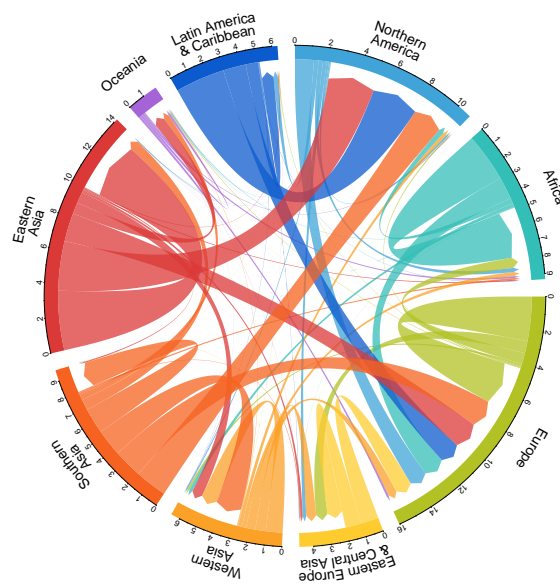
(d) Females, 2000-10. Based on UN2015 stock data

Figure 8: Estimated 10 year migrant flows by gender for selected time periods. All based on WPP2015 demographic data.
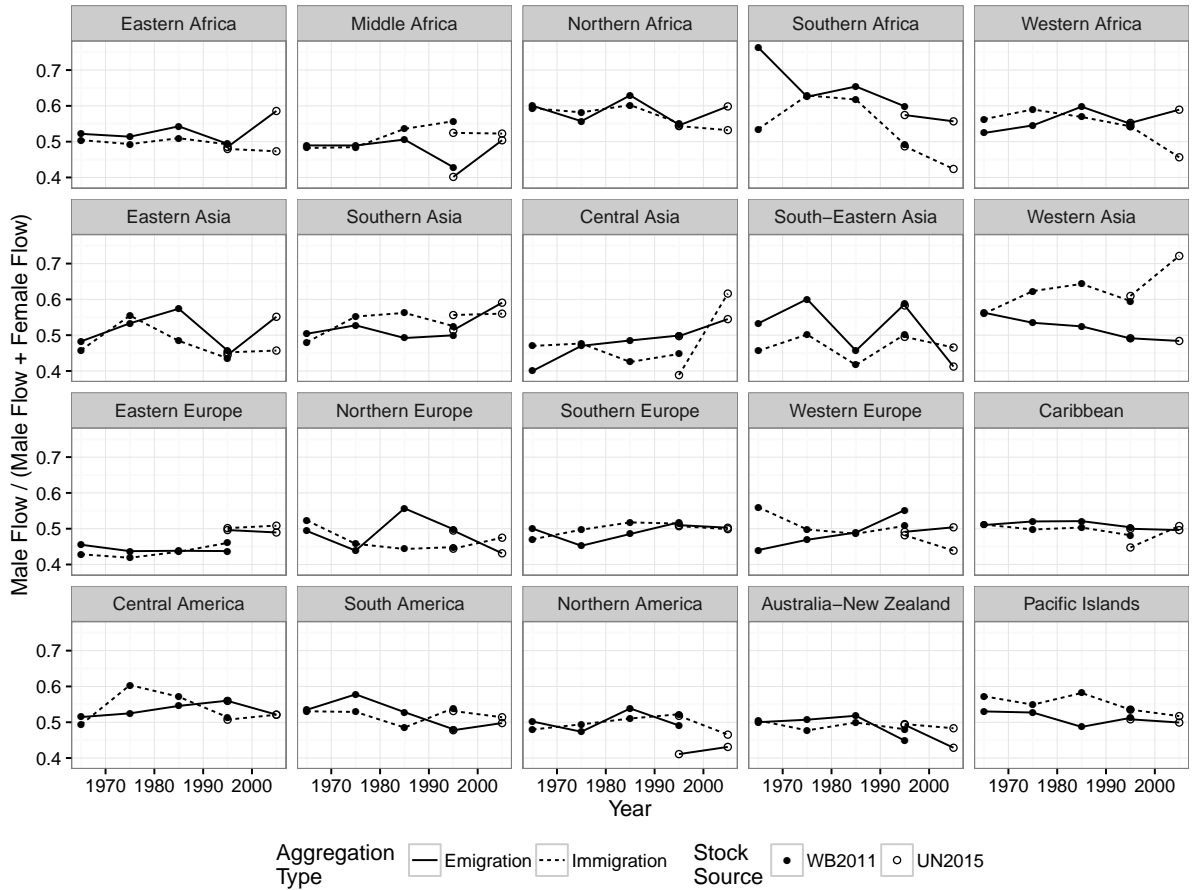
Figure 9: Male percentage of total estimated immigration and emigration flows over 10 year periods. Estimates based on aggregations of country to country bilateral flows, both World Bank and UN stock data and WPP2015 demographic data.

from Southern Asia, South-Eastern Asia, other countries 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 stock data can potentiality be sensitive to the input data used in the methodology. 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 (2013) a handful of unexpected flow estimates result from peculiarities in the input stock data. The same unexpected flows are 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.
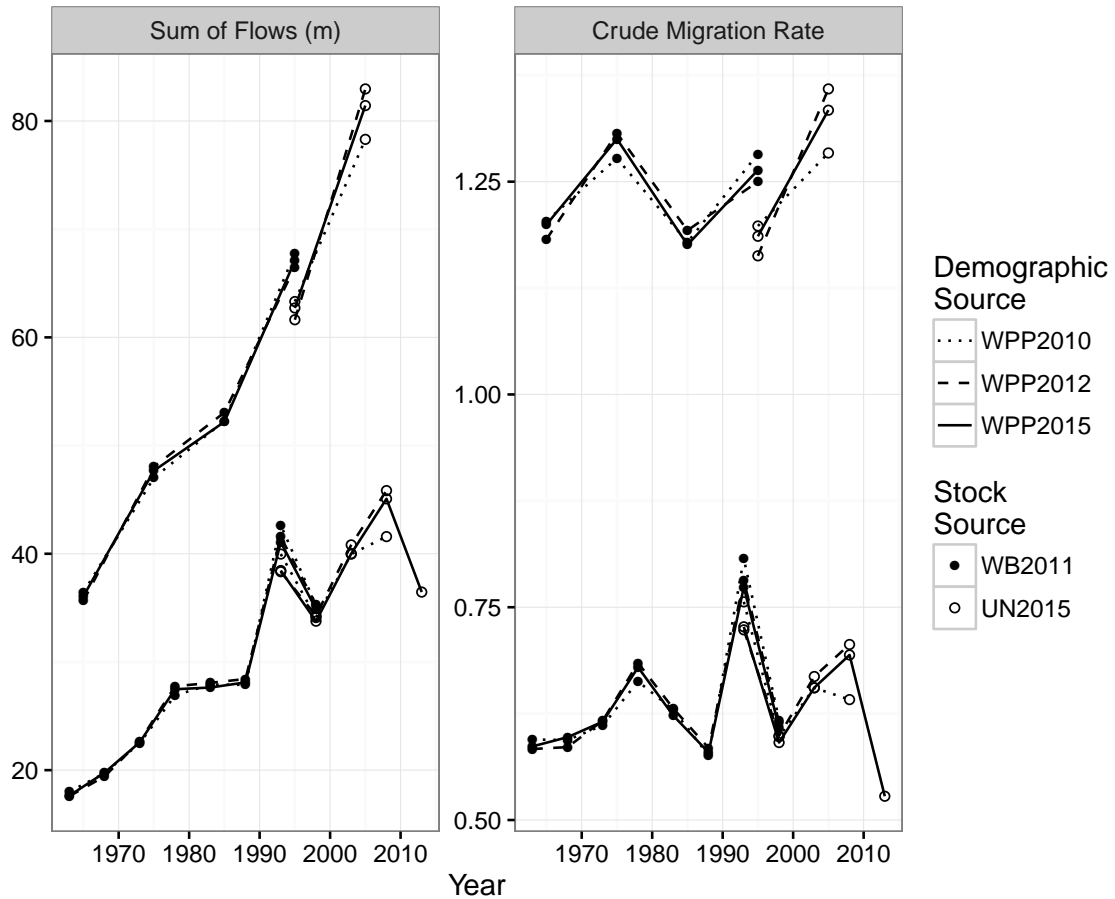
Figure 10: Total estimated migration flows and crude migration rate over five (lower lines) and ten-year (upper lines) periods from alternative demographic data sources. Estimates based on aggregations of country to country bilateral flows and both World Bank and UN stock data (WB or UN).

As shown in Figure 2 and 3 there are some differences between the summary statistics estimated from previous version of the UN stock data. However, as shown in the Appendix, the bilateral patterns do not alter significantly as the input UN stock data varies for the same period. 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 studied. This is followed by a comparison between the flows presented in the previous section with those adjusted for changes in political geography.

5.1 Demographic Data Source

The UN Population Division updates demographic estimates for all countries every two to three 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 10.

The total flows from the WPP2015 data are given by the solid line and match those in Figure 2. The updated demographic data have a noticeable effect on the total estimated flows
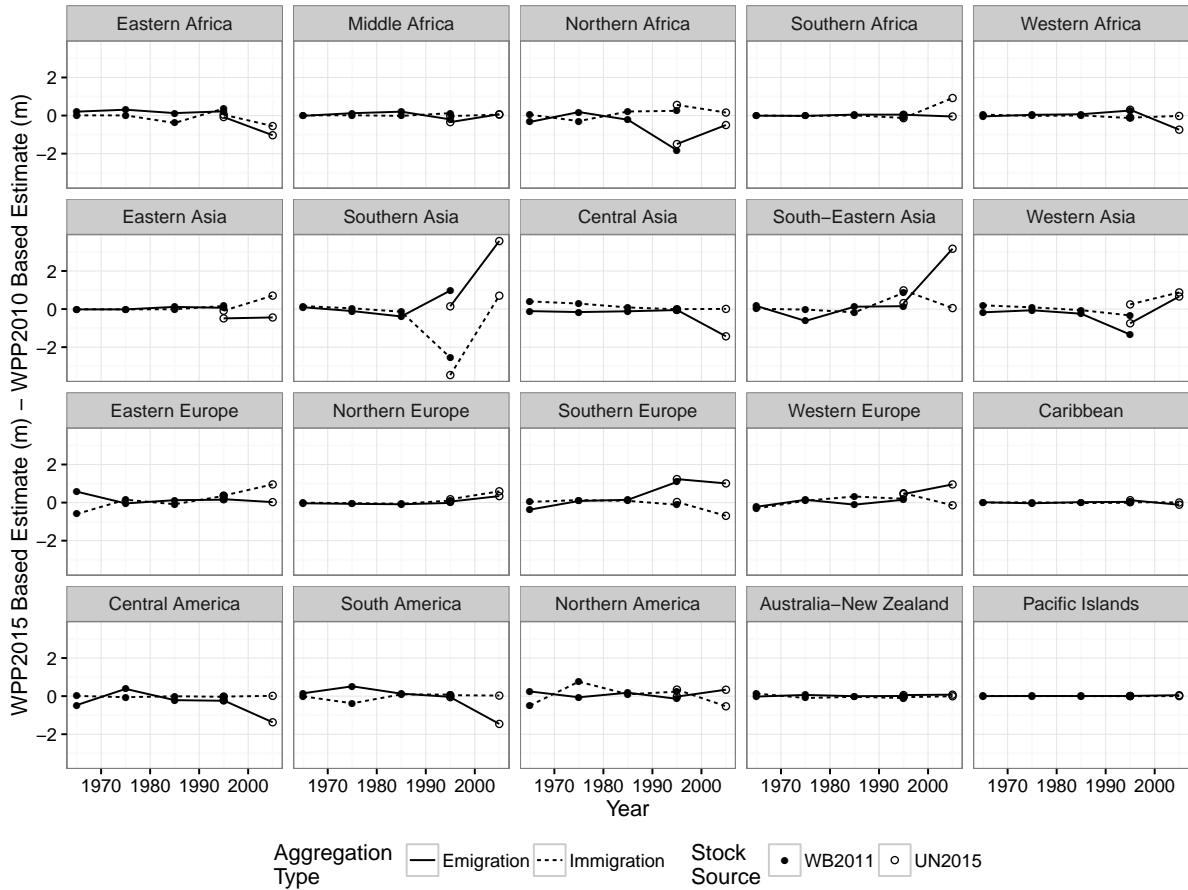
Figure 11: Differences in estimated migration flows over a 10 year period from alternative demographic data sources. Estimates based on aggregations of country to country bilateral flows and both World Bank and UN stock data (WB or UN).

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 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 2000-10 estimate based on UN2015 stock data and WPP2015 involves 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 11 plots differences of both immigration and emigration by region.

Some of the largest difference between the estimates from alternative demographic data sources appear in the estimated migrants flows from Southern Asia during 2000-10 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, 2000-10, 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 11 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 are also sensitive to alternative demographic. One of the most prominent examples are the flows into and out of Russia. In the WPP2010 data, Russia had a positive net migration of 2.7 million over the ten-year period whilst 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 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 (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 2005-10 the demographic data 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 countries 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 to the UK 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 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 and hence the results are not presented in the remainder of this paper.

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 used in the UN stock data only provides information on new countries in 2010. As a result, separate estimates into, out and between both Serbia and Montenegro and Sudan and South Sudan can not be obtained as there were no foreign born stock data for these new countries in previous decades[6][7].

In order to analyse 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 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, 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 12 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 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 are less smooth. Instead, global migrant flow numbers remain somewhat level during the late 1970's up until the late 1980's. 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 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 paper is difficult. Further, as discussed in Abel and Sander (2014) the estimated net migration for each country match those from the WPP as data on population, births and deaths are accounted for within the methodology. Nev-

---

[6]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

[7]Demographic data, not provided for the unified areas were obtained by combining data from the separate countries
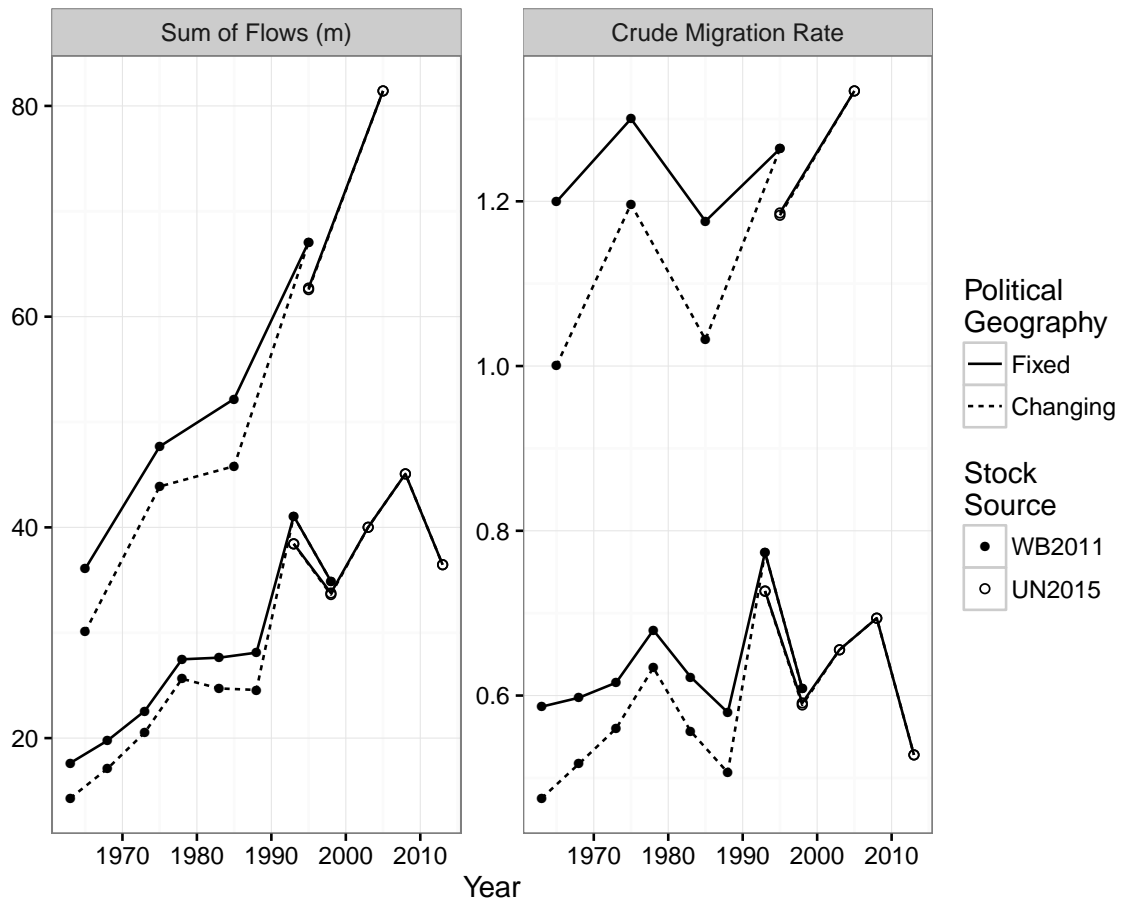
Figure 12: Total estimated international migrant flows by length of migration period in each decade. Estimates based on aggregations of country to country bilateral flows, both World Bank and UN stock data (WB or UN) and WPP2015 demographic data. Country definitions are either fixed according the union of stock and demographic data or changing to reflect altering political geographies.

ertheless, for a small amount of countries past data on immigration and emigration flows exist. In the top panel of Figure 13 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 is 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 perfect agreement of the proportions in the reported data and estimates calculated in the previous section.

The immigration flow data on the vertical axis is 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 relative long history. However, as noted earlier, there are a number of issues 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.
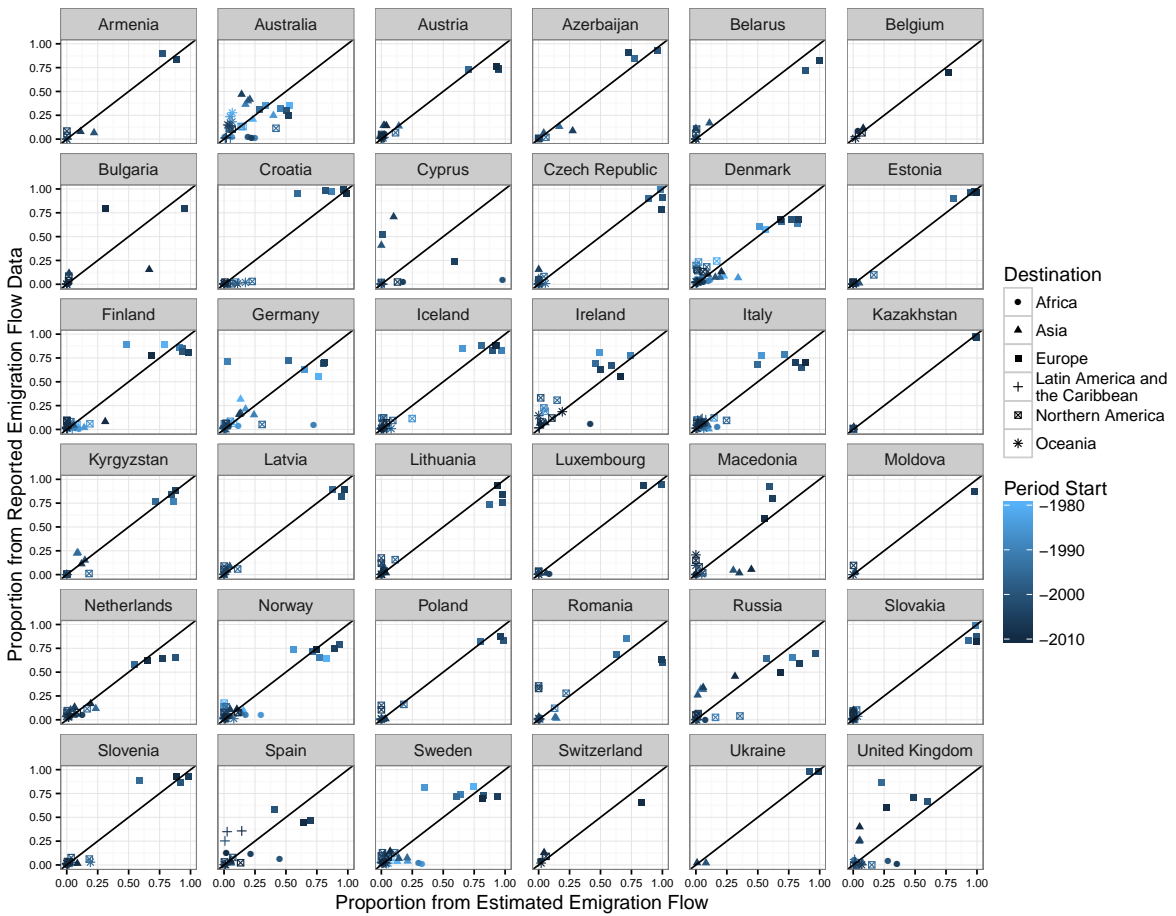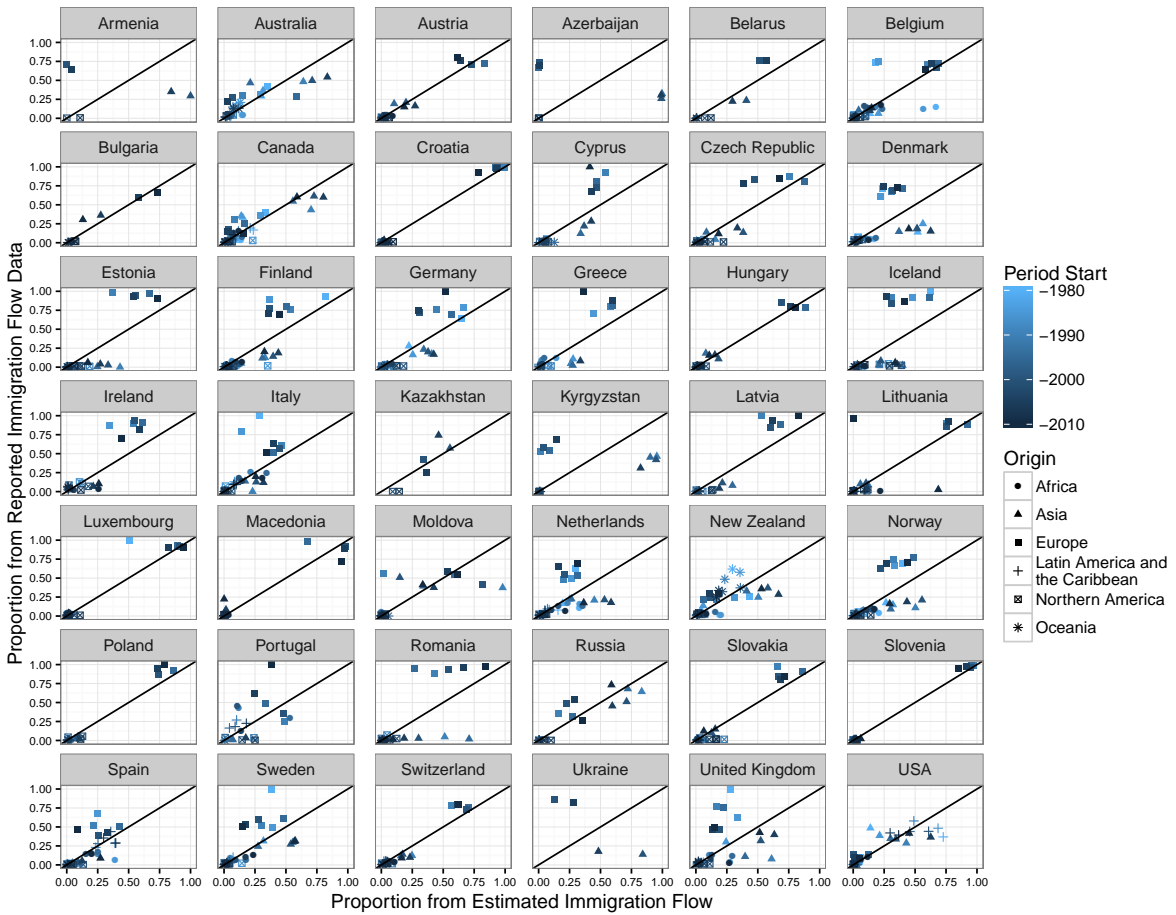
Figure 13: Comparisons of the proportion of estimated flows from and to each continent with reported immigration (top) and emigration (bottom) data.

Further, in some countries there are large counts of immigrant flows with unknown origins but are included in a total flow figure. For the data in Figure 13, 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 13. 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 the lower panel of Figure 13 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 less 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.

7   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 migration flows provide a more robust basis to understand 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 region (the estimates for all countries are available from the author or 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 0.65 and 1.25 percent of the global population 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.

During 2010-15 the estimated global migration flow fell from previous five year periods. At first glance this finding is counter to 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 individuals times. As illustrated in this paper, flow estimates based on changes in the stock data between 2010 and 2015 indicate a decreasing number of moves out of Latin America and Asia towards 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 neighbouring countries.

Overall the trends in global migration flows for each gender followed similar paths, with slightly higher males shares throughout. In recent periods such as 2000-10, the male flows grew relatively larger than female flows during the same period which also increased, but not as dramatically. This imbalance at the global level flows is predominantly to due to large increases of male migrant stocks of South Asians in oil rich Gulf States. It stands in contrast to the common narrative regarding the feminisation of international migration which was based on older migrant stock data, where no global migration flow estimates were previously available. These difference could be further explored in future work, using regression modelling techniques based on the estimates presented in this paper and potentially related explanatory variables. Information on the growth and size of specific movement types such skilled labour, 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 and appendix, the estimated flows in this paper 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 non-zero flows were skewed towards small counts driven by the large numbers of zero bilateral migrant stocks in both the World Bank and UN stock data. There is no empirical measure to help inform the estimation method on both the number of non-migrants in each periods and migrant stock to allow for a higher measures 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 are not be captured, such as away from a persons 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 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 and 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, 2015c). Fifth, in countries and time periods where reported migration flow data exists, comparisons are made with the estimates. There is a broad correspondence between the proportion of moves to and from each continent in the selected countries.

The choice of input data used were found to have an 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 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 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. In the Appendix the role of the assumption in the methodology for the number of stayers and the auxiliary data are further explored.

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, 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 the 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 deficiency 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. 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 (Artuc 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.

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. R Package Version 1.6.*

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.

Artuc, E. et al. (2015). "A Global Assessment of Human Capital Mobility: The Role of Non-OECD Destinations". In: *World Development* 65, pp. 6–26.

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. Doquier, 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.

Docquier, F., B. L. Lowell, and A. Marfouk (2009). "A gendered assessment of highly skilled emigration". In: *Population and Development Review* 35.2, pp. 297–321.

Docquier, F. et al. (2012). "Are Skilled Women More Migratory than Skilled Men?" In: *World Development* 40.2, pp. 251–265.

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

Dumont, J.-C., G. Spielvogel, and S. Widmaier (2010). *International migrants in developed, emerging and developing countries: An extended profile.* Tech. rep. Paris, France: OECD.

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.

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, Switzerland: International Labour Organization.

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.

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.

Parsons, C. R. et al. (2007). "Quantifying international migration: A database of bilateral migrant stocks". In: *World Bank Policy Research Working Paper* 41.4165, pp. 1–41.

Piper, N. (2005). *Gender and Migration*. Tech. rep. September. Policy Analysis and Research Programme of the Global Commission on International Migration, p. 55.

Piper, N. (2013). *New perspectives on gender and migration: Livelihood, rights and entitlements*. Ed. by N. Piper. New York, New York, USA: Routledge.

R Development Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Ratha, D. and W. Shaw (2007). "South-south migration and remittances". In: *World Bank Working Paper* 102, pp. 1–68.

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., G. J. Abel, and P. W. F. Smith (2007). "Combining census and registration data to estimate detailed elderly migration flows in England and Wales". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 170.4, pp. 891–908.

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

Rogers, A. (1990). "Requiem for the Net Migrant". In: *Geographical Analysis* 22.4, pp. 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*. New York, New York, USA: ACM Press, p. 265.

Truong, T.-D. et al. (2014). *Migration, Gender and Social Justice*. Ed. by T.-D. Truong et al. Hexagon Series on Human and Environmental Security and Peace. Berlin, Heidelberg: Springer Berlin Heidelberg.

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

United Nations Population Division (2012). *Trends in International Migrant Stock: Migrants by Destination and Origin (United Nations database, POP/DB/MIG/Stock/Rev.2012).* 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.

United Nations Population Division (2014). "World Urbanization Prospects: The 2014 Revision, Highlights". New York, USA.

United Nations Population Division (2015a). *International Migration Flows to and from Selected Countries: The 2015 Revision.* New York, USA.

United Nations Population Division (2015b). *Trends in International Migrant Stock: The 2015 Revision.* New York, USA.

United Nations Population Division (2015c). *World Population Prospects: The 2015 Revision. DVD Edition.* New York, USA.

United Nations Population Division (2016). "International Migration Report 2015 Highlights". 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.

World Bank (2016). "Migration and Remittances Factbook 2016". Washington D.C., USA.

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.

Zlotnik, H. (2003). *The global dimensions of female migration.* Washington, DC.

Appendix

A   Representing Bilateral Migrant Stock Data in Flow Tables

Bilateral migrant stock data can be represented using 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 for the stock data used in this paper 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 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 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 people with dual citizenship or nationalities can be counted twice.

		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.

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 for all individuals. 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 from Raymer, Abel, and Smith (2007) to account for the known margins and diagonal elements over the multiple tables. The parameters can be obtained using an iterative

Birthplace=A	Destination						Birthplace=B	Destination				
		A	B	C	D	Sum			A	B	C	D	Sum
Origin	A	*70*	20	0	10	**100**	Origin	A	*20*	0	0	0	**20**
	B	0	*10*	0	0	**10**		B	0	*55*	0	0	**55**
	C	0	0	*10*	0	**10**		C	5	5	*10*	5	**25**
	D	0	0	0	*0*	**0**		D	0	0	0	*10*	**10**
	Sum	**70**	**30**	**10**	**10**	**120**		Sum	**25**	**60**	**10**	**15**	**110**

Birthplace=C	Destination						Birthplace=D	Destination				
		A	B	C	D	Sum			A	B	C	D	Sum
Origin	A	*10*	0	0	0	**10**	Origin	A	*20*	0	0	0	**20**
	B	0	*40*	0	0	**40**		B	0	*25*	0	0	**25**
	C	0	0	*140*	0	**140**		C	10	10	*0*	0	**20**
	D	0	15	0	*50*	**65**		D	10	10	0	*180*	**200**
	Sum	**10**	**55**	**140**	**50**	**255**		Sum	**40**	**45**	**0**	**180**	**265**

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

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 a log-linear model which includes an offset term. This term 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. For the flow estimates derived in this paper, 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 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 shown in the example data of Table 4, where new data are provided in time period

	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: Example of place of birth data with births and deaths over the interval

$t+1$

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[8]. This step is illustrated on 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 the count on the number of 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 residual of the total population from UN demographic data and the foreign born population from the World Bank migrant stock data

---

[8]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 fewer deaths to 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						Place of Residence ($t+1$)			
		A	B	C	D			A	B	C	D
Birthplace	A	20.0	0.8	2.1	0.0	*Birthplace*	A	10.0	8.9	0.0	0.0
	B	4.0	4.2	5.1	0.4		B	0.0	16.1	0.0	0.0
	C	2.0	3.1	28.7	0.4		C	0.0	14.3	25.0	0.0
	D	4.0	1.9	4.1	7.3		D	0.0	10.7	0.0	60.0
	Sum	**30**	**10**	**40**	**10**		Sum	**10**	**50**	**25**	**60**

Step 2: Estimated Altered Stocks

		Place of Residence (t)							Place of Residence ($t+1$)				
		A	B	C	D	Sum			A	B	C	D	Sum
Birthplace	A	80.0	9.2	7.9	0.0	97.2	*Birthplace*	A	65.0	16.1	5.0	0.0	85.7
	B	16.0	50.8	19.9	9.6	96.3		B	20.0	28.9	30.0	15.0	90.2
	C	8.0	36.9	111.3	62.6	218.8		C	30.0	25.7	125.0	60.0	240.2
	D	16.0	23.1	15.9	192.7	247.7		D	30.0	19.3	20.0	170.0	238.9
	Sum	120.0	120.0	155.0	265.0	660.0		Sum	145.0	90.0	180.0	245.0	655.0

Step 3: Re-estimated Altered Stocks

		Place of Residence (t)							Place of Residence ($t+1$)				
		A	B	C	D	Sum			A	B	C	D	Sum
Birthplace	A	76.7	8.2	6.8	0.0	91.6	*Birthplace*	A	68.6	17.4	5.6	0.0	91.6
	B	17.0	49.8	18.8	9.5	95.1		B	19.5	29.0	31.4	15.2	95.1
	C	9.2	39.4	114.4	66.7	229.8		C	27.3	24.1	121.9	56.5	229.8
	D	17.0	22.6	15.0	188.8	243.5		D	29.6	19.5	21.1	173.3	243.5
	Sum	120.0	120.0	155.0	265.0	660.0		Sum	145.0	90.0	180.0	245.0	660.0

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

is estimated to be 0.54 million in 1970. The number of births over the ten-year period prior to 1970 is given in WPP2010 as 0.59 million. If the number of births are subtracted from the native born population, a negative value is obtained for the bilateral stock, adjusted 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.

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 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 by the numbers in bold type face in the top panel of Table 6. The diagonal cell values for the non-movers in italics are obtained using the procedure illustrated in Table 3, allowing the iterative proportional fitting algorithm of Abel (2013) to estimate the unknown flows in the non-diagonal cells. 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 flows shown in Table 1 and Table 3 as they are formed from a different set of migrant stock data in $t + 1$.

C   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 do not move. The population of stayers are represented on the diagonal elements of birthplace specific migration flow tables such as Table 3. 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 top left array in Table 3) is 50. Any lower, and the estimated flows, conditional on the margins would need to become negative.

The other extreme assumption is to minimise the number of stayers can be made. In Table 7 all stayers in the diagonal cells are set to their minimum possible values (shown in italic font)

Estimates of Origin Destination Place of Birth Flow Tables:

Birthplace=A

			Destination			
		A	B	C	D	Sum
	A	*68.6*	8.1	0.0	0.0	**76.7**
Origin	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
	A	*17.0*	0.0	0.0	0.0	**17.0**
Origin	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
	A	*9.2*	0.0	0.0	0.0	**9.2**
Origin	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
	A	*17.0*	0.0	0.0	0.0	**17.0**
Origin	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
	A		8.1	0.0	0.0	15.7
Origin	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, controlled for natural change, arranged as birthplace specific origin-destination migrant flow tables (top) with resulting flow estimates (bottom).
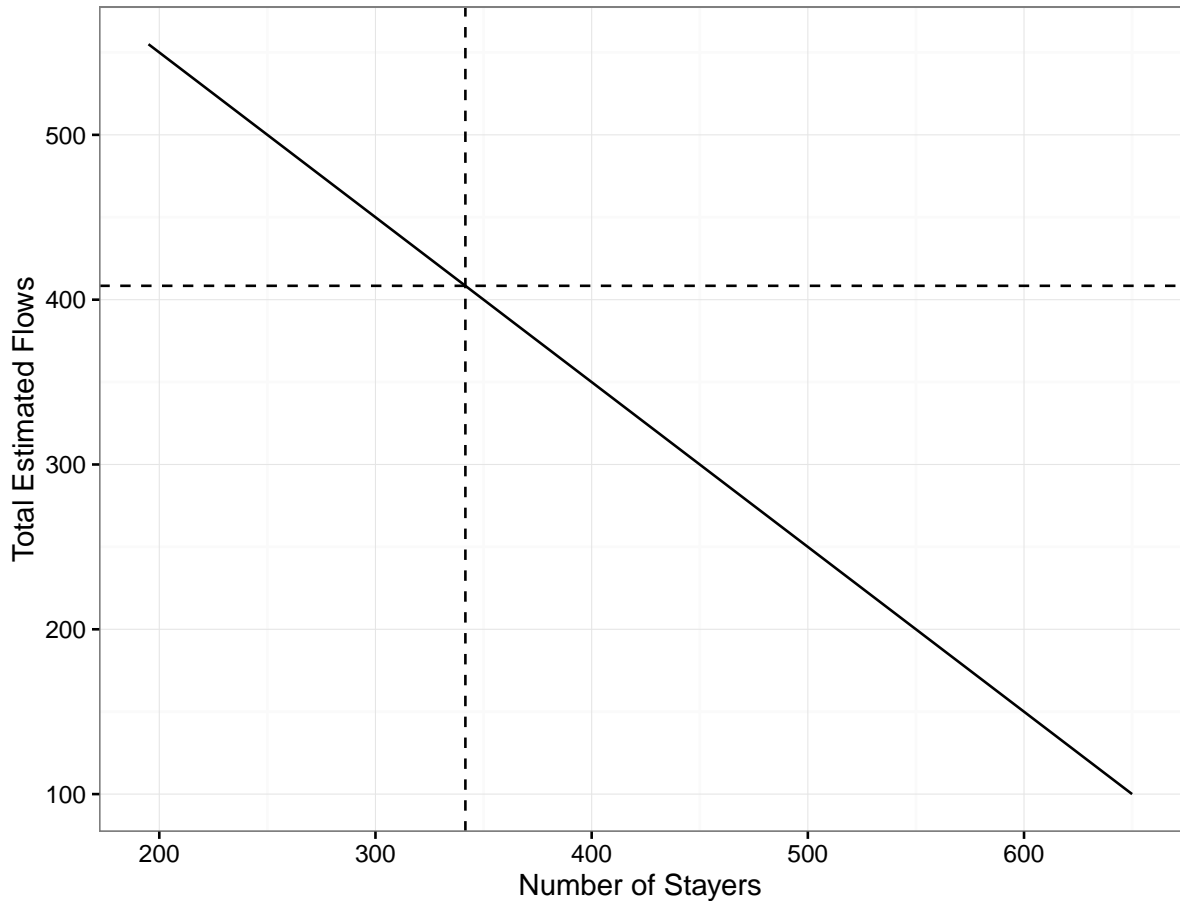
Figure 14: Total estimated flows from stock data in Table 2 under various assumptions for the number of stayers.

conditional on the marginal stock information (shown in bold font) which is the same as in Tables 2 and 3. 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 flows (555) from the minimum diagonal assumption is far greater than the total flows from the maximum diagonal assumption (100) shown in 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 14 shows this relationship. On the horizontal axis is the total number of stayers, ranging between the minimum of 195 and the maximum of 650 in (shown in Tables 7 and 3 respectively). On the vertical axis is the estimated total flow given the corresponding number of assumed stayers. Unsurprisingly the total estimated flows falls as the number of stayers increases.

Whilst 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 14 represents the total estimated flows from the independent model (where the cost to 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

Origin-Destination-Birthplace Flow Tables:

Birthplace=A

		Destination				
		A	B	C	D	Sum
Origin	A	*50*	30	10	10	**100**
	B	10	*0*	0	0	**10**
	C	10	0	*0*	0	**10**
	D	0	0	0	*0*	**0**
	Sum	**70**	**30**	**10**	**10**	**120**

Birthplace=B

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

Birthplace=C

		Destination				
		A	B	C	D	Sum
Origin	A	*0*	0	10	0	**10**
	B	0	*0*	40	0	**40**
	C	10	55	*25*	50	**140**
	D	0	0	65	*0*	**65**
	Sum	**10**	**55**	**140**	**50**	**255**

Birthplace=D

		Destination				
		A	B	C	D	Sum
Origin	A	*0*	0	0	20	**20**
	B	0	*0*	0	25	**25**
	C	0	0	*0*	20	**20**
	D	40	45	0	*115*	**200**
	Sum	**40**	**45**	**0**	**180**	**265**

Origin-Destination Flow Table:

		Destination				
		A	B	C	D	Sum
Origin	A		50	20	30	100
	B	35		50	40	125
	C	20	80		70	170
	D	40	55	65		160
	Sum	95	185	135	140	555

Table 7: Top: Bilateral stock data of Table 2 arranged as birthplace specific origin-destination migrant flow tables with minimal stayers assumption. Bottom: Corresponding origin-destination flow table.

the diagonal must be on the right hand side of this line, when the number of stayers exceeds 346.

In reality, when working with real data, there are a couple of logical arguments to assume the maximum the number of stayers. First, any other level of stayers would require some empirical information in each time period and each migrant stock on the probability staying in their same country of residence. Such data is not available for the past periods and is unlikely to be forthcoming in the near future. If such data were available, then it is likely that detailed information on migration flows is also available, and hence we would not need to estimate migration indirectly. Second, international migration is known to be a rare event. If the data for which Figure 14 were real, then the number of stayers would likely be much closer to the maximum value than from the independence model as 1) there are many barriers to international moves and 2) estimated flows in this paper are measures of migrant transitions over relative long periods (five and ten years). During these periods migration becomes an even rarer event, as short term moves for returning migrants during the period are not be considered as a migration transition.

D   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. For the results presented in the paper a distance measure was used to weight larger imputation between countries that are closer. The remaining parameters in the log-linear model ensure imputed values match the known marginals and diagonal elements in the origin-destination-birthplace arrays. In order to gauge the sensitivity in estimation methodology from auxiliary data, a four dimensional distance matrix based on sets of a Pythagorean triple (3,4,5) shown in the left hand panel in Table 8, is used to estimate the a new set of flows based on changes in the stock data shown in Table 2.

Distance Matrix:					Origin-Destination Flow Table:					
		Destination					Destination			
	A	B	C	D		A	B	C	D	Sum
A		5	50	500	A		20.0	0.0	10.0	30.0
B	5		45	495	B	0.0		0.0	0.0	0.0
C	50	45		450	C	15.2	14.8		5.0	35.0
D	500	495	450		D	9.8	25.2	0.0		35.0
					Sum	25.0	60.0	0.0	15.0	100.0

Table 8: Distance matrix and corresponding flow estimates based on the sequential stock table given in Table 2

The estimated flows, given in the right hand side of Table 8, can be directly compared to those in Table 1 which were based on a distance matrix where all values are set to one. 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

(a) Distance auxiallary data.

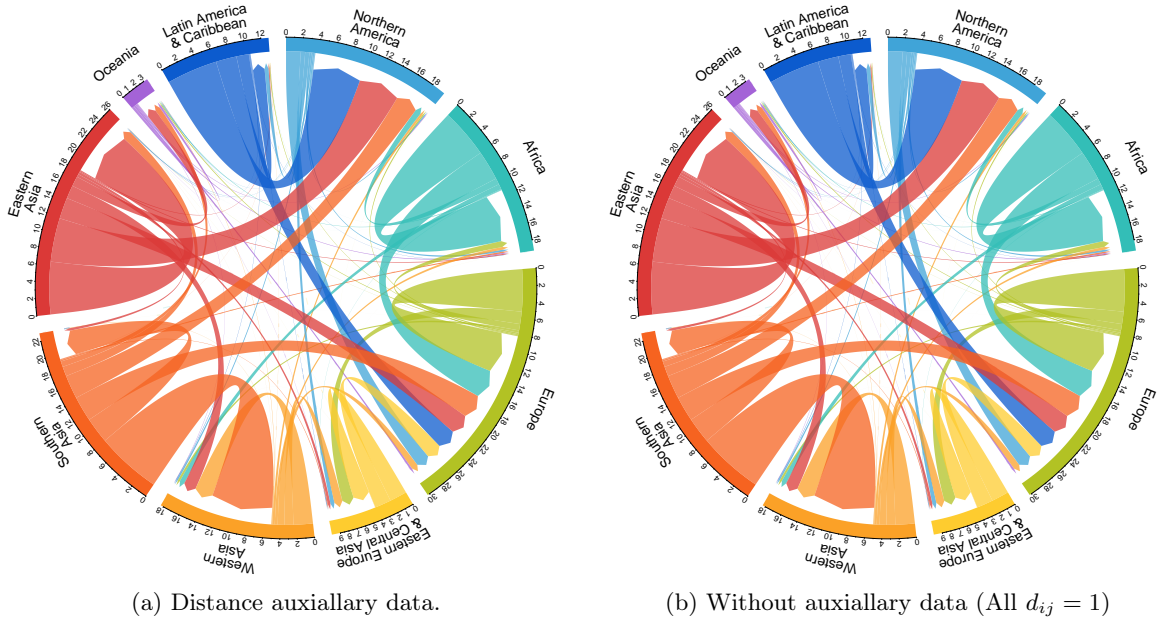(b) Without auxiallary data (All $d_{ij} = 1$)

Figure 15: Estimated 10 year global migrant flows during 2000-10 based on UN2015 stock data, WPP2015 demographic data and varying auxiliary data in the offset terms.

unchanged. Only a small difference occurs in the origin-destination flow estimates, where for example slightly larger number moves are estimated to leave from D and arrive in A (9.8 as opposed to 10) and a corresponding reduction in the moves from D to B. A similar result is also found in the estimated global migration flows. For example, Figure 15 plots side by side circular migration flow plots between 2000-10. In the left hand panel are the estimated flows based on a distance function in the offset term. It is identical to that presented in (f) of Figure 4. In the right hand panel are the estimated flows for the same stock and demographic data but with no offset term (all distances were set to one) in the model. The size of the plots outer sectors, representing the sum of immigration and emigration flows in each region are unchanged. The sizes of all the major estimated flows vary a little, but when plotted no discernible differences are apparent.
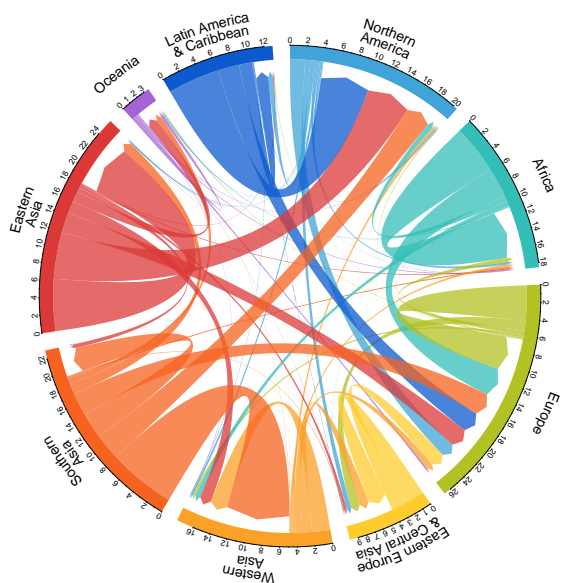
E   Alternative Stock Data

As shown in Figure 16, the bilateral patterns from alternative UN migrant stock data do not vary considerably (b-d). However, as discussed in the main text, distinct differences occur when comparing estimates from the World Bank (a) to those from the UN.
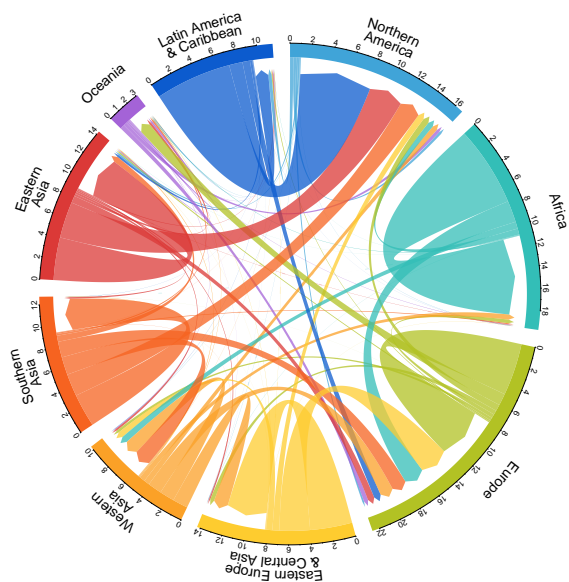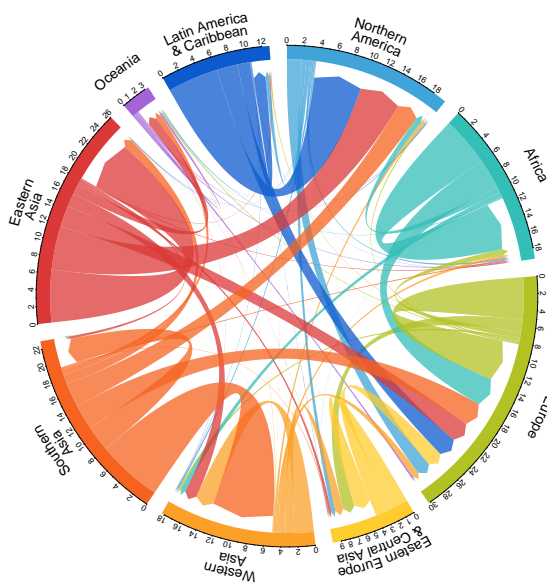
(a) 2000-10. Based on WB2011 stock data
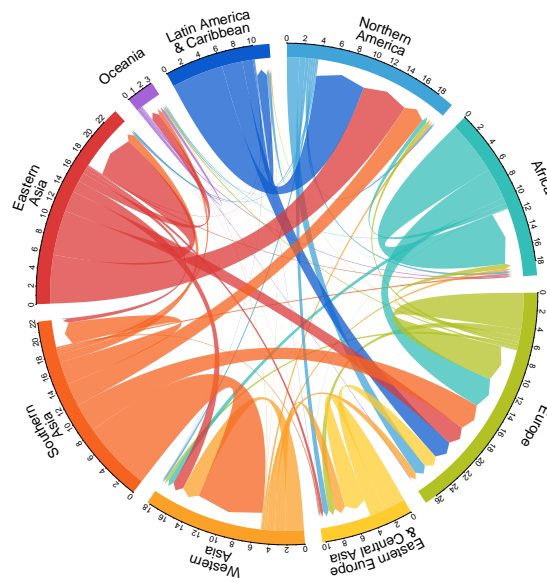
(b) 2000-10. Based on UN2012 stock data

(c) 2000-10. Based on UN2013 stock data

(d) 2000-10. Based on UN2015 stock data

Figure 16: Estimated 10 year migrant flows during 2000-10 based on alternative UN stock data. All based on WPP2015 demographic data.

VIENNA INSTITUTE OF DEMOGRAPHY

Working Papers

Testa, Maria Rita, Valeria Bordone, Beata Osiewalska and Vegard Skirbekk, *The Relation between Mother's Socio-Economic Status and Daughter's Fertility Intentions in Austria, Italy, Bulgaria, and Norway*, VID Working Paper 1/2016.

Hoffmann, Roman and Raya Muttarak, *A Tale of Disaster Experience in Two Countries: Does Education Promote Disaster Preparedness in the Philippines and Thailand*, VID Working Paper 9/2015.

Klotz, Johannes and Richard Gisser, *Mortality Differentials by Religious Denomination in Vienna 1981-2002*, VID Working Paper 8/2015.

Steiber, Nadia and Barbara Haas, *Overworked or Underemployed? Actual and Preferred Household Employment Patterns in the Context of the Economic Crisis*, VID Working Paper 7/2015.

Beaujouan, Eva, Zuzanna Brzozowska and Krystof Zeman, *Childlessness Trends in Twentieth-Century Europe: Limited Link to Growing Educational Attainment*, VID Working Paper 6/2015.

Abel, Guy, *Estimates of Global Bilateral Migration Flows by Gender Between 1960 and 2010*, VID Working Paper 5/2015.

Spijker, Jeroen, *Alternative Indicators of Population Ageing: An Inventory*, VID Working Paper 4/2015.

Kuhn, Michael and Klaus Prettner, *Population Structure and Consumption Growth: Evidence from National Transfer Accounts*, VID Working Paper 3/2015.

Sobotka, Tomáš, *Low Fertility in Austria and the Czech Republic: Gradual Policy Adjustments*, VID Working Paper 2/2015.

Buber-Ennser, Isabella, *Aspects of Gender Mainstreaming of Family and Work in Austria*, VID Working Paper 1/2015.

Buber-Ennser, Isabella and Ralina Panova, *Attitudes towards Parental Employment across Europe, in Australia and in Japan*, VID Working Paper 5/2014.

Kumar, Abhishek, Valeria Bordone and Raya Muttarak, *Influence of Older Generation's Fertility Behaviours on Daughter's Desired Family Size in Bihar, India*, VID Working Paper 4/2014.

The Vienna Institute of Demography Working Paper Series receives only limited review. Views or opinions expressed herein are entirely those of the authors.