RevistaEconomică
Journal of economic-financial theory and practice

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

A COMPARATIVE ANALYSIS OF LINEAR ECONOMETRIC AND MASCHINE LEARNING APPROACHES TO GLOBAL CLIMATE-INDUCED MIGRATION FLOWS

Christian ROGALSKI[1]

[1]Alexandru Ioan Cuza University, Iaşi, Romania, 0009-0007-6144-0631

Abstract: *This study compares linear econometric and machine learning approaches to understanding climate-induced migration from a global perspective, utilizing an expanded dataset covering the period from 1960 to 2020. The analysis contrasts a fixed-effects panel model with a Random Forest machine learning model, each designed to capture the influence of climate factors such as average temperature and precipitation on migration flows using the same underlying data. In the linear model, these climate variables interact with socioeconomic indicators like GDP per capita and agricultural dependency, as well as governance and infrastructure quality, to explain historical migration trends and highlight how institutional and structural resilience can mitigate climate pressures. The Random Forest model, in turn, uncovers non-linear interactions and threshold effects that the linear specification cannot directly assume, yet it confirms the relevance of the chosen variables and interactions by producing a similar level of explanatory power. Together, these approaches show that while climate-related changes significantly shape migration patterns, their impact is highly context-dependent. The findings underscore the advantages of combining traditional econometric modelling with machine learning methods to achieve a more comprehensive understanding of global climate-induced migration flows.*

Keywords: *climate change, migration, global, panel data, fixed effects model, machine learning, random forest*

JEL classification: *C23, F22, J61, O13, O15, Q54*

1. Introduction

The relationship between climate change and human mobility is one of the central challenges in global transformation research. Numerous studies suggest that rising temperatures, altered precipitation patterns, droughts, floods, and a gradual decline in agricultural productivity can influence migration flows (Black et al., 2011; Cattaneo & Peri, 2016; Hoffmann et al., 2020). Although it is widely recognized that certain environmental changes can promote migration, empirical findings reveal a highly heterogeneous picture. While some regions experience relatively low emigration pressure despite significant climatic burdens, others react sensitively to even the smallest climatic changes. These variations indicate that climatic factors alone are not enough to explain migration. Rather, the institutional and socio-economic context determines the extent to which people respond to climate change by deciding to emigrate. In view of this complexity, more recent work emphasizes that climate-induced migration should not be considered in isolation, but rather as a process dependent on multiple socio-economic and institutional conditions. Systematic reviews and cross-regional analyses suggest that the magnitude and nature of climate-related migration responses vary substantially depending on structural vulnerabilities and local adaptive capacities (Borderon et al., 2019; Hoffmann et al., 2020).

Institutional quality, measured for example by governance indicators, the availability and reliability of infrastructure, and a country's general level of prosperity, is an important moderator in this process (Beine & Jeusette, 2018; Missirian & Schlenker, 2017; Coronese et al., 2019). Societies with

[1] *rogalski.christian@pm.me*

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

strong governance, economic diversification and well-developed infrastructure are often better able to mitigate negative climate stress through internal adaptation strategies, technological innovation or social safety nets. In agrarian, poorer and institutionally weaker contexts, even moderate climate changes can increase economic instability and cause a larger proportion of the population to consider emigrating from the country (McLeman, 2018; Rigaud et al., 2018).

The current research landscape frequently relies on linear econometric models to investigate climate-related migration patterns. Such models are interpretable and allow for hypothesis-driven interactions between climate and contextual variables. However, they reach their limits when capturing complex, nonlinear effects and hidden interactions. Machine learning methods such as random forests promise theoretical advantages in this respect, as they work without predefined functional forms and are particularly suitable for recognizing nonlinear patterns (Breiman, 2001). However, it remains unclear whether these methods offer added value over classical econometric approaches when it comes to highly aggregated, global panel data sets that are characterized by incomplete data and the need for imputations. In this context, the present study has the following main objectives: It compares a fixed effects panel model and a random forest model in their ability to explain climate-related migration at the global level between 1960 and 2020. The integration of new, internationally harmonized data sets that include updated indicators of migration flows and governance (e.g., Abel & Cohen, 2022; World Bank, 2021; IPCC, 2022), infrastructure data, and climate variables provides a comprehensive analytical basis. Particular attention is paid to whether a random forest model, which is able to identify interactions and nonlinearities without explicit specification, offers significantly higher explanatory power than a classical fixed-effects model. Furthermore, the focus is on understanding the role of governance, infrastructure and wealth levels, as well as sectoral economic structure in attenuating the climate-emigration link, and how the results are consistent with the existing literature, including Cattaneo & Peri (2016), Missirian & Schlenker (2017), and recent IPCC reports (IPCC, 2022).

This analysis therefore contributes to the debate on the complexity of climate-induced migration by comparing linear, theory-based models with machine learning methods. Traditional econometric approaches are used to obtain interpretable, hypothesis-based coefficients. On the other hand, the Random Forest provides insights into the global significance of individual variables without imposing formal restrictions on functional relationships. Although both approaches may explain similar variance, they can offer different perspectives on data and hypotheses. This multi-dimensionality can help reconcile some of the conflicting findings in the literature and provide a more nuanced basis for policy decisions and long-term development strategies.

2. Background and Theoretical Framework

The theoretical debate on climate-induced migration integrates approaches from migration economics, development economics, environmental economics, and political science. Traditional migration models interpret emigration as the outcome of weighing push and pull factors, where negative conditions in the country of origin (such as declining incomes, resource scarcity, or political instability) encourage individuals to migrate, while better labor market opportunities, wages, or security in destination countries serve as attractive pull factors (Todaro & Maruszko, 1987). Within this framework, climate change can be considered an additional push factor if rising temperatures, altered precipitation patterns, or extreme weather events (Black et al., 2011) undermine the economic foundations of the country of origin. However, climatic changes do not operate in a vacuum. Their relevance for migration crucially depends on the socioeconomic and institutional context. Societies with a high share of agricultural employment are especially vulnerable to climate shocks since crop failures directly affect incomes and food security (Cattaneo & Peri, 2016). Dependence on agriculture thus increases vulnerability to climate change, while economically diversified countries are more likely to find internal adaptation options.

Institutional quality and governance structures also play a key role. Good governance, transparency, and political stability create conditions in which states can respond to climate stressors

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

with adaptive measures. In contrast, where corruption, weak rule of law, and inefficiency dominate, effective political instruments and social safety nets are lacking (Neumayer & Barthel, 2015; McLeman, 2018). Weak governance thus raises the probability that climate change, in the absence of adequate domestic compensation measures, becomes a more potent driver of emigration.

Infrastructure is equally crucial. Missing transportation networks, inadequate electricity supply, or insufficient storage and cooling facilities exacerbate the negative climate impacts on local markets, prices, and livelihoods. Conversely, well-developed infrastructure can facilitate adaptation, through easier transportation of goods, improved market access, or disaster preparedness. and thereby mitigate the climate-induced pressure on emigration (Beine & Jeusette, 2018; Missirian & Schlenker, 2017).

These relationships are not necessarily linear. Climate changes may only produce significant effects after certain thresholds are exceeded, or institutional buffers may work effectively up to a certain point and then suddenly fail. Such potential nonlinearities and thresholds complicate modeling efforts. While conventional linear models introduce interactions to approximate these effects, flexible machine-learning methods like Random Forests can operate without predefined functional forms (Breiman, 2001).

Overall, the theory suggests that climate change primarily increases migration where vulnerable structures, agricultural dominance, weak institutions and inadequate infrastructure are present. Wealth alone is not a sufficient buffer if key adaptation mechanisms are lacking. This theoretical foundation thus motivates the simultaneous consideration of climate, governance, infrastructure, and wealth variables, as well as their interactions and potential nonlinearities. These considerations form the basis for the empirical analysis, in which both a fixed-effects model and a random forest approach are employed to test the theoretically expected relationships and assess the importance of individual factors.

3. Data

The foundation of this study is a global country dataset spanning the period from 1960 to 2020 in decadal intervals. Aggregating data into ten-year periods is a pragmatic approach to capture long-term trends and reduce short-term fluctuations and data gaps. For shorter intervals, less comprehensive information would be available, and extreme events or volatile weather patterns would carry more weight. By using decadal averages, structural developments can

3.1 Migration Data

The migration-related information is based on bilateral estimates of international migration flows. While the data foundations for earlier decades (e.g., before 1990) are often thin and rely on indirect methods (such as stock-differencing or demographic accounting approaches), more refined estimates are available from 1990 onward. A key source here is Abel & Cohen (2022), whose updated migration flow estimates provide a significantly improved data basis. The flow data initially come in five-year intervals to capture global mobility patterns between countries. Decadal averages are then calculated from these five-year flow data.

By combining migration flows with population figures (from the World Development Indicators, World Bank), emigration rates are derived by dividing the total flow over a decade by the country's average population during the same period. These rates indicate the share of the population that emigrated over the course of a decade. The focus on gross migration rates instead of net rates allows for a more comprehensive recording of overall mobility without prior offsetting of inflows and outflows. This maintains a clearer picture of the dynamics behind emigration decisions.

The quality and availability of migration data vary by region and period. Early decades are often associated with higher uncertainties. Nonetheless, integrating these data is indispensable to obtain a global long-term perspective. In selecting the sample, care was taken to include as many countries as possible for which sufficient information is available for the relevant decades. By combining older and newer data, a panel is formed covering 1960 to 2020 (older data from Abel (2017) for earlier decades and newer data from Abel & Cohen (2022) for later decades). The periods from 1960–1985 are from the

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

old dataset, 1990–2020 from the new one. All flows were aggregated into decadal intervals to capture long-term trends and mitigate data gaps.

3.2 Climate Data

The climatic variables, particularly average temperature and average precipitation are derived from global, freely accessible climate datasets such as the CRU TS (Climate Research Unit Time-Series) Version 4.07 (Harris et al., 2020). The CRU data are based on a dense network of meteorological stations whose information is spatially interpolated. Monthly gridded data are available, from which annual and, finally, decadal country means are calculated.

Decadal aggregation reduces short-term weather noise. Average temperature (in °C) and precipitation (in mm per period) are considered key indicators because they are crucial for agricultural production conditions. While rising temperatures increasingly stress crops and worsen growing conditions in the long run, precipitation directly influences water availability for crops, livestock, and associated sectors. Changes in these climatic indicators are seen as potential drivers of shifts in the economic stability of rural areas, which in turn can affect migration decisions.

3.3 Socioeconomic and Institutional Indicators

In addition to climate and migration data, the incorporation of socioeconomic and institutional variables is essential to understand the influence of climatic changes on emigration in context. To this end, information on GDP per capita (constant 2015 US$), employment structure with a focus on the share of agriculture, and the agricultural sector's contribution to GDP from the World Development Indicators of the World Bank are used. These measures indicate the degree to which an economy depends on agriculture and how prosperous it is. A strong agricultural dominance increases vulnerability to climatic fluctuations since crop failures or price changes have immediate impacts on incomes and living conditions. In more diversified economies, alternative livelihood opportunities exist domestically.

To capture institutional quality, governance indicators such as government effectiveness, political stability, rule of law, and control of corruption are drawn from the Worldwide Governance Indicators (WGI) (Kaufmann et al., 2010). Since comprehensive data for these indicators are only available from the 1990s onward, an analytical trick is employed: The distribution of these indicators from the 2010–2019 decade serves as a reference to define categories such as "very_low," "low," "high," or "very_high," which are then retroactively applied to earlier decades. This creates a temporally constant benchmark, even if it simplifies historical developments. A similar approach is used for infrastructure indicators, such as electricity access or the Logistics Performance Index. Here too, the 2010–2019 decade provides the reference distributions for categorization. This methodological simplification is necessary to establish a uniform reference framework over the entire analysis period, allowing countries to be compared in terms of governance and infrastructure.

Additionally, a constant OECD definition (as of 2024) was adopted for all decades to classify countries consistently as OECD or non-OECD states over time. While this approach may distort historical realities, given that OECD memberships changed over time, it offers a stable framework for assessing wealth and development levels. Data availability also varies significantly between countries and decades. Not all countries have complete data for all desired variables. The Random Forest model enables limited imputations of missing values to maximize the sample size, while the Fixed-Effects model typically uses only complete datasets.

The resulting database encompasses a global sample of countries that differ markedly in terms of development level, governance, infrastructure, and economic structure. This includes highly industrialized economies with strong governance and good infrastructure, as well as poorer, agriculturally oriented states with weak institutions. Such diversity is important to empirically test the theoretically posited interactions between climate and contextual factors. The decadal aggregation, retrospective classifications, and constant OECD definition are all compromises made to create a stable comparative basis over six decades in the face of incomplete and heterogeneous data. Although these

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

methodological decisions complicate a finely tuned historical analysis, they render a globally comparative perspective on climate-related migration and its institutional-economic moderator variables feasible in the first place.

3.4 Data Analysis

In order to better understand the data for the methodological approach, the data was processed and presented graphically and analysed from different perspectives. This already allows initial insights.

Table 1 presents a set of descriptive statistics, namely the mean and standard deviation of emigration rates, temperature, and precipitation, organized by the various categories employed throughout this analysis. Each row corresponds to a specific grouping, such as a particular wealth status (e.g., “poor,” “lower_middle,” “upper_middle,” “rich”), an infrastructure category (“high,” “low,” “very_high,” “very_low”), an employment or GDP share in agriculture category (“high,” “low,” “very_high,” “very_low”), a governance index category (“high,” “low,” “very_high,” “very_low”), or a region (e.g., East Asia, Sub-Saharan Africa, Latin America & Caribbean). For each grouping, the table indicates the number of observations (n), followed by the average (mean) and standard deviation (sd) of both the emigration rate and the two principal climate variables (temperature and precipitation).A closer inspection reveals that “poor” countries, for example, register mean emigration rates around 0.031 with a standard deviation of 0.040, whereas “rich” countries display higher mean rates at approximately 0.077 and a comparatively smaller standard deviation of 0.023. Meanwhile, “lower_middle” countries have an average temperature of about 20.63 °C, whereas “poor” countries display higher temperature levels (around 22.77 °C), reflecting the tendency for lower-income nations to be located in hotter zones.

Table 1: Descriptive Statistics

Variable	n	mean_emigration	sd_emigration	mean_temp	sd_temp	mean_precip	sd_precip	Grouping
lower_middle	250	0.039	0.042	20.628	6.875	105.501	80.618	Wealth_Status
poor	248	0.031	0.042	22.771	6.054	94.866	52.657	Wealth_Status
rich	84	0.077	0.063	24.177	3.281	95.537	82.486	Wealth_Status
upper_middle	194	0.042	0.051	18.278	7.954	105.114	70.767	Wealth_Status
high	187	0.040	0.044	18.912	8.086	112.342	80.406	Infrastructure_Category
low	241	0.040	0.047	18.687	7.687	94.269	71.218	Infrastructure_Category
very_high	94	0.090	0.071	23.220	4.912	104.357	78.388	Infrastructure_Category
very_low	254	0.026	0.027	24.245	3.552	97.569	57.082	Infrastructure_Category
high	247	0.029	0.031	20.695	7.441	103.965	78.234	Employment_%_Agriculture
low	185	0.050	0.049	19.036	7.545	95.914	72.013	Employment_%_Agriculture
very_high	248	0.034	0.049	21.982	6.282	108.803	53.366	Employment_%_Agriculture
very_low	96	0.076	0.063	23.917	3.826	82.420	81.832	Employment_%_Agriculture
high	251	0.037	0.042	19.953	7.501	105.660	72.618	Agriculture_GDP
low	170	0.046	0.041	17.566	7.737	90.131	57.659	Agriculture_GDP
very_high	253	0.033	0.046	23.279	5.284	106.264	67.880	Agriculture_GDP
very_low	102	0.067	0.069	24.479	3.107	94.034	87.646	Agriculture_GDP
high	200	0.045	0.046	20.019	7.269	94.577	71.716	WGI_Category
low	222	0.036	0.039	21.066	7.621	109.272	66.321	WGI_Category
very_high	72	0.080	0.071	22.574	5.628	128.418	91.417	WGI_Category
very_low	248	0.032	0.046	21.709	6.355	90.292	66.967	WGI_Category
East Asia	18	0.041	0.069	9.851	9.653	82.648	72.528	Region
East Europe & Central Asia	48	0.046	0.036	8.746	4.716	40.070	23.134	Region
Europe	68	0.046	0.035	9.223	5.356	61.083	21.476	Region
Latin America & Caribbean	138	0.053	0.050	23.736	2.962	147.849	41.490	Region
North Africa	37	0.023	0.018	22.762	3.161	11.334	8.575	Region
Oceania	42	0.051	0.052	24.366	1.809	213.389	55.500	Region
South Asia	42	0.035	0.074	18.945	6.835	103.151	55.314	Region
South-East Asia	60	0.026	0.041	25.444	1.229	181.433	49.894	Region
Sub-Saharan Africa	253	0.031	0.037	24.135	3.401	97.017	54.537	Region
West Asia	70	0.071	0.070	22.672	4.364	17.227	14.967	Region

Similarly, the “very_low” infrastructure category features a mean temperature of 22.84 °C, contrasting with the “high” infrastructure category, at 18.91 °C on average. Regions exhibit likewise notable differences: South Asia and Sub-Saharan Africa appear in generally warmer conditions, while East Europe & Central Asia show much lower mean temperatures. In addition, standard deviations

RevistaEconomică
Journal of economic-financial theory and practice

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

underscore the internal diversity within each grouping. Even among countries labeled "poor" or "very_low" infrastructure, temperature and precipitation can vary widely. The data also suggest that infrastructure and governance categorizations align with particular climatic profiles, likely because higher-income or better-governed nations often lie in temperate areas. Overall, the metrics in Table 1 provide a quantitative baseline against which one can contextualize the figures illustrating emigration and climate interactions. By contrasting the mean and standard deviation for each category, the table indicates that emigration rates and climate variables cannot be viewed in isolation; rather, they are partially shaped by each group's economic conditions, institutional capacity, and geographical setting.

Figure 1: Average Emigration Rates by Wealth Status

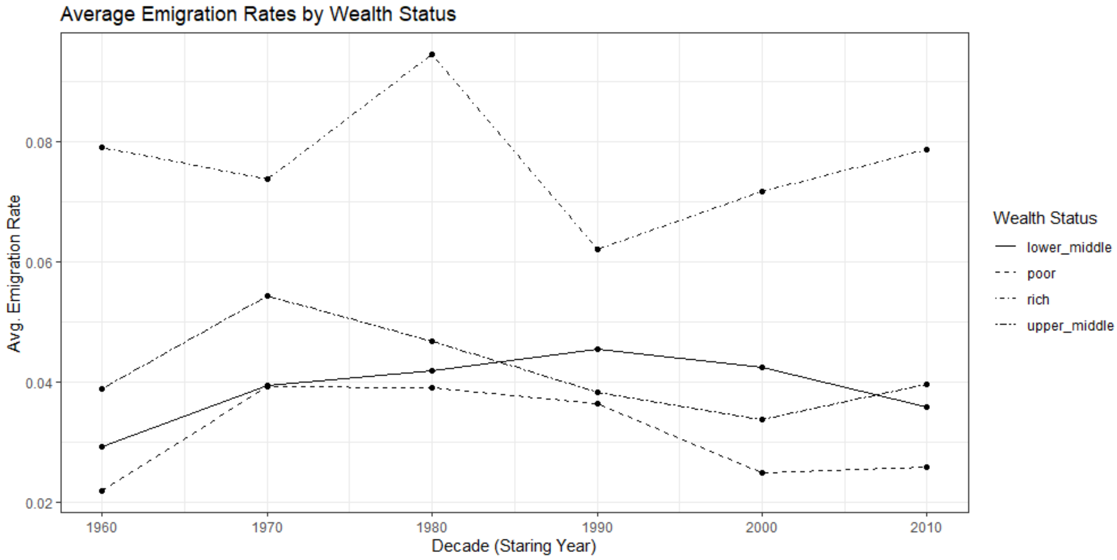


Figure 2: Average Emigration Rates by Infrastructure Category

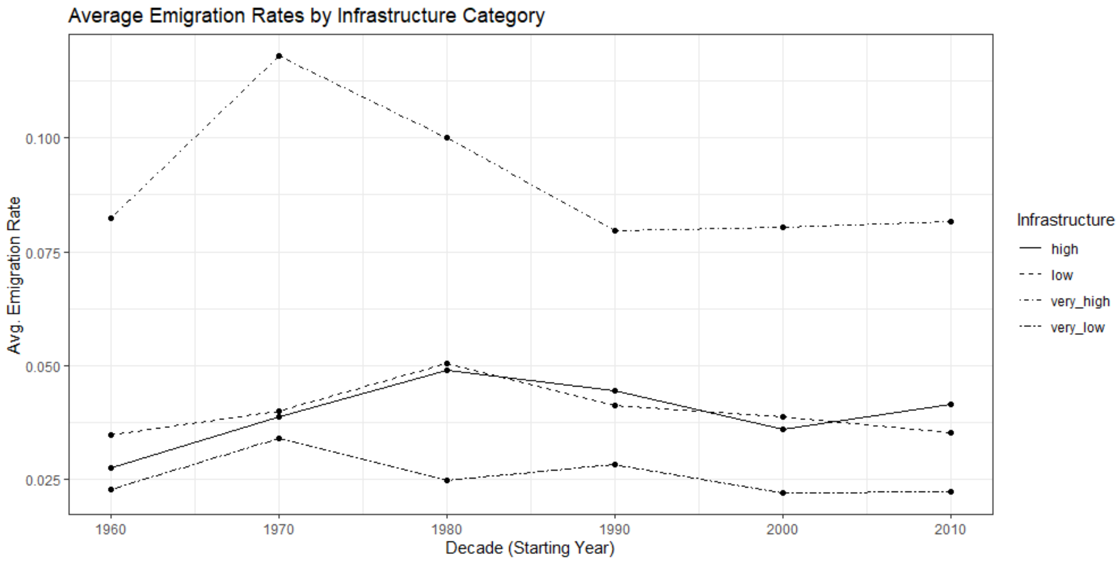


Figure 1 (Average Emigration Rates by Wealth Status) illustrates how the share of people leaving their country each decade differs among four income-based groups: "poor," "lower_middle," "upper_middle," and "rich." In this chart, each category's emigration rates are tracked over time in decadal intervals, allowing one to discern whether poorer countries see higher baseline outflows or if middle-income states experience stronger fluctuations. A first observation might be that the "poor" group, which generally faces greater resource constraints, sometimes shows more volatile patterns;

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

however, these can also reflect disruptive events—be they economic, political, or environmental. The "lower_middle" category may, in some periods, display growing emigration rates, suggesting that a modest increase in national income can actually enable more individuals to move abroad if underlying push factors (like diminishing agricultural yields or limited job opportunities) persist. Meanwhile, "rich" countries often register lower emigration rates overall, which can indicate greater adaptive capacity to both economic volatility and climate stress. Taken together, Figure 1 suggests that income level provides only part of the explanation for international mobility; while severe poverty can limit movement due to lack of resources, slight improvements in wealth sometimes foster higher mobility, and genuinely high-income contexts often see relatively stable or low rates of out-migration.

Figure 2 (Average Emigration Rates by Infrastructure Category) focuses on how the quality of a nation's infrastructure—rated "very_low," "low," "high," or "very_high"—relates to decadal emigration trends. Infrastructure includes elements such as transportation networks, electricity supply, and logistics capacity, all of which can shape how individuals interact with both local and international markets. Countries with "very_high" infrastructure levels may retain a larger share of their population because improved transport, communication, and utilities can stimulate local opportunities; yet better infrastructure can also facilitate outbound travel if other push factors remain unaddressed. Conversely, "very_low" or "low" infrastructure might exacerbate vulnerability to environmental hazards—if roads and irrigation systems fail, agricultural communities could face accelerated livelihood loss, and those who manage to gather resources may opt for international migration. Consequently, Figure 2 highlights that while robust infrastructure can mitigate some push factors through better living standards, it might simultaneously enable outflows if fundamental economic or climatic pressures persist.

Figure 3: Average Emigration Rates by Employment % in Agriculture

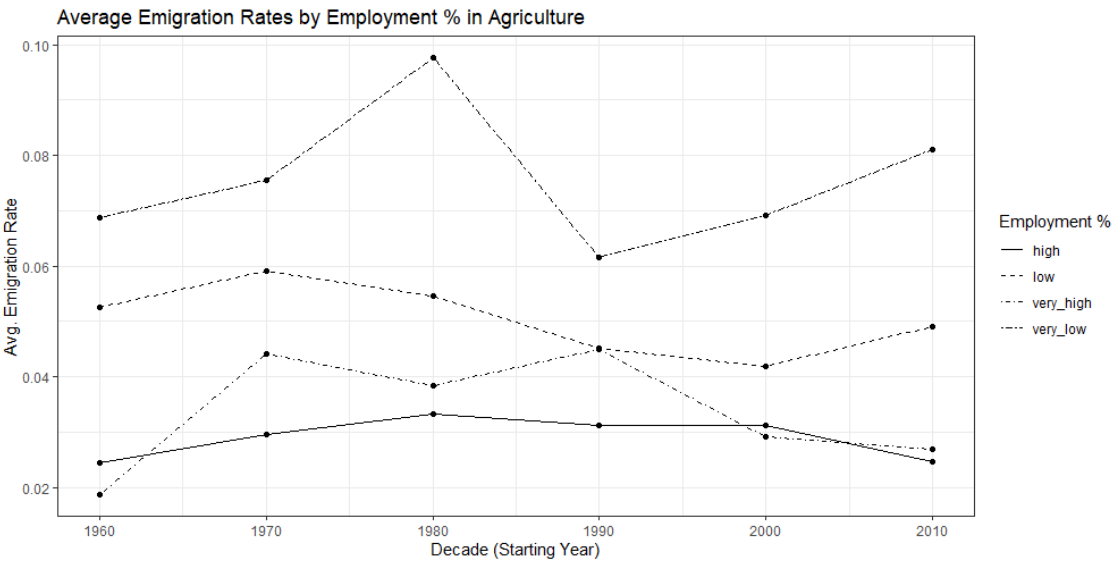


Figure 3 (Average Emigration Rates by Employment % in Agriculture) presents how strongly a country's labor force is tied to agriculture and how that correlates with people deciding to leave over a given decade. Where agricultural employment is high, communities often depend heavily on stable weather, water availability, and market conditions for their livelihoods. Repeated climate shocks—like droughts or floods—can prompt out-migration if families cannot sustain income from farming. On the other hand, extremely high agricultural employment might also correlate with lower overall resources and weaker mobility networks, limiting the feasibility of international migration for some. Therefore, this figure suggests that agriculture-reliant economies may be particularly prone to climate-driven push factors—though the actual movement observed also hinges on whether households possess the financial and informational means to emigrate.

RevistaEconomică
Journal of economic-financial theory and practice
Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

Figure 4 (Average Emigration Rates by Governance Indicators) highlights how institutional quality and political stability shape a country's decadal out-migration. The graph categorizes countries by governance tiers ("very low," "low," "high," "very high"), often derived from measures like rule of law, government effectiveness, and control of corruption. Better governance can facilitate internal adaptation to economic or environmental pressures, potentially reducing the impetus to leave. In contrast, prolonged governance deficits may intensify the impact of climate shocks if public services, disaster response, and social safety nets are ineffective, thereby amplifying out-migration pressures. This figure thus underscores the notion that climate-related stress does not operate in a vacuum; rather, it interacts with institutional capacity, which can either buffer or exacerbate the decision to migrate abroad.

Figure 4: Average Emigration Rates by Governance Indicators

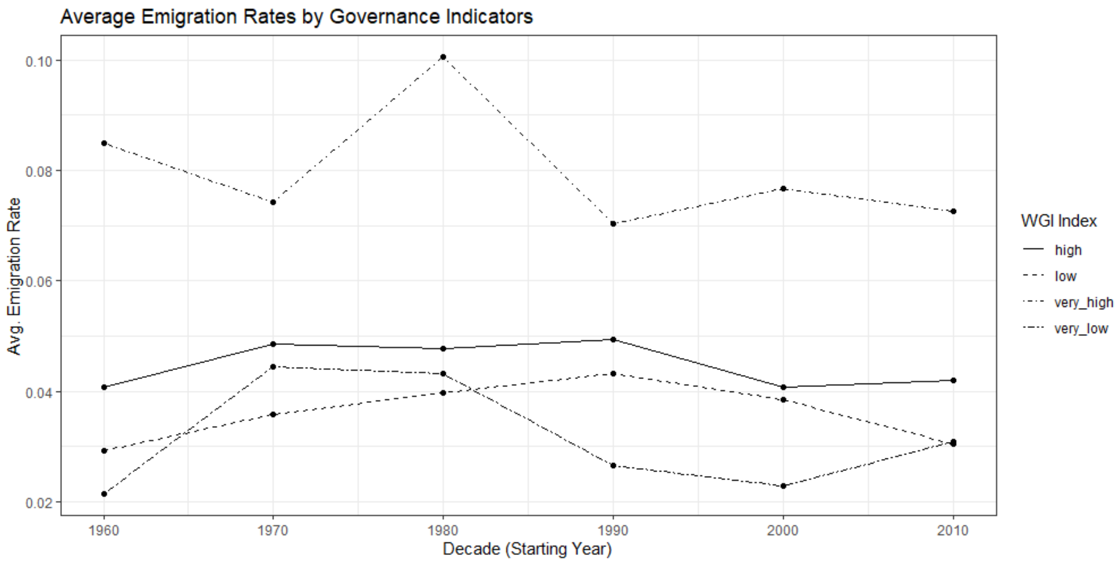


Figure 5: Average Emigration Rates by Agriculture % GDP

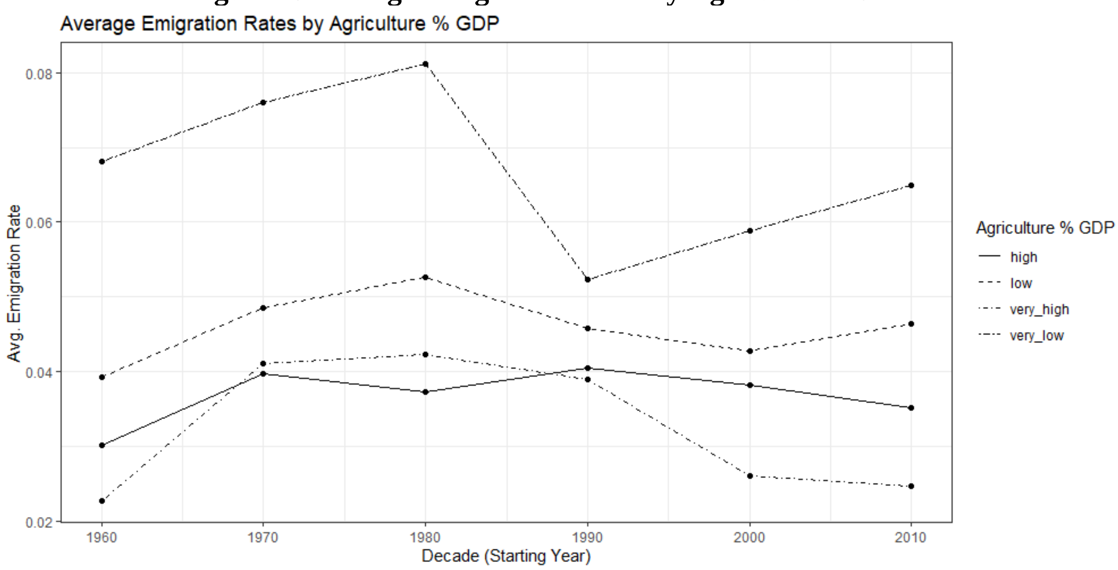


Figure 5 (Average Emigration Rates by Agriculture % GDP) adds another layer to the agricultural dimension by depicting how countries differ based on the share of agriculture in their overall economic output. While Figure 3 examined labor forces, Figure 5 focuses on agriculture's weight in GDP. Countries where farming contributes heavily to national income might see heightened emigration when global commodity prices drop or when local harvests fail due to weather extremes. If the

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

agricultural sector remains under-mechanized or vulnerable to rainfall variability, entire communities may lose their primary income source, prompting individuals to seek livelihoods elsewhere. Conversely, economies with low agricultural GDP shares and more diversified industries might cushion climate or economic shocks, yielding comparatively lower out-migration rates. By contrasting Figures 3 and 5, one can also discern whether labor-intensity or overall sectoral contribution better aligns with emigration outcomes.

Figure 6: Average Emigration Rates by Region

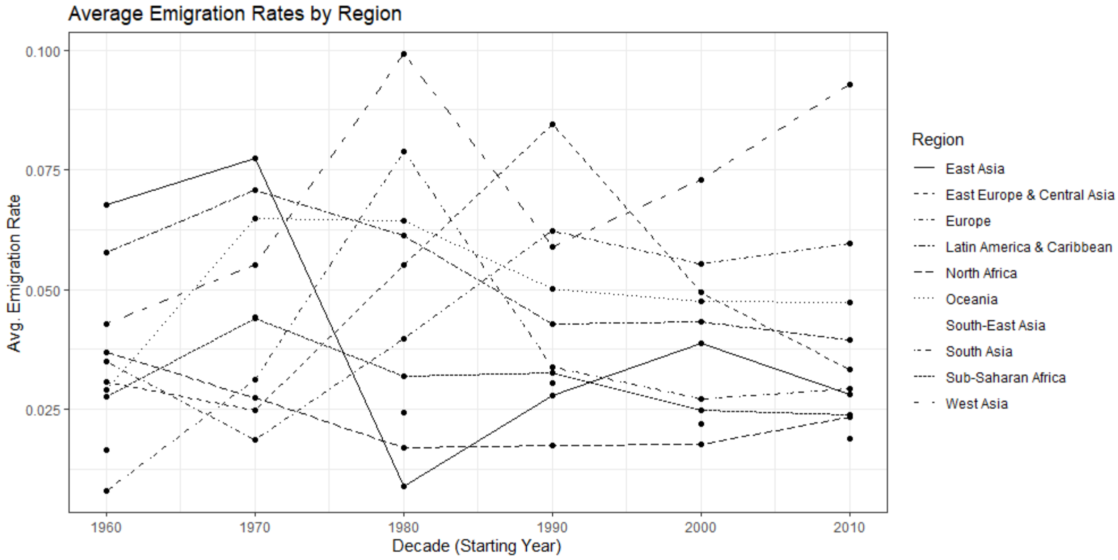


Figure 6 (Average Emigration Rates by Region) compares decadal out-migration rates on a broader geographical scale, grouping countries into major world regions such as Africa, Asia, Latin America, or Europe. This perspective reveals large-scale patterns that might reflect both regional climatic conditions and region-specific historical trajectories in migration networks. Regions like Sub-Saharan Africa or parts of South Asia, for instance, can exhibit higher or more variable emigration rates because of compounding factors: rapid population growth, limited non-farm employment, and frequent environmental events. Meanwhile, other regions with more stable or temperate climates could report lower average rates unless they confront acute upheavals, political crises, for example, can overshadow purely climate-driven motives. Thus, the figure prompts a focus on how socio-political and geographical contexts interplay with evolving environmental conditions to produce distinct regional migration patterns.

RevistaEconomică
Journal of economic-financial theory and practice
Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

Figure 7: Average Temperature by Wealth Status

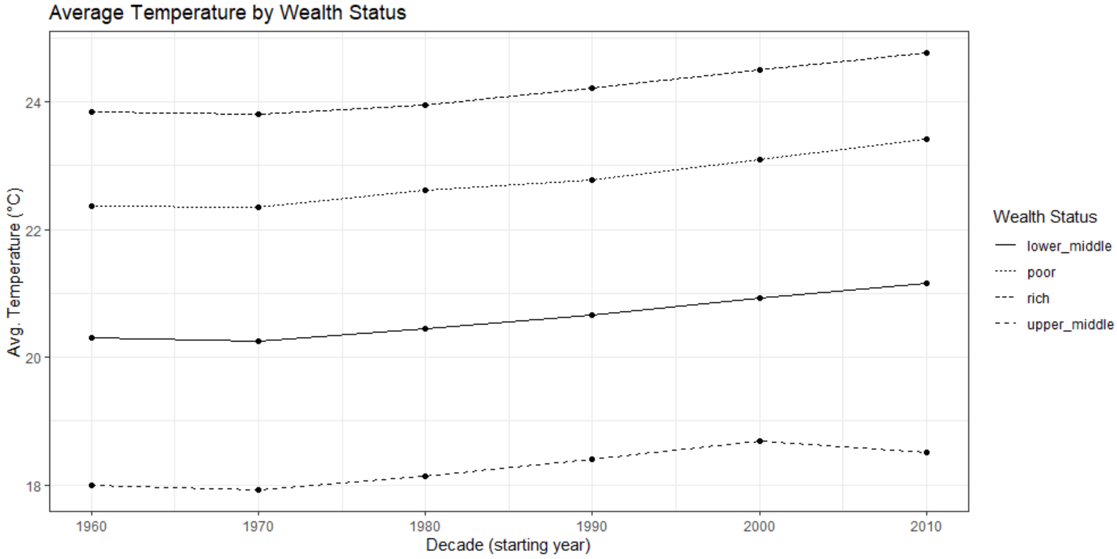


Figure 7 (Average Temperature by Wealth Status) shifts attention from emigration rates to climate variables, illustrating the decadal average temperatures in the same four wealth categories used in Figure 1. This comparison reveals whether poorer nations, which might already struggle with adaptation, are also situated in hotter or rapidly warming zones. If “poor” countries are indeed experiencing higher baseline temperatures or more pronounced warming trends over time, it raises the concern that they face a disproportionate share of climate stress while lacking extensive coping mechanisms. Conversely, “rich” or “upper_middle” groups, often located in cooler temperate zones, may be insulated from some of the immediate heat effects—although long-term global warming could still challenge even these countries’ agricultural and infrastructure resilience. Overall, Figure 7 supplies a crucial piece of the puzzle by situating different income groups within distinct thermal landscapes, hinting at how temperature overlaps with socio-economic contexts to shape potential migration pressures.

Figure 8 (Average Temperature by Region) provides a regional climate lens, showing decadal average temperatures for world regions over time. It allows for a straightforward comparison of which parts of the globe are inherently hotter, which regions are warming the fastest, and how these temperature profiles might coincide with the emigration patterns in Figure 6. Regions in tropical latitudes (e.g., parts of Africa or Southeast Asia) may show persistently high temperatures, reinforcing concerns about crop viability and water scarcity, while more temperate regions might only recently have begun experiencing noticeable warming trends. By aligning the regional temperature curves with the regional emigration rates, one can hypothesize whether climate variables plausibly contribute to mobility decisions. Ultimately, this figure bolsters the argument that climate stress—particularly in regions with inadequate adaptation capacity—can be a significant component in shaping longer-term migration flows, though the precise impact still depends on economic structure, governance quality, and existing migration corridors.

RevistaEconomică
Journal of economic-financial theory and practice
Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

Figure 8: Average Temperature by Region

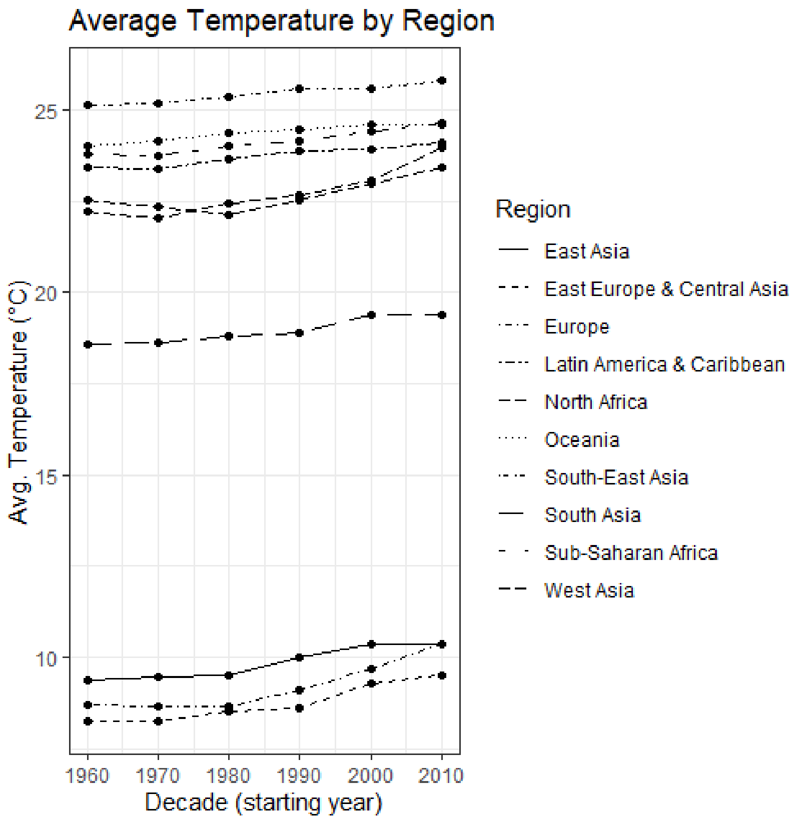


Figure 9 (Changes in Emigration Rates and Temperature – Middle-Income Countries) shows a scatter plot in which each point represents a country's change in the log of its average temperature on the x-axis and the corresponding change in the log of its emigration rate on the y-axis, focusing exclusively on countries in the "middle-income" category. A fitted trend line with a shaded confidence interval suggests a positive though modest association between rising temperature and increasing emigration. In other words, for middle-income countries, an incremental rise in temperature appears linked to a slight upward shift in emigration rates. However, the points are somewhat scattered around the regression line, indicating that while temperature changes may align with migration trends, other factors (such as governance, infrastructure, or macroeconomic conditions) likely modulate the ultimate decision to emigrate. Nonetheless, the positive slope supports the view that modestly higher incomes, compared to very low-income contexts, might enable populations to move abroad when confronted with environmental stresses.

Figure 10 (Changes in Emigration Rates and Temperature – Poor Countries) presents the same scatter-plot concept but for "poor" countries only. Again, the horizontal axis shows changes in log(Temperature), and the vertical axis shows changes in log(Emigration Rate). Here, the fitted trend line is flatter: while it leans slightly upward, the wide confidence interval indicates a weak or statistically uncertain correlation. This outcome may suggest that in very poor countries, rising temperatures do not straightforwardly translate into higher emigration. One plausible explanation is the well-known "migration paradox": extremely low income can constrain mobility, individuals most affected by environmental deterioration might also lack the resources or networks to migrate internationally. Alternatively, poor governance conditions, conflict, or domestic displacement could overshadow a simple temperature–migration link. Overall, the figure underscores that the climate–migration nexus varies substantially depending on a country's economic capacity.

RevistaEconomică
Journal of economic-financial theory and practice

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

Figure 9: Changes in Emigration Rates and Temperature (Middle Income)

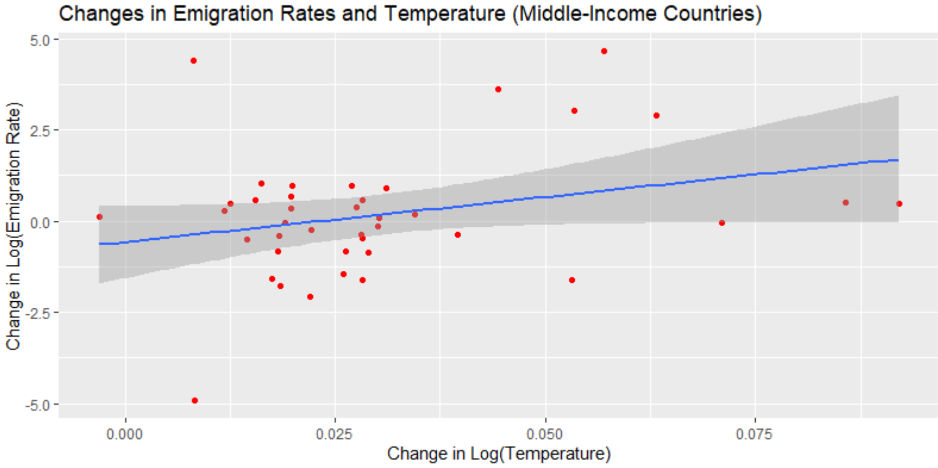
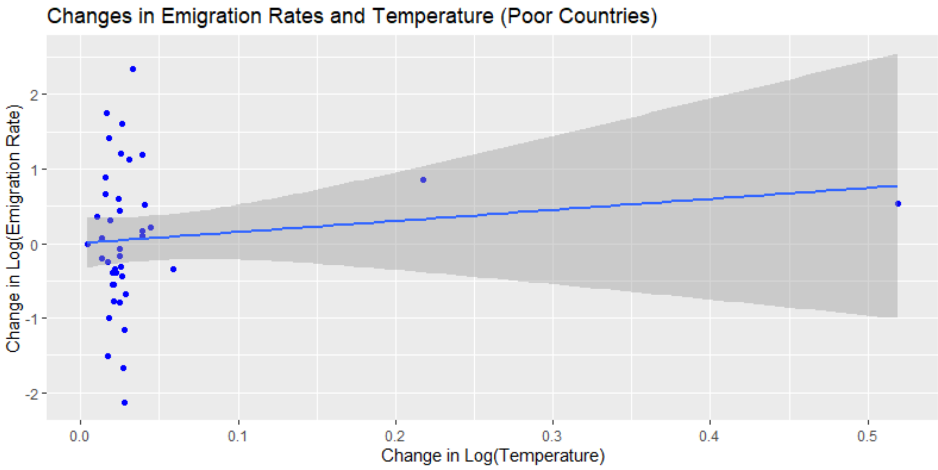


Figure 10: Changes in Emigration Rates and Temperature (Poor Income)



The descriptive evidence from Figures 1–10 and Table 1 underscores the significant heterogeneity in how climate factors, wealth status, governance, infrastructure, and agricultural dependence interact to shape migration flows. These patterns highlight the need for analytical approaches that can capture the complexity of nonlinearities, threshold effects, and rich interactions in the data. A purely linear specification may overlook crucial dynamics, which justifies the parallel use of a Fixed-Effects econometric model and a Random Forest machine-learning method in the subsequent analysis. Table 1's baseline distributions further provide critical context for interpreting any resulting coefficients or variable-importance measures, ensuring that findings are grounded in an accurate understanding of underlying group differences.

Several overarching insights emerge clearly. First, wealth status is a key moderator of climate-driven migration. While poor countries often face severe environmental stress, constrained resources can trap populations, limiting their capacity to move. In contrast, middle-income countries may exhibit more pronounced outflows when confronted with rising temperatures or erratic precipitation, as their populations possess slightly greater means and networks to migrate. Second, infrastructure quality and agricultural dependence mediate climate impacts in multifaceted ways: vulnerable infrastructure or heavy reliance on agriculture can heighten emigration incentives during climate shocks, but technological advancement or robust governance structures can mitigate these pressures. Third, regional variation reveals that baseline temperature differences, historical migration patterns, and demographic pressures also influence how climate-induced risks translate into international outflows. Finally, scatterplots linking changes in emigration rates to changes in temperature illustrate that the relationship can be positive and moderately strong in some middle-income contexts, yet substantially weaker in poor contexts, reflecting the nonlinearity of the climate–migration nexus.

In conclusion, these descriptive patterns strongly suggest that climate change does shape migration, but that its influence operates through multiple channels and is profoundly conditioned by each country's socio-economic and institutional environment. Poor governance, weak infrastructure, and heavy agricultural reliance can amplify climate-related risks, whereas extremely limited means can simultaneously inhibit external mobility. Meanwhile, societies with moderate wealth and institutional capacity are more prone to experience a measurable uptick in international migration as temperatures rise or precipitation patterns shift. Building on this foundation, the upcoming formal analyses, employing both Fixed-Effects panel regressions and Random Forests, will probe the statistical significance and relative importance of climate variables in comparison to governance, infrastructure, and economic factors, allowing for a deeper understanding of the complex interplay driving global migration flows.

4. Methodological Framework

The methodological approach of this study relies on two complementary techniques to investigate how climatic changes, socioeconomic structures, and institutional conditions interact to shape international emigration patterns. First, a classical econometric model is employed, capturing linear relationships and allowing for theoretically specified interactions between climate variables and contextual factors such as governance, infrastructure, and income levels. This approach enables transparent hypothesis testing and provides interpretable coefficients that clarify how each variable contributes to migration outcomes. Second, a random forest model serves as a machine-learning tool designed to detect nonlinearities and hidden patterns in the data without necessitating rigid functional forms. By pairing the hypothesis-driven clarity of econometric modeling with the flexibility of a random forest, this dual framework yields both interpretable statistical estimates and data-driven insights into the complex mechanisms linking climate change to global migration flows.

4.1 Fixed-Effects-Modell

The econometric model is based on a fixed-effects panel approach with time effects to control for unobserved heterogeneity between countries and global trends over the decades. The dependent variable is the emigration rate, defined as the proportion of individuals who emigrated during a given decade relative to the country's average population in that period. This regression model considers 133 countries, as sufficiently complete datasets for the required variables are available for these countries across the relevant decades. Formally, the fixed-effects model can be represented as follows. Let i denote the country and t the decade. The regression equation is:

$$(1)\ EmigrationRate_{i,t} = \alpha_i + \gamma_t + \beta_1 AvgTemperature_{i,t} + \beta_2 AvgPrecipitation_{i,t} + \sum_k \delta_k X_{i,t}^{(k)} + \sum_j \theta_j \left(Z_{i,t}^{(j)} \times Climate_{i,t}\right) + \varepsilon_{i,t}$$

Here, α_i represents country-specific effects (fixed effects per country), and γ_t denotes decade fixed effects that capture global temporal trends and shocks. $AvgTemperature_{i,t}$ and $AvgPrecipitation_{i,t}$ are the climatic indicators described in the data section, which record decadal average values of temperature and precipitation. The vectors $X_{i,t}^{(k)}$ include socioeconomic indicators such as GDP per capita (in constant 2015 US$), the share of employment in agriculture, and the share of agriculture in GDP. These variables quantify the agricultural dependence and overall wealth level of the

RevistaEconomică
Journal of economic-financial theory and practice
Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

country. A high degree of agricultural dependence makes countries more vulnerable to climate fluctuations because crop failures or water shortages lead directly to income losses.

Additionally, governance and infrastructure categories are incorporated into the model. The vector $Z_{i,t}^{(j)}$ represents categorical indicators of institutional quality and infrastructure availability, defined during the reference decade 2010–2019 and then applied retroactively to earlier decades as explained in the data section. For example, $Z_{i,t}^{(j)}$ can indicate the country's governance status (e.g., "low" or "very_high") or the infrastructure category. Interaction terms such as $Z_{i,t}^{(j)} \times Climate_{i,t}$ examine whether rising temperatures or changing precipitation have stronger or weaker effects on the emigration rate in certain institutional or infrastructural contexts. These interactions are based on the theoretical expectation that climate does not act in isolation but is particularly migration-inducing where poor governance or inadequate infrastructure make adaptation more difficult.

The estimation is carried out using panel regression, with $\varepsilon_{i,t}$ as the error term, and cluster-robust standard errors are used to control for potential heteroskedasticity and autocorrelation in the country data. Country fixed effects α_i eliminate time-invariant country-specific characteristics, and decade fixed effects γ_t f capture global temporal trends and shocks. This model allows direct interpretation of the coefficients: for example, a significant positive interaction term between high agricultural dependence and rising temperature suggests that climate change leads to increased emigration in such countries.

4.2 Random-Forest-Model

While the fixed-effects model focuses on theoretically guided interactions and linear relationships, the second methodological approach, the random forest, is rooted in data-driven pattern recognition. This ensemble model, constructed from numerous decision trees, analyzes the same explanatory variables (climate, socioeconomic indicators, governance, and infrastructure categories) but imposes no assumptions about their functional relationships to the emigration rate. Instead, it generates multiple trees from bootstrap samples and randomly selected variables at each split, estimating the emigration rate separately and averaging the outcomes to produce a final prediction. Formally, the Random Forest model can be defined as follows:

Let $\{(x_i, y_i)\}_{i=1}^{n}$ be the dataset of n observations, where y_i denotes the emigration rate (the dependent variable) for the *i*-th country and $x_i \in R^P$represents the vector of explanatory variables (temperature, precipitation, governance indicators, etc.). A random forest is composed of *B* individual decision trees, each trained on a different bootstrap sample of the data and evaluating a random subset of predictors at each split. From the full dataset, each tree $\hat{f}_b$ (for *b* =1,…, *B*) is trained on a bootstrap sample drawn with replacement. At each split in a decision tree, only a randomly selected subset of the pp available variables is considered as possible split candidates. Each decision tree $\hat{f}_b(x)$ then yields a prediction $\hat{f}_b$ for an input vector x. The final random forest prediction is the average over all *B* trees:

$$(2)\ \widehat{EmigrationRate}_{RF}(x) = \frac{1}{B}\sum_{b=1}^{B} \hat{f}_b(x)$$

Here, $\hat{f}_b(x)$ is the estimate of the emigration rate given by the *b*-th decision tree. In regression tasks (such as predicting a rate), each leaf node of a decision tree typically holds the mean of the target variable y_i (here, the emigration rate) for the observations in that leaf. Formally, if R_{bj}denotes the *j*-th leaf region in tree *b* and $\underline{y}_{bj}$ is the mean of y_i in that region, then:

$$(3)\ \hat{f}_b(x) = \sum_{j=1}^{M} \left(\underline{y}_{bj} 1_{\{x \in R_{bj}\}}\right),$$

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

where $1_{\{x \in R_{bj}\}}$ is an indicator function that equals 1 if x falls into region R_{bj} and 0 otherwise, while M is the number of leaf nodes in tree b.

The key advantage of the random forest lies in its ability to capture complex nonlinearities, interactions, and threshold effects without predefined interaction terms. Unlike the linear model, it requires no assumptions about the shape of the underlying relationships. Moreover, the random forest provides variable-importance metrics that indicate how much each explanatory factor contributes to reducing the prediction error. This makes it possible to identify particularly influential elements in the global dataset, even if those influences would only emerge through intricate interactions in a linear specification.

For the random forest approach, a slightly larger sample can be employed because limited imputation of missing values is feasible. As a result, this method involves 159 countries, compared to the 133 used in the regression model, thereby increasing the likelihood of detecting meaningful patterns. Its performance is evaluated by measuring the explained variance on a test sample. A portion of the data is devoted to training and the remainder to testing, creating a framework for comparing the predictive power of the random forest, gauged by explained variance or adjusted R^2 with that of the linear model. The objective is not to deem the random forest superior or inferior overall but rather to harness the strengths of both approaches. Without linear constraints, the random forest may reveal additional relevant variables or patterns, and certain factors may consistently exhibit strong influence, independent of theoretical assumptions.

Both models incorporate a common set of explanatory variables, including climate factors (temperature, precipitation), economic indicators (GDP per capita, agricultural employment share, agriculture's contribution to GDP), institutional quality (governance indicators from WGI data), and infrastructure characteristics (electricity access, transport connectivity, Logistics Performance Index). These variables are supplemented by regional and temporal controls to account for heterogeneity across countries and time periods.

While the Fixed-Effects regression model emphasizes theoretically grounded interactions, particularly between climate, governance, and infrastructure, the Random Forest model excels at identifying nonlinear relationships and key predictors without the need for predefined interaction terms. Together, these methods provide complementary insights, the econometric model offers interpretable, hypothesis-driven results, while the machine learning approach detects hidden patterns and complex interdependencies.

By integrating both approaches, this study examines whether nonlinear effects go undetected in the linear framework, whether the Random Forest confirms the significance of key variables (e.g., agricultural employment share) identified in the regression model, and whether the most influential predictors in the machine learning approach align with those found significant in the econometric model. This dual-method strategy strengthens the interpretative foundation and provides a more comprehensive understanding of climate-induced migration drivers.

Ultimately, this combined approach not only validates theoretically expected relationships but also uncovers new and sometimes unexpected patterns in the data. The Fixed-Effects model offers a structured, hypothesis-driven assessment, while the Random Forest model refines and extends these findings by highlighting the predictive importance of key variables on a global scale. Together, they enhance the depth, robustness, and credibility of the empirical analysis.

5. Results

After presenting the methodological approach and providing a descriptive overview of the data, the following section offers the presentation and interpretation of the central findings. These encompass the results of a fixed-effects panel model (Table 2) as well as those from a random forest model (Table 3). Subsequently, robustness tests and sensitivity analyses are discussed to support the stability and validity of the conclusions.

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

5.1 Interpretation of the Fixed-Effects Results

The fixed-effects regression results (Table 2) confirm that both temperature and precipitation exert a statistically significant influence on emigration rates. Specifically, the variable Avg_Temperature is significant at the 5% level, while Avg_Precipitation also produces a significant coefficient. These findings align with the theoretical expectation that climatic changes—rising temperatures or altered precipitation patterns, affect long-term economic and social conditions in ways that can heighten the likelihood of emigration. For example, reduced agricultural productivity due to heat stress or rainfall anomalies may drive income losses and thus intensify out-migration pressures.

Beyond these main climate variables, the model reveals that several contextual factors moderate the climate–migration relationship. Categories such as Agriculture_High_Very_High or Employment_High_Very_High as well as markers of limited infrastructure (Infrastructure_Low_Very_Low) or weak governance (WGI_Low_Very_Low), interact with temperature and precipitation in ways that either amplify or diminish the net effect on emigration. Notably, one interaction term, Avg_Temperature:Employment_High_Very_High, is both highly significant and negative. This does not necessarily mean that warming temperatures lower emigration in these heavily agricultural contexts; rather, it indicates that the marginal effect of temperature on emigration is less pronounced - or potentially reversed - relative to the baseline category. In practice, this outcome could reflect specific characteristics of agrarian communities, such as reliance on subsistence production or communal coping strategies, which might delay or modify how climate stress translates into migration decisions.

The presence of multiple interaction terms underscores that climate-induced migration is context-dependent rather than uniform across all countries. Indeed, migration arises from the interplay of climatic stressors, economic structures, governance levels, and regional-historical patterns. While the number of included interactions can introduce multicollinearity, a recognized challenge that complicates the isolated interpretation of individual coefficients, this complexity ultimately mirrors the real-world environment in which many overlapping factors shape migration. Consequently, the model's moderate R^2 values (around 0.33, with an adjusted R^2 near 0.25) are realistic in light of the diverse forces driving international mobility. Significance in key core variables and interactions nonetheless indicates that the specification captures critical dynamics.

It is also important to emphasize that statistical significance does not automatically convey effect size. Even if a coefficient is highly significant, its real-world impact might be modest. Conversely, a larger but less precisely estimated coefficient might have meaningful implications for certain countries or time periods. Interpreting these results—particularly for complex interaction terms—therefore benefits from considering both the estimated magnitude of the effects and their standard errors. Finally, it will be informative to compare these econometric findings with evidence from the random forest model and subsequent robustness tests, so as to gauge whether nonlinearities or threshold effects might refine or reinforce the conclusions drawn here.

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

Table 2: Fixed-Effects Panel Regression

	Dependent variable:
	Emigration_Rate
Avg_Temperature	0.003**
	(0.001)
Avg_Precipitation	0.0003**
	(0.0001)
Agriculture_High_Very_High	-0.100***
	(0.037)
Infrastructure_Low_Very_Low	-0.018
	(0.049)
Wealth_Poor_Lower_Middle	0.052
	(0.047)
WGI_Low_Very_Low	0.021
	(0.025)
Employment_High_Very_High	0.116***
	(0.040)
Avg_Temperature:Agriculture_High_Very_High	0.003*
	(0.002)
Avg_Precipitation:Agriculture_High_Very_High	-0.0002
	(0.0001)
Avg_Temperature:Infrastructure_Low_Very_Low	0.001
	(0.002)
Avg_Precipitation:Infrastructure_Low_Very_Low	-0.0001
	(0.0001)
Avg_Temperature:Wealth_Poor_Lower_Middle	-0.001
	(0.002)
Avg_Precipitation:Wealth_Poor_Lower_Middle	-0.0001
	(0.0002)
Avg_Temperature:WGI_Low_Very_Low	-0.001
	(0.001)
Avg_Precipitation:WGI_Low_Very_Low	-0.0001
	(0.0001)
Avg_Temperature:Employment_High_Very_High	-0.005***
	(0.002)
Avg_Precipitation:Employment_High_Very_High	0.0002*
	(0.0001)
Observations	776
R^2	0.330
Adjusted R^2	0.248
F Statistic	4.304*** (df = 79; 691)
Note:	*p<0.1; **p<0.05; ***p<0.01

Overall, the results suggest that temperature and precipitation shifts play an important role in shaping emigration decisions, but that this influence is heavily mediated by each country's economic orientation, institutional capacity, and infrastructure quality. As such, policies aiming to address or mitigate climate-related migration must account for these contextual differences, ensuring that interventions are tailored to the specific vulnerabilities and adaptation options present in different settings.

5.2 Random Forest Results and Comparison

Table 3 displays the random forest results for Emigration_Rate as the target variable, alongside two key importance measures—IncNodePurity and IncMSE_Percent—as well as standard metrics such as mean squared error (MSE) and explained variance. The MSE of approximately 0.002 and an explained variance of roughly 24 % align closely with the performance of the fixed-effects model. This suggests that, despite its flexible, data-driven nature, the random forest arrives at a broadly comparable explanatory power. In other words, the interactions and variables identified in the econometric approach also appear meaningful under a more agnostic modeling strategy. In Table 3, each row corresponds to a variable used in the random forest, including climate indicators (e.g., *Avg_Temperature*, *Avg_Precipitation*), economic measures (*GDP_Per_Capita_2015*, *Agriculture_GDP_Percentage*), institutional factors (*WGI_Index*, *Infrastructure_Index*), and time/region dummies. Two columns highlight how each predictor influences the model's accuracy:

IncNodePurity: Often associated with the improvement in splitting criteria (e.g., Gini impurity in classification or variance reduction in regression) each time a given variable is used in a decision node. Higher values suggest that the variable plays a larger role in reducing prediction errors across the ensemble of trees.

IncMSE_Percent: Represents the percentage increase in mean squared error if the model no longer has access to that particular variable. A high percentage indicates that excluding the variable substantially worsens predictive performance, underscoring its importance to the overall model.

From these measures, it becomes evident that Employment_Agriculture_Percentage, Infrastructure_Index, and Agriculture_GDP_Percentage are the top three drivers of the random forest's predictive accuracy. All three show relatively high IncNodePurity and large IncMSE_Percent values, confirming that agricultural and infrastructural factors significantly shape global emigration patterns, even in a model with no predefined linear constraints. The predictors can be interpreted as follows:

Employment_Agriculture_Percentage (top-ranked predictor): This finding aligns with the fixed-effects regression, where high agricultural employment often interacted significantly with climate variables. A plausible interpretation is that agrarian livelihoods are particularly vulnerable to climate stress, making the share of agricultural workers a strong indicator of potential out-migration under deteriorating environmental conditions. However, as the negative interaction terms in the panel model suggested, the relationship can be context-dependent—some heavily agrarian regions may not necessarily translate climate shocks into large-scale emigration if resource constraints limit travel.

Infrastructure_Index (second-highest): Consistent with the fixed-effects results, infrastructure quality strongly correlates with emigration outcomes. Inadequate roads, electricity access, or logistics might amplify the negative effects of climate stress on local livelihoods, compelling individuals to relocate internationally. Conversely, robust infrastructure can strengthen local adaptation measures, though it can also facilitate overseas migration by reducing the transaction costs of moving. The random forest's high importance ranking suggests that these infrastructural dynamics interact in complex, possibly nonlinear ways with other socio-economic variables.

Agriculture_GDP_Percentage (third-highest): Similar to agricultural employment, a large agricultural share of GDP highlights dependence on climate-sensitive sectors. This factor's prominence in the random forest mirrors the fixed-effects model's finding that countries heavily reliant on agriculture are generally more exposed to climatic fluctuations, though, once again, the precise direction of the effect could vary by local context.

WGI_Index and Avg_Temperature: Although their IncMSE_Percent values are lower than those of the top agriculture- or infrastructure-related predictors, they still appear reasonably influential. This suggests that governance quality (captured by the WGI_Index) and climatic conditions (in this case, average temperatures) do play meaningful roles, as the fixed-effects approach also indicated. In short, better governance may buffer populations from climate hazards, while rising temperatures add incremental pressure to agrarian economies.

The overall takeaway that agricultural dependence and institutional/infrastructural capacity are central correlates of emigration resonates strongly with the panel regression results. Both methods underscore that climate factors such as temperature and precipitation do not operate in isolation; rather, they shape migration decisions more forcefully when layered atop structural vulnerabilities—e.g., high agricultural shares and weak infrastructure. Consequently, the random forest corroborates the interpretive framework suggested by the linear model: climate-induced migration emerges from a combination of environmental stress and limited adaptive capacity.

Moreover, the random forest's comparable explained variance (24 % vs. roughly 25–30 % in the fixed-effects model) hints that the complex interactions the linear model explicitly posits (through numerous interaction terms) are indeed present in the data. The data-driven splits within the random forest confirm that such multi-variable interactions meaningfully boost predictive power.

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

Despite the broad agreement in top predictors, the random forest does not provide the straightforward coefficient estimates and significance tests available in fixed-effects regression. Instead, it highlights which variables matter most for predictive accuracy, potentially uncovering nonlinear relationships that a linear model might miss or have to specify manually. For instance, the partial dependence of emigration on *Employment_Agriculture_Percentage* could take a U-shaped curve, indicating different dynamics at low vs. high levels of agricultural employment, patterns that require extra steps to capture in a standard regression.

Additionally, the random forest does not explicitly control for unobserved country-specific effects; it instead treats each data point at face value and relies on repeated splits to isolate relevant patterns. This approach is great for detecting complexity but provides fewer causal inferences. By contrast, the fixed-effects model is better suited for claiming that "an X °C rise in temperature, ceteris paribus, corresponds to Y % greater emigration," given its within-country focus over time. Hence, each method yields unique benefits: one excels at interpretability and causal inference, the other at flexibility and pattern discovery.

In sum, Table 3 reaffirms the context-dependent nature of climate-driven migration, highlighting agricultural and infrastructural variables as pivotal in explaining observed emigration rates. The approach does not rely on predefined functional forms yet arrives at a similar overall explanatory power as the fixed-effects model and flags broadly similar drivers of emigration. This convergence bolsters confidence that the core relationships, particularly those centered on agriculture and governance, are not statistical artifacts but genuine patterns. It likewise underscores the value of pairing a traditional econometric framework with a machine-learning method: while the panel regression pinpoints linear and interaction effects with direct interpretability, the random forest uncovers latent nonlinearities and verifies that crucial variables consistently emerge as key predictors.

5.3 Defending the Model Specification and Robustness Checks

The comprehensive specification of the fixed-effects model may be seen as challenging due to the large number of interaction terms, making interpretation more complex. However, this complexity is theoretically grounded. There are compelling reasons to expect that climate impacts on migration vary with agricultural dependence, infrastructure quality, governance conditions, and overall economic development. Similarly, historical and regional path dependencies justify the inclusion of decade and regional dummies as well as their interactions to capture time- and place-specific effects.

Instead of simplifying the model and potentially missing key relationships, it is more appropriate to estimate a rich and detailed specification that provides a more realistic view of migration dynamics. Although the resulting multicollinearity and interpretive difficulties are acknowledged, the presence of complex patterns is neither accidental nor arbitrary. The data themselves likely contain intricate, interwoven relationships that warrant careful exploration. The random forest model indirectly supports this decision. Despite not relying on predefined interaction terms or linear functional forms, it achieves comparable explanatory power, indicating that the elaborate specification of the linear model is not merely overfitting random noise, but rather capturing genuine complexity.

While the study's core specifications are already extensive, some preliminary attempts were made to assess how the results would change with certain modifications, such as adjusting interaction terms or employing alternative dummy codings. In these tentative robustness checks, the central conclusions remained broadly consistent: Climate variables are significant, and their influence depends critically on the socioeconomic and institutional context. Although the magnitude of effects or the significance of certain interaction terms might vary in simplified versions, the fundamental insight that climate effects are context-dependent stands.

These initial robustness considerations lend some credibility to the conclusions, even if a systematic and exhaustive robustness analysis was not fully documented here. The similarity in variable importance identified by the random forest model also hints that the fixed-effects model's complexity is well-founded. Taken together, the detailed model specification and the preliminary robustness checks suggest that the results are not unduly driven by a particular set of modeling decisions. Instead, they reflect an underlying empirical reality in which climate, institutions, infrastructure, and economic structure jointly shape migration patterns in a complex and interdependent manner.

6. Discussion and Conclusion

The results presented in this study offer a detailed and context-specific portrayal of how climatic changes, in conjunction with socioeconomic and institutional factors, shape international migration flows. Both the fixed-effects panel model and the random forest analysis underscore that climate alone is not a monocausal driver of migration; rather, it acts through existing vulnerabilities such as those related to agricultural dependence, limited infrastructure, and weak governance. This aligns with earlier research (for instance, Black et al., 2011, and Cattaneo & Peri, 2016) and extends it by comparing traditional econometric methods to a flexible, nonlinear machine learning approach.

A striking outcome is how closely the two models converge on identifying agricultural variables (for example, Employment_Agriculture_Percentage or Agriculture_GDP_Percentage) and infrastructural measures (Infrastructure_Index) as critical predictors of migration. Despite relying on fundamentally different assumptions - one rooted in a linear, hypothesis-driven framework and the other in a data-driven ensemble algorithm - both approaches reveal that temperature and precipitation shifts affect migration most strongly where structural conditions involving agrarian livelihoods, poor infrastructure, and institutional fragility amplify climate stressors. This convergence substantially bolsters confidence that the relationships captured are robust rather than artifacts of any single modelling technique.

Yet, each method also contributes unique strengths to the analysis. The fixed-effects model offers interpretive clarity; coefficients can be directly read as semi-elasticities, facilitating policy-relevant questions such as “How does a 1 °C rise in temperature correlate with out-migration?” The random forest, on the other hand, excels at uncovering nonlinearities and threshold effects that might otherwise remain overlooked or improperly specified. For instance, it can detect whether emigration rates spike beyond certain critical values of temperature or agricultural dependence. By demonstrating that these more flexible patterns do not detract from the overall explanatory power, the random forest provides implicit validation of the fixed-effects model’s inclusion of numerous interaction terms. Consequently, model pluralism emerges as a valuable strategy because fixed-effects regression delivers theoretical grounding and cleaner causal inferences, while the random forest reveals hidden interactions and confirms that these complexities genuinely matter.

From a policy perspective, these findings reinforce the notion that tackling climate-induced migration requires a multi-dimensional approach. Investing in infrastructure emerges as a priority. Improving transportation networks, electricity access, water supply, and logistics can reduce the immediate shocks of climate hazards such as flooding or drought on rural livelihoods. At the same time, robust infrastructure may enable more people to migrate if conditions deteriorate significantly, illustrating that infrastructure alone does not eliminate migration pressures but may offer communities both better adaptation options and safer, more organized paths to relocate.

Institutional quality and governance consistently appear pivotal. Effective governance can deliver social safety nets, enforce property rights, manage resources, and deploy disaster response mechanisms. Where governance is weak, climate shocks more easily translate into livelihood crises, thereby pushing vulnerable populations across borders. Policy measures thus need to integrate capacity building in local institutions, ensuring transparency, corruption control, and administrative competence so that environmental pressures do not become unmanageable push factors.

Agricultural modernization remains critical in reducing vulnerability to temperature and precipitation fluctuations. Encouraging crop diversification, scaling up irrigation, and promoting climate resilient seed varieties can mitigate the direct impact of climate variability on rural households. Such efforts, however, should be coupled with economic diversification beyond agriculture to create broader employment opportunities. This diversification expands local adaptation capacity and curbs distress

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

driven out migration, aligning with the result that countries heavily reliant on agriculture are particularly susceptible to climate stress.

Climate policies such as reducing emissions or enhancing carbon sinks are central to the global effort against climate change, and long term development strategies that include education, healthcare, and inclusive governance are equally essential. The idea that "climate does not act alone" highlights how intertwined environmental risks are with broader developmental challenges. Consequently, climate migration policies must be interwoven with strategies that raise incomes, strengthen institutions, and invest in human capital, ensuring that migration becomes less of a forced reaction to climate extremes and more of a viable choice among many.

Although the two modeling strategies provide convergent results, several limitations remain. One concerns data granularity; aggregating variables at the country level conceals local heterogeneities, including urban rural disparities or intra country climate variations. Regions within the same country may experience drastically different climate regimes, economies, and governance standards, which can lead to distinct migration responses. Future research could employ subnational data, perhaps district or province level, to capture these nuances more precisely.

A second limitation involves temporal resolution. Using decadal averages reduces noise and addresses data gaps but may mask short term climate shocks or multi year adaptation patterns. More granular analyses that track annual or seasonal data could better identify acute events such as floods, hurricanes, or droughts, shedding light on how households or communities respond over shorter time scales. Such refinement might highlight rapid onset versus slow onset migration decisions and reveal whether machine learning methods detect different thresholds at finer time intervals.

A third limitation arises from the parallel use of a fixed-effects model and a random forest, which addresses many methodological questions but does not fully solve causal attribution of climate effects. In principle, exogenous instruments for climate variables or natural experiments could further bolster claims about causality. Additionally, expanding the model to include network effects such as diaspora communities or migration policy indicators in the host countries might clarify why some destinations attract more climate affected migrants than others.

Lastly, the study centers on international emigration, yet internal migration can be a dominant form of climate mobility, particularly in contexts where cross border movement is restrictive or financially unfeasible. Incorporating internal displacement data or city level records of in migration could provide a fuller account of the total migration response to climate stress.

In spite of these caveats, the combination of a theory informed fixed-effects approach and a flexible, data driven random forest provides robust evidence that climate driven migration is highly contingent on economic and institutional environments. When both methods converge on the critical role of agriculture, infrastructure, and governance, the result is a compelling argument for addressing climate resilience through broader development measures. Neither method on its own fully captures the multifaceted nature of migration, and this is why a pluralistic modeling perspective is recommended.

Hence, the overarching conclusion is that climate change amplifies existing structural vulnerabilities, prompting out migration where communities lack adaptive capacity. Better governance, higher infrastructure quality, and more diversified economies can mitigate the worst effects of climate pressure, though they may also empower individuals to move if local coping strategies prove insufficient. Policy efforts to tackle climate induced migration must therefore acknowledge this duality, integrating immediate adaptation measures with systemic governance and development reforms that address the root causes of vulnerability.

References

- Abel, G. J. (2017). Estimates of global bilateral migration flows by gender between 1960 and 2015. *International Migration Review, 51*(3), 828–852. https://doi.org/10.1111/imre.12327

Vol. 76, issue 4 Year 2024
DOI: 10.56043/reveco-2024-0037

- Abel, G. J., & Cohen, J. E. (2022). Bilateral international migration flow estimates updated and refined by sex. *Scientific Data, 9*(173), 1–13. https://doi.org/10.1038/s41597-022-01271-z
- Beine, M., & Jeusette, L. (2021). A meta-analysis of the literature on climate change and migration. *Journal of Demographic Economics, 87*(3), 293–344. https://doi.org/10.1017/dem.2019.22
- Black, R., Adger, W. N., Arnell, N. W., Dercon, S., Geddes, A., & Thomas, D. (2011). The effect of environmental change on human migration. *Global Environmental Change, 21*(Suppl. 1), S3–S11. https://doi.org/10.1016/j.gloenvcha.2011.10.001
- Borderon, M., Sakdapolrak, P., Muttarak, R., Kebede, E., Pagogna, R., & Sporer, E. (2019). Migration influenced by environmental change in Africa: A systematic review of empirical evidence. *Demographic Research, 41*, 491–544. https://doi.org/10.4054/DemRes.2019.41.18
- Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324
- Cattaneo, C., & Peri, G. (2016). The migration response to increasing temperatures. *Journal of Development Economics, 122*, 127–146. https://doi.org/10.1016/j.jdeveco.2016.05.004
- Coronese, M., Lamperti, F., Keller, K., Chiaromonte, F., & Roventini, A. (2019). Evidence for sharp increase in the economic damages of extreme natural disasters. *Proceedings of the National Academy of Sciences, 116*(43), 21450–21455. https://doi.org/10.1073/pnas.1907826116
- Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data, 7*, 109. https://doi.org/10.1038/s41597-020-0453-3
- Hoffmann, R., Dimitrova, A., Muttarak, R., Cuaresma, J. C., & Peisker, J. (2020). A meta-analysis of country-level studies on environmental change and migration. *Nature Climate Change, 10*, 904–912. https://doi.org/10.1038/s41558-020-0898-6
- IPCC. (2022). *Sixth Assessment Report (AR6).* Intergovernmental Panel on Climate Change. https://www.ipcc.ch/report/ar6/wg2/
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2010). *The Worldwide Governance Indicators: A summary of methodology, data and analytical issues* (World Bank Policy Research Working Paper No. 5430). World Bank. https://doi.org/10.1596/1813-9450-5430
- McLeman, R. (2018). Thresholds in climate migration. *Population and Environment, 39*(4), 319–338. https://doi.org/10.1007/s11111-017-0290-2
- Missirian, A., & Schlenker, W. (2017). Asylum applications respond to temperature fluctuations. *Science, 358*(6370), 1610–1614. https://doi.org/10.1126/science.aao0432
- Rigaud, K. K., de Sherbinin, A., Jones, B., Bergmann, J., Clement, V., Ober, K., Schewe, J., Adamo, S., McCusker, B., Heuser, S., & Midgley, A. (2018). *Groundswell: Preparing for internal climate migration*. World Bank. http://hdl.handle.net/10986/29461
- Todaro, M. P., & Maruszko, L. (1987). Illegal migration and US immigration reform: A conceptual framework. *Population and Development Review, 13*(1), 101–114. https://doi.org/10.2307/1972122
- World Bank. (2021). *World Development Indicators.* The World Bank. https://data.worldbank.org/