

Context matters: Local Determinants of International Emigration Patterns from Colombia

Alessandra Carioli,* Alessandra.carioli@ec.europa.eu European Commission – Joint Research Centre (JRC), Italy

Joaquin Recano, Universitat Autònoma de Barcelona and Centre d'Estudis Demogràfics (CED), Spain

1. Introduction and background

International migration is a multifaceted demographic process shaped by intertwined economic, social, and geopolitical factors. Individuals move across borders to seek improved living conditions, reunite with family members, or escape adverse circumstances such as conflict or environmental stress. According to the International Organization for Migration (2024), an estimated 281 million people reside in a country other than their country of birth—an upward trend from 173 million in 2000 and 102 million in 1980. Colombia fits well within this global context of increasing mobility. It has historically been both a country of origin and destination, currently hosting about 2.8 million immigrants while an estimated 3.3 million Colombians live abroad. Additionally, Colombia faces the world's highest number of internally displaced persons (7.7 million), resulting from decades of armed violence and natural disasters.

Despite this complexity, most empirical research continues to focus on national-level migration patterns, often overlooking the substantial heterogeneity that exists across subnational territories. Yet, migration determinants are inherently spatial: economic opportunities, social networks, and security conditions vary markedly across regions, generating distinct local configurations of emigration behaviour. In this context, Geographically Weighted Regression (GWR) provides a powerful analytical framework to capture spatial non-stationarity in the relationship between migration and its determinants. Unlike global models, GWR allows parameter estimates to vary locally, thus identifying territorial differences in the strength and direction of migration drivers

Building on this methodological approach, this paper examines the local determinants of international emigration from Colombia, with a specific focus on gender differences. We analyse male and female Colombian-born migrants aged 20–49 who left the country between 2017 and 2019, using municipal-level data (1,122 units) from the 2018 Colombian Census. The database integrates socio-economic indicators, migration history, poverty, forced displacement, and armed violence.

By providing a spatially explicit and gender-disaggregated analysis based on robust local regressors and precise measures of international emigration, this study offers an unprecedented contribution to understanding the local dynamics of international migration in a middle-income country with pronounced territorial inequalities.

2. Data and Methods

2.1. Data

We linked individual records from the Registro Estadístico de Migración Internacional (REMI) of the Departamento Administrativo Nacional de Estadística (DANE), which captures Colombian emigration flows between 2017 and 2019, with records from the 2018 Census. Through this linkage, we identified the places of residence and birth of individuals who emigrated from Colombia. This information allows us to estimate the emigration rate by place of residence for years 2017–2019. The linkage is carried out using shared identifiers and it is further validated with the variables of sex and date of birth. Subsequently, we compiled information on various variables. The aggregated data used as independent or predictive variables in the regression models initially included 21 variables and are derived from the 2018 Census microdata. The Unmet Basic Needs (UBN) index for 2005 and 2018, the Multidimensional Poverty Index, and the data on individuals employed in the formal sector (RELAB) were provided by the DCD of DANE. We also estimated local emigration rates for the population born in Colombia aged 20–49, by municipality of residence, whose denominators were obtained from an exploitation of the 2018 Census microdata. The administrative register of entries and exits maintained by Migración Colombia is a border control system based on information provided by travellers entering or leaving the country through Colombia's 100 air, land, maritime, and fluvial border posts.

The data compiled and linked for this study consists of 61 fields containing demographic, geographic, socio-economic, and migratory information. For the entire period from 2004 to 2020—the years for which data was available at the time of the study—the structure of the register adheres to the standard requirements for the estimation of migration statistics: demographic characteristics, temporal detail of migratory events, and geographic identification of the origin, destination, and place of residence. In the estimation of emigration and immigration flows developed within the REMI framework, the

recommendations of the United Nations have been applied adapted here to the Colombian context (United Nations, 1999). This information allows to aggregate up to the municipality of residence (n=1122) of the respondents.

2.2. Statistical methods

This study employs a geographically weighted regression model, (GWR) to explore gender specific migration patterns. The GWR was introduced by Fotheringham et al. (1996, 2002) to include spatial variation relationships in regression models. In this study we assume that different municipalities located in different parts of Colombia, identified by distinct features, have a different push effect in term of international migration. This feature is delineated as a non-stationary process, with relationships between variables varying geographically, and producing a different response (e.g. migration) in different parts of the same study area. The GWR allows to investigate the variation between covariates and the dependent variable over space (Fotheringham, et al. 2002). The Geographically Weighted Regression (GWR) model applied to the study of Colombian international migration considers as dependent variables the share of men and share of women aged 20–49 who migrated abroad during the 2017–2019 period, measured at the municipal level. The explanatory variables encompass seven key dimensions that reflect local demographic, social, and territorial conditions. a) The rural–urban dimension captures the proportion of men and women aged 20–49 residing in urban or rural municipalities; b) Migration history is represented by the share of women aged 80 and over with children currently living abroad, serving as a proxy for consolidated family migration networks; c) Education includes the share of men and women aged 20–49 with primary, secondary, or higher education levels; d) Female empowerment is measured by the proportion of women aged 20–49 employed in the formal economy; e) Economic data encompass employment, unemployment, and informal work rates for men and women aged 20–49; f) Poverty is assessed using the Unmet Basic Needs (UBN) indices for 2005 and 2018, as well as the Multidimensional Poverty Index (MPI) and g) Finally, violence is captured by the homicide rate (2010–2019 per 100,000 inhabitants) and the level of internal displacement due to conflict (1999–2012). Together, these variables allow for a spatially detailed assessment of how structural, economic, and contextual factors interact to explain the intensity and gender differences in Colombian international emigration.

3. Preliminary results

Figure 1 illustrates the spatial heterogeneity of key explanatory variables considered in the GWR models. The maps reveal strong territorial contrasts reflecting Colombia's uneven social and economic geography. Panel A identifies the areas with consolidated migration networks, concentrated in the Coffee Axis, Valle del Cauca, and Caribbean regions. Panel B shows higher female labour participation along the Andean urban corridor, while activity rates remain low in the Pacific and Amazon regions. Panel C highlights the concentration of secondary education among women in and around metropolitan centres such as Bogotá, Medellín, and Cali. Panel D, representing unmet basic needs in 2005, underscores persistent poverty in the Pacific and Caribbean regions, contrasting with better conditions in central Andean departments. Panel E shows that male unemployment is particularly elevated in several Caribbean and northwestern municipalities, while Panel F displays the spatial distribution of female formal employment, which remains limited outside urban and industrial areas. Together, these variables capture the complex spatial foundations of Colombia's migration system, demonstrating how local social and economic structures condition gendered emigration patterns.

We focus on the preliminary results of the female emigration model (Table 1 and Figure 2) and the male emigration model (Table 2), highlighting the spatial variability of its determinants and the local patterns shaping Colombian international migration across the 1122 municipalities.

Tables 1 and 2 present the results of the Geographically Weighted Regression (GWR) models explaining international emigration rates across 1,122 Colombian municipalities. The models achieve a strong explanatory capacity, with an R^2 of 0.86 and 0.85, and an adjusted R^2 of 0.83 and 0.82 for women and men respectively, indicating that more than 80% of the spatial variability in female emigration is accounted for by the included variables. The AIC and BIC values also confirm the model's robustness compared to a global specification.

The Multidimensional Poverty Index (2018) shows predominantly negative coefficients (the median is negative), implying that higher poverty levels tend to restrict international migration opportunities for both men and women, but poverty has a stronger impact on women's international mobility, though a few municipalities display a slight positive association, as shown by the wide range reached by the explanatory variable.

The share of women with tertiary education exhibits a consistently positive effect, confirming the strong role of educational selectivity in female emigration, less so for men. The proportion of men aged 20–49 not identifying as Indigenous or of African descent shows moderate, mostly positive coefficients, reflecting social composition effects on migration opportunities.

Being employed in the formal economy has a stable, negative relationship with emigration, suggesting that labour market stability discourages out-migration for women. In contrast, the share of elderly women with children abroad—a proxy for migration networks—has a strong and spatially heterogeneous positive effect, reinforcing the significance of family linkages and migration history for both men and women’s migration models. The rural residence variable displays a wide range (from negative to positive), underscoring spatial contrasts, where rurality either constrains or facilitates migration depending on local conditions of the origin municipality. Finally, the homicide rate maintains a uniform positive effect, indicating that violence consistently acts as a push factor for female emigration throughout the country. Figure 2 displays the spatial distribution of scaled residuals and local goodness-of-fit (R^2) for the female emigration model. The residual map (left) shows that most municipalities fall within ± 1.5 standard deviation, indicating a good overall model fit and limited spatial bias. Only a few clusters of over- or underestimation appear, mainly in the Amazonian and eastern plains regions, suggesting areas where additional local factors may influence migration patterns not fully captured by the current set of regressors. The local R^2 map (right) reveals strong spatial heterogeneity in explanatory power, with higher values (above 0.65) concentrated in the Andean and Pacific regions, where socio-economic contrasts and migration networks are more intense, and lower values in the eastern and southern territories, reflecting limited data variability and weaker model performance in sparsely populated zones.

Overall, the geographically weighted regression confirms that the determinants of international female emigration in Colombia are deeply spatial and context dependent. By integrating high-resolution demographic and socio-economic data, this study provides new evidence on how education, labour participation, poverty, and family networks interact differently across regions. The results underscore the need to adopt spatially explicit approaches when analysing gendered migration processes in countries marked by strong territorial inequalities, contributing a novel empirical framework to the study of local migration determinants in Latin America.

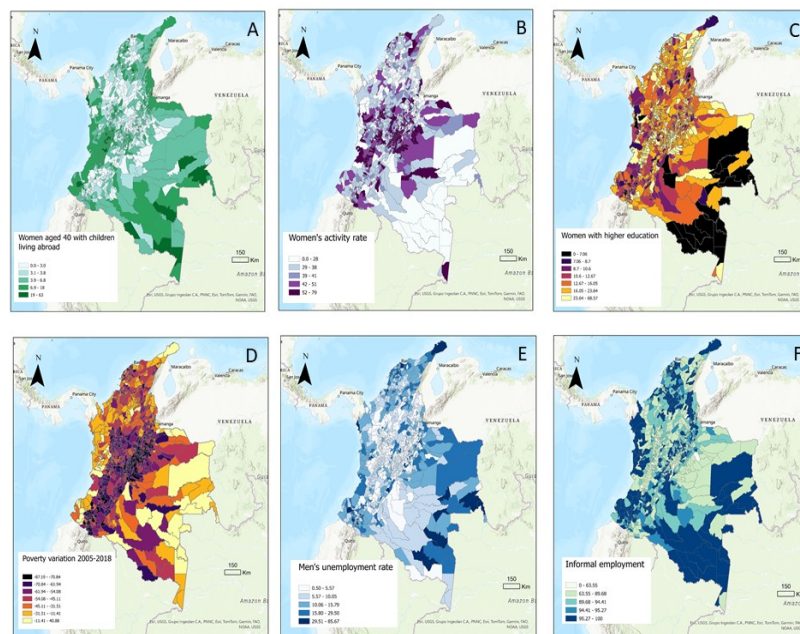


Figure 1 Selected variables clockwise: A) total share of women aged 80 and above with children residing abroad; B) total share of women aged 20-49 activity rate; C) total share of women age 20-49 with secondary education D) total share of population with unmet basic needs for year 2005; E) total share of unemployed men aged 20-49 F) proportion of women aged 20-49 employed in the formal economy.

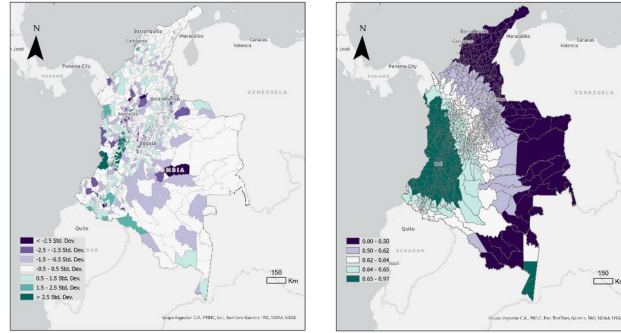


Figure 2 Scaled residuals (left) and local R^2 values (right) reported for the female migration model.

Table 1 GWR results for female migration.

| Number of data points: 1122 | Min. | 1st Qu. | Median | 3rd Qu. | Max. |
|---|------------------------------|------------------|--------|---------|--------|
| <i>Intercept</i> | -0.182 | 1.721 | 1.987 | 2.433 | 7.818 |
| Multilevel Poverty Index (2018) | -1.095 | -0.525 | -0.308 | 0.001 | 0.207 |
| Women aged 20-49 with tertiary education | 0.081 | 0.275 | 0.479 | 0.633 | 1.593 |
| Share of men aged 20-49 who do not identify as indigenous or of African descent | -0.131 | 0.056 | 0.134 | 0.240 | 0.336 |
| Women aged 20-49 employed in the formal economy | -0.017 | -0.016 | -0.016 | -0.015 | -0.014 |
| Share of women aged 80 and older with children living abroad | -0.162 | 0.597 | 1.028 | 2.707 | 3.676 |
| Share of women aged 20-49 who reside in a rural municipality | -6.586 | -0.706 | -0.115 | 0.047 | 2.994 |
| Homicide rate per 100 th. (2000-2018) | 0.464 | 0.466 | 0.466 | 0.466 | 0.467 |
| AICc value 4364 | AIC value 4170 | BIC value 3931.7 | | | |
| R-square value 0.86 | Adjusted R-square value 0.83 | | | | |

Table 2 GWR results for male migration

| Number of data points: 1122 | Min. | 1st Qu. | Median | 3rd Qu. | Max. |
|---|------------------------|-----------------|--------|---------|--------|
| <i>Intercept</i> | -0.183 | -0.056 | -0.012 | 0.036 | 0.337 |
| Basic Needs Index (2005) | -0.103 | -0.103 | -0.102 | -0.102 | -0.102 |
| Multilevel Poverty Index (2018) | -0.437 | -0.149 | -0.074 | 0.012 | 0.054 |
| Share of men aged 20-49 who do not identify as indigenous or of African descent | -0.452 | -0.211 | -0.034 | 0.041 | 0.161 |
| Men aged 20-49 with secondary education | -0.006 | -0.005 | -0.005 | -0.005 | -0.005 |
| Proportion of women aged 80 years old and older with children residing abroad | -0.045 | 0.113 | 0.326 | 0.738 | 0.940 |
| Share of forcibly displaced population (1992-2012) | -2.808 | 0.011 | 0.037 | 0.069 | 0.083 |
| Homicide rate per 100 th. (2000-2018) | 0.133 | 0.133 | 0.133 | 0.133 | 0.134 |
| Share of men aged 20-49 who reside in a rural municipality | -2.060 | -0.143 | -0.026 | 0.001 | 0.562 |
| AICc value 1360 | AIC value 1167 | BIC value 926.5 | | | |
| R-square value 0.85 | Adjusted R-square 0.82 | | | | |

4. References

- Fotheringham, A. S., Charlton, M., & Brunsdon, C. (1996). The geography of parameter space: An investigation of spatial non-stationarity. *International Journal of Geographical Information Systems*, 10(5), 605–627. <https://doi.org/10.1080/02693799608902100>
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Chichester, UK: John Wiley & Sons.

Khoudour-Castéras, D. (2007). ¿Por qué emigran los colombianos? Un análisis departamental basado en el censo de 2005. *Revista de economía institucional*, 9(16), 255-271.

Silva, A. C., & Massey, D. S. (2015). Violence, networks, and international migration from Colombia. *International Migration*, 53(5), 162-178.