INITIATIVE ON
Fragility, Conflict, and Migration

Understanding Migration within Countries

A Global Perspective

Jean-François Maystadt, Silvia Peracchi, Ella Sargsyan, Liangzhi You

Motivation

The majority of migrants remain within their own borders. While 281 million people have migrated in another country in 2020, the number of internal migrants is estimated to stand at around 763 million (UN DESA, 2016; IOM, 2021). At the global level, the existing literature has mainly focused on understanding the drivers of international migration (Karemera et al., 2000; Mayda, 2010; Kim and Cohen, 2010). Understanding the dynamics of internal migration is key for conjecturing future prospects in terms of poverty, food insecurity, urbanization, within-country inequalities and geopolitical instability.

We contribute to the literature focusing on the movements of people within countries. Existing studies have either focused on one single country or one specific channel (e.g. Cattaneo et al. (2019) and Hoffmann et al. (2021) for a review on climatic shocks; Crippa et al. (2024) on conflict). In this brief, we offer a comprehensive analysis of the drivers of internal migration. We first build a large dataset on bilateral migration flows within 50 countries during the period 1992-2014. Guided by the literature, we then implement a gravity model, a standard approach in the economics of migration (Beine et al., 2016). With a focus on drought and aridity, Hoffmann et al. (2024) follow a similar approach. Our analysis reveals that socio-economic, geographic, environmental, and political factors significantly shape migration patterns. Socio-economic factors, measured by population, nightlight intensity, crop price shocks, and health indicators, influence both outmigration and in-migration, with better conditions attracting migrants. Geographic proximity and access to major cities facilitate migration, while environmental factors like droughts, high temperatures, and precipitation levels at both origins and destinations play a crucial role. Political instability, reflected by conflict fatalities, strongly drives outmigration but shows limited influence at destinations. We then move to a Machine Learning (ML) approach more likely to deal with a high-dimensional problem. We compare the predictive performance of traditional Poisson Pseudo-Maximum Likelihood (PPML) regression with two versions of Random Forest and Gradient Boosted Trees models, in forecasting internal migration flows. Gradient Boosted Trees emerge as the best-performing model, revealing that socio-economic, geographic, and climatic factors are the most critical predictors, while conflict plays a limited role except in specific regions like Africa. Further analysis highlights demographic and regional heterogeneity: lower-educated and younger migrants are more influenced by climatic factors, while geographic factors are more significant for older individuals and urban destinations. Regional variations show socio-economic factors dominate in Asia, while geographic and socio-economic factors are equally crucial in the Americas.

Overall, our results highlight the importance of economic shocks and migration costs for policy recommendations. We also conclude with the ability to pursue our exercise to better qualify the consequences of internal migration in low- and middle-income countries.

Understanding Internal Migration with a Standard Gravity Model

We first construct bilateral migration flows within 50 countries between 1992 and 2014 (see Box 1). We then apply a standard gravity model at the subnational level (GADM1). There is a long tradition in economics to rely on the gravity model to understand migration decisions, with impressive predictive power (Beine and Parsons, 2015). The gravity model has long constituted the main empirical tool to identify “push” and “pull” factors in the decision to migrate. Based on the literature, we group these factors into four main categories: socio- economic, geographic, environmental and political factors.

First, **economic factors** should encompass economic incentives to move from and to both rural and urban areas. To start with, one of the most robust predictions of the gravity model is that higher mobility of factors (including labor) is expected between large economies. The standard ingredient of the gravity model is therefore the size of population at origin and destination. We then proxy for economic activities using nightlight densities defined at the regional level. Given the urban bias of such a proxy (Chen and Nordhaus, 2011; Keola et al., 2015), we then complement these economic factors by three additional indicators: infant mortality, the percentage of malnourished children under 5-years-old, and a measure of the main crop price shock (details in box 1).[1]

Second, **geographical factors** aim at capturing the costs associated with migration. Migration costs are often approximated by the distance and the contiguity between origin and destination.[2] Specific to internal migration, the distance to the international border can also provide a proxy for the costs associated with moving to the neighboring country. Within countries, proximity to the major city remains an attraction force, while the area of the subnational region matters in proxying for within-region migration costs.[3]

Third, the **environmental factors** are chosen based on a recent but fast-expanding literature that has established a relationship between climatic shocks and migration, both between countries (Marchiori et al., 2012; Beine et al., 2016) and within countries (Gray and Mueller, 2012; Mueller et al., 2014, 2020;

---

[1] The migration literature has also tested the so-called social magnet effect by incorporating the access to transfers, public goods, or social public spending. Within countries, we should not expect much variation, or at least, limited by the introduction of nightlight, malnutrition and infant mortality.

[2] Common language or common colonial history are usually included in gravity model for international migration (Beine et al., 2011; Grogger and Hanson, 2011). Within countries, it should matter less.

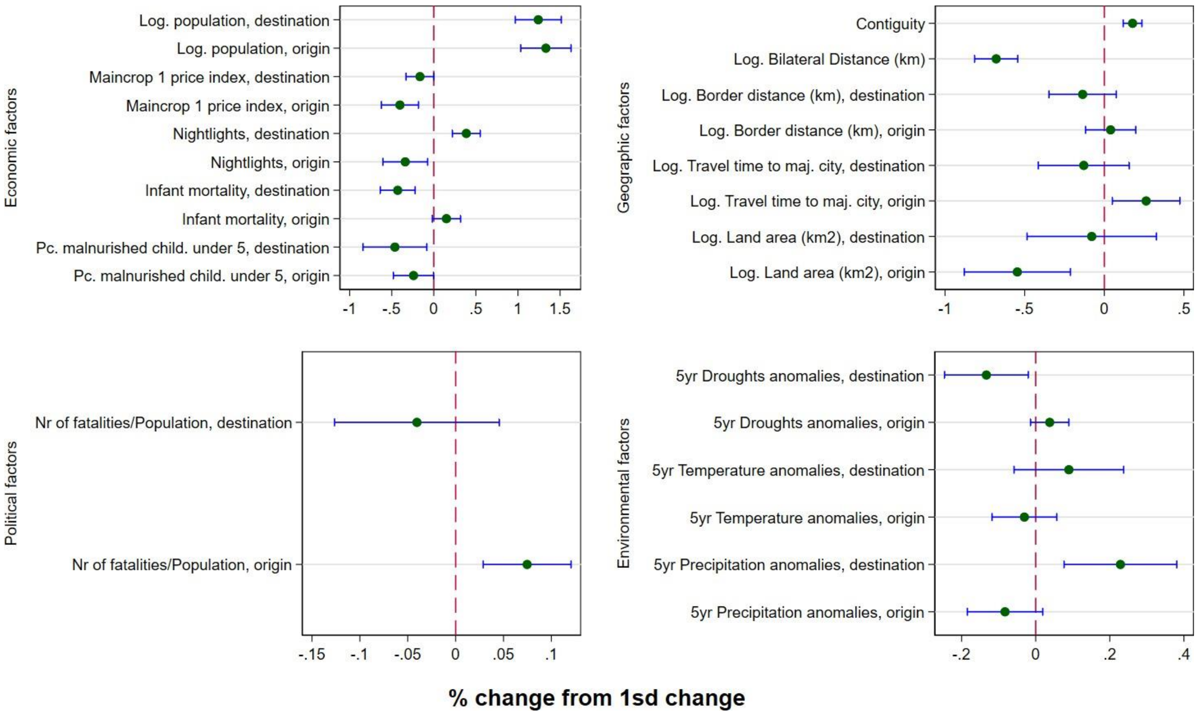
[3] We should acknowledge our grouping is far from perfect, since for instance, the distance to the capital can be seen as a proxy for access to the central political power.

Chen and Mueller, 2018; Liu et al., 2023; Quinones et al., 2023). Reviewing the recent literature, Maystadt et al. (2024) argue that following a climatic shock, migration constraints may be more important in Africa than in Latin America or South Asia.[4] Slow-onset events like droughts are also much more likely to affect migration than rapid-onset events like floods (Maystadt et al., 2024). We will therefore include in our gravity model not only precipitation and temperature anomalies but also the severity of droughts.

Fourth, **political factors** are mainly approached by the risk of organized violence (Bohra-Mishra and Massey, 2011; Melander and Oberg, 2007; Adhikari, 2013). We therefore add the number of fatalities per inhabitant among our explanatory variables.[5]

All variables are further described in Box 1. We estimate the gravity model using the Poisson Pseudo-Maximum Likelihood (PPML) estimator proposed by Silva and Tenreyro (2006). The PPML estimator has become the standard method to deal with abundant zeroes in the migration flows and the resulting heteroscedastic error terms. Estimated coefficients can be interpreted as in a log-linear model, but we present our main results for standardized coefficients in Figure 1 to make effect sizes comparable.

Figure 1: Standardized coefficients from PPML regression



Note: This figure illustrates the standardized coefficient estimates along with their 95% confidence intervals, obtained from the Poisson Pseudo-Maximum Likelihood (PPML) estimator. The specification includes country and year fixed effects, as well as a vector of land use control variables.

---

[4] It is difficult to cover all papers on environmental migration, but comprehensive reviews have been offered by Cattaneo et al. (2019), Hoffmann et al. (2021) and Millock (2015), to name a few.

[5] We further augment our specification with land use characteristics, but those variables have little explanatory power. We will, nonetheless, present this category in our ML exercise.

Results

Socio-economic factors: Consistent with existing literature, our PPML regression results indicate high mobility between large economies. This is reflected in the positive and statistically significant coefficients for population size at both origin and destination locations. Economic activity at the origin, measured by nightlight intensity (to capture activity in urban areas) and main crop price shocks (to represent rural areas), shows a negative correlation with migrant outflows, as expected. At the destination, the same measures reveal a significantly positive correlation for nightlight intensity, indicating stronger economic activity attracts migrants, while main crop price shocks exhibit a negative but moderately significant correlation, suggesting higher rural economic volatility discourages in-migration. This latter finding may reflect the impact of higher living costs driven by elevated food prices, coupled with the challenges of establishing new agricultural activities in an unfamiliar location to take advantage of these higher prices. Other measures of economic activity, such as infant mortality and the percentage of malnourished children under five years old, act as deterrents for migration to a destination. However, these factors have differing effects at the origin. High infant mortality is positively associated with outmigration, likely reflecting poor living conditions that push individuals to seek better opportunities elsewhere. In contrast, a higher percentage of malnourished children at the origin shows a negative correlation with outmigration, potentially indicating that lower-income households and those in poor health have fewer resources and opportunities to migrate.

Geographical factors: Among the geographic factors, our results identify contiguity between subnational units and bilateral distance as the most significant determinants of migration decisions. Consistent with the literature, contiguity increases migration flows between subnational units by facilitating movement, while distance, serving as a proxy for migration costs, exhibits a negative correlation with outmigration. Distance to the international border and travel time to a major city do not show significant results, except for travel time to a major city at the origin, which exhibits a significant positive correlation with outmigration. This suggests that people are more likely to leave areas with poor access to major cities. Finally, land area at origin shows a significant negative correlation with outmigration, indicating a preference for remaining in larger agglomeration units.

Environmental factors: To account for the severity of climatic shocks, we include both linear and quadratic terms for droughts, temperature, and precipitation measures in our specification. Our results indicate that droughts at the destination significantly influence migration decisions. High temperatures at the origin are a strong driver of outmigration. For precipitation, the linear term at the destination is positive, suggesting that higher levels of precipitation attract people. However, the negative coefficient on the quadratic term indicates that this effect diminishes beyond a certain threshold, after which excessive precipitation may act as a deterrent, driving people away.

Political factors: Our results show that a higher number of fatalities per inhabitant plays a crucial role in the decision to out-migrate at the origin. Concurrently, the correlation at the destination is negative, though it is not statistically significant.

As mentioned earlier, we also control for land use variables in our specification; however, we omit the interpretation of these variables to conserve space. Additionally, we examine the heterogeneity of internal migration by demographic characteristics of migrants and by continent. In Section 3, we provide graphical representations of group importance of different factors and discuss the observed group differences from our ML approach. Further details and all PPML tables are available upon request.

Box 1

Our final dataset consists of 5-year bilateral migration flows and their determinants within 50 countries during the period 1992–2014. Below, we detail the data sources and construction process.

The bilateral migration flows were derived using census microdata from the IPUMS-International (IPUMSI) database. Similar to Sorichetta et al. (2016), we utilized the harmonized variable *GEOMIG*1 5, which identifies the major-administrative unit of residence five years prior to the survey, focusing exclusively on internal migration. The 5-year migration question was selected for its broader coverage compared to the 1- and 10-year alternatives. Microdata were aggregated to compute gender-, age-, education-, and urbanity-specific 5-year migration flows from subnational unit *i* to subnational unit *j* at time *t*, using the following formula:

$$nmig_{ijgt} = \sum_{n}^{N_{ijgt}} PERWT_{ngt} \qquad (1)$$

where $PERWT_{ngt}$ is the personal weight representing the number of individuals each respondent represents. To improve on country coverage, we constructed an additional variable, *nmig_est*, by leveraging two alternative variables such as *GEOMIG*1_5 (previous administrative unit of residence) and *MIGRY S*1 (years residing in the current locality). A migration flow was recorded if *MIGRY S*1 ≤ 5.

We sourced most socio-economic, geographic, and environmental variables from the PRIO-GRID dataset, with the exception of main crop price shocks, bilateral distances, and contiguity measures. The PRIO-GRID dataset provides grid-level static and yearly variables, which we aggregated to the subnational level and computed as 5-year averages for the yearly data.

To calculate main crop price shocks, we used the international producer price index for crops from the FAOSTAT Agricultural Producer Price domain, combined with data on the main crop and harvest area from PRIO-GRID. For each subnational unit, we identified the main crop with the largest harvest area and matched it to the 5-year average price index for that crop and survey year. The bilateral distances and contiguity measures were calculated using GIS software, based on shapefiles of subnational units obtained from the IPUMSI database.

Political factors, specifically the number of fatalities per inhabitant, were derived from the Uppsala Conflict Data Program (UCDP). The UCDP provides georeferenced, event-level data on conflicts and fatalities worldwide. We aggregated the number of fatalities for each subnational unit and survey year, calculating a cumulative total for the preceding five years, which we then divided by the population of each subnational unit.

We use the below specification for all our PPML analyses:

$$Migration_{ij|t,c} = \alpha_t \alpha_c (Econ_{i|j,t,c})^{\beta_1} (Geo_{i|j,t,c})^{\beta_2} (Env_{i|j,t,c})^{\beta_3} (Pol_{i|j,t,c})^{\beta_4} (Land_{i|j,t,c})^{\beta_5} \epsilon_{ij|t,c} \qquad (2)$$

Extending the Gravity Model with Machine Learning approaches

With the growing popularity of ML methods, researchers interested in modeling bilateral outcomes have been exploring the potentials of using these tools. In particular, migration scholars have employed these methods in tasks that involve predicting (Sorichetta et al., 2016; Simini et al., 2021) or forecasting (Boss et al., 2023; Carammia et al., 2022; Ruzicska et al., 2024; Gu et al., 2024) different types of migration, or making inference on the effects of a rich set of attributes (Aoga et al., 2024; Micevska, 2021) on migration-related outcomes. These outcomes range from cross-country asylum-seeker flows like in Boss et al. (2023); Carammia et al. (2022); Micevska (2021) to city-based commuting (Simini et al., 2021; Gu et al., 2024), to migration intentions (Aoga et al., 2024).

We refine our analysis by comparing the predictive performance of the standard Poisson Pseudo-Maximum Likelihood (PPML) estimator with two tree-based models that are widely applied in supervised ML tasks: Random Forest regression models and Gradient Boosted trees. We evaluate how different models perform in predicting flows with unseen data. On the one side, PPML is the golden standard method for statistical inference on bilateral migration. On the other side, in applications that involve forecasts or imputing data where actual migration is not observed, ML models could potentially be a preferable alternative, as they are specifically designed to predict well on unobserved data. In this application, we investigate how these types of models compare for the prediction of internal migration flows within countries. To do so, we hold out 10% of the total sample and define it as our test set. The remaining observations constitute our training set and are used to estimate the parameters. To look at performance within countries, we stratify this sampling so that all countries appear both in the training and test set. Among the potential advantages, the ML-based models under consideration can flexibly handle non-linear relationships between the regressors and the outcome and can incorporate observations with missing values for some variables. As explained in Box 2 and in Table 1, we observe that Gradient Boosted trees constitute the best-performing models. Therefore, we use this estimator to study the contribution of different variables' dimensions in predicting internal migration.

Box 2

The proposed ML methods are based on regression trees. Regression trees rely on the process of splitting data into different partitions, starting from a root node and adding child nodes at each partition. Splits are performed based on the predictors X and cutoff points established for their values. At each node, cutoffs are chosen to minimize the model's loss (i.e., a function of the difference between the prediction and the actual outcome values). While splitting data can potentially progress until all observations belong to individual nodes, this process is usually stopped before reaching that stage, trees that are too deep might cause overfitting. An overfit model perfectly predicts all variation in the training sample, absorbing most or all of the data noise. This spikes the variance of predictions on unseen observations, that get very inaccurate. A limitation of regression trees is their high sensitivity to small changes within the training dataset. To overcome this instability, ensemble methods exist to combine results from several regression trees. Among these, bagging methods aggregate regression trees by bootstrapping samples and averaging across results.

Random forests (RF) constitute a variation of bagging that, on top of bootstrapping, also limits the set of *features* (i.e., explanatory variables) used to build the individual trees. For every bootstrapped sample, random forests use a random subset of regressors. Because of this, not all variables and observations are used in each tree, which makes the model able to use observations with missing data in trees that do not use variables with missing values. Modeling requires choosing *hyperparameters* such as the maximum depth of trees, the number of trees, and the number of features used in each tree. As standard in ML applications, the choice of these parameters is based on cross-validation of the training set.

Gradient boosting (XGBoost) is a different approach to bagging that instead combines information from previous trees sequentially (hence *boosting).* The method involves fitting trees iteratively, over the residuals of the previous step. Each new tree is made to decrease the loss function of the existing (ensemble of) trees. This way, the model *learns* and residuals gradually diminish at each iteration. We can then study the derivative of the loss function with respect to the number of trees (hence the name gradient) and stop when adding a tree does not marginally improve the model. The learning rate is a *hyperparameter* that weights the contribution of each new tree to the ensemble. If larger, the model converges fast but is less precise. Other important *hyperparameters* involve the maximum depth of trees, the number of trees, the ratio of features used in each tree, the ratio of the total sample to use in each tree, and the minimum sample size in a node.

We evaluate the models in Table 1. We report performance both in sample (*Train* columns) and out of sample (*Test* columns), but we focus on the test set to select the best model. We report three evaluation statistics common to regressions in ML contexts: the root mean squared error (RMSE), the R^2, and the mean absolute error (MAE). For the ML methods, we also include an alternative model that is rerun by only including a subset of the most important variables for the model outcome. We call this version *trimmed* in the table. We find XGBoost to be the best model in terms of all performance metrics, as it has the largest test R^2, and the lowest RMSE and MAE. Therefore, we select this model to study the importance of the different variable dimensions in predicting internal migration flows.

Table 1: Full baseline sample, out of sample performance

Model	RMSE		R2		MAE		N	
	Train	Test	Train	Test	Train	Test	Train	Test
PPML	19531.24	12628.274	0.396	0.64	2225.203	2268.296	44619	4926
RF	6841.253	13702.045	0.926	0.576	692.855	1921.282	44619	4926
RF, trimmed	6762.42	13390.594	0.928	0.595	686.247	1887.827	44619	4926
XGBoost	1799.607	9847.761	0.995	0.781	209.322	1543.129	44619	4926
XGBoost; trimmed	2765.987	10609.663	0.988	0.746	219.127	1571.396	44619	4926

Note: Hyperparameters for ML models are cross-validated using 5 folds of training sample. In the cross validation, we select the best hyperparameters combination based on the best cross-validated MAE scores.

In Figure 2, we quantify to what degree different dimensions contribute to the prediction of internal migration flows. To do so, we study the importance of each group of variables, where groups are Conflict, Climatic factors, Socio-economic Factors, Geographic factors, Land Use factors, Country Fixed Effects and Year Fixed Effects. We use the grouped permutation feature importance proposed in Au et al. (2022). Permutation feature importance is a post-estimation measure of importance based on reshuffling a regressor and calculating how the prediction performance of the model decreases from this permutation. The group-level alternative reshuffles a group of features jointly and therefore it maintains the variables' relationships unchanged within the group. The intuition is that if a group of variables is not useful in predicting an outcome, then randomly permuting the values of its variables will not affect the model's performance. From Figure 2, we observe overall that socio-economic, geographic, and climatic variables are the most important factors, followed by land use, in predicting internal migration flows. Our results certainly tone down the role of violent conflict in determining internal mobility within countries. Internally displaced people stay a vulnerable fringe of the population. However, internal migration remains largely an economic decision, magnified by climatic shocks and variations.

Next, we analyze whether these aggregate results hide some heterogeneity by demographic characteristics of internal migrants and by continent. First, in Figure 3, we consider if the importance of variables' groups varies by gender, educational attainment, age, and urban or rural status of the destination community where migrants reside. The top of Figure 3 plots gender-specific group importance. We do not see a visible difference between males (plot a) and females (plot b). We observe some differences by educational attainment. Climatic variables affect lower-educated internal migrants (plot c) more than higher-educated ones (plot d). In contrast, geographic variables appear to matter slightly less for low-educated migrants, though the intensity of the last pattern depends on the model's performance indicator. Climatic variables also appear to be more crucial in explaining the flows of younger people (e and f) than for the older ages (g). Older individuals appear to be more sensitive than the young to geographic characteristics. As geographic characteristics include bilateral distance and contiguity, these patterns might suggest that younger and less educated people are more willing to relocate further away. Finally, we observe that individuals migrating to urban destinations are more sensitive to geographic factors and less to climatic shocks than those relocating to rural areas. In all subsets of migrants, socio-economic factors are the most relevant drivers of internal migration. Conflict variables do not bear a strong importance, though this substantially varies by continent. Figure 4 shows that in Africa, conflict factors bear a substantially greater weight than in the other regions. Land-use variables also appear more relevant than elsewhere. Climatic factors are less important for internal migration in Africa, while they are for Asia. However, in Asia, socio-economic factors appear to be the strongest drivers of migration flows. This could be driven by the presence of countries with large populations like China and India, as population at origin and destination are a strong socio-economic determinant of bilateral flows. Finally, for the Americas, geographic and socio-economic variables are equally strong determinants of internal migration flows, while other dimensions are more negligible.

Figure 2: Grouped importance of the variables

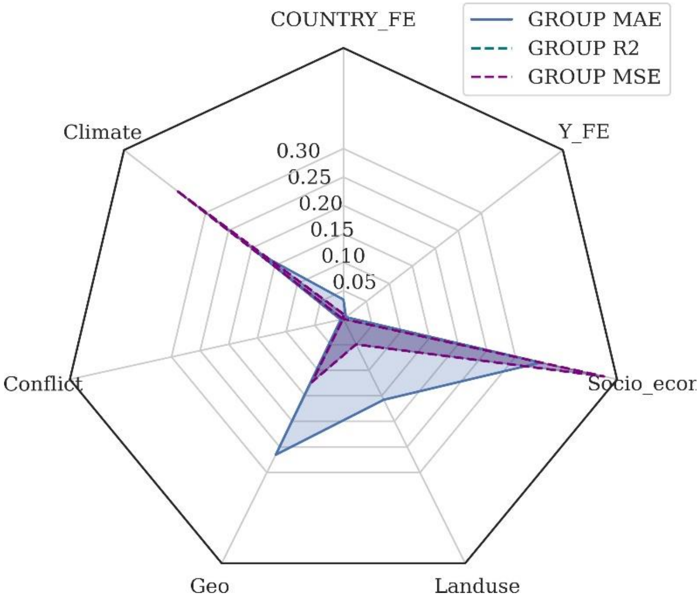
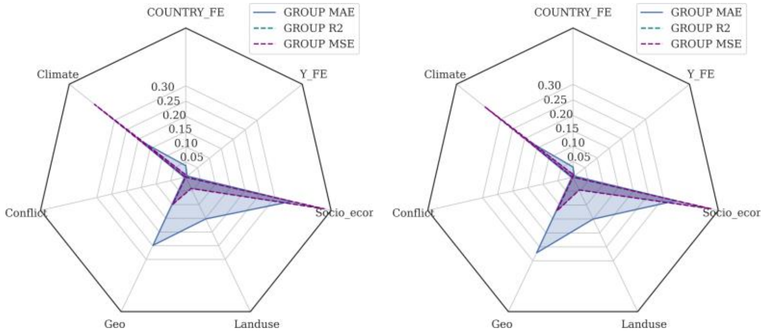
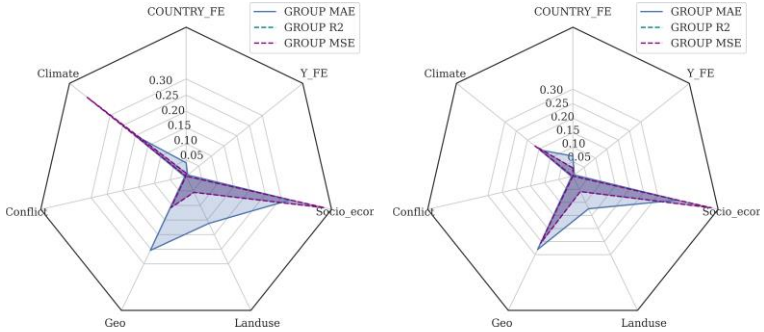


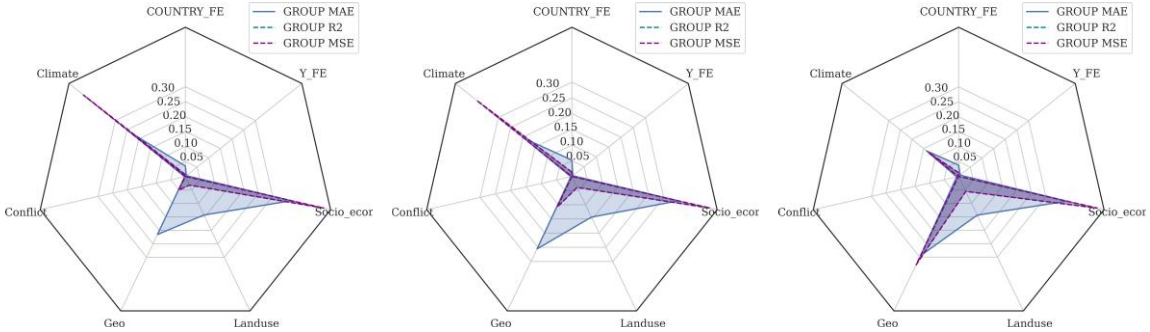
Figure 3: Grouped importance by migration type



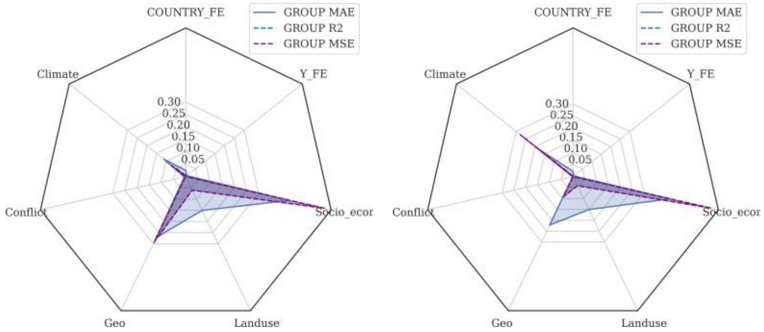
(a) Males (b) Females



(c) Lower Education (d) Higher Education

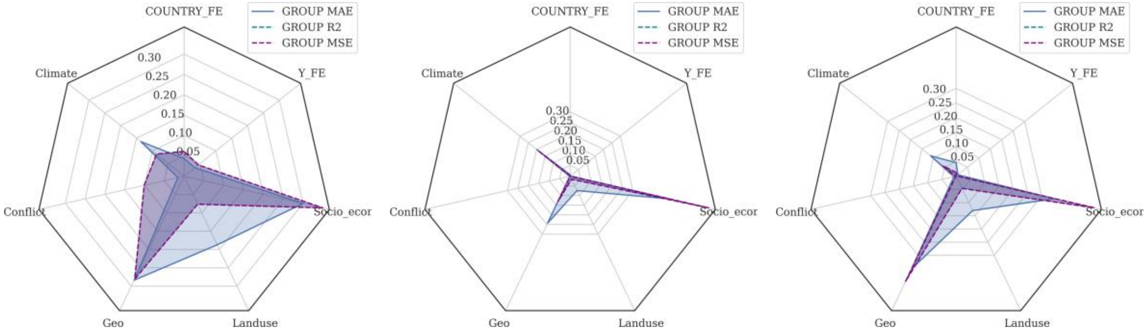


(e) Younger (f) Mid (g) Older



(h) Urban (i) Rural

Figure 4: Grouped importance by continent



(a) Africa (b) Asia (c) Americas

Policy and Research Challenges

Forced displacement has certainly been on the rise over the last 2 decades (UNHCR, 2022). However, our analysis reminds us that internal migration remains primarily motivated by economic motives. Migration constitutes an important coping strategy to deal with economic shocks and a central pathway to improve welfare in low- and middle-income countries. Conflict and violence play an important role as a push factor for migration but that needs to be put in perspective. Even in Africa, where conflict plays the strongest role compared to Asia and the Americas, socio-economic factors remain the main predictors of internal migration. Similarly, despite the strong attention given to climate migration over the last decade, climatic factors explain much less variation in internal migration than geographic and socio- economic factors. Of course, it does not mean that climate change does not matter. We show that climatic factors play a stronger role for the young and relatively low-educated population. That fact alone should increase the pressure on urbanization in the years to come. We do not find strong differences in the drivers of migration between males and females. From a policy point of view, the selectivity and partial mobility as a response to climate change and the associated weather shocks may constitute a source of concern.

From a research point of view, we see the proposed analysis as a first step towards the use of “new analytical approaches such as machine learning and unconventional data sources” (Hernandez et al., 2023, 78) to better understand the drivers and consequences of internal migration. On the drivers, we should first acknowledge that ML techniques are promising to detect general patterns but not necessarily to infer causal inference. In that respect, our results showing the negative effect of price shocks both at origin and destination call for further investigation in distinguishing the income effect from the uncertainty effect that may result from price fluctuations. That distinction is also key to better understand the role of social protection programs (Gazeaud et al., 2023). The importance of price shocks also calls for investigating the role of local and global supply chains to enhance resilience capacities in low- and middle-income countries. Not surprisingly, we also find that distance as a proxy of migration costs remains one key determinant of internal migration. We still miss strong evidence on how insecurity affects such migration costs and leads to sub-optimal migration choice or leads to costly “forced immobility” (Hernandez et al., 2023).

On the consequences, the existing studies either focus on high-income countries (Dustmann et al., 2016; Edo et al., 2020) or on forcibly displaced people in specific countries (Verme and Schuettler, 2021; Maystadt et al., 2019, 2024). Our census-based dataset at the global level paves the way for a new approach to assess the consequences of internal migration across different regions. Notably, migrants have often been blamed to spur political instability, without much solid evidence (Mach et al., 2020; Xie et al., 2024). For forcibly displaced people crossing borders, the link has been dismissed (Zhou and Shaver, 2021) or at least qualified (Coniglio et al., 2023; Bertinelli et al., 2025). More research is needed to identify causal pathways between migration (including climate migration) and conflict, as well as the context factors under which these pathways may materialize.

ABOUT THE AUTHORS

Jean-François Maystadt (jean-francois.maystadt@uclouvain.be) is Research Associate & Professor at Fonds de la Recherche Scientifique - FNRS & UCLouvain, Belgium. He is also Senior Lecturer at Lancaster University, UK.

Silvia Peracchi (silvia.peracchi@uclouvain.be) and Ella Sargsyan (ella.sargsyan@uclouvain.be) are both Postdoc Researchers at UCLouvain, Belgium.

Liangzhi You (l.you@cgiar.org) is a Senior Research Fellow in IFPRI's Foresight and Policy Modeling Unit, Washington, DC.

REFERENCES

Adhikari, P. (2013). Conflict-induced displacement, understanding the causes of flight. American Journal of Political Science 57 (1), 82–89.

Aoga, J. O. R., J. Bae, S. Veljanoska, S. Nijssen, and P. Schaus (2024, January). Impact of Weather Factors on Migration Intention Using Machine Learning Algorithms. Operations Research Forum 5 (1), 8.

Au, Q., J. Herbinger, C. Stachl, B. Bischl, and G. Casalicchio (2022, July). Grouped feature importance and combined features effect plot. Data Mining and Knowledge Discovery 36 (4), 1401–1450.

Beine, M., S. Bertoli, and J. F.-H. Moraga (2016). A practitioners' guide to gravity models of international migration. The World Economy 39 (4), 496–512.

Beine, M., F. Docquier, and C. Ozden (2011). Diasporas. Journal of Development Eco- nomics 95, 30–41.

Beine, M. and C. Parsons (2015). Climatic factors as determinants of international migration.

Scandinavian Journal of Economics 117 (2), 723–767.

Bertinelli, L., R. C¨omertpay, and J.-F. Maystadt (2025). Ethnic diversity and conflict in sub-saharan Africa: Evidence from refugee-hosting areas. Journal of Development Eco- nomics 172, 103393.

Bohra-Mishra, P. and D. S. Massey (2011, 05). Individual decisions to migrate during civil conflict. Demography 48 (2), 401–424.

Boss, K., A. Groeger, T. Heidland, F. Krueger, and C. Zheng (2023, March). Forecasting Bilateral Refugee Flows with High-dimensional Data and Machine Learning Techniques. Working Papers. Number: 1387 Publisher: Barcelona School of Economics.

Carammia, M., S. M. Iacus, and T. Wilkin (2022, January). Forecasting asylum-related migration flows with machine learning and data at scale. Scientific Reports 12 (1), 1457. Publisher: Nature Publishing Group.

Cattaneo, C., M. Beine, C. J. Fr¨ohlich, D. Kniveton, I. Martinez-Zarzoso, M. Mastrorillo,

K. Millock, E. Piguet, and B. Schraven (2019). Human migration in the era of climate change. Review of Environmental Economics and Policy 13 (2), 189–206.

Chen, J. and V. Mueller (2018). Coastal climate change, soil salinity and human migration in Bangladesh. Nature Climate Change 8 (11), 981–985.

Chen, X. and W. D. Nordhaus (2011). Using luminosity data as a proxy for economic statistics. Proceedings of the National Academy of Sciences 108 (21), 8589–8594.

Coniglio, N. D., V. Peragine, and D. Vurchio (2023). The effects of refugees' camps on hosting areas: Social conflicts and economic growth. World Development 168, 106273.

Crippa, A., G. d'Agostino, J. P. Dunne, and L. Pieroni (2024, 09). Conflict as a cause of migration. Oxford Economic Papers, gpae037.

Dustmann, C., U. Sch¨onberg, and J. Stuhler (2016, November). The impact of immigration: Why do studies reach such different results? Journal of Economic Perspectives 30 (4), 31–56.

Edo, A., L. Ragot, H. Rapoport, S. Sardoschau, A. Steinmayr, and A. Sweetman (2020, November). An introduction to the economics of immigration in OECD countries. Canadian Journal of Economics 53 (4), 1365–1403.

Gazeaud, J., E. Mvukiyehe, and O. Sterck (2023, 01). Cash transfers and migration: Theory and evidence from a randomized controlled trial. The Review of Economics and Statistics 105 (1), 143–157.

Gray, C. and V. Mueller (2012). Drought and population mobility in rural Ethiopia. World Development 40 (1), 134–145.

Grogger, J. and G. H. Hanson (2011, May). Income maximization and the selection and sorting of international migrants. Journal of Development Economics 95 (1), 42–57.

Gu, X., X. Tang, T. Chen, and X. Liu (2024, February). Predicting the network shift of large urban agglomerations in China using the deep-learning gravity model: A perspective of population migration. Cities 145, 104680.

Hernandez, M. A., O. Ecker, P. Laderach, and J.-F. Maystadt (2023). Forced migration: Fragility, resilience, and policy responses. In Global Food Policy Report 2023: Rethinking Food Crisis Responses. Chapter 7. Washington, DC: International Food Policy Research Institute (IFPRI).

Hoffmann, R., G. Abel, M. Malpede, R. Muttarak, and M. Percoco (2024). Drought and aridity influence internal migration worldwide. Nature Climate Change.

Hoffmann, R., B. Sˇedov´a, and K. Vinke (2021). Improving the evidence base: A methodological review of the quantitative climate migration literature. Global Environmental Change 71, 102367.

IOM (2021). World Migration Report.

Karemera, D., V. I. Oguledo, and B. Davis (2000). A gravity model analysis of international migration to North America. Applied Economics 32 (13), 1745–1755.

Keola, S., M. Andersson, and O. Hall (2015). Monitoring economic development from space: Using nighttime light and land cover data to measure economic growth. World Development 66, 322–334.

Kim, K. and J. E. Cohen (2010). Determinants of international migration flows to and from industrialized countries: A panel data approach beyond gravity. International Migration Review 44 (4), 899–932.

Liu, M., Y. Shamdasani, and V. Taraz (2023). Climate change and labor reallocation: Evidence from six decades of the Indian census. American Economic Journal: Economic Policy 15 (2), 395–423.

Mach, K. J., W. N. Adger, H. Buhaug, M. Burke, J. D. Fearon, C. B. Field, C. S. Hendrix,

C. M. Kraan, J.-F. Maystadt, J. O’Loughlin, P. Roessler, J. Scheffran, K. A. Schultz, and

N. von Uexkull (2020). Directions for research on climate and conflict. Earth’s Future 8 (7), e2020EF001532. e2020EF001532 2020EF001532.

Marchiori, L., J.-F. Maystadt, and I. Schumacher (2012). The impact of weather anomalies on migration in sub-Saharan Africa. Journal of Environmental Economics and Management 63 (3), 355–374.

Mayda, A. M. (2010). International migration: A panel data analysis of the determinants of bilateral flows. Journal of Population Economics 23 (4), 1249–1274.

Maystadt, J.-F., K. Hirvonen, A. Mabiso, and J. Vandercasteelen (2019). Impacts of Hosting Forced Migrants in Poor Countries. Annual Review of Resource Economics 11, 439 – 459.

Maystadt, J.-F., A. K. Mishra, V. Mueller, and M. Smoldt (2024, October). The Causes and Policy Responses to Forced Migration and Environmental Degradation in Africa. Annual Review of Resource Economics 16 (1), 301–322.

Melander, E. and M. O¨ berg (2007). The threat of violence and forced migration: Geographical scope trumps intensity of fighting. Civil Wars 9 (2), 156–173.

Micevska, M. (2021, December). Revisiting forced migration: A machine learning perspective.

European Journal of Political Economy 70, 102044.

Millock, K. (2015). Migration and environment. Annual Review of Resource Eco- nomics 7 (Volume 7, 2015), 35–60.

Mueller, V., C. Gray, and K. Kosec (2014). Heat stress increases long-term human migration in rural Pakistan. Nature Climate Change 4 (3), 182–185.

Mueller, V., G. Sheriff, X. Dou, and C. Gray (2020). Temporary migration and climate variation in eastern Africa. World Development 126, 104704.

Quin˜ones, E. J., J. Nobles, F. Riosmena, and R. Nawrotzki (2023). Anticipatory migration responses to rural climate shocks. AEA Papers and Proceedings 113, 367–371.

Ruzicska, G., R. Chariag, O. Kiss, and M. Koren (2024). Can Machine Learning Beat Gravity in Flow Prediction? In L. Matyas (Ed.), The Econometrics of Multi-dimensional Panels: Theory and Applications, pp. 511–545. Cham: Springer International Publishing.

Silva, J. M. C. S. and S. Tenreyro (2006, 11). The log of gravity. The Review of Economics and Statistics 88 (4), 641–658.

Simini, F., G. Barlacchi, M. Luca, and L. Pappalardo (2021, November). A Deep Gravity model for mobility flows generation. Nature Communications 12 (1), 6576. Publisher: Nature Publishing Group.

Sorichetta, A., T. Bird, N. Ruktanonchai, E. Erbach-Schoenberg, C. Pezzulo1, N. Tejedor,

I. C. Waldock, J. D. Sadler, A. J. Garcia, L. Sedda, and A. Tatem (2016). Mapping internal connectivity through human migration in malaria endemic countries. Sci Data (3), 160066.

UN DESA (2016). Cross-national comparisons of internal migration: An update on global patterns and trends. United Nations Department of Economic and Social Affairs. Technical Paper 2013/1, 157–181.

UNHCR (2022). Unhcr Mid-Year Trends 2022.

Verme, P. and K. Schuettler (2021). The impact of forced displacement on host communities a review of the empirical literature in economics. Journal of Development Economics 150, 102606.

Xie, X., M. Hao, F. Ding, J. Scheffran, T. Ide, J.-F. Maystadt, Y. Qian, Q. Wang, S. Chen,

J. Wu, K. Sun, T. Ma, and D. Jiang (2024, nov). The impacts of climate change on violent conflict risk: a review of causal pathways. Environmental Research Communications 6 (11), 112002.

Zhou, Y.-Y. and A. Shaver (2021). Reexamning the Effect of Refugees on Civil Conflict: A Global Subnational Analysis. Amerian Political Science Review 115 (4), 1175–1196.

This work is part of the CGIAR Research Initiative on Fragility, Conflict, and Migration and the CGIAR Initiative on Climate Resilience (ClimBeR). IFPRI and UCLouvain prepared this publication.. We would like to thank all funders who supported this research through their contributions to the CGIAR Trust Fund: https://www.cgiar.org/funders/

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
A world free of hunger and malnutrition

IFPRI is a CGIAR Research Center

1201 Eye Street, NW, Washington, DC 20005 USA | T. +1-202-862-5600 | F. +1-202-862-5606 | Email: ifpri@cgiar.org | www.ifpri.org | www.ifpri.info