

Authors: Jasper Tjaden and Eduardo Acostamadiedo

Extended abstract:

Do We Need New Theories of Irregular Migration?

Abstract

Micro-level research on migration decision-making has traditionally focused on regular migration. However, with the recent surge in irregular migration in major destination regions (US, EU), there is renewed interest in understanding irregular migration as a distinct phenomenon. Systematic theory testing for irregular migration is still limited, leaving the mechanisms behind why some individuals select irregular migration poorly understood. This paper evaluates the predictive performance of classic migration theories for explaining irregular migration intentions, using the Todaro and Maruszko (1987) model as a benchmark. Following the rational-choice tradition, this model conceptualises migration as a trade-off between expected earnings at home, weighted by unemployment risks, and expected income abroad, weighted by unemployment risks and journey risks. The paper compares its predictive performance with other classic migrations theories, including income maximisation (Sjaastad 1962), network theory, and cultural norms (Massey et al. 1993). The article draws on survey data from over 10,065 respondents in Nigeria, Senegal, and The Gambia. The study applies multinomial logistic regression and machine-learning approaches to predict intentions to migrate irregularly. Predictive accuracy is assessed through ten-fold cross-validation using AUC-ROC metric, and variable importance is estimated with SHAP values. The study contributes to cumulative theory testing in migration research and provides empirical evidence on which theoretical mechanisms best account for irregular migration aspirations across different social groups.

1. Introduction

1.1. Research problem

Do we need new theories of irregular migration? Micro-level research on migration decision-making has not traditionally distinguished between regular and irregular migration. However, the recent rise in irregular migration to the United States and the European Union has reignited calls to conceptualise irregular migration as a distinct phenomenon. Recent studies suggest that the drivers and self-selection patterns of irregular migration differ from those of regular migration (Dennison 2022; Restelli 2023; Aksoy and Poutvaara 2021). However, systematic theory testing remains limited in the migration literature. This offers an important opportunity to revisit and empirically test classic migration theories in the specific context of irregular migration.

1.2. Contribution to the migration literature

This study advances the migration literature by systematically comparing the predictive performance of classic migration theories, thereby assessing their relative empirical accuracy in the context of irregular migration. In their seminal paper, Massey et al. (1993) developed a comparative framework of migration theories, noting that they had developed in a fragmented and isolated manner. The authors urged scholars to evaluate theoretical explanations empirically. Building on this foundation, Riosmena (2024) revisited and reviewed Massey et al.'s framework after three decades of research, showing that while empirical advances have been made, further cross-theory empirical evaluation remains essential to advance the field. The scholar argues that, while it is generally recognised that migration models are not mutually exclusive, they may differ in their predictive power, including how they fit to different social groups. This study contributes to the cumulative effort of testing migration theories by identifying the predictive power of theories on irregular migration.

2. Previous literature: Empirical testing of theories

There is a rich body of empirical evidence on migration models. The accumulated knowledge highlights both overlap and heterogeneity in motivations to migrate. In line with traditional social science research, most of the studies reviewed here have primarily focused on identifying causal mechanisms of migration theories.

The studies can be broadly grouped according to the theories included in the analysis. A first group of articles is oriented at testing rational-choice cost-benefit models (Restelli 2023; Constant and Massey 2002; Yang 2006; Lindstrom and López-Ramírez 2010). A second group comprises more recent analyses that focus on assessing multiple migration drivers, including rational-choice cost-benefit models, networks, and culture (Abel et al. 2019; Hagen-Zanker, Carling, Caso and Rubio 2025; Czaika and Weisner 2025; Dennison 2022; Garip 2012; McKenzie and Rapoport 2007; Migali and Scipioni 2019). In addition, several meta-analyses have been conducted (Aslany, Carling, Mjelva and Sommerfelt 2021; Cipollina, De Benedictis and Scibè 2024; Soto Nishimura and Czaika 2024).

Most of the previous studies used observational data. In contrast, another group of studies tested migration theories by treating drivers in an exogenous manner through survey experiments (Batista and McKenzie 2023; Detlefsen, Heidland and Schneiderheinze 2022; Hager 2021; Tjaden, Liebe and Bruscoli 2025). With the exception of Ruhnke and Rischke (2024), whom analysed aspirations to migrate among Syrian migrants and host populations in Lebanon and Turkey using a data-driven machine learning approach, all previous studies focused on explanatory rather than predictive approaches.

3. Method

In contrast to the previous body of evidence, this study evaluates the predictive power of migration theories in explaining irregular migration intentions. This study differs from earlier work because it explicitly evaluates out-of-sample predictive accuracy, which most prior research overlooks.

Social science prediction has become an increasingly consolidated approach for evaluating theories and models (Hofman, Sharma and Watts 2017; Lazer et al. 2020; Hofman et al. 2021). These authors argue that prediction is not an alternative to explanation (inference) but a complementary tool that allows researchers to assess how well theories account for observable outcomes. As Shmueli (2010) emphasises, predictive accuracy provides an empirical benchmark that is independent of the internal logic of the models, thereby offering a common standard for comparing competing theoretical frameworks.

3.1. Empirical strategy

The theories to be tested are the following. The Todaro and Maruszko (1987) model follows the rational-choice tradition. This model conceptualises migration as a trade-off between expected earnings at home, weighted by unemployment risks, and expected income abroad, weighted by unemployment risks and journey risks. The paper will compare its predictive performance with other classic migrations theories, including income maximisation (Sjaastad 1962), network theory, and cultural norms (Massey et al. 1993). The empirical analysis is divided in four parts.

First, the analysis begins with a traditional inferential observational analysis of the theories. Separate multinomial logistic models will be estimated for each migration theory to assess if the data lend support to the theoretical mechanism. The outcome variable of the multinomial model has three categories: respondents not considering migration, those considering migration, and those willing to travel via irregular routes. To account for the survey design, country fixed effects will be included, and standard errors will be clustered at the enumeration area level.

Second, the analysis performs a “horse race” to identify the model with the highest predictive accuracy. Although the models are multinomial, the main focus is on predicting intentions to migrate irregularly. To estimate uncertainty, ten-fold cross-validation will be applied to the training data, comprising 75% of the original sample. The primary performance metric is the Area Under the Receiver Operating Characteristic Curve (AUC-ROC), which represents the probability that a randomly selected true positive receives a higher predicted score than a randomly selected true negative. Since only 16% of respondents reported intentions to migrate irregularly, the Area Under the Precision–Recall Curve (AUC-PR) will be estimated too, as it prioritises the correct identification of respondents with such intentions. This will be implemented on a later stage. To test whether differences in predictive performance across models are statistically meaningful rather than due to random variation, an equivalence test will be applied (Kuhn and Johnson 2024).

The first comparison assesses which theoretical model achieves the highest performance using a multinomial logistic regression. The second comparison evaluates the performance of a theory-driven model, including only theory-relevant variables, against a data-driven model that adds additional predictors beyond theory, such as: household characteristics, religion, ability to save, subjective economic situation, trust

in local authorities and community, belief in COVID-19, standard of living, attitudes to education, and use of social media. The aim is to assess the additional predictive gain obtained when models incorporate potentially omitted information. These two specifications will be tested using both multinomial logistic regression and machine learning algorithms, including Random Forest and XGBoost.

Third, variable importance will be computed using absolute SHAP values (Molnar 2025) for the theory- and data-driven models across all three modelling approaches (multinomial logistic, Random Forest, XGBoost), identifying which variables exert the greatest influence on predicting willingness to migrate irregularly.

Finally, predictive accuracy will be disaggregated by sex and education to evaluate how explanatory power varies across social groups.

4. Data

4.1. Population

The study uses baseline survey data collected to evaluate the Migrants as Messengers (MaM) programme, an information campaign implemented by the International Organization for Migration (IOM) between 2021 and 2022 in The Gambia (Upper River Region), Senegal (Thiès Region), and Nigeria (Lagos State) (Hebie, Sessou, and Tjaden 2023). The campaign consisted of community-based, peer-to-peer awareness sessions, in which trained returned migrants shared personal testimonies and information about the risks of irregular migration with potential migrants in their local communities. These regions were selected because of their high migration rates and limited previous exposure to migration-related information campaigns, according to expert assessments by IOM country offices.

The settlements sampled in The Gambia are characterised by limited access to infrastructure and public services, and most residents depend on subsistence agriculture and small livestock farming. In Nigeria, the sampled areas have better access to infrastructure and public services, are in close proximity to major urban centres, and most residents work in trade, transport, manufacturing, or informal services. In Senegal's Thiès region, access to infrastructure is moderate; the area lies close to, but not contiguous with, Dakar, and local livelihoods combine agriculture, artisanal activities, and small-scale commerce.

4.2. *Sampling*¹

Survey data for were collected between June and August 2021. The questionnaire was programmed with XLSForms and was administered face-to-face at respondents' homes using SurveyCTO via electronic tablets. The interviews were conducted by IOM staff in Senegal and by professional enumerators from private firms in Nigeria and The Gambia.

A total of 213 enumeration areas were randomly sampled across the regions of the three countries using the latest available sample frame of enumeration areas and maps provided by the respective national statistics agencies. The evaluation team conducted a household listing exercise within each selected area to update household information and identify individuals aged 18 to 30, the target population of the awareness-raising campaign.

Households were randomly sampled from this updated list, with the number of selected households proportional to the size of the 18–30-year-old population. Within each household, enumerators were instructed to interview up to two individuals, one female and one male, if available, both aged 18 to 30. In total, 10,065 individuals provided informed consent and completed the baseline questionnaire.

4.3. *Sample description*

Overall, 28 per cent of the respondents did not have intentions to migrate, 55 per cent were considering migration and 16 per cent were willing to do so irregularly. The average age of the respondents in the sample was 22 years old. Fifty-three percent were women. 38 percent reported no formal education or some primary education, 52 per cent reported some or completed secondary education and 9 per cent had some or completed tertiary education.

Migration norms was operationalised in a single variable index of positive attitudes towards irregular migration measured at the individual level using the average of ten dummy variables (see annex A for details on the operationalisation). The index goes from zero, meaning negative attitudes, to 1, meaning positive attention. With an average index of 0.6, most of the respondents held positive attitudes towards irregular migration.

¹ For additional details, please refer to Hebie, Sessou, and Tjaden (2023).

Migration networks mechanism was operationalised with three classic questions regarding having previous migration experience, knowing a returnee and having contacts abroad. About seven per cent of respondents had migration experience, about a third knew a returnee and 42 per cent had contacts with people abroad.

The classic rational choice theory was measured with expected monthly income abroad in Europe and current monthly income at home. The questions asked in the local currency and for the analysis were transformed using 2021 USD adjusted by purchasing power parity (ppp) to allow comparisons across the three countries in the sample. On average, respondents earned about USD 123 monthly and if they were to migrate to Europe, they expected to earn about USD 2400 per month.

Finally, the Todaro-Maruszko theory, in addition to current at home and expected abroad incomes, they were asked about probability of employment at home using a three-level categorical variable. They were generally optimistic saying as 76 per cent said that economic conditions would improve in their countries. Respondents were asked about the likelihood of arriving to the destination in Europe and finding a job there, if they were to do so irregularly using a 0 to 1 scale. On average, both indices are in 0.5.

Table 1. Sample Characteristics.

Migration theory/ mechanism	Characteristic	N = 10,065 ¹
Dependent variable	Migration intentions	
	Considering, %	5,580 (55%)
	Irregular migration, %	1,624 (16%)
	No migration, %	2,861 (28%)
Socio demographic information	Education	
	No education/Some primary, %	3,867 (38%)
	Some or secondary completed, %	5,254 (52%)
	Some or tertiary completed, %	944 (9.4%)
	Age, years	22.0 (19.0, 26.0)
	Female, %	5,369 (53%)
Migration norms / culture	Positive irregular migration norms, 0-1 index	0.60 (0.50, 0.80)
Migration networks	Migrated before, %	756 (7.5%)
	Knows a returnee, %	3,640 (36%)
	Has contacts abroad, %	4,248 (42%)
Rational choice / Todaro- Maruszko	Monthly income at origin, USD ppp	123 (102, 219)
	Expected monthly income abroad, USD ppp	2,443 (1,018, 4,072)
Todaro-Maruszko	Irregular migration journey cost, USD ppp	4,065 (2,249, 6,388)
	Probability of employment at home	
	Same	1,405 (14%)
	Improved	7,697 (76%)
	Worsened	963 (9.6%)
	Probability of employment abroad, 0-1 index	0.50 (0.40, 0.70)
	Probability of reaching destination irregularly, 0-1 index	0.50 (0.30, 0.60)

¹n (%); Median (Q1, Q3)

Table 2 presents the sample characteristics by migration intentions. Respondents willing to migrate irregularly tended to be younger, less educated, and more often male. They also expressed more positive attitudes towards irregular migration, had more contacts abroad, greater migration experience, and were more likely to know a returnee than those merely considering migration or not considering it at all. Compared to the other groups, they were less optimistic about their employment prospects at home but more confident in their chances of reaching Europe and finding a job there, if they were to migrate irregularly. Their income at home and expected income abroad in Europe were broadly similar to those of respondents who were only considering migration or who did not intend to migrate.

Table 2. Sample Characteristics by intentions to migrate.

Variable	Mean			Differences (T-test p-values)		
	No migration	Considering	Irregular	Irregular minus Considering	Considering minus No	Irregular minus No
No education/Some primary, %	0.49	0.31	0.46	0.000	0.000	0.054
Some or secondary completed, %	0.45	0.57	0.50	0.000	0.000	0.003
Tertiary completed, %	0.06	0.12	0.05	0.000	0.000	0.018
Age, years	22.85	22.49	22.47	0.888	0.000	0.002
Female, %	0.67	0.53	0.31	0.000	0.000	0.000
Monthly income at origin, USD	187.89	198.47	203.20	0.582	0.122	0.074
Expected monthly income abroad, USD	3099.45	3373.96	3259.70	0.258	0.000	0.123
Irregular migration journey cost, USD	4471.66	4667.48	5752.08	0.000	0.003	0.000
Same probability of work at home	0.13	0.15	0.12	0.000	0.016	0.139
Improved probability of work at home	0.80	0.74	0.77	0.031	0.000	0.016
Worsened probability of work at home	0.07	0.11	0.11	0.351	0.000	0.000
Probability of employment abroad, 0-1 index	0.51	0.52	0.54	0.000	0.431	0.000
Probability of reaching destination irregularly, 0-1 index	0.44	0.45	0.58	0.000	0.019	0.000
Migrated before, %	0.05	0.07	0.14	0.000	0.000	0.000
Knows a returnee, %	0.28	0.36	0.52	0.000	0.000	0.000
Has contacts abroad, %	0.32	0.44	0.55	0.000	0.000	0.000
Positive irregular migration norms, 0-1 index	0.60	0.61	0.68	0.000	0.362	0.000

5. Preliminary results

5.1. Inferential analysis

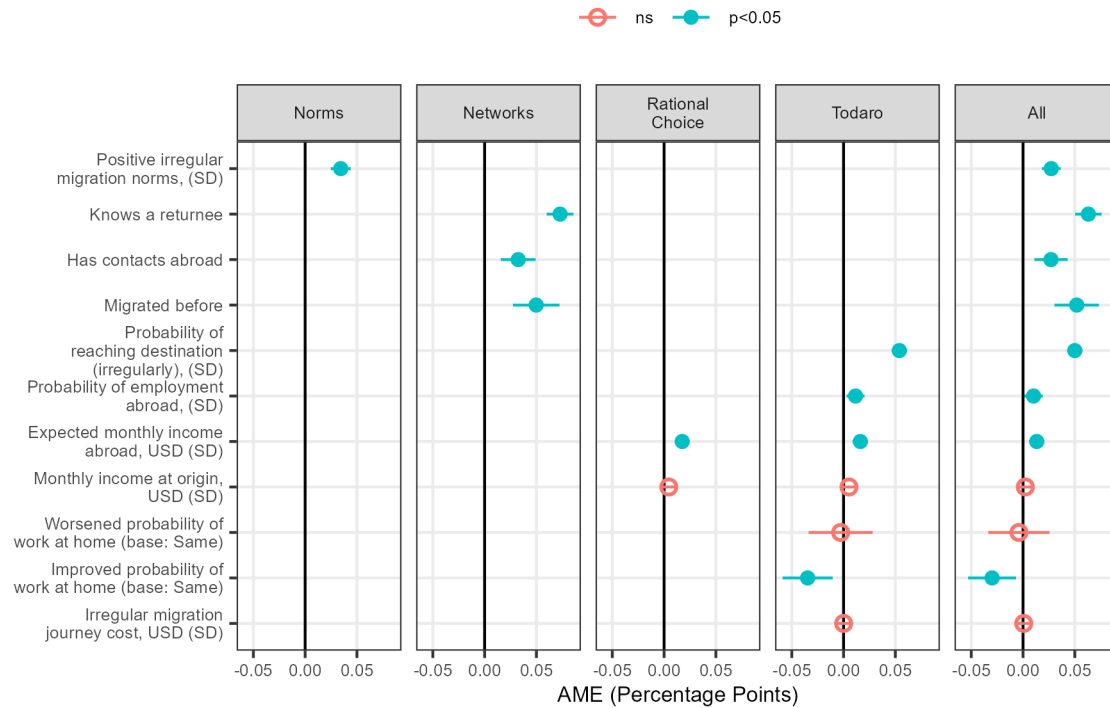
Figure 1 presents the results of the inferential analysis. The coefficients presented refer to reporting willingness to migrate irregularly, and the reference level is no migration. The figure shows five multinomial logistic models. Each model corresponds to a different theoretical mechanism. The Norms model includes a single index of positive attitudes towards irregular migration. The Networks model incorporates three variables

capturing migration experience, knowing a returnee, and having contacts abroad. The Rational-choice model includes current monthly income at home and expected income abroad (in 2021 USD, PPP-adjusted). The Todaro–Maruszko model extends this framework adding an interaction between income at home and perceived probability of employment at home, and a second interaction between the expected income abroad, probability of employment abroad, and probability of successfully reaching the destination. The model includes a final variable of perceived irregular migration costs. Finally, the Full model includes all variables from the previous specifications.

For simplicity, the coefficients of considering migration are omitted here. Annex B has the full table of coefficients. All models have country fixed effects, and the standard errors are clustered at the enumeration area level. The coefficients are presented in average marginal effects (AME), with 95 per cent confidence intervals, to simplify the interpretation of interactions in the Todaro-Maruszko model. Continuous variables were standardised (marked with SD in the figure) to facilitate the comparison of coefficients.

Overall, the inferential analysis supports the theoretical expectations. Positive migration norms, a migration background, and more optimistic expectations of income, arrival success, and employment in Europe are all positively associated with reporting a willingness to migrate irregularly. In the full model, a one–standard deviation increase in the perceived likelihood of successfully reaching Europe irregularly is associated with a five–percentage–point increase in the probability of expressing irregular migration intentions. Similarly, a one–standard deviation increase in expected income in Europe corresponds to a 1.3–percentage–point increase. By contrast, believing that economic conditions at home will improve is associated with a three–percentage–point reduction in the likelihood of reporting irregular migration intentions.

Figure 1. Multinomial logistic estimates of migration intentions



5.2. Predictive accuracy

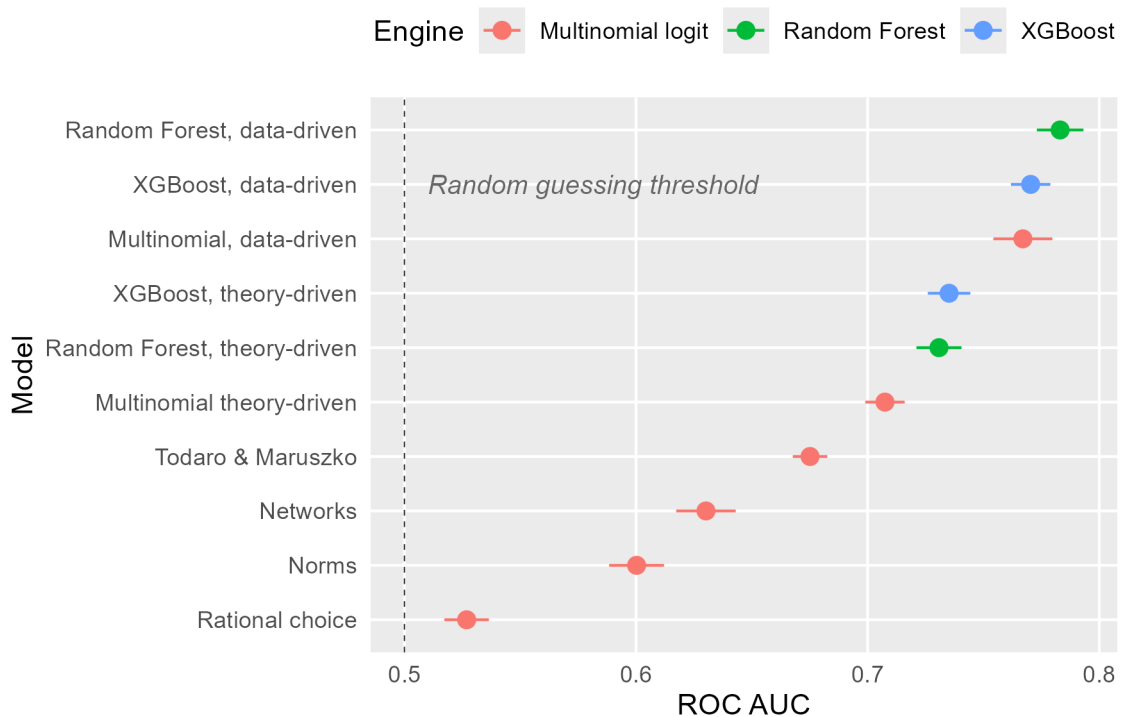
Figure 2 presents the results of the predictive accuracy analysis for irregular migration intentions. In total, ten models are compared. The first four are theory-based specifications, rational choice, norms, networks, and Todaro–Maruszko, each estimated with a multinomial logistic regression. A fifth multinomial model combines all theory variables into a single specification. A sixth, data-driven multinomial model extends this by adding demographic and contextual predictors, including age, sex, education, religion, household composition, ability to save, subjective economic situation, trust in local institutions and community, beliefs about COVID-19, perceived living standards, attitudes towards education, and use of social media. Finally, two machine learning algorithms, Random Forest and XGBoost, are trained using both the theory-driven and data-driven variable sets.

Predictive accuracy is measured using the Area Under the Receiver Operating Characteristic Curve (ROC AUC). The ROC AUC represents the probability that a randomly selected respondent who intends to migrate irregularly (a true positive) receives a higher predicted probability than a randomly selected respondent who does not (a true negative). A value of 1 indicates perfect prediction, while 0.5 reflects random guessing. The ROC AUC metric is estimated through a ten-fold cross-validation.

Predictive accuracy (ROC AUC) was estimated using ten-fold cross-validation. Mean AUC values and 90% confidence intervals were computed from the distribution of ten-fold specific AUCs, assuming a normal distribution.

There are two main takeaways. First, among the theory-based models, the Todaro–Maruszko model performs best, as expected, given that it was originally formulated to explain irregular migration. Second, incorporating additional contextual and demographic predictors in the data-driven specifications substantially improves predictive performance, particularly for the Random Forest and XGBoost models. This pattern suggests that classic theories capture key mechanisms but only partial aspects of irregular migration willingness, whereas data-rich machine-learning models provide a more multifaceted representation of the factors associated with irregular migration intentions.

Figure 2. Model predictions of irregular migration intentions



5.3. Variable importance

Figure 3 presents the variable importance for the best-performing model, the Random Forest, comparing the theory-driven and data-driven specifications. Variable importance is estimated using absolute SHAP values, averaged across all observations. To ease interpretation, raw absolute SHAP values are expressed as the percentage

contribution of each variable to the model's predictive performance.

Three main findings emerge. First, in the data-driven model, sex and enumeration area fixed effects—which capture local infrastructure, access to public services, and general economic conditions—are the most influential predictors. These variables are not part of classic migration theories, suggesting that contextual and demographic factors play a substantial role in shaping irregular migration intentions. Second, both the theory-driven and data-driven Random Forest models display a similar overall ranking of theoretical variables, with risk perceptions (probability of arrival and employment abroad), migration networks, norms, and migration costs among the most influential predictors. Third, no single theoretical mechanism dominates; instead, a combination of risk, network, and normative factors jointly explain irregular migration intentions.

Figure 3. Variable importance for the Random Forest models

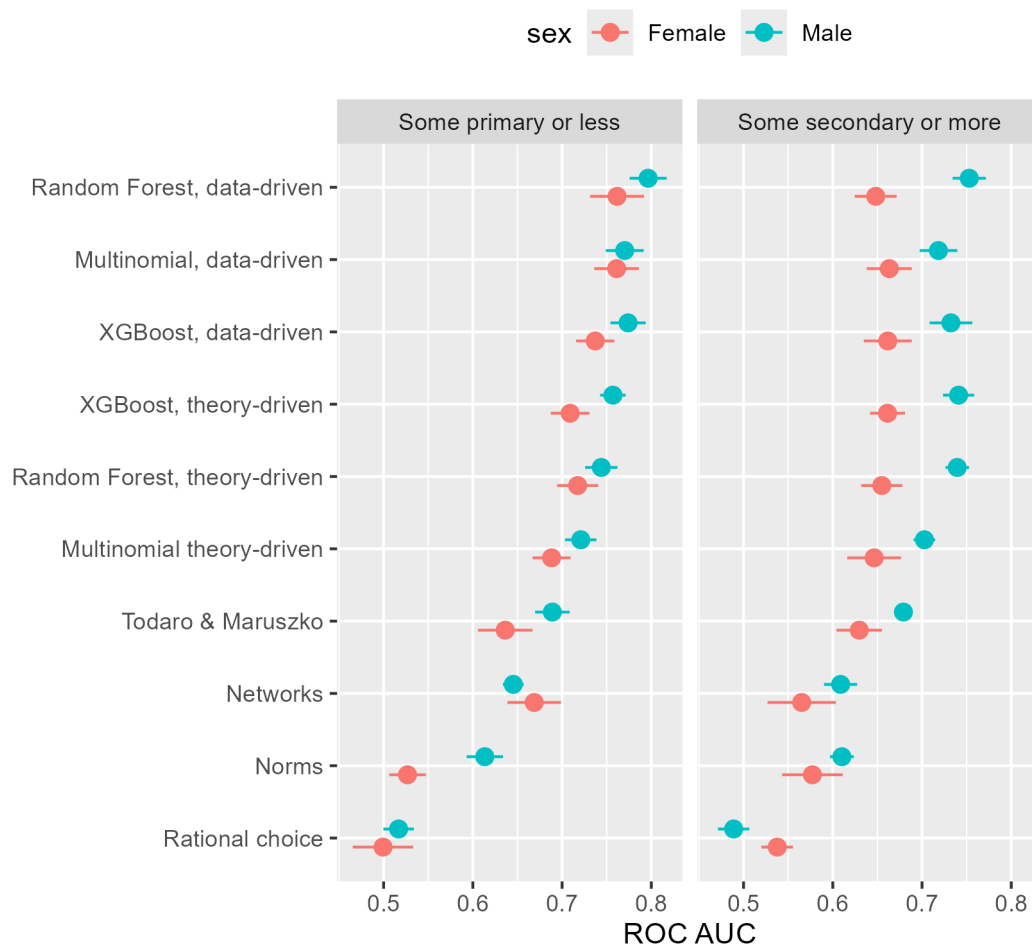


5.4. Predictive accuracy for demographic subgroups

Figure 4 presents the predictive accuracy of the models disaggregated by sex and education level. The figure compares the ROC AUC scores across men and women with primary or lower education (left panel) and secondary or higher education (right panel). First, across most models and education levels, predictions for men are more accurate than for women, particularly in the theory-driven specifications. This suggests that the mechanisms captured by classic migration theories, such as economic incentives, risk perceptions, and social networks, explain men’s intentions to migrate irregularly more

effectively than women's. For women, additional contextual or gender-specific factors not modelled here, such as family responsibilities, social norms, or safety concerns, may shape migration intentions in ways that are less predictable by standard theoretical variables. Second, models generally achieve higher predictive accuracy among respondents with lower education compared to those with secondary or higher education. This pattern implies that migration intentions among less-educated individuals are more closely aligned with traditional economic and network-based explanations, while more-educated respondents may be influenced by a broader range of motivations, such as career prospects, further education, or lifestyle aspirations, that are not captured by the theory variables.

Figure 4. Predictive accuracy by education level and sex



6. Preliminary conclusions

Do we need new theories of irregular migration? This article examined the predictive performance of classic migration theories using observational data from three African

countries. It introduced an innovative approach by evaluating theories through their predictive power, going beyond a purely inferential perspective. Specifically, the study applied out-of-sample testing, which assesses how well models generalise to new data not used during model training. This approach shifts attention from the traditional focus on coefficient significance and direction towards evaluating how accurately theories predict individuals' willingness to migrate irregularly.

The results highlight the value of combining different theoretical mechanisms and assessing their relative contributions to predictive accuracy. Classic theories explain a substantial share of irregular migration intentions, but their predictive limits become evident when broader social and contextual factors are considered. Data-driven machine-learning models show clear gains in predictive power, suggesting that established migration theories may have blind spots, particularly in capturing gendered dynamics and other social factors that shape migration decisions. Gender, in particular, emerges as a central dimension that existing migration theories have largely overlooked and that future theory-building should integrate more systematically. Overall, the findings underscore the importance of integrating multiple theoretical perspectives while also broadening their scope to account for how different social groups operate within theoretical frameworks.

Bibliography

Abel, G. J., Raymer, J., and Guan, Q. 2019. Driving factors of Asian international migration flows. *Asian Population Studies*, 15(3), 243-265.

Aslany, M., Carling, J., Mjelva, M. B., and Sommerfelt, T. 2021. Systematic review of determinants of migration aspirations. *Changes*, 1(18), 3911-3927.

Aksoy, C., and Poutvaara, P. 2021. Refugees' and irregular migrants' self-selection into Europe. *Journal of Development Economics*, 152, 102681.

Batista, C., & McKenzie, D. 2023. Testing classic theories of migration in the lab. *Journal of International Economics*, 145, 103826.

Cipollina, M., De Benedictis, L. and Scibè, E. 2024. Environmental migration? A systematic review and meta-analysis of the literature. *Rev World Econ* 160, 1393–1441. <https://doi.org/10.1007/s10290-024-00529-5>

Constant, A., and Massey, D. S. (2002). Return migration by German guestworkers:

- Neoclassical versus new economic theories. *International migration*, 40(4), 5-38.
- Czaika, M., and Weisner, Z. 2025. Migration aspirations and their realisation: a configurational driver analysis of 26 African and Asian research areas. *Comparative Migration Studies*, 13(1), 8.
- Detlefsen, L., T. Heidland, and C. Schneiderheinze. 2022. “What Explains People’s Migration Aspirations? Experimental Evidence from Sub-Saharan Africa. Experimental Evidence from Sub-Saharan Africa.” <http://dx.https://doi.org/10.2139/ssrn.4238957>.
- Dennison, J. 2022. Re-thinking the drivers of regular and irregular migration: evidence from the MENA region. *Comparative Migration Studies*, 10(1), 21.
- Garip, F. 2012. Discovering diverse mechanisms of migration: The Mexico–US stream 1970–2000. *Population and Development Review*, 38(3), 393-433.
- Hagen-Zanker, J., Carling, J., Caso, N., & Rubio, M. G. 2025. The multi-level determinants of international migration aspirations in 25 communities in Africa, Asia and the Middle East. *World Development*, 185, 106774.
- Hager, A. 2021. What drives migration to Europe? Survey experimental evidence from Lebanon. *International Migration Review*, 55(3), 929-950.
- Hebie, O., M. Sessou, and J. Tjaden. 2023. “Irregular Migration from West Africa: Robust Evaluation of Peer-to-peer Awareness-raising Activities in Four Countries.” International Organization for Migration. <https://publications.iom.int/books/irregular-migration-westafrica-robust-evaluation-peer-peer-awareness-raising-activities-four>
- Hofman, J. M., Sharma, A., and Watts, D. J. 2017. Prediction and explanation in social systems. *Science*, 355(6324), 486-488.
- Hofman, J.M., Watts, D.J., Athey, S. et al. 2021. Integrating explanation and prediction in computational social science. *Nature* 595, 181–188. <https://doi.org/10.1038/s41586-021-03659-0>
- Kuhn, M., and Johnson, K. (2024). Applied Machine Learning for Tabular Data. <https://aml4td.org/>
- Lazer, D., Pentland, A., Watts, D., Aral, S., Athey, S., Contractor, N., ... & Wagner, C. (2020). Computational social science: Obstacles and opportunities. *Science*, 369(6507), 1060-1062.

- Lindstrom, D.P., and López Ramírez, A. (2010). Pioneers and followers: Migrant selectivity and the development of US migration streams in Latin America. *The Annals of the American Academy of Political and Social Science*, 630(1), 53-77.
- Massey, D. S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, and J. E. Taylor. 1993. "Theories of International Migration: A Review and Appraisal." *Population and Development Review* 19 (3): 431–66. doi:10.2307/2938462
- McKenzie, D. and Rapoport, H. 2007. Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics*, 84(1), 1–24
- Migali, S., & Scipioni, M. 2019. Who's about to leave? A global survey of aspirations and intentions to migrate. *International Migration*, 57(5), 181-200.
- Molnar, C. 2025. *Interpretable Machine Learning: A Guide For Making Black Box Models Explainable*.
- Restelli, G. 2023. Development and international migration: The effect of income on regular and irregular migration intentions to Europe. *Population and Development Review*, 49(1), 135-174.
- Riosmena, F. 2024. "Worlds in Motion Redux? Expanding Migration Theories and Their Interconnections." *Population and Development Review* 50 (3): 677–726. doi:10.1111/ padr.12630
- Ruhnke, S., & Rischke, R. 2024. Predicting mobility aspirations in Lebanon and Turkey: a data-driven exploration using machine learning. *Data & Policy*, 6, e47.
- Sjaastad, L. 1962. The costs and returns of human migration. *Journal of political Economy*, 70(5, Part 2), 80-93.
- Soto Nishimura, A., and Czaika, M. 2024. Exploring migration determinants: a meta-analysis of migration drivers and estimates. *Journal of International Migration and Integration*, 25(2), 621-643.
- Tjaden, J. 2023. "Risky Journeys – Risk and Decision-Making among Potential Irregular Migrants in Senegal and Guinea." *International Migration* 61 (2): 212–25. doi:10.1111/imig.13003
- Tjaden, J., Liebe, U., and Bruscoli, D. 2025. Explaining re-migration preferences among

assisted returnees: evidence from Discrete Choice Experiments in Sudan, Somalia, and Ethiopia. *Journal of Ethnic and Migration Studies*, 1-25.

Todaro, M. P., and L. Maruszko. 1987. "Illegal Migration and US Immigration Reform: A Conceptual Framework." *Population and Development Review* 13 (1): 101–14. doi:10.2307/1972122

Yang, D. 2006. Why do migrants return to poor countries? Evidence from Philippine migrants' responses to exchange rate shocks. *The Review of Economics and Statistics*, 88(4), 715-735.

Appendix A: Operationalisation of migration norms

Positive attitudes towards migration. The original scale is a seven-point Likert scale. Items were transformed into dummies and then averaged per respondent. To what extent do you agree with these statements: 1) "In my community, it is normal that young people migrate and live abroad for a few years"; 2) "In my community, only people who have lived abroad are fully respected"; 3) "In my community, it is easier to find a spouse for migrants than for non-migrants"; 4) "Everyone in my community admires migrants"; 5) Reversed order: "In my community, migrants are seen as having abandoned their community." 6) "In my community, the wealthiest families are those that have members abroad". For the following items, this was the introduction: Regarding returning migrants and people leaving abroad without legal documents, do you agree with the following statement: 7) Reversed order: "People from (\$ {s_country}) who leave abroad without legal documents give our country a bad name". 8) Reversed order: "Seeing migrants return back from Europe with empty hands makes me feel ashamed". 9) "Returned migrants should be proud of themselves for their courage". 10) "Returned migrants deserve our support because they have been through a lot of difficulties"

Appendix B:

Complete table of multinomial logistic estimates of migration intentions

Variable	Norms		Networks		Rational Choice		Todaro		All	
	AME	p-value	AME	p-value	AME	p-value	AME	p-value	AME	p-value
Consider										
Irregular migration journey cost, USD (SD)							0.014	0.06	0.014	0.05
Probability of employment at home										
Improved (base: Same)							-0.003	0.87	-0.001	0.93
Worsen (base: Same)							0.014	0.48	0.013	0.54
Monthly income at origin, USD (SD)					0.006	0.36	0.006	0.32	0.001	0.91
Expected monthly income abroad, USD (SD)					0.022	0.00	0.022	0.00	0.021	0.00
Probability of employment abroad, (SD)							-0.007	0.27	-0.008	0.22
Probability of reaching destination (irregularly), (SD)							-0.014	0.03	-0.018	0.00
Migrated before			0.035	0.14					0.034	0.17
Has contacts abroad			0.061	0.00					0.060	0.00
Knows a returnee			0.010	0.45					0.008	0.53
Positive irregular migration norms, (SD)	0.005	0.5							0.003	0.68
Irregular										
Irregular migration journey cost, USD (SD)							0.000	0.99	0.001	0.87
Probability of employment at home										
Improved (base: Same)							-0.035	0.00	-0.030	0.01
Worsen (base: Same)							-0.003	0.86	-0.004	0.79
Monthly income at origin, USD (SD)					0.005	0.18	0.005	0.09	0.002	0.48
Expected monthly income abroad, USD (SD)					0.018	0.00	0.016	0.00	0.013	0.00
Probability of employment abroad, (SD)							0.012	0.01	0.010	0.02
Probability of reaching destination (irregularly), (SD)							0.054	0.00	0.050	0.00
Migrated before			0.050	0.00					0.052	0.00
Has contacts abroad			0.032	0.00					0.027	0.00
Knows a returnee			0.073	0.00					0.063	0.00

Variable	Norms		Networks		Rational Choice		Todaro		All	
	AME	p-value	AME	p-value	AME	p-value	AME	p-value	AME	p-value
Positive irregular migration norms, (SD)	0.034	0.0							0.027	0.00