Sex differences in the performance of models for bilateral migration predictions

Extended Abstract

Athina Anastasiadou[1,2], Emilio Zagheni[1]

[1]Max Planck Institute for Demographic Research
[2]Netherlands Interdisciplinary Demographic Institute, KNAW, University of Groningen

November 1, 2023

Abstract

Migration is a complex process and difficult to predict. Nevertheless, predictions of migration are vital for demographic projections and policy making. At the same time, migration is known to be a highly gendered process reflecting different motivations, propensities, and outcomes of migration by gender. Such differences are mirrored in the fact that on average 44.7% of the bilateral migration corridors worldwide are male-dominated while only 28.2% are female-dominated. As theories and models have been shaped over decades around the narrative of the male migrant, this can potentially translate into discrepancies in migration predictions by sex. Using one of the most comprehensive macro-level data sets on bilateral migration flows disaggregated by sex, we aim to explore and understand such differences in migration predictions. We compare the predictive performance of a basic gravity model with demographic and geographic covariates to an extended version of that model including gender-sensitive predictors. Our preliminary findings show that indeed worse performance measures are achieved for female migration flows compared to male when applying the baseline model. Expanding the model by adding gender-sensitive indicators improves the model performance for the predictions. The next steps involve comparing other common demographic and econometric models of migration prediction. Further enhancing our understanding of the underlying mechanisms resulting in different prediction accuracy has the potential to inform the work of international organizations, researchers, and policy-makers.

1 Introduction

Migration is a complex process sensitive to the environment surrounding it, making it the most volatile demographic process and difficult to predict. Nevertheless, demographers and policymakers need reliable predictions of migration in order to anticipate migration flows and inform policies and population projections. Therefore, theories of international migration have emerged trying to explain why some people migrate and others do not. Such theories have translated into statistical methods for migration prediction of which many are applied by researchers and international organizations in their daily work. However, migration theories have been found to disregard gender-specific aspects of the migration process and therefore fall short of explaining patterns of female migration (Boyd & Grieco, 2003). The migration data landscape tells a similar story. Comprehensive sources for international migration data disaggregated by sex or gender are scarce (Hennebry, KC, & Williams, 2021).

Several works aim to evaluate the performance of models predicting migration (Welch & Raftery, 2022; Brücker & Siliverstovs, 2006). As of yet, no attempt has focused particularly on heterogeneities by sex or gender. This work strives to compare the predictive performance of the most common models in order to understand potential discrepancies by sex. In particular, we try to explore and discover differences in the models' predictive performances by sex. By adapting the models based on recent insights on gender differences in migration patterns, we want to understand why some flows are female- or male-dominated and develop statistical models that perform equally well for all sexes. Thereby, we want to shed light on sex biases, understand their structures and origins, and suggest how they can be mitigated.

2 Data

The migration data type that can provide the clearest picture of the sex composition of migration is migration flow data (Dennett, 2016). These data measure the number of people who changed their residence over a specific time interval (Yildiz & Abel, 2021). One of the few comprehensive databases providing sex-disaggregated international migration flows are the bilateral international migration flow estimates for 200 countries by G. Abel (2022a). These estimates were derived by G. Abel (2022b) based on migrant stocks and cover five year time intervals from 1990 to 2020.

2.1 Gendered patterns of international migration

Exploring some migration corridors more closely reveals that female and male migration behave differently and sometimes even exhibit diverging time trends. Overall, 44.7% of the bilateral migration corridors in the data set are male-dominated on average over time while only 28.2% are dominated by female migrants with the remaining ones being gender balanced (when looking at the Pseudo Bayes estimates produced by G. Abel (2022b). In almost half of the male-dominated corridors, the difference in the male share on the flow is larger than 25 percentage points. Two illustrative examples are the Mexico-United States migration corridor and the Kazakhstan-Russia corridor. While the former tends to be historically male-dominated, the latter tends to be dominated by female migration. According to the theory, the motivations of the migrants constituting these two flows should be fairly similar. However, they display notably different patterns (see Figure 2.1). Such observations trigger our curiosity given that researchers are trying to predict these patterns with the same types of models.

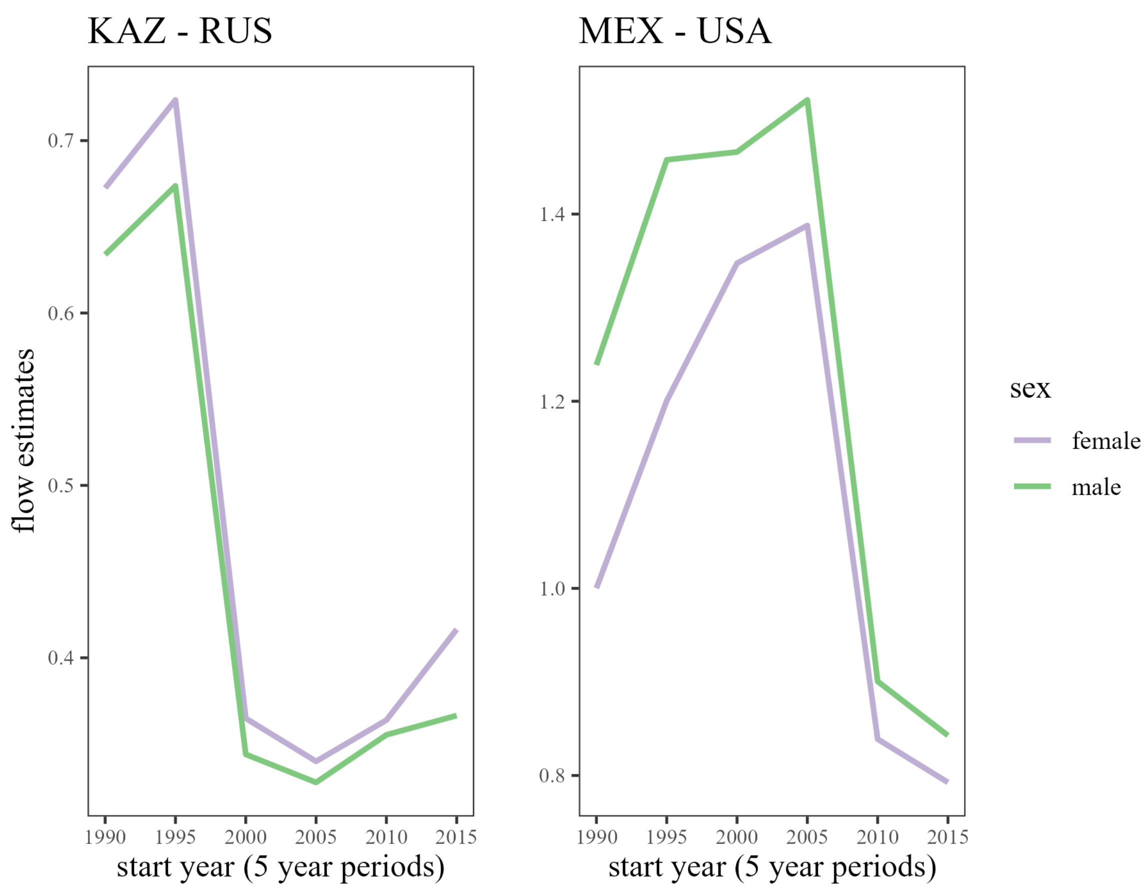

Figure 1: Migration flow count (in millions) by sex between Mexico and the US and between Kazakhstan and Russia from 1990 to 2020.

However, the picture drawn by the bilateral flow estimates should be interpreted with caution. Since the estimates were derived based on migrant stocks from time points that are five years apart, some gendered patterns might be obscured. Corridors like the ones referred to above, are known for enhanced seasonal and return migration patterns. These and other limitations of the estimates will be discussed in the following subsection.

2.2 Migration flow estimates

The methods that were applied by G. J. Abel and Cohen (2019); G. Abel (2022b) to produce the estimates are based on various assumptions.

In the usage notes, G. J. Abel and Cohen (2019) list limitations that affect all six estimation methods. One is that at least half of the estimated bilateral migration flows are zero. The estimates do not take into account additional migration events to countries outside the set of countries in the data set and return migrations during the five year time intervals. Moreover, estimates of bilateral migration flows produced by stock differencing or demographic accounting methods produce minimum estimates of flows (G. J. Abel & Cohen, 2019). Furthermore, estimates produced by the stock differencing and migration rates approaches ignore changes in the stocks due to deaths of migrants. This could potentially lead to an overestimation of migrant stocks especially in the presence of elderly migrants. While estimates produced with the demographic accounting methods take into account death and birth rates, these rates are assumed to be the same for foreign-born and native-born populations. Moreover, the validation of the measure could only be carried out for countries where migration flow data were reported, namely rich Western countries (G. J. Abel & Cohen, 2019).

G. Abel (2022b) use sex-specific inputs, namely sex-specific population counts, numbers of births and deaths where available, to derive the sex-specific bilateral flow estimates. However, for counties where information on the sex characteristics of migrants are missing, UN DESA imputed the data based on a regional or country model (G. Abel, 2022b). Not capturing gender-differences directly but inferring them might additionally lead to a general underestimation of these differences. All the assumptions have not been altered for female and male migration flows across the different estimation methods. Potential sources of gender bias in the estimation methods are therefore measurement error and imputation methods applied to the stock estimates by UN DESA.

Moreover, the estimates tend to underestimate overall migration flows due to the long time intervals. Consequently, this approach might underestimate the gender differences between the flows as well. As we know very little about South-South migration due to data scarcity, validating the sex-disaggregated estimates for non-Western countries is not possible. The determinants of emigration to wealthy countries certainly differ from determinants of emigration from wealthy countries (Massey, 2006). Therefore, we can expect the sex-composition of these flows to be different due to gendered immigration policies in Global North countries and an access to resources to migrate stratified by gender (Anastasiadou, Kim, Sanlitürk, de Valk, & Zagheni, 2023).

2.3 Demographic and geographic variables

For the following analysis, we included country-level data on the population of the countries of origin and destination POP, the share of the population residing in urban areas URB, the infant mortality ratio IMR, the potential support ratio (i.e. the number of persons aged 15-64 per person aged 65+ multiplied by 100) PSR, and the land area in squared kilometers LA. All these variables were obtained from the World Bank's World Development Indicators at the start of each period t (World Bank, n.d.). Indicators on shared land borders between countries LB, shared official language OL, and ever existing colonial relationship between the countries COL were obtained from the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) gravity database (Conte, Cotterlaz, & Mayer, 2022). Geographic coordinates were obtained from the *countryref* dataset in the *coordinate cleaner* package in R and used for calculating the bilateral distances to the countries' capital cities (Zizka et al., 2019).

2.4 Gender-sensitive migration indicators

For the extended version of the gravity model, we include gender-sensitive indicators to the estimation. The indicators were chosen based on our findings in a previous project in which we mapped out differences in the determinants, motivations, and obstacles to migrate for female compared to male migrants (Anastasiadou et al., 2023). The gender-sensitive predictors selected include an indicator for equal rights in the country of destination and origin proxied by a binary variable that equals one when men and women have the same right to divorce a spouse EQ. The indicator for equal economic opportunities is a binary variable that equals one when women and men face the same conditions when opening a business BUS. In order to proxy safety, we include a binary variable that equals one when legislation addressing domestic violence exists in the respective country SAF. An indicator for free mobility is a binary variable indicating that women and men can travel in the same way outside the country when equal to one TRV. These variables were obtained from the Gender Statistics database provided along with the World Development Indicators by the World Bank (World Bank, n.d.).

The baseline model developed by Kim and Cohen (2010) and employed by Welch and Raftery (2022) includes only demographic and geographic explanatory variables and no economic indicators. Therefore, we refrain from including gender-sensitive economic variables (like gender wage gaps, female employment rates, etc.). Such predictors would naturally have high explanatory power and obscure the picture. But we consider including them in a later stage of the analysis.

3 Methodology

In order to evaluate migration predictions by sex, we plan to produce predictions for the time period 2015-2020 by fitting a number of models for migration predictions to the data from 1990-2015. Starting with a baseline gravity model without economic predictors as developed by Kim and Cohen (2010) and later employed by Welch and Raftery (2022), we estimate gravity models for *total* flows and *female* and *male* flows separately.

The following equation describes the model specification of the baseline gravity model setup. The subscripts i and j stand for country of origin and destination respectively. The variables and their abbreviations have been described in more detail in the previous section. We fit the model to the time period 1990-2015 for each of the six different time series of estimates produced by G. Abel (2022b).

$$\begin{aligned} log(m_{i,j,t}) = \beta_0 + \beta_1 log(GDP_{i,t}) + \beta_2 log(GDP_{j,t}) + \beta_3 log(D_{i,j}) + \beta_4 log(PSR_{i,t}) \\ + \beta_5 log(PSR_{j,t}) + \beta_6 log(URB_{i,t}) + \beta_7 log(URB_{j,t}) + \beta_8 log(IMR_{i,t}) \\ + \beta_9 log(IMR_{j,t}) + + \beta_{10} log(LA_i) + \beta_{11} log(LA_j) + \beta_{12} LB_{i,j} \\ + \beta_{13} OL_{i,j} + \beta_{14} COL_{i,j} + \beta_{15}(t - 2005) + \beta_{16}(t - 2005)^2 + \epsilon_{i,j,t} \end{aligned} \tag{1}$$

The above variables and gravity model variables more generally are derived based on the theory of pull and push factors of migration and have been widely used for migration modelling since the 1940s (Lee, 1966; Pu, Zhao, Chi, Zhao, & Kong, 2019). However, literature has found that female and male migrants react differently to such factors in some cases. For instance, women have been found to be less sensitive to the distance between the origin and destination (Beine & Salomone, 2013). Moreover, labor-market-based explanations of migration do not take into account that labor markets are stratified by gender and hence fail to explain the sex composition in migration corridors as the ones illustrated in Figure 2.1 (Pedraza, 1991).

Therefore, we also estimate an extended version of the baseline gravity model including the above described gender-sensitive indicators.

$$\begin{aligned} log(m_{i,j,t}) = \beta_0 + \beta_1 log(GDP_{i,t}) + \beta_2 log(GDP_{j,t}) + \beta_3 log(D_{i,j}) + \beta_4 log(PSR_{i,t}) \\ + \beta_5 log(PSR_{j,t}) + \beta_6 log(URB_{i,t}) + \beta_7 log(URB_{j,t}) + \beta_8 log(IMR_{i,t}) \\ + \beta_9 log(IMR_{j,t}) + + \beta_{10} log(LA_i) + \beta_{11} log(LA_j) + \beta_{12} LB_{i,j} \\ + \beta_{13} OL_{i,j} + \beta_{14} COL_{i,j} + \beta_{15}(t - 2005) + \beta_{16}(t - 2005)^2 \\ + \beta_{17} EQ_{i,t} + \beta_{18} EQ_{j,t} + \beta_{19} BUS_{i,t} + \beta_{20} BUS_{j,t} + \beta_{21} SAF_{i,t} \\ + \beta_{22} SAF_{j,t} + \beta_{23} TRV_{i,j} + \beta_{24} TRV_{j,t} + \epsilon_{i,j,t} \end{aligned} \tag{2}$$

In order to evaluate biases that arise from gender-blind theories and models, our methodological approach is to comparing these two predictive models based on their performance in predicting migration flows by sex.

We rely on the out-of-sample mean absolute error for point forecasts as applied in Welch and Raftery (2022) to evaluate the predictions across models.

$$MAPE(M, \tilde{M}) = \frac{100}{F} \sum_{i \neq j} \frac{\mid m_{i,j} - \tilde{m_{i,j}} \mid}{m_{i,j} + 1} \tag{3}$$

with F denoting the total of flows contained in the flow matrix M.

4 Preliminary Findings

After fitting the two models to each of the six estimates produced by G. Abel (2022b) for the three categories total, male, and female, we can observe a weaker predictive performance of the models for the female data set. For both models, we ruled out multicollinearity issues by checking the variance inflation factors for each estimation. Tables containing detailed results can be found in the Annex 6.

4.1 Baseline gravity model

For the predictions derived based on the baseline gravity models, we can clearly see a pattern of differences by sex. The MAPE values in Figure 4.1 are the lowest for the models of total flows, while they are notably higher for the prediction by the models that were fitted to the female data. These preliminary results indicate differences in the prediction accuracy by sex which will be further explored in this paper.

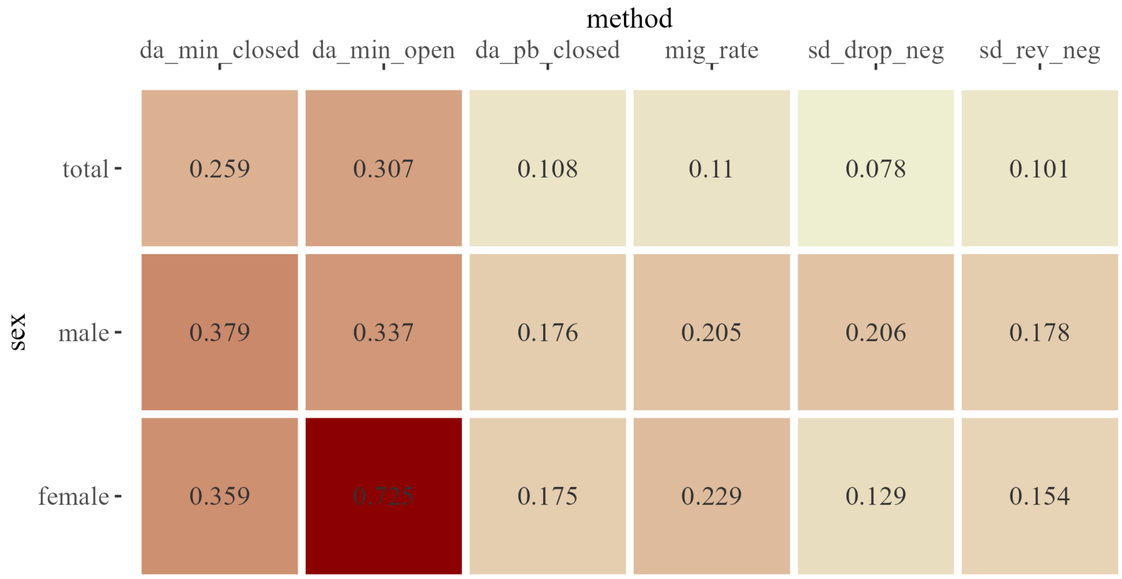

Figure 2: MAPE values for predicting migration flows in the period 2015-2020 with the *baseline gravity model* by using all different six estimators separately for male, female, and total flows.

4.2 Extended gravity model

Extending the gravity model by adding gender-sensitive variables yields an overall better model fit as can be judged by the R^2 values in the output tables in the Annex 6. Also, the majority of the MAPE values have improved slightly. The values in Figure 4.2 show a clear picture of improvement after adding the gender indicators to the model specification.

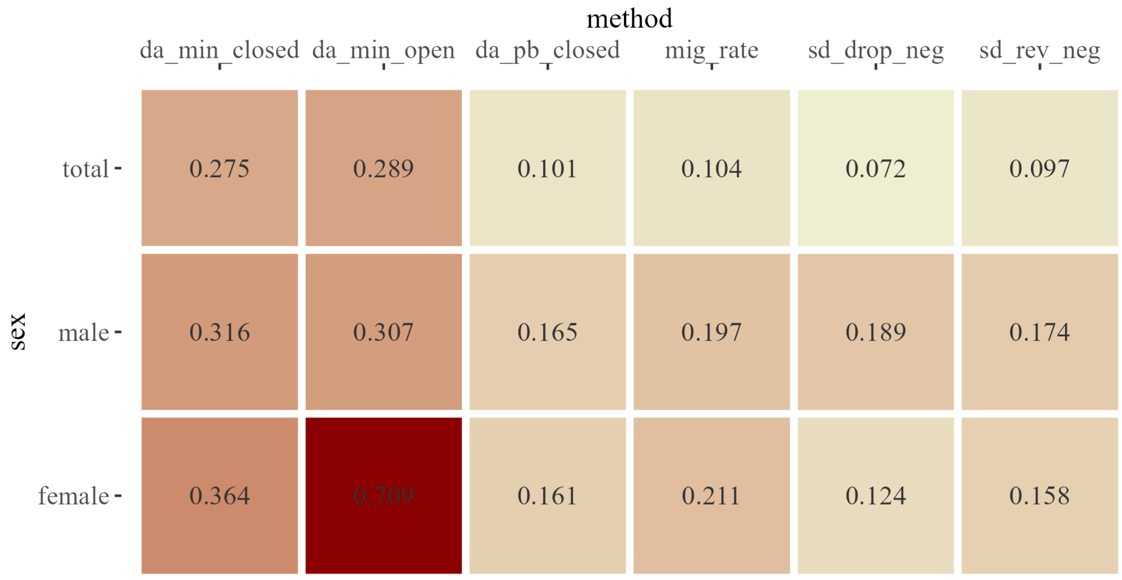

Figure 3: MAPE values for predicting migration flows in the period 2015-2020 with the *extended gravity model* by using all different six estimators separately for male, female, and total flows.

Unfortunately, the gender indicators are largely time-invariant and contain also a lot of missing values. Therefore, the number of observations included to fit the model shrank notably for the extended gravity model. Future steps will address this issue by finding better indicators with better time and regional coverage. However, further investigation is needed to understand whether the gender-sensitive indicators explain more variance regardless of the sex category or if they can improve specifically the predictions of female migration flows.

5 Preliminary Conclusion

The results of the preliminary analysis point to a discrepancy in predictive migration model performance by sex that deserves further exploration and understanding. The next steps include the evaluation of other more advanced models for migration predictions including econometric and demographic methods. For instance, models that acknowledge the count character of the data and do not assume normality like the Poisson-hurdle model or generalized linear models. This will be complemented by a thorough analysis of the potential drivers of the differences.

Knowing that for some regions the estimates produced by G. Abel (2022b) are less certain than for others, we will also assess the performance of the predictions for sub samples of the data by removing migration corridors with high uncertainty in the estimates (Azose & Raftery, 2015). We are planning to validate them against other migration flow data should the results be consistent. Moreover, other limitations of the data include the certainty about the sex classification in each corridor based on the data it was derived from. Such limitations will be addressed by focusing on corridors where data collection is more transparent and reliable than in others.

The goal of this work is to assess potential sex differences in migration predictions of common methods. Advancing our understanding of such differences and their origins can improve models of migration predictions and their outcomes. This will help researchers to understand female- and male-dominated migration corridors better and can inform migration theories. Consequently, our results can have direct implications for international organizations, researchers, and policy makers.

References

Abel, G. (2022a). *Bilateral international migration flow estimates by sex.* figshare. Retrieved from https://figshare.com/collections/Bilateral_international_migration_flow_estimates_by_sex/5800838/6 doi: 10.6084/m9.figshare.c.5800838.v6

Abel, G. (2022b, October). Bilateral international migration flow estimates by sex. Retrieved 2023-08-02, from https://figshare.com/collections/Bilateral_international_migration_flow_estimates_by_sex/5800838 (Publisher: figshare) doi: 10.6084/m9.figshare.c.5800838.v6

Abel, G. J., & Cohen, J. E. (2019, June). Bilateral international migration flow estimates for 200 countries. *Scientific Data*, *6*(1), 82. Retrieved 2021-12-21, from https://www.nature.com/articles/s41597-019-0089-3 (Bandiera_abtest: a Cc_license_type: cc_publicdomain Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Geography;Social sciences Subject_term_id: geography;social-sciences) doi: 10.1038/s41597-019-0089-3

Anastasiadou, A., Kim, J., Sanlitürk, A. E., de Valk, H., & Zagheni, E. (2023). Sex- and gender-based differences in the migration process: a systematic literature review. *MPIDR Working Papers*(WP-2023-039).

Azose, J. J., & Raftery, A. E. (2015, September). Bayesian Probabilistic Projection of International Migration. *Demography*, *52*(5), 1627–1650. Retrieved 2021-12-21, from https://doi.org/10.1007/s13524-015-0415-0 doi: 10.1007/s13524-015-0415-0

Beine, M., & Salomone, S. (2013). Network Effects in International Migration: Education versus Gender. *The Scandinavian Journal of Economics*, *115*(2), 354–380. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9442.2012.01733.x doi: https://doi.org/10.1111/j.1467-9442.2012.01733.x

Boyd, M., & Grieco, E. (2003). Women and Migration: Incorporating Gender into International Migration Theory. , 7.

Brücker, H., & Siliverstovs, B. (2006). Estimating and forecasting European migration: methods, problems and results. *Zeitschrift für ArbeitsmarktForschung - Journal for Labour Market Research*, *39*(1), 35–56. Retrieved 2021-12-21, from https://www.econstor.eu/handle/10419/158623 (Publisher: Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung (IAB))

Conte, M., Cotterlaz, P., & Mayer, T. (2022). *The CEPII Gravity database.* CEPII Working Paper No 2022-05. Retrieved from http://www.cepii.fr/DATA_DOWNLOAD/gravity/doc/Gravity_documentation.pdf

Dennett, A. (2016, May). Estimating an Annual Time Series of Global Migration Flows - An Alternative Methodology for Using Migrant Stock Data. In A. Wilson (Ed.), *Global Dynamics* (pp. 125–142). Chichester, UK: John Wiley & Sons, Ltd. Retrieved 2023-06-16, from https://onlinelibrary.wiley.com/doi/10.1002/9781118937464.ch7 doi: 10.1002/9781118937464.ch7

Hennebry, J., KC, H., & Williams, K. (2021). Gender and migration data: A guide for evidence-based, gender-responsive migration governance.

Kim, K., & Cohen, J. E. (2010, December). Determinants of International Migration Flows to and from Industrialized Countries: A Panel Data Approach beyond Gravity. *International Migration Review*, *44*(4), 899–932. Retrieved 2023-10-24, from https://doi.org/10.1111/j.1747-7379.2010.00830.x (Publisher: SAGE Publications Inc) doi: 10.1111/j.1747-7379.2010.00830.x

Lee, E. (1966). S., 1966,"a theory of migration,". *Demography*, *3*(1), 47–57.

Massey, D. S. (2006). Building a comprehensive model of international migration. *Long-Term Immigration Projection Methods: Current Practice and How to Improve It*, *5*.

Pedraza, S. (1991, August). Women and Migration: The Social Consequences of Gender. *Annual Review of Sociology*, *17*(1), 303–325. Retrieved 2022-07-21, from https://www.annualreviews.org/doi/10.1146/annurev.so.17.080191.001511 doi: 10.1146/annurev.so.17.080191.001511

Pu, Y., Zhao, X., Chi, G., Zhao, J., & Kong, F. (2019, October). A spatial dynamic panel approach to modelling the space-time dynamics of interprovincial migration flows in China. *Demographic Research*, *41*, 913–948. Retrieved 2023-10-19, from https://www.demographic-research.org/articles/volume/41/31

Welch, N. G., & Raftery, A. E. (2022, August). Probabilistic forecasts of international bilateral migration flows. *Proceedings of the National Academy of Sciences*, *119*(35), e2203822119. Retrieved 2022-08-30, from https://www.pnas.org/doi/abs/10.1073/pnas.2203822119 (Publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.2203822119

World Bank. (n.d.). *World development indicators.* Retrieved from https://databank.worldbank.org/source/world-development-indicators

Yildiz, D., & Abel, G. (2021). Migration stocks and flows: data concepts, availability and comparability. In M. McAuliffe (Ed.), *International Migration and Digital Technology* (pp. 29–41). Cheltenham, UK and Northampton, MA USA: Edward Elgar Publishing.

Zizka, A., Silvestro, D., Andermann, T., Azevedo, J., Duarte Ritter, C., Edler, D., ... Antonelli, A. (2019). Coordinatecleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*(10), -7. Retrieved from https://github.com/ropensci/CoordinateCleaner (R package version 3.0) doi: 10.1111/2041-210X.13152

6 Annex

	log(sd_drop_neg)	log(sd_rev_neg)	log(mig_rate)
	(1)	(2)	(3)
log(POPo)	0.591***	0.598***	0.625***
	(0.010)	(0.009)	(0.009)
log(POPd)	0.599***	0.522***	0.582***
	(0.010)	(0.009)	(0.009)
log(dist)	−0.990***	−1.060***	−1.122***
	(0.013)	(0.012)	(0.012)
log(PSRo)	−0.199***	−0.075***	−0.425***
	(0.026)	(0.025)	(0.023)
log(PSRd)	1.081***	0.517***	0.911***
	(0.030)	(0.027)	(0.026)
log(IMRo)	−0.026	−0.104***	−0.099***
	(0.019)	(0.018)	(0.017)
log(IMRd)	−1.248***	−0.935***	−0.974***
	(0.022)	(0.020)	(0.019)
log(URBo)	−0.016	0.065**	0.089***
	(0.031)	(0.029)	(0.027)
log(URBd)	0.227***	0.178***	0.285***
	(0.041)	(0.034)	(0.033)
log(LAo)	−0.042***	−0.017**	−0.031***
	(0.008)	(0.008)	(0.007)
log(LAd)	0.094***	0.098***	0.151***
	(0.008)	(0.008)	(0.007)
LB	1.346***	1.508***	1.562***
	(0.058)	(0.053)	(0.048)
OL	1.574***	1.506***	1.704***
	(0.031)	(0.029)	(0.027)
COL	1.236***	1.383***	1.504***
	(0.066)	(0.064)	(0.057)
t	−0.061***	−0.064***	−0.021***
	(0.003)	(0.003)	(0.003)
t2	−0.001***	−0.002***	0.001***
	(0.0003)	(0.0002)	(0.0002)
Constant	−4.427***	−0.276	−1.449***
	(0.370)	(0.344)	(0.321)
Observations	35,071	42,845	44,886
Adjusted R^2	0.460	0.449	0.517
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 1: OLS regression results for the baseline gravity model for **total** migration flows.

	log(da_min_open)	log(da_min_closed)	log(da_pb_closed)
	(1)	(2)	(3)
log(POPo)	0.728^{***}	0.609^{***}	0.608^{***}
	(0.009)	(0.009)	(0.007)
log(POPd)	0.451^{***}	0.430^{***}	0.557^{***}
	(0.009)	(0.009)	(0.007)
log(dist)	-1.429^{***}	-1.317^{***}	-1.676^{***}
	(0.013)	(0.013)	(0.011)
log(PSRo)	0.235^{***}	0.488^{***}	0.178^{***}
	(0.023)	(0.022)	(0.018)
log(PSRd)	0.080^{***}	0.272^{***}	-0.018
	(0.022)	(0.023)	(0.017)
log(IMRo)	-0.215^{***}	-0.296^{***}	-0.798^{***}
	(0.018)	(0.017)	(0.014)
log(IMRd)	-1.281^{***}	-1.331^{***}	-1.303^{***}
	(0.018)	(0.018)	(0.014)
log(URBo)	0.343^{***}	0.015	-0.015
	(0.027)	(0.027)	(0.022)
log(URBd)	0.452^{***}	0.231^{***}	0.283^{***}
	(0.030)	(0.030)	(0.023)
log(LAo)	0.014^{*}	0.090^{***}	0.186^{***}
	(0.008)	(0.008)	(0.006)
log(LAd)	0.231^{***}	0.254^{***}	0.237^{***}
	(0.007)	(0.007)	(0.006)
LB	1.249^{***}	1.325^{***}	2.110^{***}
	(0.065)	(0.070)	(0.059)
OL	1.810^{***}	1.751^{***}	2.088^{***}
	(0.029)	(0.030)	(0.024)
COL	1.620^{***}	1.781^{***}	2.602^{***}
	(0.078)	(0.083)	(0.071)
t	-0.052^{***}	-0.063^{***}	-0.078^{***}
	(0.003)	(0.003)	(0.002)
t2	0.001^{***}	0.001^{***}	0.001^{***}
	(0.0002)	(0.0002)	(0.0002)
Constant	0.777^{**}	-0.150	8.273^{***}
	(0.343)	(0.344)	(0.278)
Observations	76,383	80,887	117,990
Adjusted R^2	0.493	0.446	0.536
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 2: OLS regression results for the baseline gravity model for **total** migration flows (cont'd).

	log(sd_drop_neg)	log(sd_rev_neg)	log(mig_rate)
	(1)	(2)	(3)
log(POPo)	0.561***	0.555***	0.607***
	(0.010)	(0.009)	(0.008)
log(POPd)	0.541***	0.480***	0.542***
	(0.010)	(0.009)	(0.009)
log(dist)	−0.945***	−0.986***	−1.078***
	(0.014)	(0.013)	(0.011)
log(PSRo)	−0.152***	−0.081***	−0.385***
	(0.026)	(0.025)	(0.023)
log(PSRd)	1.115***	0.667***	0.946***
	(0.029)	(0.027)	(0.026)
log(IMRo)	0.0002	−0.071***	−0.055***
	(0.019)	(0.019)	(0.017)
log(IMRd)	−1.135***	−0.863***	−0.941***
	(0.022)	(0.020)	(0.019)
log(URBo)	0.011	0.082***	0.101***
	(0.031)	(0.029)	(0.027)
log(URBd)	0.251***	0.223***	0.326***
	(0.041)	(0.035)	(0.032)
log(LAo)	−0.045***	−0.010	−0.040***
	(0.008)	(0.008)	(0.007)
log(LAd)	0.119***	0.112***	0.172***
	(0.008)	(0.008)	(0.007)
LB	1.257***	1.427***	1.502***
	(0.058)	(0.053)	(0.047)
OL	1.403***	1.337***	1.573***
	(0.031)	(0.029)	(0.026)
COL	1.271***	1.358***	1.529***
	(0.066)	(0.064)	(0.055)
t	−0.049***	−0.051***	−0.015***
	(0.003)	(0.003)	(0.003)
t2	−0.001**	−0.001***	0.001***
	(0.0003)	(0.0003)	(0.0002)
Constant	−5.524***	−2.181***	−2.798***
	(0.370)	(0.347)	(0.315)
Observations	31,972	38,392	44,132
Adjusted R^2	0.446	0.430	0.505
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3: OLS regression results for the baseline gravity model for **male** migration flows.

	log(da_min_open)	log(da_min_closed)	log(da_pb_closed)
	(1)	(2)	(3)
log(POPo)	0.704***	0.588***	0.576***
	(0.009)	(0.009)	(0.007)
log(POPd)	0.438***	0.421***	0.555***
	(0.009)	(0.009)	(0.007)
log(dist)	−1.371***	−1.242***	−1.658***
	(0.013)	(0.013)	(0.011)
log(PSRo)	0.289***	0.497***	0.164***
	(0.024)	(0.023)	(0.018)
log(PSRd)	0.211***	0.421***	0.088***
	(0.023)	(0.024)	(0.017)
log(IMRo)	−0.222***	−0.268***	−0.761***
	(0.018)	(0.018)	(0.014)
log(IMRd)	−1.205***	−1.222***	−1.278***
	(0.018)	(0.019)	(0.014)
log(URBo)	0.299***	0.007	−0.028
	(0.028)	(0.029)	(0.023)
log(URBd)	0.489***	0.308***	0.312***
	(0.031)	(0.032)	(0.023)
log(LAo)	0.032***	0.100***	0.198***
	(0.008)	(0.008)	(0.006)
log(LAd)	0.228***	0.244***	0.226***
	(0.007)	(0.008)	(0.006)
LB	1.150***	1.191***	2.025***
	(0.065)	(0.071)	(0.059)
OL	1.750***	1.659***	2.025***
	(0.030)	(0.031)	(0.024)
COL	1.514***	1.697***	2.561***
	(0.078)	(0.085)	(0.071)
t	−0.051***	−0.054***	−0.073***
	(0.003)	(0.003)	(0.002)
t2	0.001***	0.001***	0.001***
	(0.0003)	(0.0003)	(0.0002)
Constant	−1.564***	−2.945***	7.033***
	(0.351)	(0.360)	(0.279)
Observations	68,526	71,565	115,280
Adjusted R^2	0.481	0.430	0.522

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4: OLS regression results for the baseline gravity model for **male** migration flows (cont'd).

	log(sd_drop_neg)	log(sd_rev_neg)	log(mig_rate)
	(1)	(2)	(3)
log(POPo)	0.530***	0.540***	0.560***
	(0.010)	(0.009)	(0.009)
log(POPd)	0.574***	0.492***	0.551***
	(0.010)	(0.010)	(0.009)
log(dist)	−0.918***	−0.970***	−1.047***
	(0.014)	(0.013)	(0.012)
log(PSRo)	−0.199***	−0.094***	−0.465***
	(0.027)	(0.026)	(0.024)
log(PSRd)	0.831***	0.365***	0.711***
	(0.030)	(0.027)	(0.027)
log(IMRo)	−0.034*	−0.078***	−0.118***
	(0.020)	(0.019)	(0.017)
log(IMRd)	−1.118***	−0.864***	−0.906***
	(0.022)	(0.020)	(0.019)
log(URBo)	−0.044	0.101***	0.051*
	(0.031)	(0.030)	(0.028)
log(URBd)	0.250***	0.154***	0.324***
	(0.042)	(0.036)	(0.034)
log(LAo)	−0.019**	−0.004	−0.013*
	(0.008)	(0.008)	(0.007)
log(LAd)	0.116***	0.116***	0.151***
	(0.008)	(0.008)	(0.007)
LB	1.411***	1.577***	1.706***
	(0.058)	(0.053)	(0.048)
OL	1.525***	1.433***	1.697***
	(0.031)	(0.029)	(0.027)
COL	1.178***	1.320***	1.462***
	(0.065)	(0.064)	(0.057)
t	−0.060***	−0.065***	−0.024***
	(0.003)	(0.003)	(0.003)
t2	−0.001***	−0.002***	0.001***
	(0.0003)	(0.0003)	(0.0002)
Constant	−3.794***	−0.256	−0.563*
	(0.374)	(0.351)	(0.325)
Observations	31,423	37,933	43,658
Adjusted R^2	0.451	0.435	0.490
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 5: OLS regression results for the baseline gravity model for **female** migration flows.

	log(da_min_open)	log(da_min_closed)	log(da_pb_closed)
	(1)	(2)	(3)
log(POPo)	0.698***	0.626***	0.612***
	(0.009)	(0.009)	(0.007)
log(POPd)	0.482***	0.409***	0.517***
	(0.009)	(0.009)	(0.007)
log(dist)	−1.357***	−1.277***	−1.610***
	(0.013)	(0.013)	(0.011)
log(PSRo)	0.176***	0.430***	0.119***
	(0.024)	(0.022)	(0.018)
log(PSRd)	0.008	0.079***	−0.277***
	(0.023)	(0.025)	(0.018)
log(IMRo)	−0.183***	−0.240***	−0.795***
	(0.018)	(0.018)	(0.014)
log(IMRd)	−1.228***	−1.227***	−1.214***
	(0.018)	(0.019)	(0.014)
log(URBo)	0.382***	0.131***	−0.031
	(0.028)	(0.029)	(0.023)
log(URBd)	0.438***	0.124***	0.191***
	(0.031)	(0.032)	(0.023)
log(LAo)	0.003	0.049***	0.168***
	(0.008)	(0.008)	(0.006)
log(LAd)	0.200***	0.262***	0.250***
	(0.008)	(0.008)	(0.006)
LB	1.248***	1.280***	2.209***
	(0.066)	(0.071)	(0.058)
OL	1.779***	1.789***	2.095***
	(0.030)	(0.031)	(0.024)
COL	1.532***	1.708***	2.568***
	(0.078)	(0.084)	(0.070)
t	−0.043***	−0.058***	−0.078***
	(0.003)	(0.003)	(0.002)
t2	0.002***	0.001***	0.001***
	(0.0003)	(0.0003)	(0.0002)
Constant	0.167	0.338	9.479***
	(0.353)	(0.359)	(0.281)
Observations	67,835	70,268	111,796
Adjusted R^2	0.486	0.444	0.527

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6: OLS regression results for the baseline gravity model for **female** migration flows (cont'd).