

Fleeing a Failing State: Self-selection, Earnings, and Migration Costs

Federico Maggio*

Abstract

Three-quarters of the global migrant population have fled fragile contexts, with 64 percent hosted by similarly vulnerable countries. These contexts account for about 75 percent of those living in extreme poverty. Despite the scale of this phenomenon, the extent to which the self-selection of South-South migrants differs from those migrating to more developed countries remains poorly understood. In this paper, I investigate the self-selection of Venezuelan migrants during the 2015-2021 crisis, which led to greater migration to less developed countries such as Colombia and Peru, compared to more developed ones like the United States and Chile. Using individual-level data representative of the Venezuelan population and similar data on Venezuelan migrants in these key destination countries, the study finds that migrants are generally positively selected in terms of education compared to those who remain in Venezuela, with migrants to developed countries being positively sorted with respect to those to developing countries. However, comparing the cumulative distribution functions of pre-migration predicted earnings reveals that migrants to developing countries are negatively selected relative to stayers, while migrants to developed countries are positively selected. This highlights the significant role of unobserved abilities in shaping South-South migration patterns. Furthermore, a discrete choice model shows that women and college graduates face lower migration costs, independent of expected earnings at the destination. Factors such as the distance to the destination country and pre-crisis networks also play a crucial role in shaping migration decisions.

Keywords: Venezuela, Migration, Self-Selection, Individual-level data, Education

*University of Bologna. Email: federico.maggio3@unibo.it.
<https://federicomaggio.github.io/>. Telephone number: +39 3474882214

Website:

1 Introduction

Over the last decade, there has been a consistent upward trend in the global migrant population, with figures rising from 221 million in 2010 to 281 million in 2020. This surge can be attributed to a combination of economic challenges, conflicts, and political instability, which have compelled millions of individuals to leave their homes¹. Three-quarters of the global refugee population have fled fragile contexts, and 64 percent of them have been hosted by similarly vulnerable countries². Moreover, these contexts account for about 75 percent of people living in extreme poverty³. The most extensive migration episode ever generated by a fragile, non-war country in modern history is the Venezuelan exodus. The increase in state violence and economic crisis prompted about one-fifth of the Venezuelan population to move mostly to other developing South American countries over the 2015-2021 period⁴ (Maggio and Caporali, 2024).

Although a substantial number of works have addressed the impact of Venezuelan migrants on the society and economy of neighboring countries (Olivieri et al., 2022; Roza and Vargas, 2021), the characteristics of those who are leaving the country, and the determinants driving them to choose a specific country have not yet been investigated. In a setting where individuals mostly migrate from a failing country to developing fragile countries, studying migrants' decision-making process, according

¹By the year 2050, an estimated 200 million individuals could be displaced as a result of increasingly frequent and severe weather phenomena such as wildfires, floods, and droughts (Clement et al., 2021). These climatic disruptions contribute significantly to the drivers of migration, alongside conflict and political instability. In 2020, the global scale of displacement caused by enduring conflicts and political unrest in countries such as Afghanistan, Myanmar, South Sudan, Syria, and Venezuela, as well as escalating conflicts within and beyond the borders of Ethiopia, reached a staggering total of over 34 million people (UNHCR, 2021).

²Fragility is the combination of exposure to risk and insufficient coping capacities of the state, system, and/or communities to manage, absorb and mitigate those risks (OECD, 2016).

³By the end of 2021, 89.3 million people were forcibly displaced worldwide as a result of persecution, conflict, violence, or human rights violations. This includes 27.1 million refugees, 53.2 million internally displaced people, 4.6 million asylum seekers, and 4.4 million Venezuelans displaced abroad (UNHCR, 2021).

⁴According to OECD's "States of Fragility 2018-2021" report, Venezuela since 2018 has been among the 30 most fragile countries in the world (alert countries). Colombia and Peru during the same period were ranked as elevated warning countries.

to their sociodemographic characteristics, and intended destination countries is particularly important for several reasons. First, understanding migrants' motivations helps to discern migrants forced to leave their countries from migrants seeking better opportunities. Second, the skill composition of migrants has critical implications for the reconstruction of the country of origin. The more skilled migrants are, the more difficult will be to rebuild the country at the end of the political and economic crisis. Third, knowing the level of intrinsic income determinants of migrants is useful for planning integration policies, thus contributing to the political and economic stability of receiving countries.

In the present study, I investigate whether migrants are self-selected based on observable characteristics and pre-migration earnings in a context where their own country is undergoing a severe crisis and migration patterns to fragile and less developed countries outweigh migration to developed and stable countries ^{5 6}.

To shed light on this issue, I combine unique individual-level data representative of the Venezuelan population and similar data on Venezuelan migrants residing in Colombia, Peru, Chile, and the US. To analyze the migrants' self-selection based on observable characteristics, I run a series of multivariate regression models examining the relationship between individual characteristics and migration probability to different countries. However, these regressions do not address whether migrants are selected based on unobservable characteristics.

To tackle this, I leverage unique data on the annual earnings of Venezuelans to predict counterfactual earnings in both the origin country and the four destination countries. I then compare the cumulative distribution functions of the per-migration earnings to assess the influence of unobserved abilities on migration decisions.

⁵The most common destination for Venezuelan migrants was Colombia, which received 1.8 million migrants. Peru welcomed 1.3 million Venezuelan migrants, while Chile and the United States hosted 450,000 and 460,000 Venezuelans, respectively (UNHCR).

⁶To classify the destination countries according to fragility, I use data from The Fund for Peace (FFP, 2015). The fragility index is based on 12 economic, political, and social variables. At the beginning of the Venezuelan crisis in 2015, the fragility indices were as follows: Venezuela had a score of 78.7, Colombia 82.5, Peru 72, Chile 41.5, and the United States 35.1. Throughout this paper, I consider Peru and Colombia to be fragile destination countries.

Finally, to evaluate the independent role of individual and country characteristics in shaping migration costs, I estimate a discrete choice model while controlling for counterfactual expected earnings across all alternative locations. Specifically, I employ a Conditional Logit Model (CLM) to disaggregate country-specific migration costs – such as network, employment rate, distance, and language – and a Mixed Logit Model (MLM) to disentangle costs according to individual characteristics, including education, age, and gender.

The linear probability models indicate that educated individuals are significantly more likely to migrate to Colombia, Chile, and the United States, suggesting positive selection in education among these migrants compared to stayers. In contrast, Venezuelans who moved toward Peru appear negatively selected in terms of education. However, a comparison of the cumulative distribution functions (CDFs) of Venezuelans’ pre-migration earnings with stayers’ CDFs reveals that migrants to the United States and Chile are positively selected in terms of expected income. On the other hand, migrants to Colombia and Peru are negatively selected compared to stayers. Since the unobserved abilities are factored into the prediction of pre-migration earnings, this finding provides novel evidence of the negative self-selection of South-South migrants based on unobservable characteristics, while they are consistent with the positive self-selection of South-North migrants reported in the literature.

The results of both the discrete choice model estimations confirm that income differential is a robust determinant of migration choice, and migration cost independently of the expected earnings is a relevant factor in shaping migration choices. Specifically, the CLM results show that distance to country of origin and network at destination are significant factors in shaping women’s and men’s migration choices, partially contributing to explaining why Venezuelans still prefer South-South migration routes. MLM results, on the other hand, reveal that even when controlling for expected earnings at the destination, individual characteristics have a significant effect on the likelihood of migrating to a specific country. In particular, people with college degrees face lower costs in migrating toward Colombia, Chile, and the United States compared to uneducated. In the case of Peru, however, the college coefficient is very low and not significant, suggesting that Peru, in addition to being unattrac-

tive to educated migrants due to low college wage premium, has other confounding "non-pull" factors that make it the last choice for the most educated Venezuelans. This result implies that the costs of migration, and consequently migration patterns may depend on the policies of destination countries even in a South-South migration context.

Although extensive literature exists on migrant self-selection, the focus has primarily been on South-North migration. [Borjas \(1990\)](#) laid the groundwork for self-selection theory, demonstrating that migrants from a country with high returns to education inequality who move to a country with a more equal education premium are likely to come from the upper end of the skill distribution. [Chiquiar and Hanson \(2005\)](#), in their pioneering work, merged information from the US census on Mexican migrants' characteristics with data from the Mexican census on the characteristics of stayers. They introduced the method of comparing the counterfactual wages of migrants in their origin country (as if they had not migrated) to the wages of stayers, finding that Mexican-born men were intermediately selected, while Mexican-born women were positively selected. [Sotomayor \(2009\)](#) confirmed the findings of intermediate selection for Puerto Rican migrants, while higher inequality and returns to education on the island attract migrants from the top of the distribution. Further supporting the self-selection hypothesis, [Grogger and Hanson \(2011\)](#) used aggregate data on emigrant stocks by schooling level and source country in OECD destinations to estimate an income maximization model. Their findings confirmed that countries with high returns to skills attract highly skilled migrants. Specifically, they showed that migrants were positively selected and that more educated migrants were more likely to settle in destination countries with high-skill returns. However, none of these studies addressed how the distribution of unobserved characteristics may affect the probability of migration and wages at the destination.

[Moraga \(2011\)](#) and [Kaestner and Malamud \(2014\)](#), provided evidence that part of the negative selection could be attributed to the unobservable characteristics that shaped migrants' earnings. Bertoli and co-authors ([Bertoli et al., 2013](#)), while accounting for unobservable characteristics, found that earnings significantly shape individual migration decisions, even in an episode in which Ecuadorians mostly chose

Spain where earnings were lower than in the US. While these findings could account for the pattern of positive sorting in education toward the United States, they fail to explain why most Ecuadorians opted for Spain. In the following paper ([Bertoli et al., 2011](#)), they further explain that this preference seems to reside in the Spanish visa waiver program for Ecuadorians. [Clemens and Mendola \(2020\)](#) presents estimates of emigrant selection on both observed and unobserved determinants of income. In particular, they found that in low-income countries, people actively preparing to emigrate have 30 percent higher incomes than others overall, 14 percent higher incomes explained by observable traits, and 12 percent higher incomes explained by unobservable traits.

This study is also related to the literature strand on the costs of international migration, which, as highlighted by [Hanson \(2010\)](#), is the most significant gap in the economics of migration literature. [Ortega and Peri \(2009\)](#) provided estimates of the effects of immigration policy tightening on the magnitude of migration flows in 14 OECD countries over the period 1980 to 2005. They confirmed that, while South-North migration flows increase as a function of the per capita income gap between origin and destination, they decrease significantly when destination countries adopt stricter immigration laws. [McKenzie and Rapoport \(2010\)](#) highlighted the importance of the network in shaping migration costs. They found a positive selection for Mexican migrants leaving from communities with weak migrant networks, but negative self-selection for Mexican migrants from communities with stronger networks.

Academic interest in South-South migration is relatively new, and empirical discussions to date have mostly focused on migration and development, exploring how the effects of South-South migration may differ from those of South-North movements ([Melde et al., 2014](#); [Hujo and Piper, 2010](#); [Bakewell et al., 2009](#)). This body of literature suggests that South-South migration may involve more common return migration, movements across more porous borders, smaller wage differentials, and lower remittances. Despite these insights, our understanding of the dynamics surrounding South-South migration remains relatively limited. Few studies have empirically explored the potential distinctions in self-selection and migration determinants between South-South and South-North migration. [Lanati and Thiele \(2024\)](#) found that

refugees tend to move to safe neighboring countries and respond positively to local pull factors such as relatively high per-capita income levels and the availability of education and health services.

This study aims to fill these gaps by examining the magnitude of South-South migration costs tied to individual characteristics and disentangling them from the migration premium associated with those characteristics. Isolating the specific impact of individual traits on migration decisions is crucial for shaping effective migration policies in destination countries. Furthermore, this paper extends the existing literature on the Venezuelan migration episode, which has predominantly focused on the impact of Venezuelans' arrival on destination countries ([Bahar et al., 2020](#); [Olivieri et al., 2022](#); [Doocy et al., 2019](#); [Namen et al., 2019](#); [Caruso et al., 2021](#)).

The remainder of the paper is organized as follows: Section 2 provides a brief overview of the salient features and economic determinants of the Venezuelan exodus. In Section 3, I present the various data sources and relevant descriptive statistics. Section 4 details the empirical strategy employed in the study. The main results are discussed in Section 5. Finally, Section 6 offers a brief discussion of the key findings and their potential policy implications.

2 The Venezuelan Exodus

In this section, I discuss the nature of the Venezuelan migration crisis, which has been mainly directed to Colombia, Peru, Chile, and the United States. Venezuela has been facing one of the worst socioeconomic and political crises ever experienced by a Latin American country. The beginning of this crisis dates back to the election of Hugo Chavez as president in 1998. During his presidency, Chavez introduced a socialist regime that included constitutional amendments, expropriations of land, implementation of populist social programs, nationalization, and restrictions on private companies. Nicolas Maduro, elected president of Venezuela in 2013, continued along the same lines as his predecessor, worsening the country's economic and social crisis. Populist policies, unsustainable public debt, low oil prices, and excessive and rigorous

controls and regulations on the private sector have led to a deep economic recession. In early 2018 the Venezuelan government "essentially stopped" producing inflation estimates. However, as shown in Figure 1 ⁷, by the end of 2018, the country was experiencing a hyperinflation rate of 63,151% (Hanke, 2018).

The financial and economic crisis has led to a harsh humanitarian crisis, prompting an unprecedented wave of international emigration ⁸, with nearly 5 million Venezuelans leaving over the 2015-2021 period, from a country with a total population of 28.9 million ⁹ (Figure 2). Crucially for my analysis, Venezuelans moved essentially to four main destination countries, which absorbed 70 percent of the Venezuelan migration driven by the economic crisis. The most common destination was Colombia (1.8 million migrants). Peru welcomed 1.3 million Venezuelan migrants, whereas Chile and the United States hosted 450,000 and 460,000 Venezuelans, respectively. Figure 3 plots the stock of Venezuelan migrants by year and destination country ¹⁰.

Interestingly, although this migration episode is highly concentrated in a few years and the timing is similar in all destinations, the scale differs substantially. Indeed, from 2015 to 2017 about 54,000 Venezuelans had migrated to the United States ¹¹,

⁷Figure 1 illustrates a dramatic deterioration across several economic, political, and health indicators. Notably, it shows a sharp decline in GDP per capita, which fell from \$19,000 in 2012 to \$5,000 in 2020. This economic collapse was accompanied by an increase in state capacity (tax revenues as a share of GDP per capita), indicating a constant tax burden for Venezuelan citizens despite the recession. Concurrently, the political crisis escalated violence (both civil and police-related), leading to a significant decline in the Human Development Index and human rights standards. The severe economic and political collapse also adversely affected medium- to long-term health indicators, with life expectancy dropping from 73 years in 2013 to 72 years in 2017, and child mortality rising from 1.7% in 2010 to 2.4% in 2017. Additionally, the economic downturn resulted in a reduced food supply from 2015 onward.

⁸Migration throughout South America is not a new phenomenon, and Venezuela has long been a host nation for migrants from neighboring countries. Historically, migration in the region has occurred from South to North, as the ongoing conflict in Colombia pushed millions of people to flee the violence, seeking refuge in Venezuela in the 1980s and 1990s. The direction of the migration flow changed with the first wave of Venezuelan migrants in the early 2000s under the government of Hugo Chavez. The second wave took place in 2015, while from 2017 to 2019 there was the third, most dramatic and complex Venezuelan migration wave.

⁹According to the last Population Census in 2011

¹⁰As reported by R4V (Plataforma de Coordinación Interagencial para Refugiados y migrantes de Venezuela), the remaining destination countries comprise Ecuador, Spain, Brazil, Argentina, Panama, and Mexico.

¹¹According to the Migration Policy Institute, in 2018 most Venezuelans were concentrated in the

65,000 to Chile, while only 22,000 had migrated to Peru. In comparison, Venezuelan inflow to Colombia was twenty times larger over the same period. Specifically, in 2017 Colombia had already welcomed 600,000 Venezuelan migrants.

The policy response of destination countries has also been different. Colombia has been the most welcoming country. Since the beginning of the crisis, it has allowed Venezuelans to enter with their passports, but without the need to hold a visa. Additionally, to facilitate mobility in border areas, in February 2017, the Border Mobility Card (TMF in Spanish) was introduced by the Colombian government. This document allowed Venezuelans without passports to cross the border to buy food or access benefits such as vaccines and return to Venezuela within 7 days. Since many Venezuelans were irregularly overstaying in the neighboring country, the TMF was suspended in February 2018 ¹². Peru and Chile have had a more restrictive attitude toward Venezuelan immigration. As of June 2019, only Venezuelans holding passports and humanitarian visas could enter Peruvian and Chilean territories. In October 2019, 67 percent of Venezuelans had arrived in Peru without having any residence permit or visa ^{13 14}. The number of Venezuelans who entered Chile illegally rose to 3,333, up from 1,536 in 2018 ¹⁵. The United States is undoubtedly the destination country that has implemented the most restrictive migration policy. Since the Trump administration (early 2017), the United States has been virtually closed to Venezuelans. The number of tourist visas issued to Venezuelans dropped to 47,942 after reaching 239,772 in 2015 ¹⁶. Moreover, as of 2017, the United States increased the cost per visa application from 30 (in 2015) to 160 US Dollars ¹⁷.

cities of Houston, Orlando, and especially Miami, where about 109,000 Venezuelans (2 percent of the population) lived.

¹²TMF had been approved for 1,624,915 Venezuelans until its suspension in 2018. On an average day in 2017, more than 30,000 Venezuelans used these cards to enter and leave Colombia (IOM, April 2020).

¹³22 percent had a temporary residence permit; 5 percent a tourist visa; 3 percent a permanent residence permit; and 2 percent a work or study visa.

¹⁴Encuesta Dirigida a la Poblacion Venezolana residente en el Pais 2019

¹⁵Chilean Investigative Police (PDI)

¹⁶US officials could detain Venezuelans arriving on tourist visas if inspectors suspected they intended to stay rather than visit (IOM, 2020)

¹⁷The price was prohibitive, as the average salary of a Venezuelan in 2018 was about 250 US Dollars per year

Obviously, the distance, geographical characteristics of the routes, and the military control of the borders have made irregular migration to Chile and the United States much more difficult than to Colombia and Peru ¹⁸. Willing to tackle irregular migration, countries in the region intensified programs to regularize irregular migrants already in the country. Colombia introduced the PEP (Permiso Especial de Permanencia) implemented in four different rounds (June 2017, February, August, and December 2018) ¹⁹, Peru developed the PTP (Permiso Temporal de Permanencia) in 2019 ²⁰, and Chile launched the Process of Extraordinary Regularization in July 2018 ²¹.

The lack of a history of emigration from Venezuela, combined with the sudden economic crisis and subsequent surge in emigration, provides a clean exogenous shock that rules out the possibility that education may be endogenous to the prospect of emigration.

3 Data

3.1 Data sources

The analysis requires individual-level data representative of the Venezuelan population residing in the origin country and of the four main receiving countries over the relevant period. Specifically, the full dataset was created by combining five different data sources. For Venezuela, I use the 2018 Encuesta Nacional de Condiciones de Vida (hereafter ENCOVI ²²), representative of the Venezuelan population remain-

¹⁸Refugees and migrants compelled to leave Venezuela to use diverse routes to reach their destination including by land, plane, and sea, sometimes putting their lives at risk. Many leave without documents to cross borders. They face great risks at the hands of smugglers and traffickers (UNHCR 2019).

¹⁹According to Migración Colombia, the State's immigration authority, there are approximately 600,000 Venezuelans who have a PEP.

²⁰As of Oct 29, 2019, 549,606 Venezuelans had received a PTP

²¹As of Jul 23, 2018, 31,682 Venezuelans had taken part in the regularization program.

²²The survey was carried out by the Universidad Catolica Andres Bello de Caracas in 2018. It provides information about 21,382 individuals across 22 states (The sample does not include

ing in the country of origin. As a representative sample of the Venezuelan migrants residing in Colombia, I rely on the 2021 Colombian National Household Survey (or GEIH in the Spanish acronym ²³). For Peru, I use the 2021 Encuesta Dirigida a la Población Venezolana que Reside en El País (ENPOVE ²⁴). For the United States, I chose the 2021 American Community Survey ²⁵. For Chile, I rely on the National Socioeconomic Survey 2021 (or Casen in the Spanish Acronym ²⁶).

Finally, I combined the five datasets, appropriately weighted, to account for the different coverage of the samples with respect to the populations under analysis ²⁷. The full dataset contains individual-level information on age, employment status, occupation, and labor earnings. According to the existing literature on economic migration, I opted to restrict the sample to the working-age population (18 to 60). My final sample consists of 40,031 observations provided by 12,335 stayers, 17,742 Venezuelans settled in Colombia, 7,484 migrants in Peru, 1805 migrants in Chile, and 765 migrants in the United States.

The timing of the surveys described above fits well with the need to focus on Venezuelans who left their home country driven by Venezuela’s sudden economic collapse. Furthermore, I chose to use ENCOVI 2018 as a sample representative of stayers

Amazonas and Dependencias Federales). The survey covers a wide variety of topics, including basic demographics, educational background, and labor market conditions.

²³Source: Departamento Administrativo Nacional de Estadística, Gran Encuesta Integrada de Hogares - GEIH - 2021. It was conducted at the end of 2021 and it includes 17,7428 Venezuelan migrants.

²⁴Instituto Nacional de Estadística e Informática, ”Encuesta Dirigida a la Población Venezolana que Reside en el País 2021”). The survey, performed at the end of 2021, collects information on 9.847 Venezuelan migrants residing in Peru, which is the second-largest receiving country. It was carried out by the Peruvian National Institute for Statistics (INEI) between November and December 2021. It is representative by design of the Venezuelan population residing in Peru. In particular, it was conducted in the five largest cities in the country (Lima, Tumbes, Trujillo, Cusco, and Arequipa), where 85 percent of Venezuelans reside.

²⁵Source: U.S. Census Bureau, American Community Survey (ACS), Public Use Microdata Sample (Pums), 2021. Its sample is approximately 2.5 percent of the resident population in the US.

²⁶conducted by the Ministry of Social Development. It provides information on 1.2 percent of the total population living in Chile.

²⁷The ENCOVI sample is representative of the estimated 26 million Venezuelan residents in 2018. The GEIH sample is representative of 1.8 million Venezuelan migrants residing in Colombia. The ENPOVE sample is representative of 1.3 million Venezuelans residing in Peru. While CASEN and ACS samples are representative of 448,000 and 460,000 Venezuelan migrants residing in Chile and the United States, respectively.

since most of the migrants residing outside their country of origin, moved between 2017 and 2019 (see Figure 2). Finally, the investigation concerns only Venezuelan migrants who were legal at the time of the survey. However, I cannot rule out the possibility that some Venezuelans previously entered irregularly between 2015-2021 and then obtained the required documents thanks to the regularization policies mentioned in section 2.

3.2 Descriptive statistics

Table 1 reports some basic descriptive statistics on stayers and migrants to the four destination countries. It shows that Venezuelans who emigrated to Colombia and Chile are predominantly women (53-54 percent), while in Peru and the United States, the sample is gender balanced. Consistent with the literature on migration selection, migrants are younger than stayers, regardless of gender, with the exception of the United States, where Venezuelan migrants are, on average, the same age as those who remained in the country of origin. The average age of the Venezuelans who chose not to leave the country was 37 at the time of the survey ²⁸. For both genders, migrants living in Colombia, Chile, and the USA are more educated than non-migrants. Venezuelan migrants residing in Peru are on average as educated as those remaining in Venezuela. 19 percent of them have a university degree. Venezuelans who moved to the United States are the most educated. 55 percent of them are college graduates. The share of college graduates among Venezuelan women is higher than among men (58% in the US, 55% in Chile, 26% in Colombia, and 23% in Peru).

When examining the earnings data, several aspects deserve careful consideration. First, I note a remarkable migration premium. Despite being younger, migrants earned much higher annual incomes than non-migrants. Additionally, within our sample, Venezuelan migrants are classified as short-term migrants, meaning they have departed from their home country within the past five years. This implies that some years after migration they may obtain a much higher migration premium. Second,

²⁸Some World Bank estimates show a constant increase in the average age since the beginning of the crisis, due to a lower survival rate at birth, a lower birth rate and migration of young people.

the data reveal a substantial and relatively constant gender disparity in earnings. The largest gap is observed in the Venezuelan population residing in the United States (35 percent for women with a college education and 31 percent for women without a college education), while Chile has the lowest female penalty among the countries analyzed (12 percent for college graduated and 16 percent for non-college graduated women). Third, university premiums differ greatly depending on the destination country. Peru shows the lowest return to education. On average Venezuelan migrants who settled in Peru and have a degree earn only 9% more than migrants with no university degree (4612\$ compared to 4228\$). Colombia has much higher rates of return to education than Peru: the degree on average ensures 25% higher income. Chile shows the highest return to education: a college-graduated migrant earns 50% percent more than a non-college graduate. The premium differences across destinations are not due to the length of migrants' stay. The clean "push shock" due to the sudden economic crisis led to mass migration in the short term. Most of the migrants in my sample migrated between 2017 and 2019 (Figure 2).

These descriptive statistics regarding earnings and education rates align with the findings of [Grogger and Hanson \(2011\)](#), which indicate that the absolute differences in earnings between college-educated and non-college-educated individuals are greater in high-income countries. While the descriptive statistics suggest that migrants have the potential to achieve significant financial gains by leaving their home country — unsurprising given the dire economic situation in Venezuela — it remains unclear to what extent their migration decisions are influenced by potential earnings at the destination and the migration costs associated with individual and destination country characteristics.

4 Empirical strategy

As discussed in Section 2, the Venezuelan exodus offers a unique opportunity to study the self-selection process of migrants. Nearly a quarter of Venezuela's population – approximately 6 million out of 29 million people – has emigrated, resulting in a highly

diverse migration flow in terms of skill, age, gender, and potentially in terms of unobservable characteristics ²⁹. Moreover, the Venezuelan migration is directed toward countries with varying levels of economic development and distinct migration policy responses. The diverse migration patterns, combined with rare data on the labor market conditions of both stayers and migrants in the main destination countries, allow a detailed analysis of the roles of earnings and the characteristics of individuals and destination countries in shaping migration choices.

To achieve this, I begin by running a series of multivariate regression models to examine the relationship between observable individual characteristics and migration probabilities to different countries. Next, I leverage unique data on the annual earnings of Venezuelans to predict counterfactual earnings, corrected for self-selection in unobservable characteristics, in both the origin country and the four destination countries. I then compare the cumulative distribution functions of predicted pre-migration earnings to assess the self-selection of migrants, controlling for unobservable characteristics. Finally, the counterfactual earnings are used to evaluate how migration decisions are influenced by individual and destination country characteristics, independently of the expected earnings at the destination.

4.1 Predicted individual earnings

As the migrants' surveys did not collect information on how much Venezuelans earned in their home countries before they left and, of course, each individual is only observed in one single location, the first step involves predicting individual-level earnings for all alternative locations, according to their individual observable characteristics and their earnings in the location in which they are observed. The availability of data on migrants' earnings in destination countries is of great value, especially for South-South migration episodes. Indeed, most empirical studies have had to rely on extrapolations

²⁹In terms of scale, only the Syrian exodus is comparable, with around 6 million refugees. However, the contexts differ significantly; while Syria's migration is primarily due to conflict, Venezuela's crisis is rooted in profound economic and political instability, making it the most severe crisis of any non-war failing state in modern history. Additionally, the Syrian migration has predominantly flowed towards Turkey, a country with a much higher income per capita than Colombia or Peru.

from income figures for the general population. However, as shown by McKenzie et al. (2010) in the context of Tongan migrants to New Zealand, migrants are likely to differ from non-migrants in several unobservable aspects. Even in the context of interest, there might be some unobserved characteristics that push people to move to a specific destination country and systematically bias their wage draws. For example, suppose that equally educated Venezuelans moving to the United States are less risk-averse people than Venezuelans migrating to Peru or Venezuelans who chose to stay in their home country. At the same time, they will be more likely to be hired in more risky and better-paid jobs. Therefore, unobserved heterogeneity in the propensity to migrate affecting also earnings would create a selection bias in the wage equation.

To address this concern, I apply the self-selection correction method proposed by Dahl (2002)³⁰. First, I divided the population into 18 mutually exclusive groups, based on gender, education (college versus high school versus non-educated), and age (three age groups). For each of these groups, I calculate the proportion of individuals who remain in Venezuela and work for a wage there (\hat{p}_{i1}), migrate to the US, and work for a wage there (\hat{p}_{i2}), migrate to Colombia (\hat{p}_{i3}), migrate to Peru (\hat{p}_{i4}), and migrate to Chile and work for a wage there (\hat{p}_{i5}). These terms represent the predicted probabilities of individuals belonging to specific groups choosing to work in the respective locations. Next, I estimate the Mincer regressions by adding the Dahl's correction Polynomial $f_j(\hat{p}_{i,j}, \hat{p}_{i,j'})$. As suggested by Bertoli et al. (2013), I include it as a second-order polynomial in the retention probability for stayers and a second-order polynomial in the retention and first-best probability for migrants plus an interaction term. The purpose of this polynomial term is to correct the potential bias arising from migrant workers having a higher unobserved propensity to migrate and work in specific destinations, which can impact their actual earnings. The underlying assumption of Dahl's method is that the unobserved heterogeneity within the groups is relatively small. Lastly, I predict the logarithm of earnings (\hat{w}_{ij}) in all five locations for all individuals within the sample.

³⁰The method was also used by Bertoli et al. (2013). In a South-North migration episode (from Ecuador to the US and Spain) they found that, including the Dahl parameter, the estimates of the expected earnings do not change much.

4.2 Self-selection based on pre-migration earnings

Once I obtain selection-corrected individual wages, I use them to assess the self-selection of migrants based on pre-migration earnings. To this end, I apply the method suggested by [Borjas et al. \(2019\)](#), who demonstrated that the conditions leading to positive or negative selection in terms of expected earnings also imply a stochastic dominance relationship between the counterfactual earnings distribution of migrants in the origin country and the earnings distributions of stayers.

Furthermore, since we correct the prediction of counterfactual earnings in all alternative locations for individual unobservable factors, comparing the cumulative distribution functions (CDFs) of these counterfactual earnings allows us to evaluate the relevance of unobservable factors in migrants' self-selection.

4.3 Self-selection based on country and individual characteristics

Finally, I evaluate the independent role of country and individual characteristics in the migration decision process.

Specifically, I estimate a discrete choice migration model, controlling for the estimate of predicted self-selection corrected earnings for all alternative locations ($\hat{w}_{i,j}$). The utility function of individual i in location $j = 1, 2, 3, 4, 5$ can be summarized as follows:

$$U_{i,j} = \alpha \ln(\hat{w})_{i,j} + x_i \beta_j + v_{i,j} \quad (1)$$

In practice, I estimate a conditional logit model to disentangle the effect of country-specific characteristics on location choice from the effect of expected earnings ($\hat{w}_{i,j}$). In the conditional logit model estimation, I include country-specific intercepts in the first specification. In the second specification, these intercepts are replaced by four country-specific migration cost variables: distance from the country of origin

(kilometers), percentage of Venezuelans in the population of the destination country, language similarity, and unemployment at the destination. I estimate this specification separately by gender.

I then estimate a mixed logit model, which allows me to separate the effect of individual characteristics – age, age squared, education, and gender (x_i) – from the effect of income differentials. Although individual controls have already been used in earnings prediction, I include them in the decision model to capture differences in migration costs depending on individual characteristics and the chosen destination, which are unrelated to the wage. I also include country-specific intercepts to account for differences in institutions, culture, and cost of living across the four destination countries.

5 Results

Table 2 shows the estimated effects of individual characteristics on migration probabilities to different countries. It presents evidence for the full migrant sample and separately for the four destination countries³¹. To focus on Venezuelans most likely to have the decision-making power to migrate and who have completed their education, I restrict the analysis to working-age Venezuelans (aged 18 to 60)³². In column 1, I find that, for both males and females, educated people are significantly more likely to migrate, which means that migrants are positively selected in education with respect

³¹Outcome variables are as follows: *migrant to any country* is the outcome variable in Column 1; migrants who moved to Colombia in Column 2; Venezuelan migrants in Peru in Column 3; migrants in Chile in Column 4, and migrants in the US in Column 5. In each analysis, the control group of stayers consists of ENCOVI respondents.

³²In the appendix, I also present a table that provides estimates from the same analysis but uses a sub-sample of individuals who reported non-zero earnings (Table A1). As expected, compared to Table 2, this analysis shows a stronger positive selection for male migrants residing in Colombia, Chile, and the United States, and no selection for male migrants living in Peru. These findings may reflect the positive selection of employed migrants at their destinations compared to non-employed migrants. The estimates for the female sample, however, are quite unexpected. Table A1 indicates a negative selection for female migrants living in Colombia and Peru. A possible explanation for this could be that in low-educated and low-income households of migrants, female members — who are culturally less likely to be employed compared to males — may need to find jobs out of necessity.

to stayers. However, as suggested by the two-sided tests of coefficient equality across regressions, the effect of having a college degree on the probability of migrating is different depending on the country of destination and gender. Indeed, Table 2 shows a weaker positive selection for Venezuelan females who moved to Colombia and a negative selection for Venezuelan males who moved to Peru (respectively Columns 2 and 3 of Table 2) ³³. Moreover, it shows a positive sorting of the migration flow towards developed countries (the United States and Chile) with respect to Venezuelan migrants living in Peru or Colombia ³⁴.

In general, these descriptive results lead to a first consideration: even in a situation where migrants leave a country undergoing a serious crisis and most of them move to similar southern countries, they are positively selected in terms of observable characteristics compared to stayers. However, these results are non-exhaustive for two main reasons.

First, they do not provide insights into whether Venezuelans are selected based on unobservable characteristics across origin and destination countries. Indeed, Borjas et al. (2019) suggests that in the Danish migration context, the selection process is not solely based on observable characteristics, rather the unobserved abilities play the most dominant role. While they demonstrate positive selection for both observable and unobservable characteristics in a North-North migration context, this may not apply in a situation where the origin country is undergoing a severe economic and political crisis and migration towards similar developing countries greatly exceeds North-North migration. Second, the estimates of Table 2 do not explain the extent to which this selection process depends on expected earnings at the destination or on country and individual characteristics.

³³A logistic model is also estimated, with which I measure the selection and sorting of migrants (In appendix, Table A2).

³⁴Stronger positive self-selection of people moving to more developed countries is in line with Borjas (1990), as the US and Chile have much wider income differences than Venezuela (Grogger and Hanson, 2011).

5.1 Results - Self-selection based on pre-migration earnings

Figures 4 and 5 compare the cumulative distribution functions (CDFs) of selection-corrected ³⁵ migrants' predicted income – in the origin country – with stayers' CDFs. Figures 4 and 5 indicate that male and female migrants to the US and Chile are positively selected in terms of predicted incomes. In contrast, male and female migrants to Peru and Colombia appear to be negatively selected compared to the stayers. These results confirm the negative self-selection in education observed in Section 5 for Peru but contradict the case of migration from Venezuela to Colombia. Specifically, migrants to Colombia are better educated than stayers but are negatively selected in terms of predicted income.

Two potentially coexistent factors could explain these results. The first factor is the role of the unobserved abilities of migrants, considered in the predicted counterfactual earnings but not in the estimates of Table 2. While the findings for migrants going to the US and Chile are consistent with Borjas et al. (2019), which shows positive self-selection for both observable and unobservable characteristics, the results for migrants going to Colombia and Peru suggest they are negatively selected in terms of unobserved abilities. Several explanations could account for the negative selection based on unobservable characteristics of South-South migrants. For instance, more adaptable and resilient individuals facing job loss and financial challenges might stay in their country and find a better job in a different labor sector. In contrast, less adaptable workers might decide to move to a country with a higher employment rate in their sector. However, facing financial constraints (Angelucci, 2015; Gazeaud et al., 2023) – and risk aversion – they may not be able to move to a more developed and stable country, such as the United States or Chile.

The second factor explaining the negative self-selection in predicted incomes for

³⁵Table A3 shows the results of the wage regressions with (even-numbered columns) and without (odd-numbered columns) including the Dahl correction term for self-selection in unobservables. The upper panel refers to the sample of males, while the lower panel refers to females. P-Values (F-test Dahl polyn.) suggest that unobserved individual characteristics play an important role in predicting the wage. Among destination countries, the estimates imply the lowest college premium for migrants residing in Peru (9,3% for males and 11,9% for females). This is in line with the low-skill composition of migrants in Peru.

Venezuelans going to Colombia, despite positive self-selection in education, is that the choice to move to Colombia might be strongly influenced by the "education status" – rather than the actual abilities of migrants –, which lowers migration costs. Even for migrants with lower unobserved abilities, high education status might lower migration costs thanks to more welcoming policies for educated individuals, increasing the likelihood of obtaining the required migration documents.

This finding adds further insight into the understanding of the Venezuelan exodus and South-South migration patterns: migration costs strongly influence location choice and vary greatly according to country and individual characteristics. The results described in the next subsection will help shed light on this mechanism.

5.2 Results - Self-selection based on country and individual characteristics

Table 3 shows the estimates of the conditional logit model specifications, in which the effects of individual characteristics on migration costs do not vary across alternatives. The earning coefficient α is positive and highly significant for both genders, suggesting that higher expected earnings in a particular country increase the probability of moving there. In the first two columns, the coefficients of the country intercepts are high and strongly significant. When I replace the country intercept with country-specific migration cost variables, I find that distance from the country of origin – a proxy for the cost of travel – and the network at the destination are significant factors in shaping women's and men's migration choices (columns 3 and 4). Additionally, the great variation in the country intercept across destination countries shown in columns 1 and 2 may also suggest a very heterogeneous response of the destination countries in terms of migration policies.

While the results of the conditional logit model partially contribute to the explanation of why Venezuelans still prefer South-South to South-North migration patterns – due to proximity, stronger networks, and similar language and culture – they do not provide much insight into the self-selection in education of individuals across similar

southern destination countries.

Table 4 shows the estimates for the Mixed Logit Model. The first column shows the effect of individual characteristics on the likelihood of migrating to Colombia, the second column to Peru, the third to Chile, and the last column to the United States. Again, the coefficient of expected earnings is positive and highly significant (0.651), clearly suggesting that earnings at the destination are a robust determinant of migration choices. Turning to the interpretation of the coefficients of individual controls and country intercepts, I note that for Colombia, Chile, and the US, the coefficients associated with the education variable are still positive and significant, even when controlling for earnings.

This confirms that, even in the context of migration driven by a sudden crisis in the origin country, highly educated people may face lower migration costs depending on the destination country. While higher coefficients for education in South-North destinations compared to South-South countries may be expected – due to lower rates of illegal migration, more stringent border controls, and more welcoming migration policies – the very small and not significant coefficient for Peru, compared to the high and significant one for Colombia, deserves further consideration. Peru is geographically farther and had a weaker Venezuelan network compared to Colombia. In fact, Table 3 confirms that migrating to Peru is 30 percent more expensive than migrating to Colombia. Moreover, according to the literature, better-educated individuals should also be better equipped to cope with the non-monetary costs of migration (i.e. through better use of networks and access to information about the destination and migration route).

The negative selection of Venezuelan migrants living in Peru (Table 2 and Figure 4 and 5) combined with the results in Table 4 suggest that Peru, in addition to being unattractive to more educated migrants due to its low college premium (Table A3) compared to neighboring Colombia, has other confounding "non-pull" factors that make it the last choice for the most educated.

The explanations behind these unexpected results are likely multifaceted, each requiring a thorough and up-to-date analysis. However, it is reasonable to consider

that differing migration policies between Colombia and Peru may play a significant role.

The next section provides an overview of the policy background within Latin America, offering a potential explanation for the results observed concerning immigration to Peru.

5.2.1 Immigration policies

Until the onset of the Venezuelan crisis in 2015, Latin American citizens could migrate to other countries within the continent without the need for a visa or passport (Selee and Bolter, 2020). South American governments, through institutions such as MERCOSUR, CAN, CARICOM, and UNASUR, and a series of multilateral agreements had made the continent a liberal context in terms of free movement of people (Ceriani and Freier, 2015; Fernandez-Rodriguez et al., 2020). However, faced with a backlash from Peruvian public opinion following the first wave of migration in 2017-2018 (Figure 3), Peru began to implement restrictions starting in 2018 (Arcarazo et al., 2019). In 2018 it imposed passport requirements, and, realizing that the passport requirement had not had the desired effect, in June 2019 it required visas only for Venezuelan migrants (Selee and Bolter, 2020). As shown in Figure 6, although there was a spike in migrant arrivals right before the implementation of the restrictive policy – possibly due to anticipation of it – as of July 2019, the number of people migrating to Peru started to gradually decrease.

Despite the intended decrease in immigration, in the same month, the percentage of migrants crossing the border illegally increased dramatically (Figure 7). Concurrently, as migration policies became more restrictive and the number of illegal migrants increased, the skill composition of migrants began to decline gradually from the first quarter of 2018 (Figure 8). At the same time, Figure 9 shows that the skill composition of Venezuelan migrants did not change much over time.

One plausible explanation could be that introducing the visa increased migration costs for all categories of migrants, prompting some to consider migrating irregularly.

However, the cost opportunity related to irregular migration is not constant across educational classes. An educated Venezuelan who migrates irregularly would not be able to validate their degree and/or find a job in the formal labor market that matches their skills. A poorly educated Venezuelan, on the other hand, might be able to find the same job and salary they would have if they had migrated regularly. The other costs related to irregular migration – such as the likelihood of being caught or lack of health care – are the same for both educational classes.

In conclusion, while the effect of migration policies implemented by high-capacity and developed countries seems quite clear ([Ortega and Peri, 2009](#); [Mayda, 2010](#)), the effect of restrictive migration policies introduced by neighboring and developing countries requires much more research.

6 Conclusion

In this paper, I examine migrants' self-selection patterns according to individual and country characteristics, and pre-migration earnings in a context where their own country is undergoing a serious crisis and most destination countries are similarly fragile countries. I combine individual-level data representative of the Venezuelan population with similar data on Venezuelan migrants residing in Colombia, Peru, Chile, and the US.

The results show that Venezuelan migrants are positively selected based on observable characteristics compared to stayers. Specifically, Venezuelans residing in the US, Chile, and Colombia are positively selected in terms of education, while migrants in Peru are negatively selected. Moreover, South-North migrants appear positively sorted compared to South-South migrants. By comparing cumulative distribution functions of predicted pre-migration earnings, I find that migrants located in Chile and the US are positively selected compared to stayers in terms of predicted income. South-South migrants, instead, appear negatively selected compared to Venezuelans who chose not to leave the country. Such a result implies that unobservable characteristics play a key role in identifying the level of productivity of migrants.

I also find that income differentials and migration costs are robust determinants of Venezuelan migration choice. Discrete choice models reveal that controlling for the wage migration premium, migration costs still depend greatly on individual and destination country-specific characteristics. Particularly, women and college graduates face lower migration costs. Moreover, distance to the destination country and pre-crisis network play a key role in migration choice. In the case of Peru, however, it is surprising to find that having a college degree does not ensure lower migration costs. Although the evidence confirms that Peru has higher migration costs than Colombia given the greater geographic distance and weak network, Venezuelans residing in Peru are less educated.

These findings extend the literature on migrant self-selection, bringing empirical evidence in a context where South-South migration outweighs South-North migration. Although observable characteristics explain much of the variance in wages, unobservable characteristics also play an important role in estimating the migrant's level of productivity. In addition, distinguishing each migration cost faced by Venezuelan migrants is crucial for analyzing this specific migration episode and properly implementing migration policies in both destination countries and the country of origin. Although this current work illuminates the importance of self-selection and migration costs in shaping the migration decision, further studies focusing on South-South migration and effective policy interventions are highly recommended.

References

- Angelucci, Manuela**, “Migration and financial constraints: Evidence from Mexico,” *Review of Economics and Statistics*, 2015, 97 (1), 224–228.
- Arcarazo, Diego Acosta, Cécile Alice Blouin, and Luisa Feline Freier**, “La emigración venezolana: respuestas latinoamericanas,” *Documentos de trabajo (Fundación Carolina): Segunda época*, 2019, (3), 1.
- Bahar, Dany, Meagan Dooley, and Andrew Selee**, “Venezuelan migration, crime, and misperceptions: A review of data from Colombia, Peru, and Chile,” 2020.
- Bakewell, Oliver, Hein de Haas, Stephen Castles, Simona Vezzoli, and Gunvor Jónsson**, “South-South migration and human development: reflection on African experiences,” 2009.
- Bertoli, Simone, J Fernández-Huertas Moraga, and Francesc Ortega**, “Crossing the border: Self-selection, earnings and individual migration decisions,” *Journal of Development Economics*, 2013, 101, 75–91.
- , **Jesús Fernández-Huertas Moraga, and Francesc Ortega**, “Immigration policies and the Ecuadorian exodus,” *The World Bank Economic Review*, 2011, 25 (1), 57–76.
- Borjas, George J**, “Self-selection and the earnings of immigrants: reply,” *The American Economic Review*, 1990, 80 (1), 305–308.
- , **Ilpo Kauppinen, and Panu Poutvaara**, “Self-selection of emigrants: Theory and evidence on stochastic dominance in observable and unobservable characteristics,” *The Economic Journal*, 2019, 129 (617), 143–171.
- Caruso, German, Christian Gomez Canon, and Valerie Mueller**, “Spillover effects of the Venezuelan crisis: migration impacts in Colombia,” *Oxford Economic Papers*, 2021, 73 (2), 771–795.

- Ceriani, Pablo and Luisa Feline Freier**, “Lights and shadows in Latin American immigration legislation,” *A Liberal Tide: Towards a Paradigm Shift in Latin American Migration and Asylum Policy-Making*, Londres, *Institute of Latin American Studies*, 2015, pp. 11–32.
- Chiquiar, Daniel and Gordon H Hanson**, “International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States,” *Journal of political Economy*, 2005, 113 (2), 239–281.
- Clemens, Michael A and Mariapia Mendola**, “Migration from Developing Countries: Selection, Income Elasticity, and Simpson’s Paradox,” 2020.
- Clement, Viviane, Kanta Kumari Rigaud, Alex de Sherbinin, Bryan Jones, Susana Adamo, Jacob Schewe, Nian Sadiq, and Elham Shabhat**, *Groundswell part 2: Acting on internal climate migration*, World Bank, 2021.
- Dahl, Gordon B**, “Mobility and the return to education: Testing a Roy model with multiple markets,” *Econometrica*, 2002, 70 (6), 2367–2420.
- Doocy, Shannon, Kathleen R Page, Fernando De la Hoz, Paul Spiegel, and Chris Beyrer**, “Venezuelan migration and the border health crisis in Colombia and Brazil,” *Journal on Migration and Human Security*, 2019, 7 (3), 79–91.
- Fernandez-Rodriguez, Nieves, Luisa Feline Freier, and Omar Hammoud-Gallego**, “Importancia y limitaciones de las normas jurídicas para el estudio de la política migratoria en América Latina,” *Abordajes Sociojurídicos Contemporáneos Para El Estudio de Las Migraciones Internacionales*, edited by Gandini, L.. Mexico, 2020, pp. 85–119.
- Gazeaud, Jules, Eric Mvukiyehe, and Olivier Sterck**, “Cash transfers and migration: Theory and evidence from a randomized controlled trial,” *Review of Economics and Statistics*, 2023, 105 (1), 143–157.
- Grogger, Jeffrey and Gordon H Hanson**, “Income maximization and the selection and sorting of international migrants,” *Journal of Development Economics*, 2011, 95 (1), 42–57.

- Hanke, Steve**, “IMF Produces Another Bogus Venezuela Inflation Forecast,” <https://www.forbes.com/sites/stevehanke/2018/07/31/imf-produces-another-bogus-venezuela-inflation-forecast/?sh=46a98f637dc8> 2018. Last checked on 2022/07/21.
- Hanson, Gordon H**, “International migration and the developing world,” in “Handbook of development economics,” Vol. 5, Elsevier, 2010, pp. 4363–4414.
- Hujo, Katja and Nicola Piper**, *South-south migration: Implications for social policy and development*, Springer, 2010.
- Kaestner, Robert and Ofer Malamud**, “Self-selection and international migration: New evidence from Mexico,” *Review of Economics and Statistics*, 2014, 96 (1), 78–91.
- Lanati, Mauro and Rainer Thiele**, “South-south refugee movements: Do pull factors play a role?,” *Economics & Politics*, 2024.
- Maggio, Federico and Carlo Caporali**, “The Impact of Police Violence on Migration: Evidence from Venezuela.,” *Journal of Population Economics*, 2024, 37 (1).
- Mayda, Anna Maria**, “International migration: A panel data analysis of the determinants of bilateral flows,” *Journal of Population Economics*, 2010, 23 (4), 1249–1274.
- McKenzie, David and Hillel Rapoport**, “Self-selection patterns in Mexico-US migration: the role of migration networks,” *the Review of Economics and Statistics*, 2010, 92 (4), 811–821.
- , **Steven Stillman, and John Gibson**, “How important is selection? Experimental vs. non-experimental measures of the income gains from migration,” *Journal of the European Economic Association*, 2010, 8 (4), 913–945.
- Melde, Susanne, Rudolf Anich, Jonathan Crush, and John O Oucho**, “Introduction: The south–south migration and development nexus,” in “A new perspective on human mobility in the South,” Springer, 2014, pp. 1–20.

- Moraga, Jesús Fernández-Huertas**, “New Evidence on Emigrant Selection,” *The Review of Economics and Statistics*, 02 2011, 93 (1), 72–96.
- Namen, Olga, Mounu Prem, Sandra Rozo, and Juan F Vargas**, “Effects of Venezuelan migration on education outcomes in Colombia,” *Research Prosals Inter-American Bank of Development*, 2019.
- OECD**, “States of Fragility Report 2016,” 2016.
- Olivieri, Sergio, Francesc Ortega, Ana Rivadeneira, and Eliana Carranza**, “The Labour Market Effects of Venezuelan Migration in Ecuador,” *The Journal of Development Studies*, 2022, 58 (4), 713–729.
- Ortega, Francesc and Giovanni Peri**, “The causes and effects of international mobility: Evidence from OECD countries 1980% 2005,” *NBER Working Paper*, 2009, 14833.
- Rozo, Sandra V and Juan F Vargas**, “Brothers or invaders? How crisis-driven migrants shape voting behavior,” *Journal of Development Economics*, 2021, 150, 102636.
- Selee, Andrew and Jessica Bolter**, “An uneven welcome: Latin American and Caribbean responses to Venezuelan and Nicaraguan migration,” 2020.
- Sotomayor, Orlando**, “Puerto Rican Migration Flows and the Theory of Migrant Self-Selection,” *World Development*, 2009, 37 (3), 726–738.
- UNHCR, Global**, “Forced Displacement in 2021, UNHCR. org,” 2021.

Tables

Table 1: Descriptive statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Venezuela		Colombia		Peru		Chile		U.S.	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	0.52	0.50	0.54	0.50	0.51	0.50	0.53	0.50	0.50	0.50
Age	37	12	32	10	33	10	35	10	37	11
College	0.19	0.39	0.23	0.42	0.19	0.39	0.51	0.50	0.55	0.50
Labor income, USD	249	539	2,332	1,354	4,305	2,246	8,643	5,724	31,998	27,380
Non-college graduate	241	523	2,201	1,195	4,228	2,121	6,810	2,564	26,621	23,243
College graduate	276	594	2,756	1,708	4,612	2,667	10,201	7,055	36,356	29,643
College premium	0.15		0.25		0.09		0.50		0.37	
<i>Male</i>										
Age	36.82	12.42	31.55	9.75	32.35	9.51	34.32	9.53	36.32	11.24
College graduate	0.15	0.36	0.19	0.39	0.14	0.34	0.46	0.50	0.52	0.50
Labor income, USD	268	545	2,558	1,311	4,740	2,310	9,070	5,854	37,624	29,417
Non-college graduate	254	513	2,420	1,158	4,655	2,195	7,288	2,830	30,805	24,652
College graduate	342	691	3,145	1,704	5,241	2,846	10,887	7,389	43,719	31,950
College premium	0.35		0.30		0.13		0.49		0.42	
<i>Female</i>										
Age	36.80	12.44	32.12	10.34	33.43	10.46	35.44	9.94	36.98	10.85
College graduate	0.23	0.42	0.26	0.44	0.23	0.42	0.55	0.50	0.58	0.49
Labor income, USD	216	527	2,013	1,349	3,747	2,030	8,154	5,537	25,447	23,177
Non-college graduate	211	544	1,845	1,167	3,587	1,826	6,133	1,946	21,136	20,054
College graduate	225	499	2,407	1,636	4,178	2,447	9,544	6,668	28,559	24,780
College premium	0.07		0.30		0.16		0.56		0.35	
<i>Wage Female Penalty</i>										
Non-college graduate	-0.17		-0.24		-0.23		-0.16		-0.31	
College graduate	-0.34		-0.23		-0.20		-0.12		-0.35	
Observations	12,235		17,742		7,484		1,805		765	

Source: Author's elaboration on ENCOVI 2018, GEIH 2021, ENPOVE 2021, ACS 2021 and CASEN 2021.

Notes: The first panel includes the full sample of working-age people (aged 18 to 60). The second panel includes sample of male and the third sample of females. The wage female penalty represents the difference of wages between females and males by class of education. I applied 2020 USD exchange rate.

Table 2: Self-selection of Venezuelan migrants (aged 18 - 60)

	(1) All Migrants	(2) Colombia	(3) Peru	(4) Chile	(5) US
<i>Male</i>					
College Graduate	0.068*** (0.008)	0.059*** (0.011)	-0.034** (0.014)	0.213*** (0.013)	0.157*** (0.012)
Age	0.026*** (0.002)	0.021*** (0.002)	0.030*** (0.003)	0.015*** (0.002)	0.002 (0.002)
Age sq. * 1000	-0.461*** (0.028)	-0.424*** (0.031)	-0.505*** (0.036)	-0.221*** (0.027)	-0.032 (0.021)
t-test of Equality of Means (p-value)		0.047	0.000	0.000	0.000
Observations	18,855	13,926	9,442	6,649	6,214
R-squared	0.059	0.067	0.052	0.082	0.058
<i>Female</i>					
College Graduate	0.046*** (0.007)	0.025*** (0.009)	-0.002 (0.012)	0.174*** (0.011)	0.108*** (0.009)
Age	0.018*** (0.002)	0.016*** (0.002)	0.022*** (0.003)	0.015*** (0.002)	0.005*** (0.002)
Age sq. * 1000	-0.336*** (0.026)	-0.329*** (0.030)	-0.367*** (0.036)	-0.207*** (0.026)	-0.061*** (0.020)
t-test of Equality of Means (p-value)		0.000	0.000	0.000	0.000
Observations	21,063	15,855	10,081	7,195	6,673
R-squared	0.038	0.048	0.028	0.067	0.040

Notes: Standard errors in parentheses. Asterisks denote statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Outcome variable, migrant, is equal to 1 for Venezuelans who left the origin country and 0 for respondents in Venezuela at the time of the ENCOVI survey. The sample of column 1 includes stayers and the migrants residing in all destination countries. In column 2 I restrict the sample to stayers and migrants residing in Colombia. In column 3 the sample include stayers and migrants living in Peru. In column 4, the sample includes stayers and migrants in Chile. In column 5, the sample comprises stayers and migrants in the US. The college graduate coefficients are statistically different across all specifications.

Table 3: Location choice model: Conditional Logit Model

	(1) Male	(2) Female	(3) Male	(4) Female
Earnings	1.611** (0.921)	2.939*** (0.397)	1.611** (0.921)	2.939*** (0.397)
Colombia dummy	-5.212* (2.870)	-9.588*** (1.231)		
Peru dummy	-7.418*** -2.492	-12.966*** (2.014)		
Chile dummy	-9.608** (4.238)	-16.153*** (2.800)		
US dummy	-11.920** (5.153)	-19.160*** (2.936)		
Network			0.491** (0.237)	0.917*** (0.107)
Distance			-1.018* (0.539)	-1.512*** (0.506)
Unemployment			0.079 (0.272)	0.136 (0.303)
Language			2.319 (2.098)	2.964* (1.534)
Observations	13,546	8,530	13,546	8,530
Pseudo R-squared	0.244	0.237	0.244	0.237
Log-likelihood	-26535.84	-26535.84	-26535.84	-26535.844

Notes: Standard errors in parentheses. Asterisks denote statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. This Table shows the estimates of the conditional logit model specifications, in which the effects of the individual characteristics on migration costs do not vary across alternatives. The first row indicates the effect of expected earnings at destination to migrate to that country. The higher the expected earnings, the greater the probability of migrating to that destination. In the first two columns, I only include expected earnings and country intercepts. In columns 3 and 4, I replace the country intercept with country-specific migration cost variables, such as distance to country of origin (cost of travel), network at destination, pre-crisis unemployment rate, and a language dummy (which takes value 1 if spanish is the official language at destination).

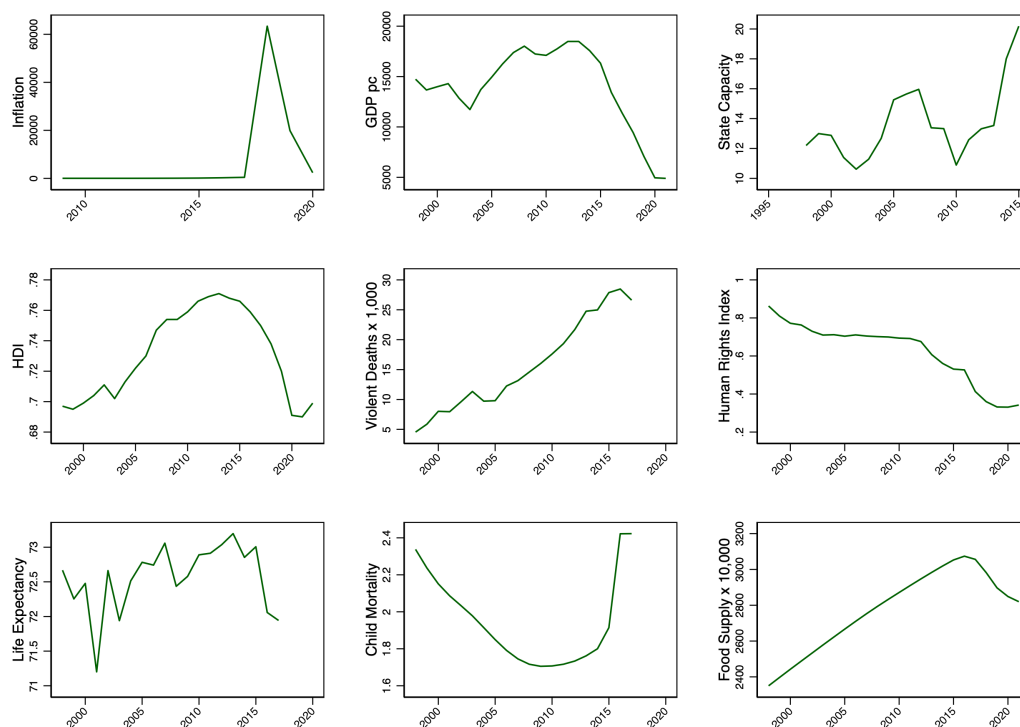
Table 4: Location choice model: Mixed Logit Model

	(1) Colombia	(2) Peru	(3) Chile	(4) USA
Earnings		0.651*** (0.108)		
College graduate	0.165*** (0.042)	0.035 (0.060)	1.395*** (0.074)	1.496*** (0.085)
Age	0.014 (0.011)	0.034** (0.015)	0.131*** (0.027)	-0.059** (0.028)
Female	0.287*** (0.035)	0.318*** (0.044)	0.070 (0.075)	0.261*** (0.084)
Age sq*1000	-0.001*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)	0.001* (0.000)
Country Intercept	-1.185** (0.482)	-3.015*** (0.586)	-7.203*** (0.722)	-5.428*** (0.797)
Number of cases		22,967		
Log-likelihood		-27337		

Notes: Standard errors in parentheses. Asterisks denote statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. This Table shows the estimates for the Mixed Logit Model. The first row indicates the effect of expected earnings at destination to migrate to that country. The higher the expected earnings, the greater the probability of migrating to that destination. The first column shows the effect of individual characteristics on the likelihood of migrating to Colombia, column 2 of migrating to Peru, column 3 to Chile, and the last column the choice to migrate to the United States. The sample includes males and females.

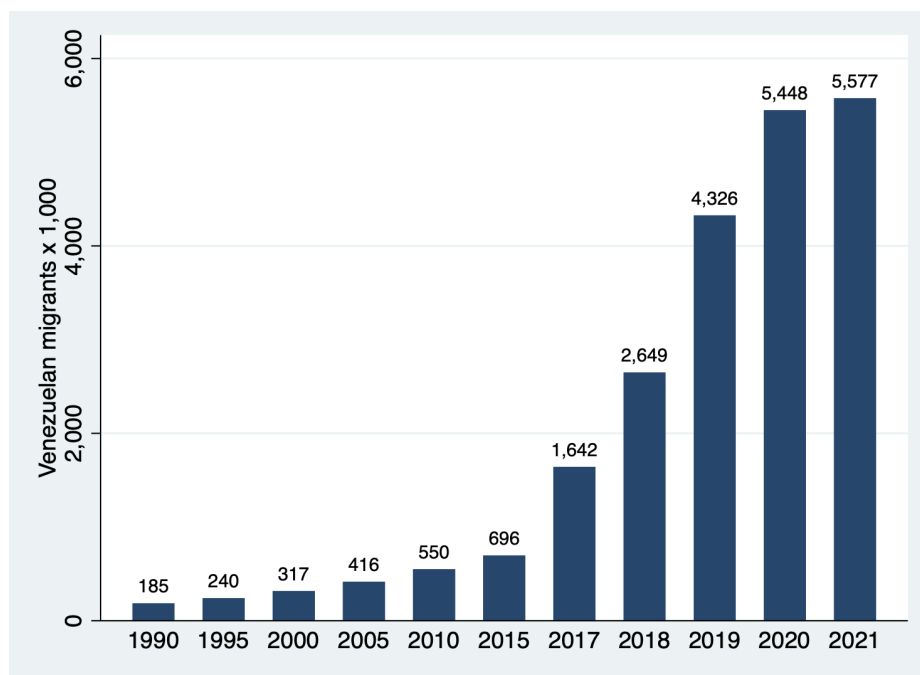
Figures

Figure 1: Venezuelan crisis shown through several economic, political, and health dimensions



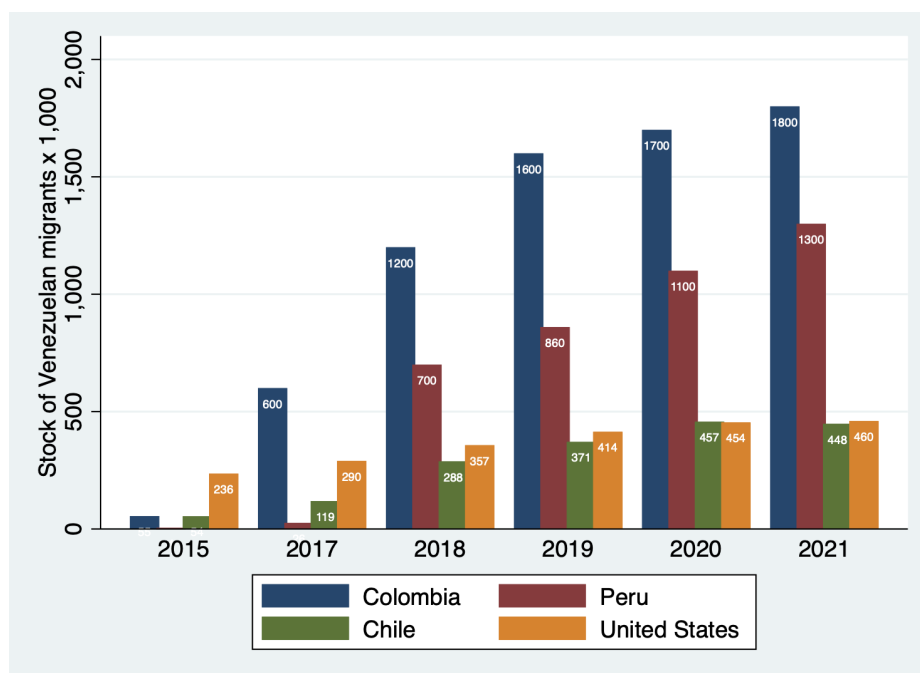
Source: World Bank (Economic Indicators), United Nations (Health and human capital indicators), OVV (Observatorio Venezolano de Violencia) (violence indicators).

Figure 2: Stock of Venezuelan migrants by year (1990 - 2021)



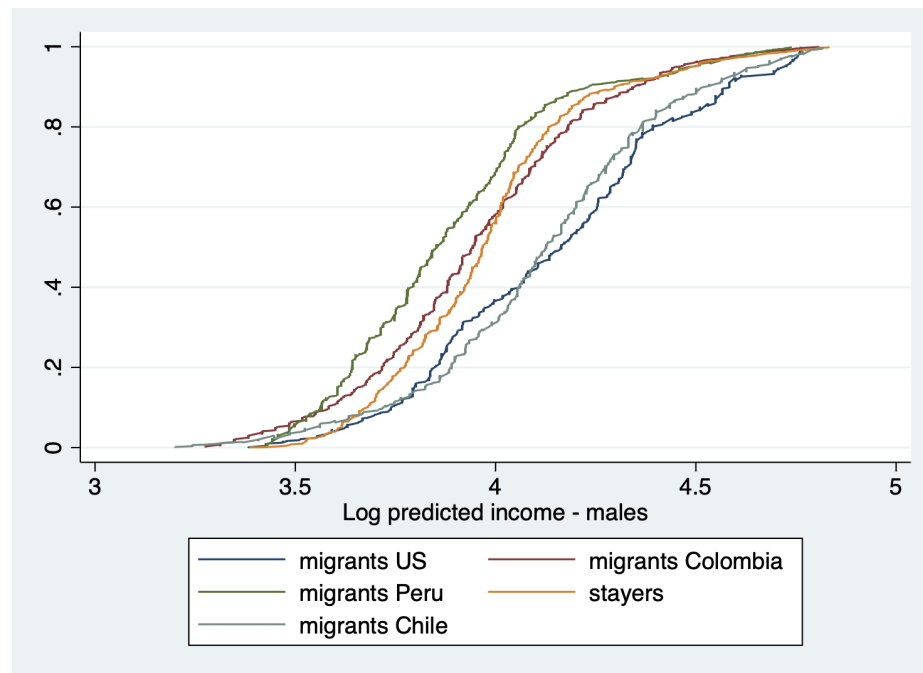
Source: Author's calculations based on R4V - Plataforma de Coordinación para Refugiados y Migrantes de Venezuela and United Nations High Commissioner for Refugees (UNHCR) data.

Figure 3: Stock of Venezuelan migrants by year and country (2015 - 2021)



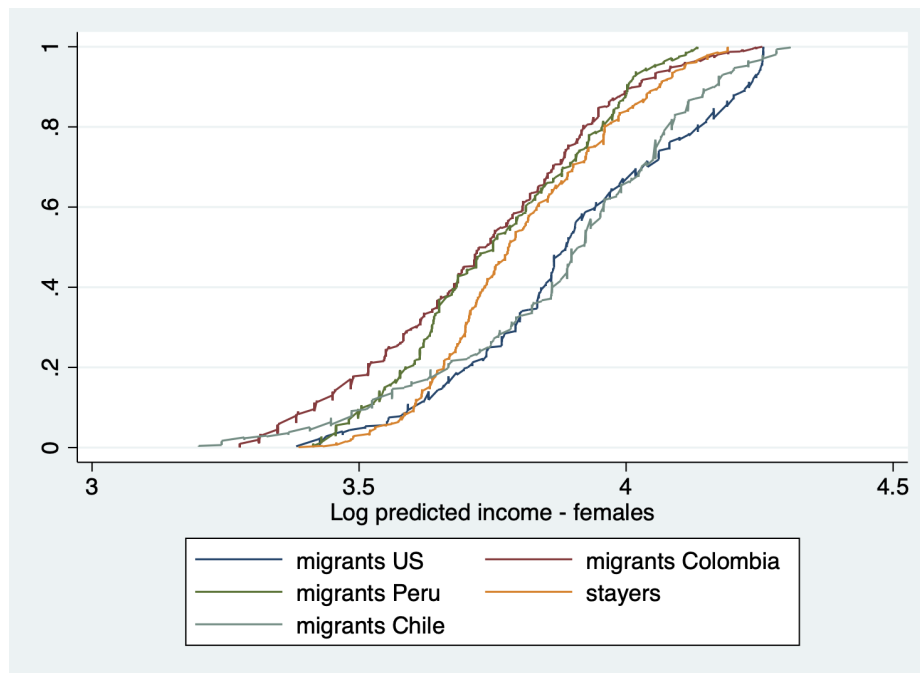
Source: Author's calculations based on R4V - Plataforma de Coordinación para Refugiados y Migrantes de Venezuela and United Nations High Commissioner for Refugees (UNHCR) data.

Figure 4: Cumulative distribution functions of predicted income in Venezuela - Males



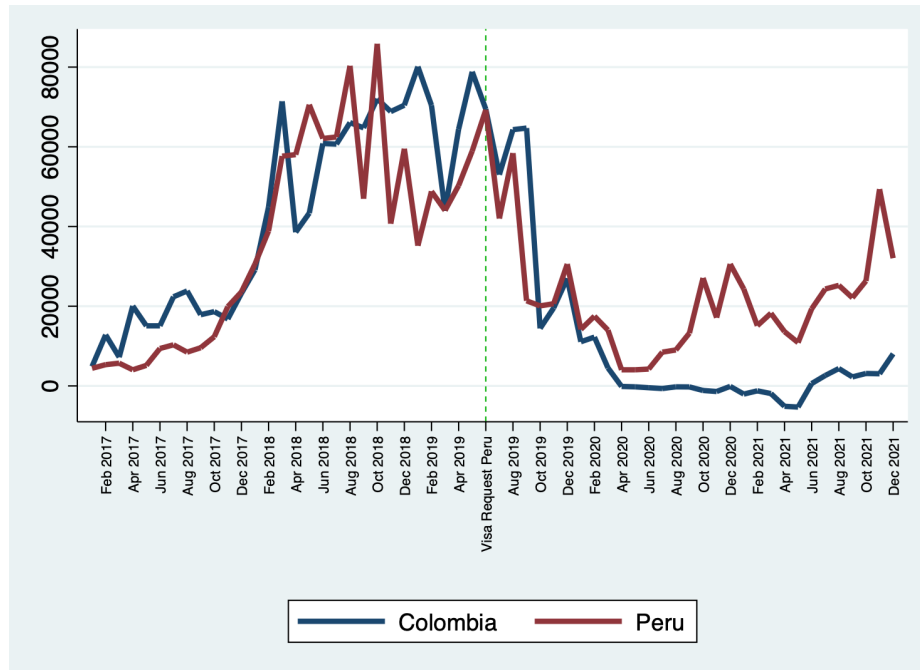
Notes: This chart presents cumulative distribution functions of Venezuelan migrants' and stayers' predicted income in the origin country (Male sample). I used the predicted earning selection-corrected à la Dahl.

Figure 5: Cumulative distribution functions of predicted income in Venezuela - Females



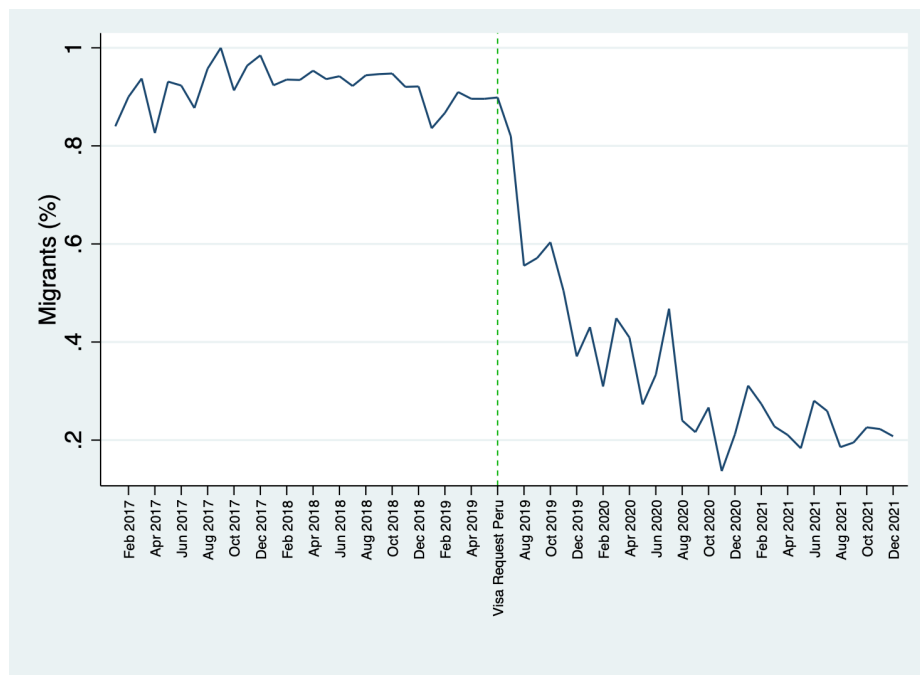
Notes: This chart presents cumulative distribution functions of Venezuelan migrants' and stayers' predicted income in the origin country (Female sample). I used the predicted earning selection-corrected à la Dahl.

Figure 6: Venezuelan arrivals in Peru and Colombia (Jan 2017 - Dec 2021)



Notes: This chart presents the arrivals of Venezuelan migrants in Peru and Colombia. Source: For Peru, I use elaborations from ENPOVE 2021. For Colombia, I extrapolate data from the Ministerio de Relaciones Exteriores.

Figure 7: Percentage of regular migrants in Peru



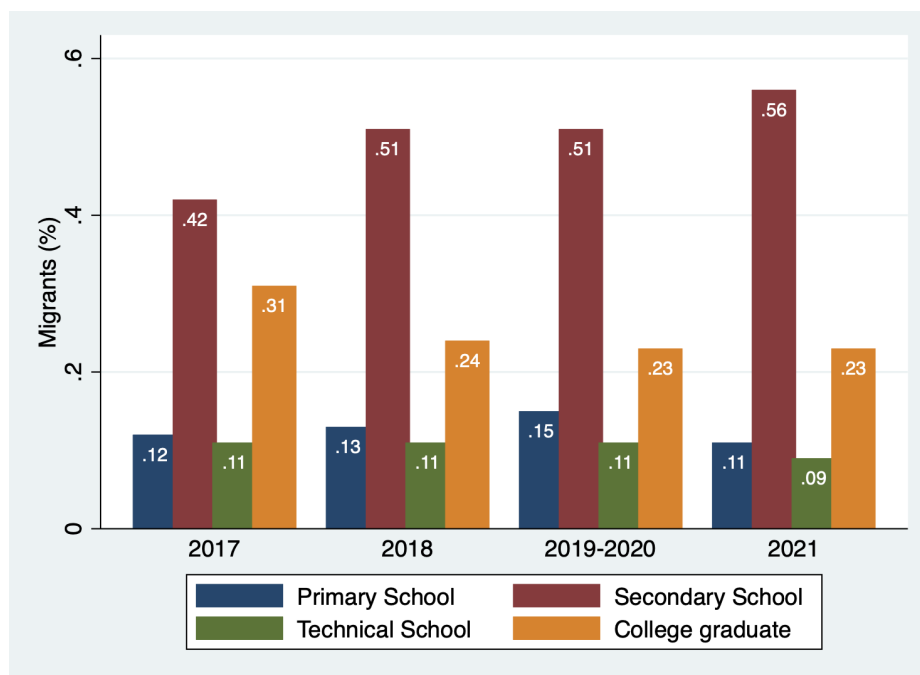
Notes: This chart presents the percentage of migrants who regularly entered Peru over time (Jan 2017 - Dec 2021) Sources: Author's elaboration of ENPOVE 2021 data.

Figure 8: Share of college-graduated Venezuelans in Peru



Notes: This chart presents the share of college-graduated Venezuelans (Jan 2017 - Dec 2021) Sources: Author's elaboration of ENPOVE 2021 data.

Figure 9: Skill composition of Venezuelan migrants over time



Notes: This chart presents the skill composition of migrants over time Sources: ENCOVI 2017, ENCOVI 2018, ENCOVI 2019 - 2020, ENCOVI 2021

Appendix

Table A1: Self-selection of Venezuelan migrants (aged 18 - 60) - Non-zero earnings

	(1) All Migrants	(2) Colombia	(3) Peru	(4) Chile	(5) US
<i>Male</i>					
College Graduate	0.077*** (0.009)	0.060*** (0.012)	-0.000 (0.018)	0.237*** (0.017)	0.212*** (0.017)
Age	0.012*** (0.002)	0.010*** (0.003)	0.002 (0.004)	0.012*** (0.003)	-0.001 (0.003)
Age sq. * 1000	-0.298*** (0.033)	-0.294*** (0.038)	-0.175*** (0.047)	-0.185*** (0.035)	0.011 (0.035)
Observations	13,507	10,324	5,873	4,143	3,979
R-squared	0.069	0.081	0.068	0.097	0.080
<i>Female</i>					
College Graduate	-0.007 (0.009)	-0.044*** (0.012)	-0.053*** (0.018)	0.152*** (0.017)	0.124*** (0.016)
Age	0.002 (0.003)	0.003 (0.004)	-0.013*** (0.005)	0.007 (0.005)	-0.010** (0.005)
Age sq. * 1000	-0.149*** (0.040)	-0.186*** (0.049)	0.009 (0.066)	-0.135** (0.059)	0.109* (0.060)
Observations	9,385	6,823	3,815	2,517	2,368
R-squared	0.049	0.066	0.079	0.053	0.032

Notes: Standard errors in parentheses. Asterisks denote statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The outcome variable, migrant, is equal to 1 for Venezuelans who left the country of origin and 0 for respondents in Venezuela at the time of the ENCOVI survey. The sample of column 1 includes stayers and migrants residing in all destination countries. In column 2 I restrict the sample to stayers and migrants residing in Colombia. In column 3 the sample includes stayers and migrants living in Peru. In column 4, the sample includes stayers and migrants in Chile. In column 5, the sample comprises stayers and migrants in the US. The college graduate coefficients are statistically different across all specifications.

Table A2: Self-selection of Venezuelan migrants (aged 18 - 60) - Logit specification

	(1) All Migrants	(2) Colombia	(3) Peru	(4) Chile	(5) US
<i>Male</i>					
College Graduate	0.357*** (0.044)	0.260*** (0.048)	-0.166*** (0.063)	1.475*** (0.079)	1.762*** (0.103)
Age	0.108*** (0.010)	0.095*** (0.011)	0.168*** (0.014)	0.186*** (0.027)	0.037 (0.033)
Age sq. * 1000	-1.976*** (0.133)	-1.870*** (0.149)	-2.760*** (0.200)	-2.760*** (0.370)	-0.575 (0.420)
Observations	16,876	12,189	9,825	6,084	6,154
<i>Female</i>					
College Graduate	0.239*** (0.037)	0.109*** (0.039)	-0.014 (0.051)	1.307*** (0.073)	1.477*** (0.103)
Age	0.076*** (0.009)	0.065*** (0.010)	0.111*** (0.013)	0.169*** (0.025)	0.095*** (0.034)
Age sq. * 1000	-1.424*** (0.123)	-1.362*** (0.135)	-1.812*** (0.176)	-2.396*** (0.332)	-1.254*** (0.438)
Observations	17,039	12,826	9,769	6,535	6,650

Notes: Standard errors in parentheses. Asterisks denote statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Outcome variable, migrant, is equal to 1 for Venezuelans who left the origin country and 0 for respondents in Venezuela at the time of the ENCOVI survey. The sample of column 1 includes stayers and the migrants residing in all destination countries. In column 2 I restrict the sample to stayers and migrants residing in Colombia. In column 3 the sample include stayers and migrants living in Peru. In column 4, the sample includes stayers and migrants in Chile. In column 5, the sample comprises stayers and migrants in the US. The college graduate coefficients are statistically different across all specifications.

Table A3: Determinants of labor earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Venezuela		Colombia		Peru		Chile		Usa	
	Standard	Selection	Standard	Selection	Standard	Selection	Standard	Selection	Standard	Selection
	Wage Eq	corrected	Wage Eq	corrected	Wage Eq	corrected	Wage Eq	corrected	Wage Eq	corrected
<i>Male</i>										
College graduate	0.391*** (0.087)	-0.182 (0.126)	0.245*** (0.019)	0.231*** (0.051)	0.061** (0.028)	0.093** (0.045)	0.287*** (0.044)	0.483*** (0.120)	0.309*** (0.100)	0.646*** (0.213)
Age	0.024 (0.019)	0.050** (0.020)	0.052*** (0.005)	0.056*** (0.006)	0.035*** (0.007)	0.034*** (0.008)	0.057*** (0.017)	0.064*** (0.019)	0.077** (0.031)	0.096*** (0.036)
Age squared * 1,000	-0.213 (0.246)	-0.182 (0.244)	-0.712*** (0.071)	-0.664*** (0.084)	-0.445*** (0.094)	-0.461*** (0.124)	-0.674*** (0.224)	-0.716*** (0.257)	-0.861** (0.396)	-0.981** (0.453)
Observations	3,604	3,604	6,720	6,720	2,269	2,269	539	539	373	373
R-squared	0.008	0.019	0.045	0.056	0.020	0.021	0.116	0.125	0.077	0.087
P-Value, F-test										
Dahl polyn.		0.000		0.119		0.770		0.046		0.282
<i>Female</i>										
College graduate	0.349*** (0.084)	0.319*** (0.090)	0.246*** (0.027)	0.322*** (0.032)	0.121*** (0.027)	0.119*** (0.030)	0.345*** (0.050)	0.369*** (0.076)	0.319*** (0.119)	-0.002 (0.236)
Age	0.080*** (0.027)	0.091*** (0.030)	0.011 (0.009)	0.015 (0.011)	0.021** (0.008)	0.031*** (0.010)	0.023 (0.020)	0.028 (0.028)	0.029 (0.038)	0.026 (0.045)
Age squared * 1,000	-0.949*** (0.342)	-1.013*** (0.349)	-0.209* (0.120)	-0.152 (0.155)	-0.313*** (0.111)	-0.528*** (0.139)	-0.344 (0.265)	-0.399 (0.375)	-0.466 (0.485)	-0.492 (0.523)
Observations	2,046	2,046	4,777	4,777	1,769	1,769	471	471	320	320
R-squared	0.015	0.016	0.021	0.028	0.020	0.027	0.110	0.110	0.031	0.043
P-Value, F-test										
Dahl polyn.		0.370		0.009		0.000		0.660		0.499

Notes: Standard errors in parentheses. Asterisks denote statistical significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The odd-numbered data columns report the results of the Standard wage equations, while the even columns show the Dahl earnings regression estimates. I include the Dahl term as a second-order polynomial in the retention probability for stayers and a second-order polynomial in the retention and first-best probability for migrants plus an interaction term. The sample has been restricted to working-age population (aged 18 to 60), who reports wage different from zero.