Education Differences in the Drivers of Internal Migration

Guy J. Abel[1,2]

1 University of Hong Kong
2 International Institute for Applied Systems Analysis

Abstract

Age, gender and education are the three largest sources of observable heterogeneity in the study of population. Understanding these demographic differentials is essential for projecting future population sizes and compositions (Lutz 2014). Exploiting the harmonised census records in the Integrated Public Use Microdata Series for over 50 countries from the past six decades, we first describe migration intensity by gender and education. Subsequently, we fit a series of weighted multilevel gravity-type spatial interaction models. We utilise a range of variables to study differences in population migration by educational attainment, examining the influence of country-specific contextual factors and regional 'push' and 'pull' factors. We find distinct patterns in migration levels across education groups, where, for example, more educated migrants are associated with longer-distance moves and away from older regions. Within educational level, gender differences are also apparent. For example, males with lower levels of education are attracted to areas with better job opportunities than their female counterparts.

1 Background

Population heterogeneity plays a vital role in shaping demographic behaviours. Age, gender and education are the three largest sources of observable heterogeneity in the study of population (Lutz 2014a). Understanding demographic differentials in population dynamics is relevant for helping explain societal changes and for projecting future population sizes and compositions (Lutz 2014b). Whilst age, gender and education dimensions are well understood in fertility and mortality across both time and space, there are considerably fewer studies by these measures for migration.

Consistencies in the age patterns of migrants are comparatively well understood compared to those in gender and education dimensions. Migration rates are generally consistent across different populations, with a higher rate of migration in early adulthood and a lower rate of migration upon exit from the labour market (Rogers and Castro 1981) although variations are known to exist (Bernard et al. 2014; Bernard and Bell 2015). Migration decisions are known to vary along the life-cycle model, where life-course transitions such as leaving parental home, marriage and employment determine patterns of movement (Aybek et al. 2015).

There are scarce empirical regularities in gendered patterns of migration. Most gendered studies of migration are carried out from an anthropological perspective, using qualitative, ethnographic, and eclectic methods (Donato et al. 2006). Evidence of gender differentials in migration patterns thus remains inconclusive. While studies of micro-level data point to a feminisation of migration (Camlin et al. 2014; Reed et al. 2010), distinct male migration corridors remain evident and, in some international cases, have increased in volume and share in recent decades (Abel 2018). The lack of clear evidence on gendered patterns of migration may depend on differences in a country's socioeconomic structure and the time period studied. The increase in international female migration in OECD countries, for instance, is due to a growing number of economic migrants among women rather than marriage migration or family reunification (Le Goff 2015).

Furthermore, there is no conclusive evidence on whether migrants are drawn from a pool of less or more educated individuals. Empirical studies show that the direction of selectivity by educational level differs across countries. (Cattaneo 2007; Gould 1982). Empirical studies at the individual level provide inconsistent evidence on the relationship between education and the propensity to migrate. On the one hand, a series of studies reported a positive effect of educational attainment on the likelihood of migration (Donato 1993; Stark and Taylor 1991; Williams 2009; Yang and Guo 1999) . On the other hand, many studies found a negative relationship between education and migration (Massey et al. 1987; Massey and Espinosa 1997; Quinn and Rubb 2005) and some studies reported no significant association at all (Adams 1993; Curran and Rivero-Fuentes 2003). The inconsistency in the effects of educational attainment on migration may partially stem from changing gender roles. Given that men are conventionally more likely than women to migrate for secure employment or to improve living conditions, it is expected that the effects of educational attainment on migration are more potent for men. Nevertheless, Williams (2009) reported that, with a rise in female labour-force participation, women in rural Nepal are increasingly migrating for job reasons rather than exclusively for marriage. Accordingly, she found that education is more important for female migration than for males.

To provide more consistent findings on migration differentials by gender and education, comparisons across multiple countries and time periods are required. Studies comparing levels and patterns of internal migration across countries are limited (Bell et al. 2002). The lack of cross-national empirical studies of migration stems from the complexities of measuring migration and defining spatial units. Comparisons of internal migration that do exist tend to be based on a handful of countries or summary measures and rankings without disaggregation by gender or education (Bell et al. 2015; Rees et al. 2016). To this end, our study aims to utilise statistical methods to control for cross-country differences and provide a systematic analysis of demographic profiles of internal migrants by gender and education across 58 countries. We begin by describing migration intensity by gender and level of education in each country and census for which data are available. Subsequently, we investigate how both country contexts and ``push'' factors at the origin and ``pull'' factors at the destination influence migration patterns of different gender and education groups. This is undertaken using a series of multilevel models, where we use weighting to control for differences in spatial units across countries and random effects to control for variation across migration definitions, countries, and census years. In the results section, we illustrate differences in the associations by education levels, where several distinct differences in the country-level contextual variables and regional push-and-pull factors are apparent. In the final section, we summarise and discuss our results.

2 Data

Migration and socio-demographic data are derived from harmonised census microdata samples from the Integrated Public Use Microdata Series International (IPUMS-I) database (Minnesota Population Centre, 2015). Each set of census microdata contains a small random sample (0.4%-10%) of unidentified private households and associated persons based on a complete census conducted by the national statistical agency in each country. The countries and years used in this study, shown in Figure 1, are based on censuses collected between 1960 and 2012. In total, we used data from 434,108,676 individual records across 58 countries and 198 censuses to inform our analysis.

The initial exploratory plot in Figure 2-1 demonstrates variation in the female share of internal migration flows across education groups in both space and time. In many countries, the proportion of female flows in higher education attainment categories is low, particularly in African and Asian countries. When data from multiple censuses are available, we observe feminisation of migration at different stages, depending on the education group. For example, in Cameroon or Thailand, the proportion of female migrants in the more educated age groups is much lower in the first available census. The share of female migrants becomes more balanced over time, first among less educated populations and then among more educated groups. Migration within European and North American countries tends to have balanced gender decompositions for each educational attainment level.

Our eligibility criteria for including countries in our analysis were based on the availability of migration, gender, and education measures in the IPUMS-I, where all three were required to derive bilateral migration flows by gender and educational attainment. Figure 2-2 illustrates a map of countries used in this study, with their sub-regional boundaries used to define migration flows.

One advantage of using the IPUMS-I database is that potential explanatory variables for migration are both available and standardised, allowing cross-country comparisons. However, the geographical detail available for each country is not uniform and depends on the sample size, population distribution, and the administrative units in place.

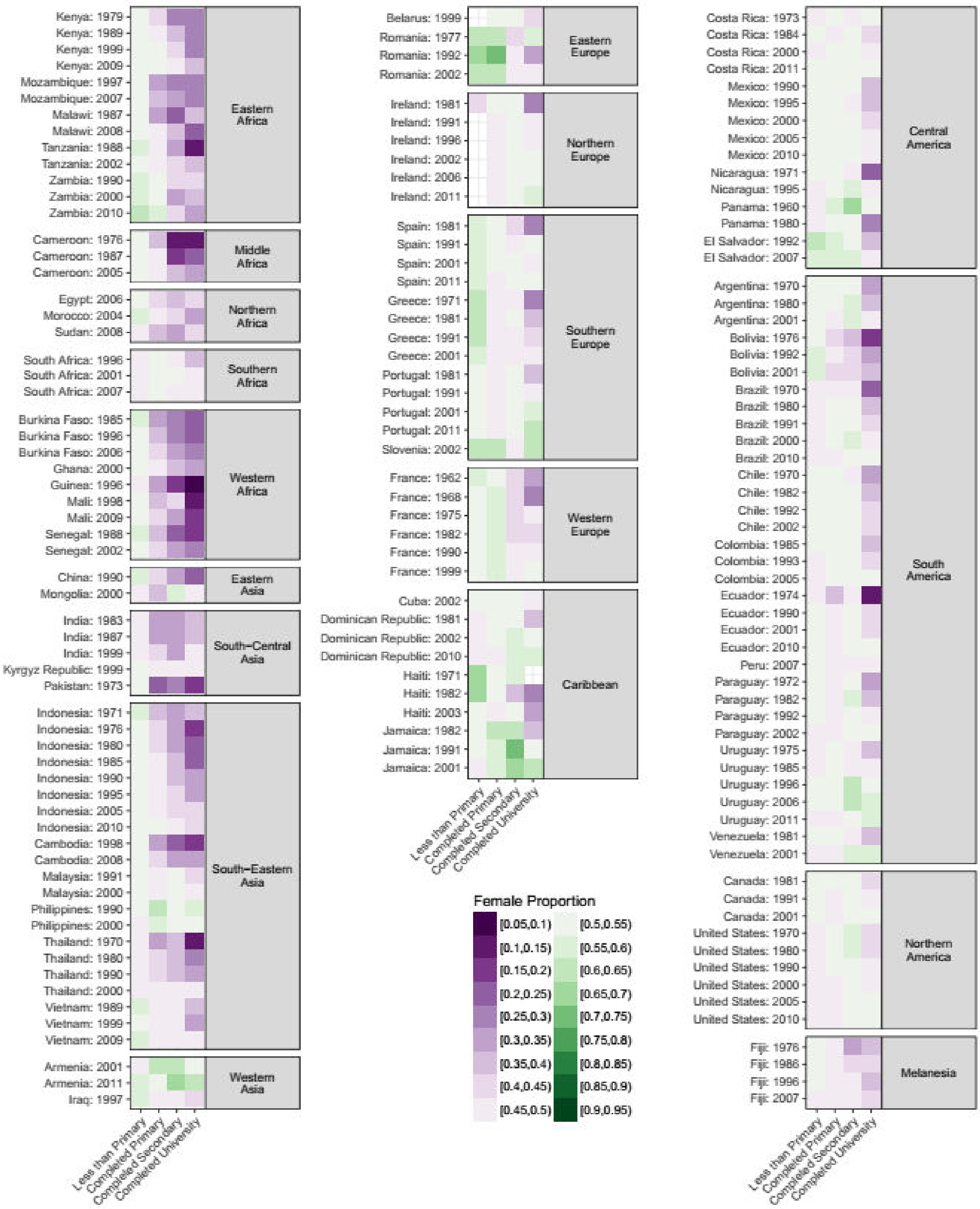

Figure 2-1 Countries and census years used in the study, along with the proportion of female internal migrants calculated from the IPUMS-I data.

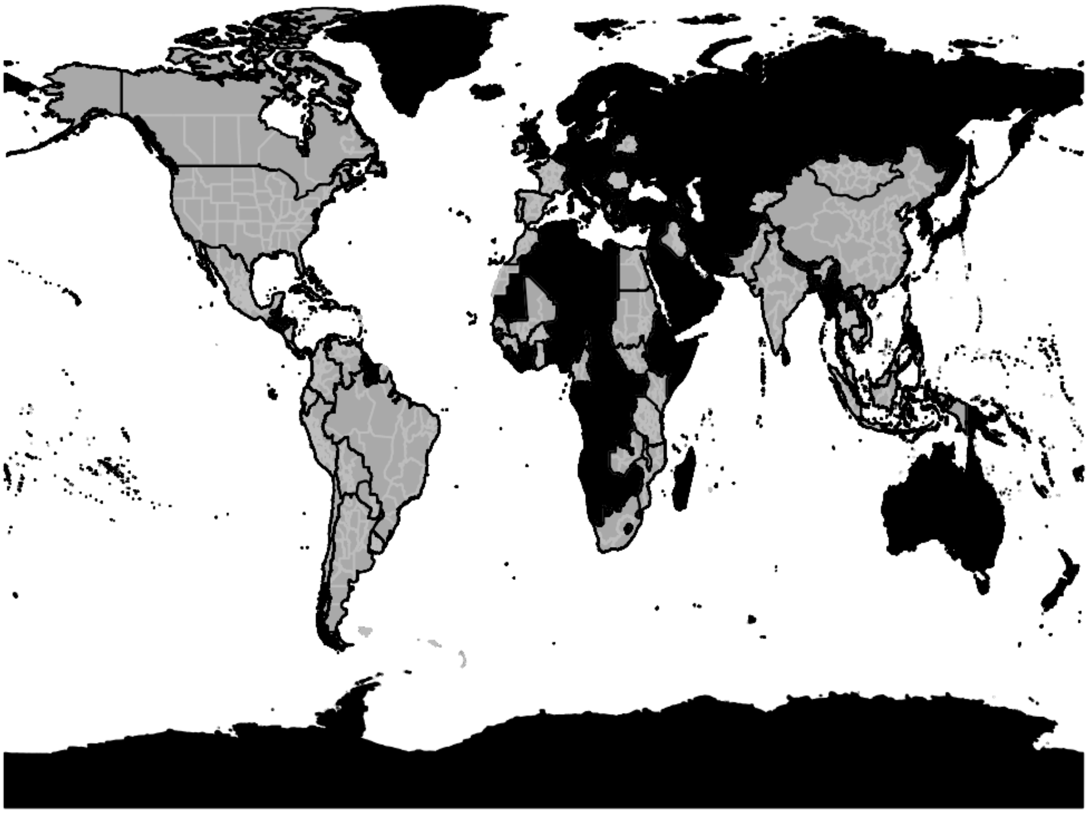

Figure 2-2 Countries used in the study and their sub-national boundaries.

3 Measurement and Model Specification

3.1 Migration Measurement

The count of migration flows between the pair of origin *i* and destination *j* was taken from the question, which asks individuals where they lived previously. These can be separated into two broad classes based on 1) a question asking respondents for their place of residence at a fixed time interval (such as place of residence five years ago) or 2) a set of three questions on the length of stay in current location, place of previous residence, and whether the previous residence was in the same administrative unit. For the latter, we used the combined responses to derive a one-year migration flow by counting those who had changed their major administrative unit in the year preceding the census date. When responses to both classes of migration questions were available, we gave preference to the fixed-time-interval measure. When multiple fixed-interval measures were available, we gave preference to the shortest interval. From these, we derived eight internal migration measures between all regions (by sex and four education levels) based on the summation of migration responses for individuals aged 15 and above.

The geographic coding of current and previous residence in IPUMS-I did not always yield the same numeric values. On occasion, they also conflicted with the IPUMSI spatial files we later used to calculate distance measures. We harmonised the coding between origin and destination units manually and, where required, matched them to the appropriate shapes in the spatial files. In some countries, we made one of two adjustments to the geographies to enable a more effective cross-national comparison. Firstly, in some countries, individuals' previous and current residences were reported at the second administrative level. This led to fewer reported migrations in some countries. In these cases, we aggregated administrative units to the first administrative level to provide a more robust estimate of migration flows. Secondly, in the IPUMS shapefiles, administrative units at the same level were aggregated, with individual responses on previous and current residence provided at the non-aggregated geography. In these cases, we used a spatial file from the Database of Global Administrative Areas, which offered a better, more detailed match to the regional geography in the migration questions from IPUMS-I.

3.2 Explanatory Variables

We derived two geographic measures to study how each factor might vary by sex and education. The first was based on the distance $DIST_{ij}$ between origin and destination regions corresponding to our migration flow measures. We used a Euclidean distance i.e. the straight-line distance between two points calculated from the centroid of each administrative unit to the centroid of all other administrative units in the same county. The second was a contiguity $CONT_{ij}$ dummy variable that equalled one if regions shared a border and zero otherwise.

Other socio-demographic and economic variables previously found to be associated with migration were also constructed from census microdata for each administrative unit. These variables include: 1) the total population of the origin and destination region (POP_i and POP_j) 2) the proportion living in an urban environment in the region ($URBAN_i$ and $URBAN_j$) 3) the median age of the population (AGE_i and AGE_j), 4) the proportion economically active (EMP_i and EMP_j) aged between 15 and 60 years old, and 5) the proportion who did not complete their primary education ($LOWED_i$ and $LOWED_j$). Summary measures of each of these variables are shown in Table 3-1. Also given in Table 3-1 are summaries of country-time-specific economic factors: GDP per capita (GDPc) and the Gini coefficient (GINIc), a standard measure of income inequality. These two variables will allow us to see how migration levels by sex and education vary across countries with different economic conditions.

Table 3-1 Summary of explanatory data used in the study

Variable	*N*	*Mean*	*Std. Dev.*	*Min.*	*Max.*
*GDP per capita (*GDP*)*	198	110,21.92	11,676.47	596.69	50,512.41
*Income Inequality (*GINI*)*	198	42.61	9.41	21.8	67.4
*Population (*POP*)*	5,604	2,311,312	7,863,921	80	160,415,921
*Distance (*DIST*)*	132,732	543.15	612.17	1.3	16,001.12
*Contiguity (*CONT*)*	132,732	0.07	0.26	0	1
*Urban Proportion (*URBAN*)*	4,356	0.37	0.31	0	1
*Median Age (*AGE*)*	5,604	22.4	7.2	12	50
*Actively Employed (*EMP*)*	5,183	0.61	0.13	0.21	1
*Low Education Proportion (*LOWED*)*	5,559	0.47	0.3	0	1

3.3 Missing Data

The *URBAN*, *EMP* and *LOWED* variables were not always available for each administrative unit in each census, with 22.2%, 7.5% and 0.8% missing, respectively. Rather than dropping all the migration flows to and from these observational units, we used a multiple imputation strategy, implemented via the *Ameila* R package (Honaker et al. 2006) to generate five alternative data sets.

3.4 Model Specification

A sequence of spatial interaction models was fit to each of the multiple imputation data sets using a cross-classified multi-level regression model. Derived from the gravity theory of migration (Zipf 1942, 1946) this model focused on the role of distance and population sizes in explaining spatial movements. We specified our model to account for additional variation generated by a wide range of origin and destination regions, the set of countries in our study, the duration intervals used to define migration, and the years of the censuses used. This was operationalised in our model using random-effect parameters that cross-classify each observed flow by these characteristics. Before fitting the gravity model, we also control for economic development and income inequality. We then use additional fixed effects parameters to study the role of the socio-demographic and geographic variables described above. To account for different administrative geographies across countries, we weighted observations by the number of possible migration corridors. This allowed migration flows within countries with fewer regions to weigh more than flows within countries with many regions.

$$
\begin{aligned}
y_{ijctd} = \beta_0 &+ u^{(2)}_{ORIG(i)} + u^{(3)}_{DEST(j)} + u^{(4)}_{COUNTRY\ (c)} + u^{(5)}_{YEAR(t)} + u^{(6)}_{DURATION(d)} + \\
&\beta_1 \log GDP_{ct} + \beta_2 GINI_{ct} + \\
&\beta_3^O \log POP_{it} + \beta_3^D \log POP_{jt} + \beta_4 \log DIST_{ij} + \\
&\beta_5 \log CONT_{ij} + \\
&\beta_6^O \operatorname{logit} URBAN_{it} + \beta_6^D \operatorname{logit} URBAN_{jt} + \\
&\beta_7^O AGE_{it} + \beta_7^D AGE_{jt} + \\
&\beta_8^O \operatorname{logit} EMP_{it} + \beta_8^D \operatorname{logit} EMP_{jt} + \\
&\beta_9^O \operatorname{logit} LOWED_{it} + \beta_9^D \operatorname{logit} LOWED_{jt} + \varepsilon_{ijctd},
\end{aligned}
$$

where $\varepsilon_{ijtcd} \sim N(0, \sigma^2)$ and $u^{(k)} \sim N\left(0, \sigma^2_{(k)}\right)$, $k = 2, \dots, 6$. This model sequence was run eight times, once for each sex-education attainment level combination, using the *lme4* package in R (Bates et al. 2015).

4 Results

4.1 Fixed Effects

The fixed-effect parameter estimates from the complete model for the first of the multiple imputation data sets are plotted in Figure 4-1. The uncertainty in the parameter effects is illustrated by error bars representing +/-1.96 times the standard error. In this section, we describe the fixed-effect parameters into three groups: country context effects, geographical factors, and social and economic characteristics of the origin and destination regions.

There are distinct differences in the country context parameters across models based on different levels of sex education attainment. The estimated parameter for the level of economic development, measured by GDP per capita, exhibits a J-shape as education levels increase. Migration of those higher levels of attained education was associated with increases in GDP per capita. For males, a 1% increase in national GDP per capita leads to a 0.15% increase in overall migration among those with completed university education (with all other factors held constant). For females with the same level of education, the effect is large: a 1% increase in GDP per capita leads to a 0.3% increase in migration. In population groups with either less than completed primary education or completed secondary education, there was little association between migration level and GDP per capita. However, for those with completed primary education, a 1% increase in GDP per capita would lead to more than a 0.1% reduction in the number of internal migrant moves.

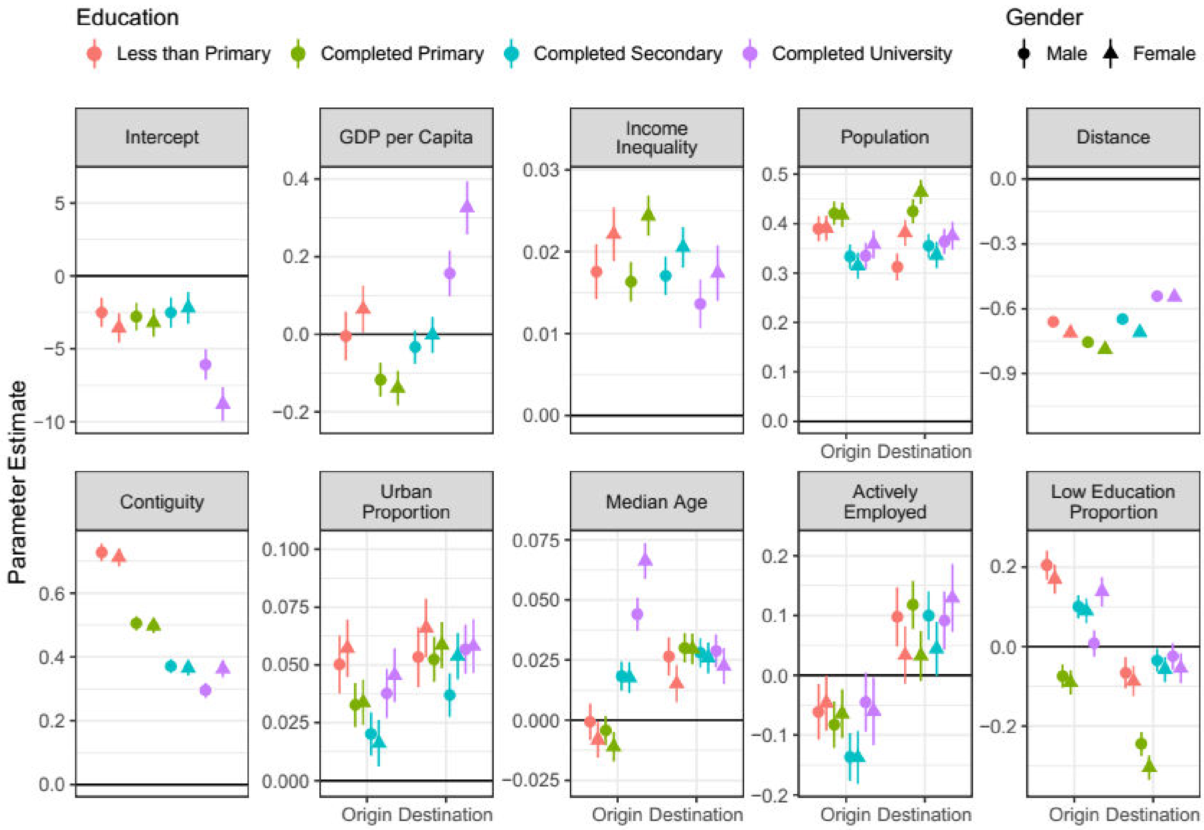

Figure 4-1 Fixed effect parameter estimates and their standard errors from the complete model on the first of the multiple imputation data sets. Note, error bars represent 1.96 times the standard error.

Migration in all sex-education groups is positively associated with income inequality, as measured using the Gini coefficient. For those with less than primary education, a unit increase in the Gini coefficient will lead to around 0.02% more migration. The positive effect of rising income inequality on migration is consistently greater for females than for males across all education groups. The migration levels of those in higher education categories tend to be less affected by income inequality. For example, males with completed university education are expected to migrate at only 0.014% more per unit increase in the country's Gini coefficient.

All estimated distance parameters are negative, indicating that regions further apart experience lower levels of migration. The sizes of the negative parameters differ by sex and educational groups. In each sex, the parameters exhibit a distinctive J-shaped pattern as education levels increase. The estimated parameters are lowest among those with completed primary education. For males and females with completed primary education, a hypothetical 1% increase in the distance between any two regions would lead to an approximate 0.75% decrease in migration flows (assuming all other factors remain unchanged). The migration levels of those with higher education are least affected by longer distances, where, for example, as distances between regions increase by 1% the migration flows fall on average by only 0.5%. At all levels of education, except among those with completed university education, female groups exhibit larger negative parameters, indicating greater sensitivity of migration flows to distance.

The contiguity parameters decline sharply as educational attainment increases. All are positive, implying expected higher migration between bordering regions. For those in the lowest education category, bordering regions are associated with, on average, 0.7% higher migration flows than non-neighbouring regions. For those with completed university education, especially males, the association between migration and contiguous regions is weaker, with neighbouring regions experiencing only a 0.3% increase in migration flows. Considered alongside the distance measures, the patterns in the estimated parameters indicate that those with higher educational attainment are more selective and willing to move farther than those with lower levels of education.

The effect of population size, classically used in gravity model analyses, is very similar across all sex-education categories. Unsurprisingly, the parameters at both the origin and the destination are positive, indicating that migration is expected to increase as the populations at both locations grow.

Regions with higher levels of urbanisation experience more in- and out-migration. Whilst there are no clear distinctions in the positive association between migration into destination regions and education or sex, there are apparent differences in moves out of origin regions. Populations at either end of the education spectrum are positively associated with higher levels of migration out of origin regions, with higher levels of urbanisation than those with completed primary or secondary education.

As with the effects of urbanisation, regions with higher median ages attract more migrants, with very little difference by sex or educational level; however, distinct differences occur between education groups in migration out of regions, depending on median age. In particular, those with completed secondary and university are far more likely to leave regions with higher median ages. The strongest association is among female migrants with completed university education, for whom a single-year increase in the median age in the origin region is expected to lead to a 0.07% increase in migration levels. Migration from the original regions of populations in lower education groups tends to be unaffected by changes in the overall population becoming younger or growing older.

Higher proportions of the population actively employed in the origin region are negatively associated with migration across all education groups. The association is strongest among those with completed secondary education, suggesting that an increase in the employment rate of the population will lead to relatively lower migration out of the origin region than among those with other educational attainment levels. The association between migration regions and the proportion of the population that is actively employed is switched when considering destination regions. In this case, in all education groups, there is a positive association between the high levels of employment and more migrants arriving in the region. The association is approximately the same for males across all education categories as for females with university education. Males in lower-educated populations are more likely to migrate to regions with higher employment levels than their female counterparts.

Populations with completed primary education are less likely to migrate from regions with high shares of the population with lower (less than primary) education. As the share of the population with lower education in origin regions increases, more migrants from all other educational groups (except males with completed education) are expected to leave. Combined, the results suggest that those with completed primary education are aware of a competitive advantage in regions with high proportions of the population that did not complete primary education and, hence, lack the incentives to out-migrate that are present for populations in other education groups. Similar distinctions are evident when considering the associations between destination regions and high shares of the population with less than primary education. All migrant groups are deterred from moving to these regions, where all parameter estimates are negative. However, the negative association is strongest for those with completed primary education. Thus, whilst those who complete primary education are strongly incentivised to remain in low-education origin regions, they are also strongly deterred from moving to similar destination regions.

4.2 Multiple Imputation

To gauge the robustness of the model fits to the multiple imputation process, each of the eight models in the sequence was fitted to the five multiple-imputation data sets. For clarity in the previous subsection, the results for the parameters based only on the first data set were used. In Figure 4-2, the parameter estimates from all five data sets, based on different imputed values for missing values, are plotted.

The consistency in parameter estimates for the variables with no missing data (population, distance, contiguity, and median age) is clearly visible. Elsewhere, there is a larger variation in the parameter estimates due to the use of alternative imputed values. This is greatest in the actively employed and urban proportion parameters, where the missingness was highest. However, almost all the parameters have overlapping error bars, indicating that we can be reasonably confident the imputation strategy has not led us to misinterpret the general direction or sign of the estimated parameter effects.

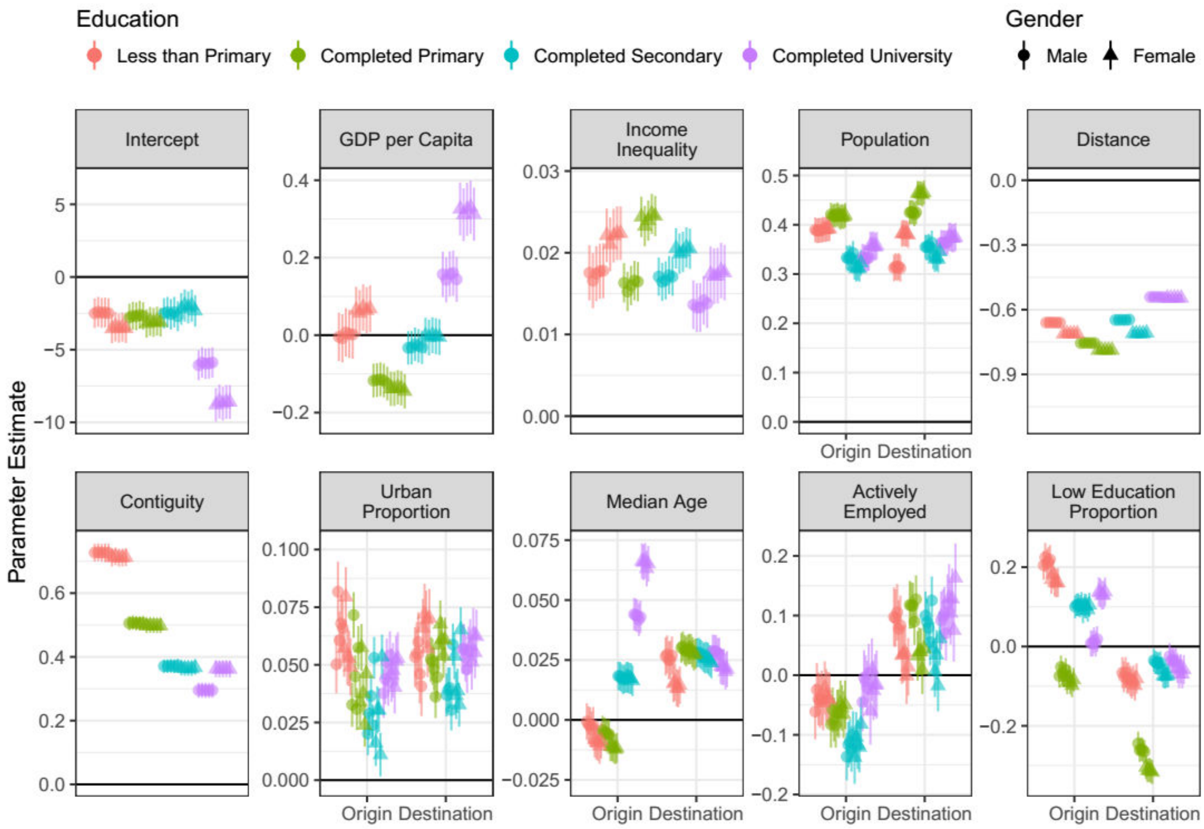

Figure 4-2 Fixed effect parameter estimates and their standard errors from the full model on all multiple imputation data sets. Note, error bars represent 1.96 times the standard error.

4.3 Random Effects

In each model, we controlled for variation from the origin, destination, country, year and interval duration definition for migration using random effects. In Figure 4-3 a plot of the standard deviations of each of the random effects is shown. The estimated standard deviations for both the origin and destination parameters indicate lower variation between regions than for the other parameters. This is due to the large number of random-effect parameters (over 1200) used for each of the origin and destination variables. In each sex-education group, the standard deviation of destination residuals is slightly greater than that of origin residuals, indicating greater overall selectivity among migrants in choosing their destination regions. The variation explained by using multiple countries accounts for one of the largest components of the model's variation. The larger country-standard deviation is due to the wide range of administrative geographies across our 58 countries. Administrative geographies directly influence migration flows: countries with more boundaries tend to generate higher flows. The variation in the standard deviation parameters for years illustrates changes in the migration in countries with multiple censuses. Some of the largest differences by educational level in parameter values here indicate higher variation in migration levels among those with less than primary education than among those with other levels of education. The standard deviations associated with the duration timing intervals are relatively large. As discussed later, larger intervals used in some countries resulted in greater overall migration. The final set of standard deviations corresponds to the residuals. Their values are considerably smaller than those of the other sources of variance, suggesting that appropriate controls for variation between the data sources have been used.

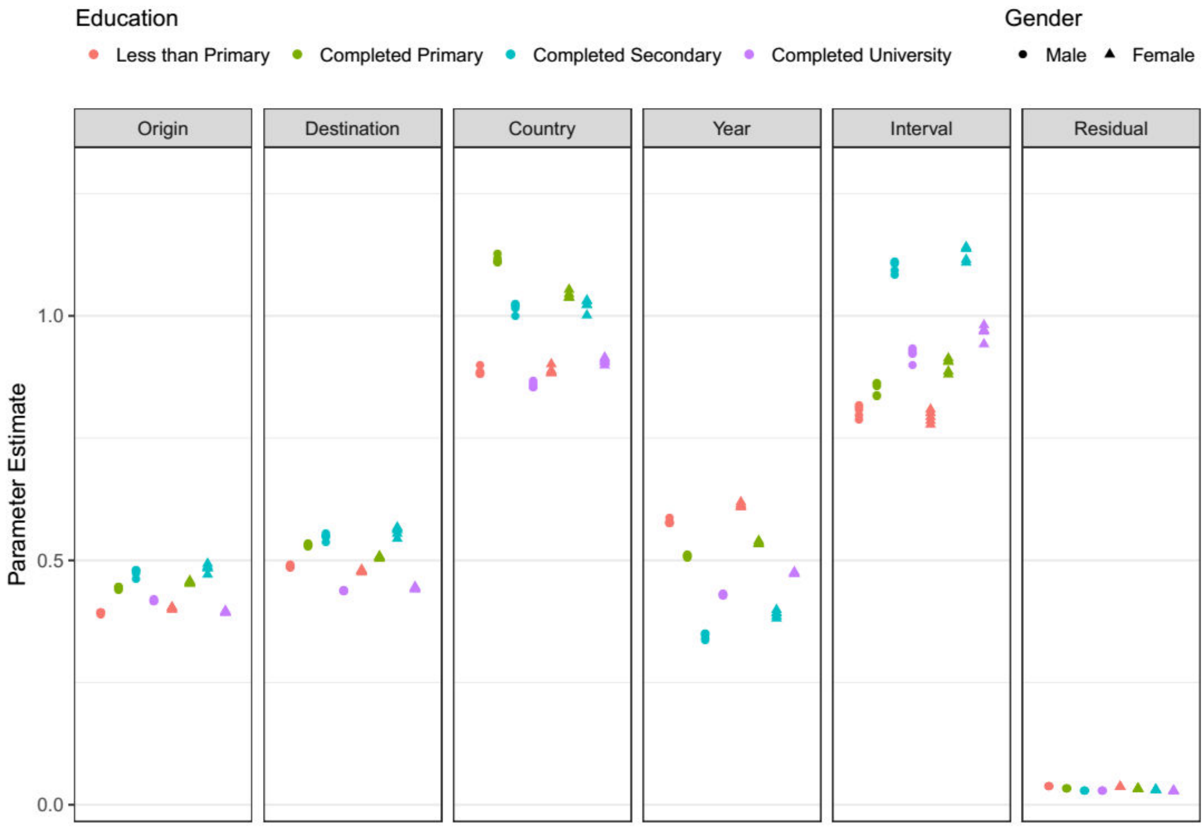

Figure 4-3 Standard deviations of random effect parameter estimates from the full model on the multiple imputation data sets

In Figure 4-4 the random effects of the country, interval and year are plotted. We do not display the random-effect parameters for over 1,200 origin-destination regions. As discussed, the standard deviation of these values reflects the level of country effects, an artefact of the administrative geography. The interval random effects show a general increase, indicating that for shorter durations, fewer migrants are recorded. The variation in the values is greater, especially over more extended periods, during which only a few countries use a fixed interval based on location from the previous census date to define migration flows. The year random effects are a by-product of differences in census timing across countries.

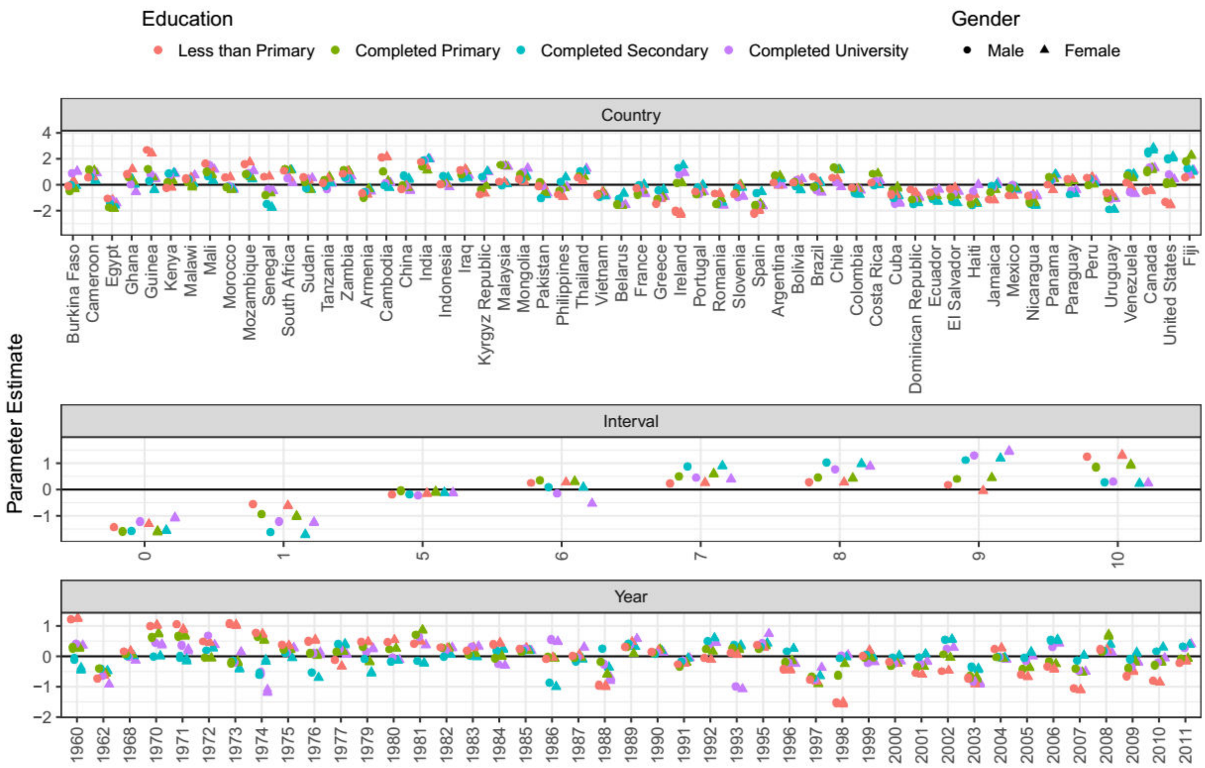

Figure 4-4 Random effect parameter estimates and their standard errors from the full model on the first of the multiple imputation data sets

5 Summary

In this paper, we have explored differences in internal migration levels by sex and education. From IPUMS-I, we constructed a near-exhaustive database of migration moves within countries and potentially related explanatory variables.

In our initial exploratory analysis, we found considerable variation in the female share of internal migration flows across education groups, both in space and over time. We identified some general patterns in the data: female shares of migration totals tended to be higher in lower-education categories, especially in earlier time periods and in less developed countries. In more recent data and in more developed countries, the sex inequality in migration was far less pronounced across all education categories. This pattern may indeed reflect the increase in female labour force participation, which, in turn, drives female labour migration. (Leibert 2016; Posel and Casale 2003).

To better understand these differences and explore regional variation, we used a series of multilevel models. In each model, we used the migration flow count on each sex-education level as a dependent variable. Our explanatory variables were categorised into four groups. The first group involved economic measures to provide a country and time-specific context on migration levels by sex and education. The effect of economic development formed an inverted J-shape (as the education levels increased). Migration of those with higher levels of attained education was associated with increases in GDP per capita. In contrast, migration of those with only completed primary education was associated with decreases in GDP per capita. Migration of those with less than a primary education was also positively related to increases in countries' GDP per capita, but to a lesser extent than that of those with a completed university education. The level of migration across all sex-specific education levels, but especially among females, was positively associated with increases in income inequality in countries.

Our second group of variables provides geographic explanations for variations in migration by sex and education. In each population, we found that lower migration levels were associated with longer distances. This negative association was less severe for those with higher education. Populations with lower levels of education showed a stronger association with movement to neighbouring regions than those from higher education strata. In all our sex-education subgroups, we found positive associations between migration levels and urbanity levels in both the sending and receiving regions.

The third group of variables helped provide a more detailed social-demographic context for the internal migration flows in our data set. We found distinct differences in the association between outward migration levels and the regional median age. Those in higher education groups were more likely to move away from older regions. In contrast, those in lower education groups, particularly females with higher education, were less likely to do so. There was little to no association between migration levels and the age of the origin region among those in lower education categories. Those with completed primary education were associated with lower out-migration when employment levels in a region increased. They were also more likely to remain in regions with large populations without a completed primary education.

The fourth group of variables were random effects to control for variation across origin and destination regions, as well as the country, year, and duration definitions used in migration. Using a weighted multilevel regression model, we were able to undertake cross-national comparisons of migration whilst accounting for differences in administrative geographies and migration definitions. We used a multiple imputation strategy to predict values for explanatory variables, thereby including as many countries as possible in our analysis.

In conclusion, migration, like other demographic components of population change, can vary considerably by sex and education level. Using the IPUMS-I database, we have conducted a near-exhaustive exploration of migration patterns by sex and education at the macro level. We found variations in migration levels by sex and education level at both the country level, in relation to economic conditions, and the regional level, in relation to geographic, economic, and social factors.

Expected Updates

This initial work was carried out pre-pandemic and has not been published. I intend to update the models for 1) new IPUMS-I data covering census samples from the most recent round and some additional countries, 2) to re-frame the explanatory variables to cover three different aspects of development (health, education and living conditions), and 3) update the literature review to cover new publications on cross-national comparisons of internal migration. I am also considering 4) removing the sex-differential aspect to concentrate solely on educational differences and 5) implementing robust checks for our results – thus far, regression models were run on only the positive migration flows, where zero counts were dropped, following a similar approach to that of (Beine et al. 2011; Kim and Cohen 2010) and assumed residual errors follow a Gaussian normal distribution – I would like to explore multi-level Pseudo-Poisson or Negative Binomial type models to include zero counts – but my recent efforts have failed (models not converging)

Bibliography

Abel, G. J. (2018). Estimates of Global Bilateral Migration Flows by Gender between 1960 and 2015. *International Migration Review*

Adams, R. H. (1993). The economic and demographic determinants of international migration in Rural Egypt. *The Journal of Development Studies*, *30*(1), 146–167. doi:10.1080/00220389308422308

Aybek, C. M., Huinink, J., & Muttarak, R. (2015). Migration, Spatial Mobility, and Living Arrangements: An Introduction. In *Spatial Mobility, Migration, and Living Arrangements* (pp. 1–19). Cham: Springer International Publishing. doi:10.1007/978-3-319-10021-0_1

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Beine, M., Docquier, F., & Özden, Ç. (2011). Diasporas. *Journal of Development Economics*, *95*(1), 30–41. doi:10.1016/j.jdeveco.2009.11.004

Bell, M., Blake, M., Boyle, P., Duke-Williams, O., Rees, P., Stillwell, J., & Hugo, G. J. (2002). Cross-national comparison of internal migration: issues and measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *165*(3), 435–464. doi:10.1111/1467-985X.00247

Bell, M., Charles-Edwards, E., Ueffing, P., Stillwell, J., Kupiszewski, M., & Kupiszewska, D. (2015). Internal Migration and Development: Comparing Migration Intensities Around the World. *Population and Development Review*, *41*(1), 33–58. doi:10.1111/j.1728-4457.2015.00025.x

Bernard, A., & Bell, M. (2015). Smoothing internal migration age profiles for comparative research. *Demographic Research*, *32*(1), 915–948. doi:10.4054/DemRes.2015.32.33

Bernard, A., Bell, M., & Charles-Edwards, E. (2014). Life-course transitions and the age profile of internal migration. *Population and Development Review*, *40*(2), 213–239. doi:10.1111/j.1728-4457.2014.00671.x

Camlin, C. S., Snow, R. C., & Hosegood, V. (2014). Gendered Patterns of Migration in Rural South Africa. *Population, Space and Place*, *20*(6), 528–551. doi:10.1002/psp.1794

Cattaneo, C. (2007). *The Self-Selection in the Migration Process: What Can We Learn?* (No. 199). *LIUC Papers in Economics*.

Curran, S. R., & Rivero-Fuentes, E. (2003). Engendering migrant networks: The case of Mexican migration. *Demography*, *40*(2), 289–307. doi:10.1353/dem.2003.0011

Donato, K. M. (1993). Current trends and patterns of female migration: evidence from Mexico. *The International Migration Review*, *27*(4), 748–771.

Donato, K. M., Gabaccia, D., Holdaway, J., Manalansan IV, M., & Pessar, P. R. (2006). A glass half full? Gender in migration studies. *International Migration Review*, *40*(1), 3–26. doi:10.1111/j.1747-7379.2006.00001.x

Gould, W. T. S. (1982). Education and internal migration: A review and report. *International Journal of Educational Development*, *1*(3), 103–111.

Honaker, J., King, G., & Blackwell, M. (2006). Amelia II: A program for missing data. *Journal of Statistical Software*, *45*(7).

Kim, K., & Cohen, J. E. (2010). Determinants of International Migration Flows to and from Industrialized Countries: A Panel Data Approach Beyond Gravity. *International Migration Review*, *44*(4), 899–932. doi:10.1111/j.1747-7379.2010.00830.x

Le Goff, M. (2015). Feminization of migration and trends in remittances. *IZA World of Labor*. doi:10.15185/izawol.220

Leibert, T. (2016). She leaves, he stays? Sex-selective migration in rural East Germany. *Journal of Rural Studies*, *43*, 267–279. doi:10.1016/j.jrurstud.2015.06.004

Lutz, W. (2014a). A Population Policy Rationale for the Twenty-First Century. *Population and Development Review*, *40*(3), 527–544. doi:10.1111/j.1728-4457.2014.00696.x

Lutz, W. (2014b). Epilogue: With education the future looks different. In W. Lutz, W. P. Butz, & S. KC (Eds.), *World Population and Human Capital in the 21st Century* (pp. 628–649). Oxford: Oxford University Press. http://ukcatalogue.oup.com/product/9780198703167.do

Massey, D. S., Alarcón, R., Durand, J., & González, H. (1987). *Return to Aztlan: The Social Process of International Migration from Western Mexico*. Berkeley; Los Angeles; London: University of California Press. http://www.jstor.org/stable/10.1525/j.ctt1ppp3j. Accessed 17 April 2017

Massey, D. S., & Espinosa, K. E. (1997). What's Driving Mexico-U.S. Migration? A Theoretical, Empirical, and Policy Analysis. *American Journal of Sociology*, *102*(4), 939–999.

Minnesota Population Center. (2015). Integrated Public Use Microdata Series, International: Version 6.4 [Machine-readable database]. Minneapolis, Minnesota, USA.

Posel, D., & Casale, D. (2003). What Has Been Happening to Internal Labour Migration in South Africa, 1993–1999? *South African Journal of Economics*, *71*(3), 455–479. doi:10.1111/j.1813-6982.2003.tb00081.x

Quinn, M. A., & Rubb, S. (2005). The importance of education-occupation matching in migration decisions. *Demography*, *42*(1), 153–167.

Reed, H. E., Andrzejewski, C. S., & White, M. J. (2010). Men's and women's migration in coastal Ghana. *Demographic Research*, *22*, 771–812. doi:10.4054/DemRes.2010.22.25

Rees, P., Bell, M., Kupiszewski, M., Kupiszewska, D., Ueffing, P., Bernard, A., et al. (2016). The Impact of Internal Migration on Population Redistribution: an International Comparison. *Population, Space and Place*, (April). doi:10.1002/psp.2036

Rogers, A., & Castro, L. J. (1981). *Model Migration Schedules*. *IIASA Research Report* (Vol. 81). Laxenburg, Austria.

Stark, O., & Taylor, J. E. (1991). Migration Incentives, Migration Types: The Role of Relative Deprivation. *The Economic Journal*, *101*(408), 1163–1178. doi:10.2307/2234433

Williams, N. (2009). Education, gender and migration in the context of social change. *Social science research*, *38*(4), 883–896.

Yang, X., & Guo, F. (1999). Gender Differences in Determinants of Temporary Labor Migration in China: A Multilevel Analysis. *The International Migration Review*, *33*(4), 929–953. doi:10.2307/2547358

Zipf, G. K. (1942). The Unity of Nature, Least Action, and Natural Social Science. *Sociometry*, *5*(1), 48–62.

Zipf, G. K. (1946). The P1P2/D hypothesis: On the intercity movement of persons. *American Sociological Review*, *11*(6), 677–686.