Immigration and Occupational Downgrading in Colombia
Click here for most recent version

Jeremy Lebow *

September 2021

Abstract

Between 2015-2019, around 1.8 million Venezuelans migrated into neighboring Colombia. Despite having similar education as native Colombians, these migrants are disproportionately employed in occupations with less-educated natives. In this paper, I study the consequences of this migrant occupational downgrading for native labor market outcomes. I estimate a model of labor demand that incorporates migrant downgrading and imperfect substitutability between migrants and natives. I find that migrants and natives are more substitutable in low-skill occupations and that substitutability across education groups as low. As a result, the model predicts large consequences of migrant downgrading for the hourly wages of less educated natives, alongside minimal benefits for more educated natives. This reflects the aggregate productivity benefits that would result from moving migrants into high-skill occupations, as well as the greater complementarity between migrants and natives in these occupations.

Keywords: immigration, occupational downgrading, informality
JEL Codes: F22, J24, J46

---

*Duke University, email: jal99@duke.edu; I thank Duncan Thomas, Daniel Xu, Arnaud Maurel, Erica Field, Rafael Dix-Carneiro, Rob Garlick, Giovanni Peri, Sandra Rozo, Jaime Tenjo, Piedad Urdinola, Jonathan Moreno and Thomas Ginn, along with participants at the IZA Annual Migration Meeting, the Econometric Society NA Summer Meeting, the World Bank Jobs & Development Conference, and the Duke Labor and Development lunch groups. I am also grateful to the Departamento Administrativo Nacional de Estadística in Colombia for generously providing the data used in this paper.

1 Introduction

Between 2015-2019, approximately 1.8 million Venezuelans fled into neighboring Colombia, increasing the Colombian population by almost 4%. Despite being well-educated on average, upon arrival these migrants entered occupations that typically require less education, which is a common characteristic of migration waves and especially episodes of forced displacement (Brell *et al.* , 2020; Chiswick & Miller, 2009). To put this in perspective, over 25% of working Venezuelan migrants in Colombia in 2019 had post-secondary education, and around 70% of them were in occupations where the mode native did not go to college. This over-education of migrants relative to Colombian natives in their occupation, which I define as occupational downgrading, has stimulated debate regarding the economic consequences of migrant occupational downgrading for low-skill natives, as well as potential lost opportunities for increased productivity.[1]

In this paper, I evaluate the consequences of Venezuelan migrant occupational downgrading for native Colombian labor market outcomes. I estimate a model that captures three key channels. First and most obviously, if migrants are disproportionately concentrated in low-skill occupations, then the effects of economic competition will be strongest for these workers.[2] Second, the impact of migrant downgrading on native labor outcomes also depends on the substitutability between migrants and natives within different types of occupations. Because migrants and natives work different jobs, complete different tasks within jobs and possess different skills, this substitutability is an empirical question that may vary across contexts and occupations. It may be that migrants in low-skill jobs, which are more routine and less specialized, are more substitutable with natives than migrants in high-skill jobs. For example, migrants working in restaurant or construction jobs (two of the most common occupations among Venezuelan migrants in Colombia) may be more substitutable with natives than migrants working as engineers or software developers. If this is the case, then the impact of migrant competition will be larger in low-skill occupations, and this will amplify the average consequences of downgrading for native labor outcomes. Finally, if native substitutability *across* skill groups is low, then the effects of migrant competition will fall more heavily on low skill workers if these workers face a disproportionate increase in labor supply.

[1]For example, the think-tank Global Americans stated in August 2021 that “the Colombian economy can benefit from the vast talent pool of Venezuelan professionals who have migrated to Colombia and currently rely on informal work when they could be more productive in other sectors of the economy” (Guzmán & Marmolejo, 2021).

[2]Throughout this paper, I define the occupation skill distribution according to the average completed schooling of natives in each occupation, even though, strictly speaking, jobs can require little education but important skills.

To capture these elements, I estimate a nested CES demand system that incorporates imperfect substitutability between migrants and natives within skill groups and imperfect substitutability across skill groups. This framework has been used to study the labor market impacts of immigration in the US and in Europe (Ottaviano & Peri, 2012; Manacorda *et al.* , 2012), but in its traditional form does not consider that migrants often compete with natives in different education groups (Dustmann *et al.* , 2016). The methodological contribution of this paper is to adapt this framework to directly incorporate occupational downgrading. I do this using data on migrants' observed occupation in Colombia, combined with a random assignment algorithm that places migrants into education groups according to the native education distribution in their observed occupation. I therefore estimate the elasticities of substitution between natives in an education group and the migrant workers who they directly compete with. This framework allows me to quantify the total wage effects of the Venezuelan migration and how it would differ in the absence of occupational downgrading.

There are various reasons why this setting is particularly interesting to study. First, genuinely supply-driven migration flows of this magnitude are rare. The causes of the Venezuelan crisis were unrelated to social and economic changes happening in Colombia at the time, instead driven by drastic increases in poverty, violence and inadequate health and education services in Venezuela that were stimulated by the sudden collapse of oil prices in 2014 (Migration Policy Institute, 2020). Colombia is the largest recipient of Venezuelans primarily because it is the country located closest to the population centers of Venezuela, with relatively few restrictions on who could enter the country. This therefore presents a unique opportunity to study the economic impacts of a massive and supply-driven migration wave.

Second, the developing country setting has received less attention in the literature on the economic impacts of migrants and forced displacement. This is despite the fact that 85% of people displaced across borders are hosted in a developing country and almost 75% in countries neighboring the country of origin (UNHCR, 2019). Developing countries tend to have large informal sectors - in Colombia, around 60% of native and over 90% of Venezuelan migrants remain informal.[3] These jobs are characterized by a lack of minimum wages, ease of entry, low hiring costs, high turnover rates, and labor supply flexibility, implying that substitutability between migrants and natives in these primarily low-skill jobs may be particularly large. The elasticity of substitution between education groups also tends to be lower in developing countries - in particular, the literature estimates much smaller substitutability between those with and without a high school degree in Latin America than in the US (Manacorda *et al.* , 2010; Fernández & Messina, 2018; Acosta *et al.* , 2019). Both

[3] I define informality according to enrollment in mandatory health and pension schemes.

of these characteristics may amplify the economic consequences of migrant downgrading for low-skill natives in the developing country context.

Third, a unique characteristic of this migration wave is that Venezuelans and Colombians share a similar culture and language, thus mitigating any effects of cultural or linguistic barriers. This may further increase migrant-native substitutability (Braun & Mahmoud, 2014). It also eliminates a key channel through which natives often upgrade occupations, in which migrants specialize in manual occupations and natives upgrade to communication-intensive occupations (Peri & Sparber, 2009; Peri *et al.* , 2020), thus decreasing substitutability across education groups.

Finally, there is substantial evidence that, in the short-run, this migration wave decreased hourly wages for Colombia natives and especially for less educated natives. I document this in a companion paper, Lebow (2021), in which I study the effects of changes in the total migrant share on economic outcomes for natives, using an instrumental variable strategy and geographic variation across 79 metropolitan areas. I also find little evidence for changes in employment, implying that most of the economic consequences for natives are realized along the hourly wage margin.[4] These results make no assumptions about the substitutability between migrants and natives, and are informative about the total effects of immigration on economic outcomes for natives overall and by demographic group (Dustmann *et al.* , 2016). Considering the characteristics of the setting described in the last three points, migrant occupational downgrading may play a uniquely important role in these wage effects that disproportionately impact less educated natives. The structural model allows me to directly study the role of migrant downgrading in this setting.

My research design exploits geographic variation in the arrival of migrants across 79 metropolitan areas constructed according to commuting patterns (Duranton, 2015). Labor market data comes from the Colombian National Integrated Household Survey (GEIH), which contains detailed labor market outcomes for both migrants and natives. Combined with the size of the migration and the sample of over 750,000 observations per year, this allows me to directly compare economic outcomes for migrants and natives across locations and over time, which is rare in the developing country setting and necessary to estimate this model.

The research design addresses two fundamental challenges highlighted by the economics

[4] These results are robust to accounting for native internal migration, use of various instruments, alternative geographic units and various other modelling choices. They are also consistent with other published and working papers studying the economic effects of Venezuelans in Colombia (Delgado-Prieto, 2021; Bonilla-Mejía *et al.* , 2020; Santamaria, 2020; Rozo & Vargas, 2021; Caruso *et al.* , 2019). See Lebow (2021) for details.

and migration literatures. First, to account for endogenous sorting of migrants into destinations within Colombia, I use an instrumental variable strategy based on the historical settlement patterns of Venezuelans (Altonji & Card, 1991; Card, 2001, 2009). The historical location of Venezuelans was determined two decades before the onset of the Venezuelan exodus, well before the election of Hugo Chávez, and I show it was for the most part uncorrelated with levels or trends in economic outcomes leading up to the current immigration (Goldsmith-Pinkham *et al.* , 2020). Furthermore, the historical number of Venezuelans was miniscule relative to the recent inflow, and there was very little immigration from Venezuela or any other country before 2014, mitigating the concern highlighted in Jaeger *et al.* (2018) regarding serial correlation in economic outcomes. At the same time, these historical flows strongly predict the current location of Venezuelans, making this instrument uniquely appropriate in this setting. Second, I use labor data matched to migration histories to test robustness of results to the spatial arbitrage of native workers (Borjas, 2003; Borjas & Katz, 2007; Monras, 2020). I find that Venezuelan arrivals caused small increases in native out-migration, and my labor market analysis is robust to accounting for this native internal migration.

My primary contribution is to estimate a production function that unpacks the relationship between migrant occupational downgrading and the labor market effects of immigration on natives across the education distribution. Demand-side models of the labor market have long been applied to study the relationship between the relative supply of skilled and unskilled workers and wage inequality.[5] These models have also been applied to study the impacts of immigration on native wages, with the adjustment of allowing for imperfect substitutability between migrants and natives, which substantially reduces estimates of the native wage effect (Ottaviano & Peri, 2012; Manacorda *et al.* , 2012; Card, 2009). However, these models do not explicitly incorporate migrant occupational downgrading, since they simply assign migrants to education groups without considering that they are effectively competing with workers in different groups (Dustmann *et al.* , 2013, 2016).[6]

I therefore adapt these models with a method of assigning migrants to education groups according to their observed occupation, directly incorporating occupational downgrading into this framework. Under certain assumptions, this allows me to construct a counterfactual in which I estimate the labor market effects of Venezuelan immigration in the absence of

[5] For example, see Katz & Murphy (1992); Katz *et al.* (1999); Goldin & Katz (2009); Goldin *et al.* (2020).

[6] One can still interpret the substitutability parameters estimated in these models as incorporating the role of downgrading, agnostic on the mechanisms that drive substitutability. However, if the goal is to think about the effects of migration under different degrees of downgrading, then we need a framework that directly incorporates downgrading separately from the other mechanisms that drive substitutability.

occupational downgrading. I also contribute to the literature by applying this framework for the first time to studying immigration in a developing country context, where, as discussed, migrants and natives may be more substitutable, and workers with and without a high school degree have been found to be less substitutable. In doing so, I shed light on a mechanism that may amplify the labor market impacts of migration in developing countries when that migration is concentrated at the bottom of the skill distribution.

I estimate migrant-native elasticities of substitution of around 6 for workers with post-secondary education, implying levels of short-run substitutability comparable with that of high-skill workers recently arrived in the UK (Manacorda *et al.* , 2010). Importantly, this elasticity increases to 15 for workers without secondary schooling - relatively large considering that most Venezuelans have been in Colombia for only a few years, and comparable in magnitude with the average long-run migrant-native substitutability estimated in the US (Ottaviano & Peri, 2012). That migrants and natives are more substitutable in low-skill occupations is consistent with patterns observed in the US (Card, 2009), and could be driven by the fact that these jobs tend to be more routine and less specialized. It may also be that these jobs are more likely to be informal - I observe that migrant-native substitutability is similar but slightly larger in the informal sector, lending some support to this hypothesis. Finally, consistent with the literature in Latin America (Manacorda *et al.* , 2010; Fernández & Messina, 2018; Acosta *et al.* , 2019), I estimate low substitutability between workers with and without secondary education, and this has important consequences for the simulated effects of downgrading.

I then calculate the total wage impacts of immigration over 2015-2019 predicted by the production function. Considering the short time-frame of the migration (the majority of migrants did not arrive until after 2018), I hold capital fixed, such that there are diminishing returns in the labor aggregate. This generates total predicted wage effects of -1.1% overall and -4% for workers without a high-school degree. I then simulate a counterfactual in which migrants are reallocated to jobs that match their observed education. This reallocation reduces the wage effect for workers without a high school degree from -4% to -3.1%. At the same time, despite facing increased within-group competition, the wage effects for natives with completed high school is unchanged, and natives with completed college are only slightly negatively impacted. This is driven by both the larger complementarity between migrants and college-educated natives, as well as the large gains in output from increasing labor supply in high-skill occupations, which are more productive and relatively under-supplied. In other words, better migrant-occupation matching directly reduces labor market competition for low-skill jobs, but it also generates productivity benefits for all workers.

The results highlight the importance of properly matching migrants to jobs given their education and experience, not only to directly benefit migrant workers, but also to mitigate negative wage effects for the most vulnerable natives and to maximize the economic gains from migration. Much of the discussion over labor market integration has focused on the importance of language training and the right to work (Clemens *et al.* , 2018; Lochmann *et al.* , 2019). The case of Colombia shows that this is not enough: Colombians and Venezuelans speak the same language, and many Venezuelan migrants in Colombia have access to legal work status, thanks to a globally unprecedented regularization program that incorporated around 60% of undocumented Venezuelans (Rozo *et al.* , 2020; Migración Colombia, 2019). Despite this regularization, over 90% of migrants remained in the informal sector at the end of 2019. Additional frictions may include a lack of networks with employers, barriers in occupational licencing, low recognition of education attained abroad, uncertainty over legal status or return migration plans, and anti-immigrant sentiment among employers. Distinguishing the importance of these mechanisms will help policy-makers tackle the sources of occupational downgrading.

My analysis excludes various other factors that are likely to further increase the benefits of reduced downgrading, such as skill-biased capital adjustments, sector-specific transfer of human capital and networks (Bahar *et al.* , 2020, 2019), or the increases in spending and taxable income that would accompany increasing migrant salaries (Graham *et al.* , 2020). There are also additional costs to economic competition that I do not model, such as its impact on anti-immigrant sentiment and reduced social cohesion (Card *et al.* , 2012; Alesina *et al.* , 2018; Lebow *et al.* , 2021). I focus on a very specific channel of labor market competition and changes in total output that is likely to underestimate the societal benefits from reducing migrant downgrading.

This paper relates to a large literature on over-education and labor market mismatch of native and migrant workers, which documents the systematic over-education of migrants relative to natives across many countries, its persistence over time and its negative consequences for migrant earnings and assimilation (Leuven & Oosterbeek, 2011; Piracha *et al.* , 2012; Chiswick & Miller, 2009; Eckstein & Weiss, 2004). This paper also fits into the broader literature studying the causal effects of large and sudden migration episodes on host country labor markets (Beerli *et al.* , 2021; Edo, 2020; Peri *et al.* , 2020; Clemens & Hunt, 2019; Peri & Yasenov, 2019; Dustmann *et al.* , 2017; Foged & Peri, 2016; Friedberg, 2001; Hunt, 1992; Card, 1990). The literature studying the economic effects of forced displacement has grown recently. A recent review by (Verme & Schuettler, 2021) highlights that the majority of these studies find insignificant effects on native employment and occupations, and when effects are significant, they tend to be negative. The positive effects are most likely to be observed in the developed country context (such as in Foged & Peri (2016), which documents increases in native task complexity in response to refugee dispersal policies in Denmark). A relevant middle-income country setting is that of Syrian refugees in Turkey, which also caused a large and sudden labor supply increase mostly in the informal sector. In this case, the literature has found some evidence for occupational upgrading by natives (Del Carpio & Wagner, 2015; Ceritoglu *et al.* , 2017; Tumen, 2016; Altındağ *et al.* , 2020), with small or negligible negative wage effects specifically for informal workers (Aksu *et al.* , 2018). As discussed, one reason that wage effects may be larger in Colombia, and occupational upgrading less prevalent, is that Venezuelan migrants speak the same language, thus increasing substitutability with natives and reducing the potential for communication-intensive occupational upgrading by natives (Peri & Sparber, 2009).

This paper proceeds as follows: Section 2 overviews the causes of the Venezuelan migration to Colombia. Section 3 discusses the data and characterizes migrant downgrading. Section 4 highlights key results regarding the total economic effects of Venezuelan migration. Section 5 outlines the model and estimation, Section 6 presents the results and Section 7 concludes.

2 The Venezuelan Exodus

Venezuela and Colombia were once part of the 19th-century "Gran Colombia" state that encompassed much of northern South America. They have historically had close relations characterized by a common culture and language, heavy trade exchanges and migratory flows. Until recently, favorable economic conditions meant there was relatively little migration out of Venezuela. In fact, Venezuela was a major recipient of Latin American migrants, who were attracted by relatively low poverty rates and generous social programs, including Colombians fleeing the decades-long civil war in Colombia.

This began to change when Hugo Chávez begin his term as President in 1999. Migration to Venezuela slowed and members of the Venezuelan upper class began to migrate away, primarily to the US and Spain, with concerns that socialist reforms were paving the way for economic instability (Freitez, 2011). Shortly after the death of Chávez in 2013 and the inauguration of his successor Nicolás Maduro in 2014, global oil prices collapsed. This, combined with a dramatically diminished private sector and weakened oil and agricultural industries, prompted an economic recession with a severity that was unanticipated. In

2016, Venezuela entered hyperinflation. By 2018, GDP had contracted by 45% since 2013, and over 90% of the population was estimated to be living in poverty. Over 20% of the population was undernourished, access to water and electricity became increasingly scarce, and an estimated 85% of essential medicines were scarce (Wilson Center, 2019; Reina *et al.* , 2018). The murder rate rose to one of the highest in the world. The Maduro administration showed little capacity to address the crisis, and used violence and intimidation to weaken the growing political opposition. Prospects for political or economic change were slim, especially after the overwhelming defeat of the opposition party in regional elections in 2017.

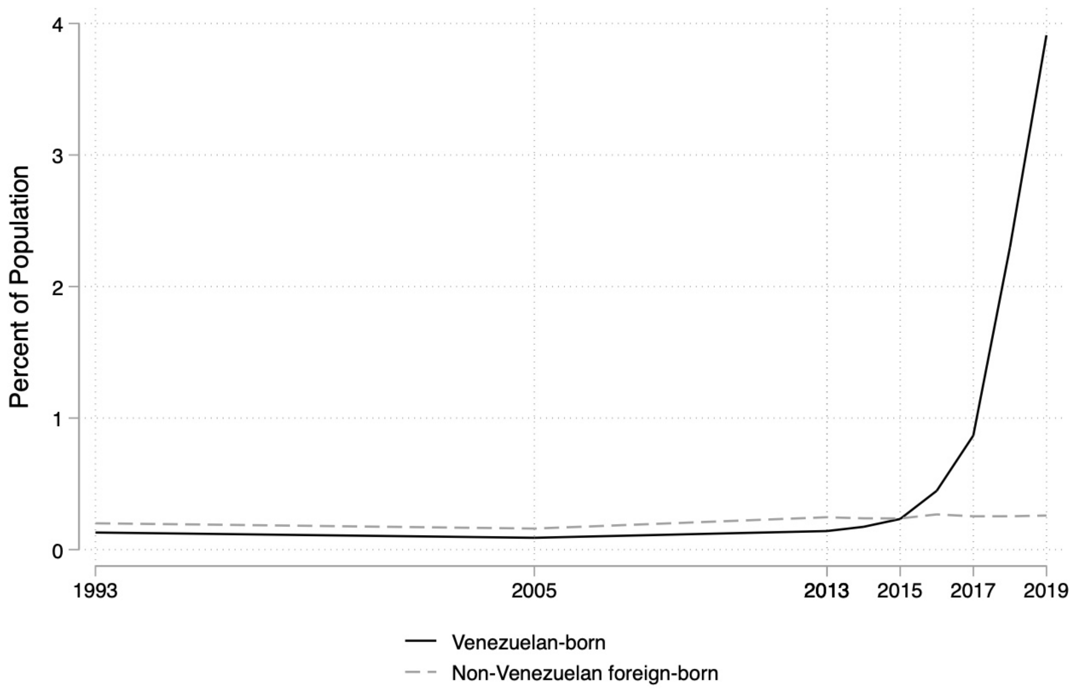
Between 2015 and 2019, 4.8 million Venezuelans fled the country, making Venezuela the second-largest country of origin for internationally displaced people after Syria. Colombia, the neighbor closest to the population centers of Venezuela, received an estimated 1.8 million of these migrants, more than any other country, and representing almost 4% of the Colombian population (Migration Policy Institute, 2020). This is the first time Colombia has received a large migration wave from another country: according to the census, .13% of the population was Venezuelan born and .2% was born in a different foreign country. Figure 1 shows that these rates remained relatively constant until the onset of the Venezuelan migration in 2015. The arrival rate increases in 2016 and again in 2017, with the majority of migrants arriving between 2018-2019. There were relatively few requirements for migrants to cross the border, and those without necessary documents could easily walk around checkpoints. This migration was therefore sudden, drastic, unprecedented and to a large extent unanticipated.

Figure 2 shows the migrant share of the population across 79 metro areas in 2019, where a migrant is defined as someone who was living in Venezuela 5 years ago. There is extensive variation in these migrant shares across Colombia. They tend to be largest closer to the Venezuelan border, in many cases exceeding 10% of the metro area population. In Cúcuta and Riohacha, two cities close to the primary entry points along the Venezuelan border, the migrant shares are around 16% and 11% respectively. In Bogotá, Medellín and Cali, the three largest cities in Colombia, the shares range between 4-5%. For other cities they remains below 1%.

3 Data and Descriptive Statistics

Data for this project comes primarily from the Colombian National Integrated Household Survey (GEIH), which is a nationally-representative survey collected by the National Department of Statistics (DANE) and is the official source for labor market indicators in

Figure 1: Foreign-Born Population in Colombia



Source: GEIH (2013-2019), Population Census (1993, 2005)

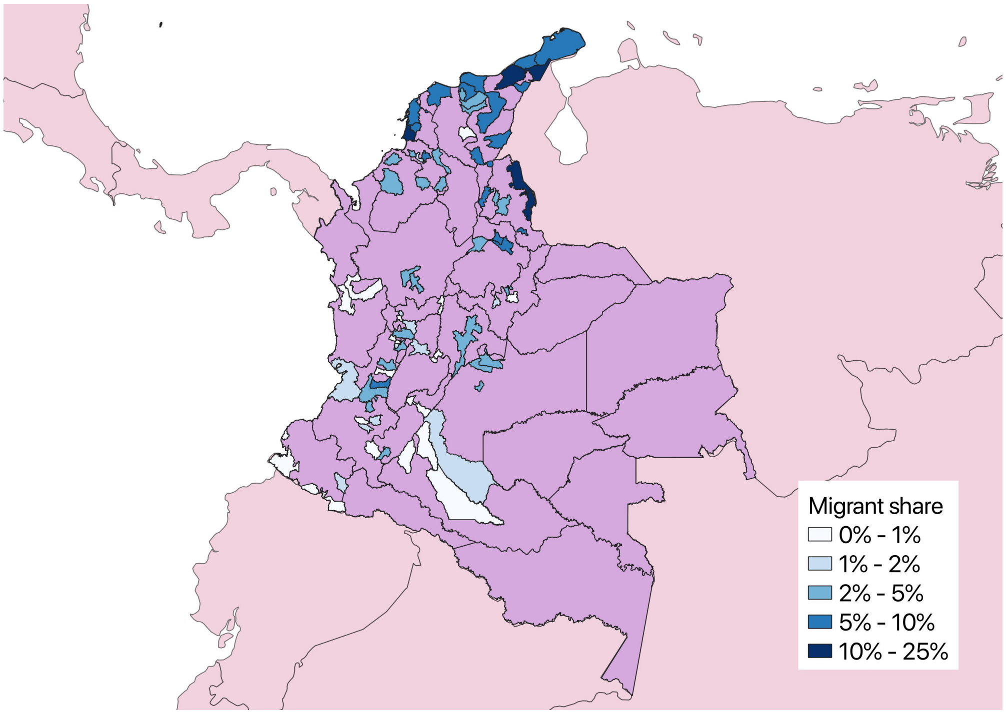
Colombia. This survey includes a migration module that records where a person was living 1 and 5 years ago, and can therefore be used to measure both migration rates and labor market outcomes for natives and migrants.[7] [8]

Colombia is divided into a capital district and 32 departments, which are further divided into 1,122 municipalities. I group municipalities into metropolitan areas according to commuting patterns. The goal is to generate a geographic unit that represents a contiguous labor market in which workers compete. I follow Duranton (2015) and use a recursive algorithm based on a 10% commuting threshold: a municipality is grouped with another municipality if over 10% of its residents commute to work in that municipality. They are then treated as a single unit in the next round of the algorithm, and this is repeated until

[7]While the GEIH is not intended to be representative of Venezuelan migrants in Colombia, it is increasingly being used to track the Venezuelan population across Colombia over time (Graham *et al.* , 2020; Tribín-Uribe *et al.* , 2020). The sample size is large enough to include a substantial number of migrants (over 21,500 migrants in metro areas in 2019) and in 2018 is closely correlated with estimates from the complete 2018 census ($\rho = .76$ across metro areas in my sample).

[8]I define migrants as anyone who was living in Venezuela 5 years ago. This includes Colombian-born return migrants, who make up around 20% of all migrants from Venezuela during this period. These migrants also increase the local labor supply and are closely correlated with the locations of Venezuelan born-migrants, and so excluding them could cause an overestimation of the labor market effects of immigration. Ideally, one could study the effects of each group separately, but these effects are empirically difficult to untangle due to the close correlation in the location of foreign-born and return migrants in Colombia.

Figure 2: Venezuelan Migrants in Colombia (2019)



Source: GEIH (2019)

no more municipalities meet this threshold. This is done using data from the 2005 census, which is the latest year before the start of the migration period in which this commuting data is available. The algorithm results in 184 metro areas with at least 30,000 residents in 2005, and these units are relatively independent of the choice of commuting threshold.[9]

The GEIH does not survey people in all of these metro areas and in others has a very small sample. I therefore restrict my analysis to the 79 metro area that contain at least 300 GEIH observations per year over the study period, representing around 80% of the Colombian population and 90% of the Venezuelan migrant population. The GEIH is designed to be representative of a smaller group of 23 metro areas that are loosely overlapping with my largest 23 metro areas, but not for areas finer than this. One might therefore be concerned about measurement error in the smaller metro areas, and I address this in two ways. First, I test the robustness of results to increasing the 300 annual observation threshold to 1,000, similar to restricting analysis to the 23 metro areas for which the survey is officially

[9]While in some cases they are identical to the officially defined metropolitan census areas (for example, in the case of Medellín), in others they are distinct (for example, the constructed areas of Bogotá, Calí and Baranquilla all include substantially more municipalities than the official census areas). The official census metro areas are politically determined in consideration of factors such as allocation of city resources and jurisdiction of city government activity, and thus are not always the most appropriate unit of economic analysis. See Duranton (2015) for a detailed discussion.

Table 1: Summary Statistics - Education and Occupation Groups

	2014 Entire Population				2019 Migrant Population			
	Total	< Secondary	Secondary	Post-Secondary	Total	< Secondary	Secondary	Post-Secondary
	265,727	111,093 (41.8%)	77,230 (29.1%)	77,404 (29.1%)	13,971	4,904 (35.1%)	5,355 (38.3%)	3,713 (26.6%)
Percent in occupation with:								
Mode < Secondary	56.0%	81.0%	57.2%	18.9%	66.4%	77.4%	67.3%	50.8%
Mode Secondary	17.4%	14.5%	26.2%	12.7%	23.4%	20.0%	26.5%	23.5%
Mode Post-Secondary	26.6%	4.5%	16.6%	68.4%	10.1%	2.7%	6.2%	25.7%
	Total	< Secondary	Secondary	Post-Secondary				
Change in labor supply under:								
Downgrading (assigned by occupation)	6.25%	8.00%	6.10%	3.90%				
No Downgrading (assigned by education)	6.25%	5.25%	8.25%	5.70%				

Data from GEIH using national population weights. Restricted to working urban residents age 15-64 with completed education. Migrant defined as anyone in 2019 who was living in Venezuela 5 years ago. Occupations grouped according to the mode education of native workers in that occupation pre-2014. Migration-induced changes in labor supply under downgrading calculated by splitting each migrant observation into education groups according to the pre-2014 native distribution in their occupation.

representative. Second, in my analysis I use an instrumental variable based on migration shares from the complete 2005 National Census, which will isolate measurement error in the yearly metro-level migration rates.

Migrants have a similar gender composition to natives and are moderately younger, with an average age of 31.5 compared to 36.5 for natives (see Table A1). They have greater labor force participation rates but also higher unemployment rates such that the total working share is only slightly higher for migrants. Most notably, migrants have substantially lower wages and almost 90% are concentrated in the informal sector. As I show in Table 1, this is despite having similar rates of education.

Table 1 shows the education and occupation distribution for non-migrants in 2014 (defined as the entire population) and Venezuelan migrants in 2019 (defined as arriving from Venezuela in the past 5 years), among metro area workers. The first row shows the total share of each population with less than secondary, completed secondary and any post-secondary education. Migrants are slightly less likely to have post-secondary and more likely to have completed secondary. However, over 25% of migrants have post-secondary, demonstrating that this is not a low-skill migration wave. The next three rows show the concentration of workers in each education group across occupation skill groups, which are defined according to the mode education of natives in that occupation between 2010-2014. Among all education groups, migrants are less likely to be observed in post-secondary occupations. Among those with post-secondary education, 25.7% of migrants and 68.4% of natives are in occupations where the mode native has post-secondary, demonstrating the striking level of downgrading among well-educated migrants. Similar downgrading, though less severe, is observed for migrants with a secondary degree.

To compare downgrading in Colombia with downgrading in other migration waves around the world, I calculate the share of migrants who have a level of education greater than the mode in their occupation.[10] This results in an estimate of 45.5% of migrants, relative to 29.2% of natives, who are over-educated. These estimates are large relative to immigrant over-education numbers calculated using the same method in other countries, which range from 16% in Denmark to 24% in the UK and 28% in the US (Piracha *et al.* , 2012). However, they are not unheard of - estimates are as high as 39% in Spain, 41% among established male migrants in New Zealand and 52% among those who arrived in the past five years (Sanroma *et al.* , 2008; Poot & Stillman, 2010). Therefore, the extent of migrant over-education in Colombia is large but within the range of estimates observed in other countries.

The bottom of Table 1 presents the group-specific increase in labor supply from immigration as a share of the 2014 population. To understand the extent to which migrant downgrading impacts labor competition by skill group, I first assign migrants to education groups based on the native distribution between 2010-2014 in their observed occupation. For example, 14% of migrants are in the "restaurant workers" occupation, which includes waiters, bartenders and chefs. Among natives in this occupation, 61% did not complete secondary, 27% completed secondary and 12% have post-secondary. Migrants in this occupation are thus assigned to education groups with these proportions. This is what I refer to as the change in labor supply "under downgrading". As I discuss in Section 5, this is the assignment mechanism that I use when I estimate the elasticity of substitution between migrants and natives within education group, thus capturing the effect of competition between natives and migrants who compete within their education group. I then show how labor supply would change "under no downgrading", where migrants are assigned to education groups according to their actual level of education.

The bottom of Table 1 shows that the total growth in the population due to immigration over this period is 6.25% when restricted to the sample of urban workers age 15-64. Under downgrading, natives in each education group experience an 8%, 6%, and 4% increase in labor supply respectively, demonstrating the disproportionate impact of immigration on supply of less educated workers. If migrants were instead competing with workers of their own education, the supply shifts heavily towards completed secondary and post-secondary workers - both experience around a 2pp increase in labor supply. One of the key questions I explore is whether such a reallocation would simply shift the economic effects of competition to better-educated workers, or whether there are aggregate gains in wages under the components of the imperfect substitutability framework.

[10]The mode is preferred over the mean for this exercise because it has been shown to generate over-education rates more comparable with those based on classifications of occupation skill requirements and self-reported worker over-skilling when such data is available (Leuven & Oosterbeek, 2011).

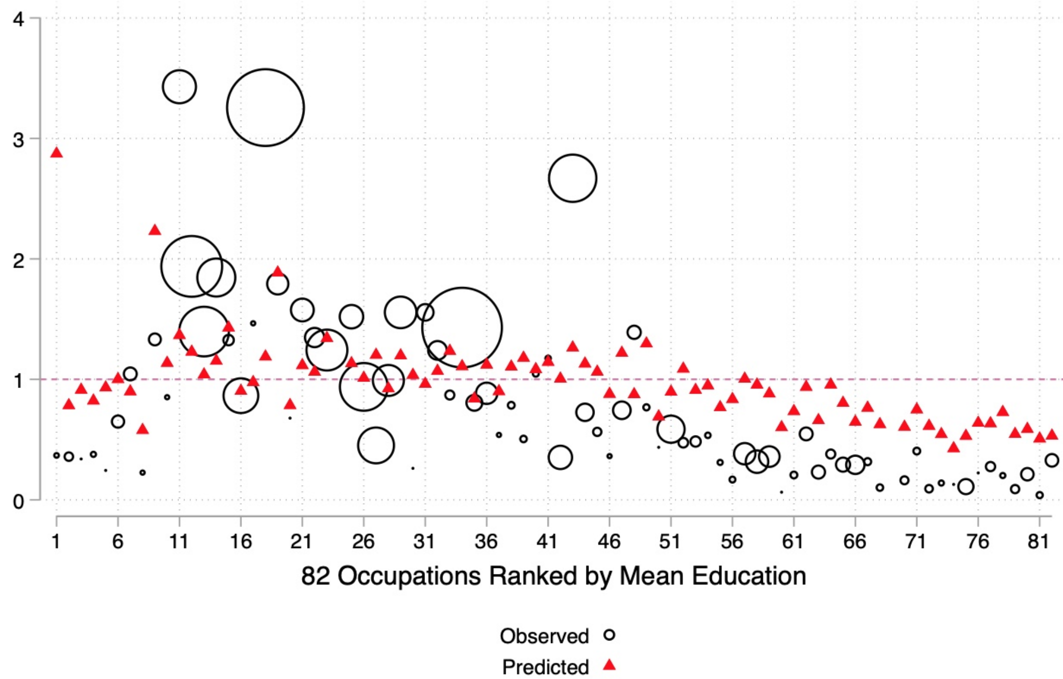
When I define “no downgrading” as migrants competing with workers of similar education, I implicitly assume that schooling quality in Colombia and Venezuela are comparable. To assess this, I first look at the 2009 Programme for International Student Assessment (PISA), which evaluates 15-year-olds' knowledge and skills in reading, mathematical and scientific literacy and included both Colombia and the Venezuelan state of Miranda (which contains around 10% of the Venezuelan population). Out of 74 countries, Colombia and Venezuela ranked, respectively, 52nd and 56th in reading literacy, 64th and 60th in math literacy, and 60th and 53rd in science literacy (Walker, 2011). At least in terms of primary and lower secondary, the education systems appear to be comparable in 2009, six years before the onset of the migration. To build further confidence in this assumption, I also look at the wage returns to schooling for Venezuelans and Colombians in various settings. First, in panel 1 of Table A2, I show that Venezuelan migrants who have been in Colombia for at least five years display similar returns to secondary and post-secondary education as natives controlling for age and gender. However, it is possible that some of these migrants received schooling in Colombia, which I do not observe in the data. In panels 2 and 3, I estimate the returns to schooling for Colombian and Venezuelan migrants in the 2015 US census and the 2010 Panama census, which are two of the most common migrant destinations for each country.[11] Returns to schooling are comparable and if anything larger for Venezuelans - wage returns to secondary schooling are 77% and 61% larger for Venezuelans in the US and Panama respectively, and wage returns to post-secondary in the US are 117% for Venezuelans compared to 103% for Colombians. If it is the case that Venezuelan schooling is more valuable than Colombian schooling within Colombia as well, then any exercise that measures downgrading according to Venezuelan migrants' observed schooling will underestimate the true level of occupational downgrading.

For an alternative characterization of migrant occupational downgrading, Figure 3 displays the occupation concentration of migrants and how it would differ if migrants worked the same jobs as natives with similar observable characteristics. On the horizontal axis, the 82 ISCO-68 occupations in the GEIH are ranked by mean native education. The black markers plot the concentration of migrants in each occupation divided by the native concentration in that occupation. The size of each dot is proportional to the total number of migrants. This ratio is mostly above one for lower-skill occupations and below one for higher-skill occupations, demonstrating that migrants tend to be more concentrated relative to natives in low-skill occupations.[12] To calculate predicted migrant concentrations in

[11] Census data was accessed via IPUMS International (IPUMS, 2019).

[12] The four most common migrant occupations (corresponding to the four largest circles in the graph) are street vendors and merchants, restaurant workers, construction workers and domestic servants, all of which

Figure 3: Relative Concentration of Migrants Across Ranked Occupations



Sample: GEIH 2019, urban workers age 15-64. Black markers measure the migrant-native ratio of the share in each occupation (size is proportional to total Venezuelan population). Red triangles plot the predicted relative shares according to age group, education, gender and metro area.

the absence of downgrading, I assign each migrant the composition of natives within their 5-year age bin, education group, gender and metro area. These relative concentrations are clustered much closer to the line at $y = 1$, though they remain below this line at high-skill occupations due to migrants' slightly lower post-secondary attainment. Overall, the results imply that, if migrants worked the same jobs as natives with similar characteristics, they would be substantially less concentrated in low-skill occupations.

4 Motivating Results from Non-Structural Analysis

In a companion paper to this paper, Lebow (2021), I study the effects of migration on native economic outcomes using variation in the total migrant share across metropolitan areas and over time. The econometric specification is

$$Y_{ct} = \beta M_{ct} + \gamma_c + \delta_t + \epsilon_{ct} \tag{1}$$

are occupations below the median skill level and in which migrants are over-represented.

The endogenous variable, M_{ct} is the migrant population as a share of the 2014 population in metro area c and year t. I also include metro and year fixed effects and weight observations by the population within each metro-year cell. This approach makes no assumptions about the substitutability between migrants and natives or economic mechanisms, and going forward I refer to these as "non-structural results". This approach is informative about the total effects of migration for natives overall and by demographic group (Dustmann *et al.* , 2016). It therefore provides a useful benchmark for the structural analysis, which should at least partially replicate the non-structural results when used to predict total effects of migration.[13]

To account for endogenous migrant sorting, this analysis uses an instrumental variable based on historical migration rates similar to the one used in this paper (discussed extensively in section 6). These results are robust to accounting for native internal migration, various choices of instrument, and alternative specifications that are not sensitive to biases introduced by two-way fixed effects with dynamic treatment effects (Callaway & Sant'Anna, 2020; Goodman-Bacon, 2021; Borusyak & Jaravel, 2021; De Chaisemartin & d'Haultfoeuille, 2020). They are also broadly consistent with results from other papers that used similar methods to study the same migration wave (Delgado-Prieto, 2021; Bonilla-Mejía *et al.* , 2020; Santamaria, 2020; Rozo & Vargas, 2021; Caruso *et al.* , 2019). For a detailed discussion of this analysis, alternative specifications, robustness checks, and a comparison of results with the literature, refer to Lebow (2021). I now review a few key results from this non-structural analysis that motivate the structural model presented in the following section:

Motivating Result 1: Venezuelan migration to Colombia had little effect on native employment and negative effects on native wages.

I find that a 1 percentage point (pp) increase in the metropolitan area migrant share causes a 1.05% decrease in native hourly wages (95% CI ranging from -1.48 to -.62), with similar effects by gender and age group. At the same time, there are little to no effects on unemployment and only small decreases in labor force participation concentrated among workers under age 25, both overall and within education groups. This motivates a model in which employment is held fixed while wages react to movements along the demand curve.

Motivating Result 2: Less educated natives experienced a disproportionate decreases in wages.

A 1pp increase in the migrant share causes a -1.42%, -.86% and -.75% drop in wages respectively for workers without completed secondary, with completed secondary and with any

[13] My primary analysis is "structural" in that I specify the structure of production and the parameters of interest are economic primitives. However, some may prefer to call it "semi-structural" because I only model one side of the labor market.

post-secondary. While the magnitude of the total wage effect varies across specifications, this pattern of wage effects by education group is extremely robust across specifications. Thus, we would like a structural model that is consistent with larger negative wage effects for less educated workers.

Motivating Result 3: There was little native occupational upgrading or movements into the formal sector. In some cases, we observe natives upgrading to higher-skill occupations in response to immigration (Peri & Sparber, 2009; Peri *et al.* , 2020). This is less likely when migrants and natives speak the same language, but could still occur via mechanisms other than communication specialization. Though the data does not allow me to study job transitions, I can study changes in total employment by occupation skill groups (categorized according to the mean education of natives in that occupation). I find that Venezuelan migration caused small movements from from lower- to middle-skill occupations among workers with completed secondary (.15pp from a 1pp increase in the migrant share), and small movements from upper- to middle-skill occupations among workers with post-secondary (.12pp from a 1pp increase in the migrant share). These are small effects: with around a 6pp increase in the migrant share over this period, these magnitudes predict less than a 1pp shift for each of these groups. I also find little evidence for upgrading into the formal sector, and if anything, immigration pushed natives *out* of the formal sector, consistent with results from (Delgado-Prieto, 2021). Thus, while I would like my structural framework to allow for natives to switch between occupation groups, I would not expect occupational upgrading to lead to a high elasticity of substitution across education groups.

5 Nested CES Framework

5.1 Theoretical Framework

To model the economic response to the migration, I consider a model of CES labor demand disaggregated by education group and nativity status. Even in a basic disaggregation (as I use here) with only six cells, a fully flexible production function with minimal structure would require estimating 30 cross-group elasticities. To reduce the number of parameters estimated, I use a nested CES structure, following an extensive literature in labor economics that has used this framework to study relative changes in labor supply and inequality (Katz & Murphy, 1992; Card & Lemieux, 2001). Following the examples of Ottaviano & Peri (2012) and Manacorda *et al.* (2012), I extend this framework to allow for imperfect substitutability between migrants and natives, which was shown to be very relevant for estimates of the total wage effects of immigration.

Firms produce output according to a constant-returns-to-scale Cobb-Douglas production function, which is widely applied in the macro growth literature:

$$Y = AK^{1-\zeta}L^{\zeta} \tag{2}$$

where K is capital, ζ is the labor share of income and A is a skill-neutral technology parameter. Labor is a CES aggregate of three education types: less than high school ($e = 1$), high school graduates ($e = 2$), and any post-high school ($e = 3$):

$$L = \left(\alpha_3 L_3^{\frac{\sigma_3-1}{\sigma_3}} + \alpha_{-3} L_{-3}^{\frac{\sigma_3-1}{\sigma_3}}\right)^{\frac{\sigma_3}{\sigma_3-1}} \tag{3}$$

$$L_{-3} = \left(\alpha_2 L_2^{\frac{\sigma_2-1}{\sigma_2}} + \alpha_1 L_1^{\frac{\sigma_2-1}{\sigma_2}}\right)^{\frac{\sigma_2}{\sigma_2-1}} \tag{4}$$

where σ_3 measures the elasticity of substitution between workers with and without a college degree, and σ_2 measures the elasticity of substitution between workers with and without a high school degree. α_e is the relative productivity of group e, standardized so that $\alpha_1 + \alpha_2 = \alpha_3 + \alpha_{-3} = 1$.

I then assume that each education-specific cell is a CES combination of native (n) and migrant (m) labor:

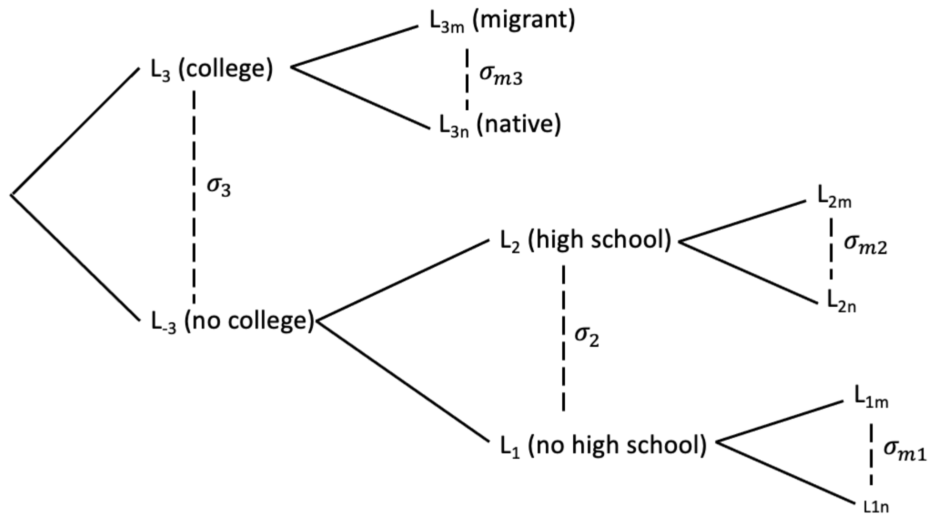
$$L_e = \left(\alpha_{em} L_{em}^{\frac{\sigma_{em}-1}{\sigma_{em}}} + \alpha_{en} L_{en}^{\frac{\sigma_{em}-1}{\sigma_{em}}}\right)^{\frac{\sigma_{em}}{\sigma_{em}-1}} \tag{5}$$

where the elasticity of substitution between natives and migrants σ_{em} is allowed to vary by education group, as are the productivity parameters that are again standardized such that $\alpha_{em} + \alpha_{en} = 1$. This is to allow for the possibility that migrants and natives are more substitutable in low-skill jobs. The nesting structure is presented in Figure 4. [14]

I assume perfect competition, such that wages are set equal to the marginal product of labor, and this generates the following wage equation by nativity status $j \in \{m, n\}$ and

[14] In Section A3, I estimate a generalized production function that allows the substitutability parameters to vary across all education group pairs. While results are imprecise, I find that native substitutability between education groups 1-3 and 2-3 are comparable, while that of groups 1-2 is substantially lower, motivating the chosen nesting structure.

Figure 4: CES Nesting Structure



education groups $e \in \{1, 2\}$:

$$\begin{aligned} \ln W_{ej} = \ln\left(AK^{1-\zeta}\zeta\right) + \ln\alpha_{-3} + \ln\alpha_e + \ln\alpha_{ej} + \left(\zeta - 1 + \frac{1}{\sigma_3}\right)\ln L \\ + \left(\frac{1}{\sigma_2} - \frac{1}{\sigma_3}\right)\ln L_{-3} + \left(\frac{1}{\sigma_{em}} - \frac{1}{\sigma_2}\right)\ln L_e - \frac{1}{\sigma_{em}}\ln L_{ej} \end{aligned} \tag{6}$$

and for $e = 3$:

$$\ln W_{3j} = \ln\left(AK^{1-\zeta}\zeta\right) + \ln\alpha_3 + \ln\alpha_{3j} + \left(\zeta - 1 + \frac{1}{\sigma_3}\right)\ln L + \left(\frac{1}{\sigma_{3m}} - \frac{1}{\sigma_3}\right)\ln L_3 - \frac{1}{\sigma_{3m}}\ln L_{3j} \tag{7}$$

According to these equations, wages are a function of group-specific productivity and the labor supplies in each group. The final term in each equation captures the fact that wages are decreasing in the labor supply in ones own skill-nativity group (L_{ej}), and less so as the substitutability between natives and migrants in skill group e increase. In this sense, greater substitutability between migrants and natives "spreads out" the wage impact across both groups. Likewise, greater substitutability across skill groups "spreads out" the wage impact across skill groups, and wages are decreasing in skill-specific labor supply (L_e) as long as the within-skill group migrant-native substitutability is greater than the across-skill group substitutability with the next group in the nest (in the case of $e \in \{1, 2\}$, when $\sigma_{em} > \sigma_2$). Finally, the term $\frac{1}{\sigma_3}\ln L$ reflects the fact that, in a model with imperfect substitutability, all workers benefit from increases in the total labor supply. This benefit is mitigated by the term $\zeta - 1$, which captures the diminishing returns in the labor aggregate when capital is held fixed. This is discussed further in Section 6.

We can use equations (6) and (7) to express the wage change experienced by natives as a function of the change in migrant supply while holding the native labor supply fixed. This results in, for $e \in \{1, 2\}$:

$$d\text{ln}W_{en} = \left(\zeta - 1 + \frac{1}{\sigma_3}\right) d\text{ln}L + \left(\frac{1}{\sigma_2} - \frac{1}{\sigma_3}\right) d\text{ln}L_{-3} + \left(\frac{1}{\sigma_{em}} - \frac{1}{\sigma_2}\right) d\text{ln}L_e \tag{8}$$

and for $e = 3$:

$$d\text{ln}W_{3n} = \left(\zeta - 1 + \frac{1}{\sigma_3}\right) d\text{ln}L + \left(\frac{1}{\sigma_{3m}} - \frac{1}{\sigma_3}\right) d\text{ln}L_3 \tag{9}$$

A major take-away from this framework is that, because of imperfect substitutability, native wages depend on the labor supply in all skill groups, not just one's own. This implies that if migrants are moved to compete with more educated natives, this does more than mechanically shift competition from low- to high-skill groups; it also adjusts the benefits that workers receive from complementarities across groups. Whether the wage effect is negative or positive for a skill groups depends on the tradeoff between these across-group and within-group effects. A "skill-cell" approach, in which own-group wages are regressed on own group labor supply and group fixed effects, would exclude these across-group benefits from the group-specific effect (Ottaviano & Peri, 2012; Dustmann *et al.* , 2016).

Finally, note that the change in labor supply within an education group can be expressed as a function of the total migrant share of the population M[15]:

$$d\text{ln}L_e = \frac{s_e^M}{s_e^N} M \tag{10}$$

where s_e^M is the share of incoming migrants in education group e and s_e^N is the share of the native population in group e. Furthermore, in the nested CES framework, the change in labor supply for an aggregate group can be expressed as sum of the supply change in each subgroup weighted by each group's wage bill share. Thus, the wage equations (8) and (9) can be expressed as a linear function of M, with a coefficient that is a function of the share of the migrant population going to each education group, the wage bill of each group, and the remaining parameters of the model. In this sense, the non-structural estimates of the total wage effects estimated with equation (**??**) make no assumptions about the degrees of

[15] A similar modelling approach is taken in Edo (2020) and Dustmann *et al.* (2017).

substitutability between migrants and natives or across skill groups.

5.2 Estimation

5.2.1 Estimating Substitutability Parameters

Taking the ratio between $\ln W_{em}$ and $\ln W_{en}$ yields the following equation, which can be used to estimate σ_{em} as well as the relative productivity terms:

$$\ln\left(\frac{W_{em}}{W_{en}}\right) = \ln\left(\frac{\alpha_{em}}{\alpha_{en}}\right) - \frac{1}{\sigma_{em}}\ln\left(\frac{L_{em}}{L_{en}}\right) \tag{11}$$

Intuitively, as migrants and natives become increasingly substitutable, wages equalize across groups as the relative labor supplies change, and the sensitivity of the wage ratio to the labor ratio approaches zero.

This is estimated via OLS ans 2SLS with the following regression:

$$\ln\left(\frac{W_{m}}{W_{n}}\right)_{ect} = \lambda_{ec} - \frac{1}{\sigma_{em}}\ln\left(\frac{L_{m}}{L_{n}}\right)_{ect} + \epsilon_{ect} \tag{12}$$

where e, c, and t represent education groups, metro areas and years respectively. Average wages and number of workers within each of the 1,422 education-year-metro groups (3 education groups X 79 metro area X 6 years covering 2014-2019) are calculated among migrants and natives age 15-64, where a migrant is defined as anyone who was living in Venezuela 5 years ago.[16] Each regression is run separately by education group, so the migrant-native substitutability parameter is allowed to vary by education group. Cells are weighted by sample size to down-weight small-sample cells. Finally, I hold the native population fixed to its 2014 value to mitigate any effects driven by changes in this variable. In order to account for endogenous migrant locational sorting, I use an instrument based on the historical settlement of Venezuelans in Colombia that is discussed extensively in Section 5.2.3.

[16] Hourly wages are calculated for the primary job as past-month income over 4.2 times hours worked in a typical week in the past month, including all overtime, benefits and other transfers. This is calculated for all workers regardless of self-employment or formality status. There is a concern about measurement error of self-employment income when there is cyclical variation in earnings throughout the year. However, because the analysis is conducted at the annual level, rather than the monthly level, monthly fluctuations in self-employment income will average out across the year. Wages are residualized from a regression on gender, age, age-squared, to absorb effects of the demographic composition of workers within education groups. Wages are also winsorized at the top and bottom .5% within each year. I choose not to use the GEIH sampling weights since the survey is not designed to be representative at this level, but results are not sensitive to this decision.

Geographic variation is essential in my context, considering the large geographic spread and short time period of the Venezuelan migration to Colombia. There has been debate over the value of cross-city comparisons with this framework, considering that migrants and natives can endogenously sort across locations in response to changing labor market outcomes (Borjas *et al.* , 1997). Nonetheless, after carefully accounting for geographic sorting, cross-city comparisons have proved useful for identifying labor market impacts of immigration and produce estimates that are generally consistent with time series-based estimates (Card, 2009; Ottaviano & Peri, 2012). As mentioned, because of the unique data that matches labor market outcomes to migration histories in Colombia, I confirm in my robustness checks that endogenous native internal migration does not bias my results.

The education-metro fixed effects absorb the relative wage factors that are constant within metro area and can be used to construct α_{em} and α_{en}.[17] Thus, identification is driven by changes in the wage and labor ratios within metro area over time, and anything that effects the wage ratios across metro areas (for example, wage discrimination against Venezuelans) will not impact estimates of $\widehat{\sigma_{em}}$ as long as it does not vary over time. Importantly, because equation (12) uses ratios and not levels, any changes in wages that equally affect migrants and natives within the same education-metro-year cell (for example, aggregate productivity shocks) will also not impact estimates of $\widehat{\sigma_{em}}$.[18]

In the final steps of the estimation, I estimate σ_2 and σ_3. Armed with the estimates of $\widehat{\sigma_{em}}$, $\widehat{\alpha_{em}}$ and $\widehat{\alpha_{en}}$, I use equation (5) to calculate the education-specific "effective" labor supplies and estimate the next step of the model, which is derived using the population-weighted mean of migrant and native waves within each group:

$$\ln\left(\frac{W_2}{W_1}\right)_{ct} = \lambda_{2,c} + \lambda_{2,t} - \frac{1}{\sigma_2}\ln\left(\frac{L_2}{L_1}\right)_{ct} + \epsilon_{2,ct} \tag{13}$$

Likewise, the estimates from regressions (12) and (13) can now be used to calculate the terms needed to estimate σ_3:

$$\ln\left(\frac{W_3}{W_2}\right)_{ct} = \lambda_{3,c} + \lambda_{3,t} - \frac{1}{\sigma_3}\ln\left(\frac{L_3}{L_2}\right)_{ct} + \epsilon_{3,ct} \tag{14}$$

[17] $\widehat{\alpha_{em}} = \frac{e^{\widehat{\lambda_{ec}}}}{1+e^{\widehat{\lambda_{ec}}}}$ and $\widehat{\alpha_{en}} = \frac{1}{1+e^{\widehat{\lambda_{ec}}}}$

[18] Ideally, this regression could also include year fixed effects to absorb trends in demand that differentially impact migrants and natives. In practice, adding year fixed effects substantially reduces power. In Table A5, I show that, when estimating (12) at the aggregate level (grouping education groups together), the inclusion of year fixed effects has a minimal effect on the point estimates but increases the standard errors by a factor of three. In the 2SLS model, the first stage Kleibergen-Paap Wald statistic becomes weak. Thus, while it is not possible to include these controls in the full model, I can at least be confident that they do not drastically change the OLS results (which are very similar to the 2SLS results) in the aggregate sample.

The same intuition applies for these equations, in that variation in the ratio of wages and labor supplies is used to identify the substitutability parameters, holding constant metro and year fixed effects. In this case, I can compare my estimates of σ_2 and σ_3 to other estimates from the literature in Colombia and Latin American.

5.2.2 Incorporating Occupational Downgrading

As it stands, this framework does not account for the occupational downgrading of migrants, in that it places migrants into groups according to their observed education (Dustmann *et al.* , 2017). In order to estimate an elasticity of substitution between migrants and natives who are competing in the same jobs, I develop a method to assign migrants to education groups according to their observed occupation. Within each occupation, I calculate the share of natives in each education group between 2010-2014. Next, I randomly assign migrants into education groups with probability equal to these shares for their observed occupation. Intuitively, if a high-skill migrant is working in a job that is mostly worked by natives in education group 1, they will be assigned to this group with high probability. I then calculate cell-specific average wages and labor supplies and I estimate the entire model. I repeat this for 100 random assignments and take the average parameter values. Standard errors are bootstrapped with 100 samples and clustered at the metro level.[19]

An assumption of this framework is that educated and uneducated migrants have the same productivity within a given occupation. This would be violated if, for example, migrant restaurant workers with post-secondary were more productive than those without post-secondary. This is an assumption that I can test empirically by estimating the education wage premium for migrants and natives across occupation groups. Table A3 shows that, holding constant gender, age and city, migrants face a substantially smaller education premium than natives in occupations where the mode native worker has no secondary: 6.8% for having completed secondary and 14.5% for any post-secondary, relative to 15.8% and 35.4% respectively for natives. The education wage premia are similarly smaller in high skill occupation groups, generally one-half to two-thirds of the premia experienced by natives. Though small, these premia add a bias to my estimates of σ_{em}. I therefore adjust the migrant labor supplies using the estimates from Column 2 of Table A3. For example, if a migrant with post-secondary is assigned to the "less than secondary" group, then I multiply their labor supply by 1.146, reflecting the wage premium for post-secondary migrants in less-than-secondary occupations. In practice, this has little impact on the results.

[19]I drop the very small share of resample where a first-stage Kleibergen-Paap Wald statistic is weak, to ensure that this does not bias downstream steps in the estimation.

5.2.3 Instrumenting for Migration

Equation (12) is biased if migrant locations $L_{m,ect}$ are correlated with the error term. Specifically, the concern regards migrant sorting into locations according to the migrant-native wage ratio, as opposed to the traditional concern of sorting according to native wages. To deal with this potential endogeneity, I construct the following instrument for $\ln\left(\frac{L_m}{L_n}\right)_{ect}$:

$$Z_{ect} = \ln\left(\frac{L_{Ven}}{L_{Col}}\right)_{e,c,2005} * \ln\left(\frac{L_m}{L_n}\right)_{e,Nat^{-c},t} \tag{15}$$

The first term is the log ratio of Venezuelan-born to Colombian-born in the complete 2005 census in education group e and metro area c, which is a strong predictor of the present migrant-native log ratio (see Figure A1).[20] The second term is the national log ratio of migrants to natives in education group e and year t, excluding migration into metro area c to reduce the impact of large inflows into cities that are correlated with changes in economic outcomes in those cities (Card, 2001; Tabellini, 2020). This instrument is the standard "shift-share" instrument based on historical migrant populations, adjusted to match the functional form of the endogenous covariate.[21] I use education-specific migrant shares to generate independent variation in predicted labor ratios by education group, exploiting the fact that migrants are more likely to have networks with migrants of similar levels of education.

The national change in migration in year t is assumed to be exogenous, driven primarily by push factors in Venezuela that are uncorrelated with any economic or other changes occurring in Colombia, reflecting the fact that Venezuelans were primarily migrating to escape poverty and violence in Venezuela. Of greater concern is the potential endogeneity of the 2005 migrant-native ratio, which could be correlated with *changes* in economic outcomes between 2014-2019. This concern is mitigated by the fact that historical migrant shares were determined two decades before the onset of the Venezuelan exodus, well before the election of Hugo Chávez. While I use the 2005 census to construct the instrument, results using the 1993 census are very similar. There was little subsequent immigration over this period and the correlation between the migrant share in these years is .89 in the metro areas in my sample.[22] The instrument would also be invalid if immigration before 2005 stimulated

[20] 2005 Colombian census data was provided by DANE.

[21] By matching the functional form of the endogenous variable, the log transformation gives me substantially more power in the first stage regression, but it is not necessary to replicate my primary results. I also normalize the 2005 log-share to be positive by adding it to its minimum value (this ensures that the migrant share is increasing in the instrument), and I add one to the 2005 Venezuelan population to make the log-ratio non-missing in places with no Venezuelans in 2005.

[22] I prefer to use the 2005 census to construct the instrument because I have access to the complete census

dynamic economic responses or subsequent immigration correlated with current economic trends (Jaeger *et al.* , 2018). However, there was almost no immigration into Colombia between 2005-2015, and the 2005 Venezuelan share was miniscule relative to the current immigration: in no metropolitan area in 2005 was this share greater than 1%.

It remains possible that migrants historically sorted into locations that had different economic trends unrelated to the mass migration. While this cannot be formally tested, I conduct various checks and robustness tests recommended by the literature (Goldsmith-Pinkham *et al.* , 2020). First, I check for a correlation between the 2005 shares and pre-period economic outcomes. Table A4 shows a regression of the 2005 Venezuelan-Colombian log-ratio by education on a set of metro area characteristics measured in 2014, before the onset of the migration. The coefficients are mostly small and insignificant. A 1% increase in 2014 wages is associated with a .012 SD decrease in the 2005 log-ratio among working with post-secondary, and the magnitudes are even smaller for other education groups. Most notably, a 1% increase in the unemployment rate is associated with around a .06SD decrease in the 2005 log-ratio for workers with completed secondary or post-secondary. However, none of these coefficient are significant at the 10% level once we account for multiple hypothesis testing using sharpened false discovery rate q-values with 15 tests.[23]

It is arguably more relevant to test if historical migrant shares are correlated with pre-trends in the outcomes leading up to the migration. While there are not enough migrants before 2015 to estimate pre-trends in the migrant-native wage ratio, I estimate pre-trends in native wages with the following event-study model:

$$W_{ect} = \sum_{y=2010...2019, y\neq 2015} [\sigma_{ey} \ln \left(\frac{L_{Ven}}{L_{Col}}\right)_{e,c,2005} * (t = y)] + \gamma_{ec} + \delta_{et} + \epsilon_{ect} \tag{16}$$

where σ_{ey} measures the effect of a change in native log wages interacted with the 2005 Venezuelan-Colombian log-ratio in education group e relative to the excluded year of 2015, the year proceeding the onset of the migration. This is presented in Figure A2 with 95% confidence interval based on heteroscedasticity-robust standard errors. Within each education group, native hourly wages do not exhibit a pre-trend associated with the 2005 log ratio. In education group 2, there is a negative pre-trend before 2013, but this levels out in the 2 years before 2015. Only after 2015 is the 2005 log ratio associated with a wage decline for natives in education group 1, and less so in education groups 2 and 3, consistent with the

in this year, which helps to minimize measurement error, and to mitigate concerns regarding changing metropolitan boundaries before 2005.

[23]Sharpened false discovery rate q-values can be interpreted as the expected portion of rejections that are type-1 errors (Anderson, 2008).

effects on native wages discussed in Section 4.

6 Results

6.1 Parameter estimates

Table 2 shows the estimated elasticities of substitution between migrants and natives within each education group. The OLS and 2SLS results are almost identical, indicating little sorting of migrants according to migrant-native wage ratios. The first-stage Kleibergen-Paap Wald statistic (averaged over the 100 migrant random assignments) remains over 50 for each education group, and the coefficients from the first-stage regression are all positive and significant.[24] Below each 2SLS coefficient, I also present the negative inverse of each coefficient which corresponds to the elasticity of substitution, using the Delta method to calculate the standard errors.[25]

The estimates imply substitutability parameters of 14.8, 12.9 and 6.4 for those without secondary, completed secondary and any post-secondary respectively. All of these estimates are smaller than the migrant-native substitutability estimated in the US of around 20 (Ottaviano & Peri, 2012). An important reason is likely that those are long-run elasticities estimated across decades, while my estimates are very short-run responses to a sudden migration surge. A more comparable reference is the estimated elasticity of around 5 for recently arrived high-skill workers in the UK (Manacorda *et al.* , 2012). This closely matches my estimated substitutability of 6 for college educated natives with migrants in high skill occupations. Less educated workers are more substitutable with migrants, with elasticities closer to the long-run estimates from the US. This will inflate the distributional consequences of occupational downgrading.

Why are migrants and natives more complementary in high-skill work and substitutable in low-skill work? This may reflect the fact that low-skill occupations tend to be more routine and require less skill specialization. These occupations are also more likely to be informal, which as discussed may increase substitutability with migrants due to fewer wage rigidities. I test this directly by breaking the sample up according to formality status. Unlike education,

[24] The Kleibergen-Paap LM test tests the null hypothesis that the structural equation is underidentified. With a single endogenous regressor, this reduces to a standard first-stage F-statistic that is heteroskedasticity robust.

[25] Directly bootstrapping the inverse function is not possible because it is discontinuous at zero, which generates negative elasticities in extreme resamples.

Table 2: Migrant-Native Substitutability Parameters

	$-\frac{1}{\sigma_{1m}}$ (1)	$-\frac{1}{\sigma_{2m}}$ (2)	$-\frac{1}{\sigma_{3m}}$ (3)
OLS	-0.069	-0.077	-0.135
	(0.011)	(0.008)	(0.027)
2SLS	-0.068	-0.077	-0.155
	(0.009)	(0.009)	(0.035)
2SLS (σ)	14.75	12.92	6.44
	(1.90)	(1.50)	(1.46)
Mean Kleibergen-Paap Wald stat	59.5	70.9	92.0
Mean First-Stage Coef.	0.33	0.37	0.40
	(0.03)	(0.03)	(0.03)

Means presenteed from 100 random assignments of migrants to education groups. Metro-clustered bootstrapped standard errors with 100 resamples in parenthesis. Standard errors for the elasticities calculated using the Delta method. Kleibergen-Paap Wald statatic is clustered at the metro level. See paper for estimation details.

formality status is not fixed for native workers and endogenous sorting by natives could therefore bias these estimates. With this caveat in mind, I show in Table A6 that, across all skill groups, migrant-native substitutability is slightly higher in informal jobs, and this difference is largest for college-educated workers. The standard errors are wide such that difference across groups cannot be rejected - in particular, they become very large for low-skill formal jobs in which the migrant population is very small. Overall, this indicates that formality may explain at least part of the higher migrant-native substitutability for low-skill workers. In Section A4, I use a two-step selection procedure to confirm that these results are not substantially biased by the selection of natives into formality status, and if anything the gap between formal and informal parameters increase when this selection is accounted for.

In Table 3, I present the OLS estimates for the elasticities of substitution across education groups. The elasticity of substitution between workers with secondary and less than secondary, σ_2, and between workers with secondary and post-secondary, σ_3, are 3.6 and 1.2 respectively. It is possible that these labor shares are the result of sorting along education-specific wage premia, and I do not attempt to instrument for them. However, recall that the native population is held fixed in 2014, such that changes in these labor ratios are driven solely by the arrival of migrants to metro areas. The standard errors are also large, reflecting the uncertainty introduced into the bootstrap by the first steps of the estimation. However, it is worth noting that the point estimates are very close to existing estimates from Colombia and other countries in Latin America. Across 16 countries in Latin America between

Table 3: Substitutability Across Education Groups

	$-\frac{1}{\sigma_2}$ (1)	$-\frac{1}{\sigma_3}$ (2)
OLS	-0.278	-0.831
	(0.134)	(0.358)
OLS (σ)	3.60	1.20
	(1.73)	(0.52)

See notes to previous table.

1991-2013, Acosta *et al.* (2019) estimate values of σ_2 and σ_3 of 3.5 and 2 respectively. Over the same period, in Argentina Brazil and Chile, Fernández & Messina (2018) estimate values of 2.3 and 1.5. While to my knowledge there are no existing estimates of σ_2 specifically in Colombia, various papers have estimated values of σ_3 in Colombia that consistently land between 1.3-1.5 (Medina & Posso, 2010; Santamaría *et al.* , 2004; Núñez & Sánchez, 1998). Thus, these estimates are very reasonable given the existing literature.

It is notable that σ_2 is much smaller than estimates of this parameter in the US, which range from around 30 to perfect substitutability (Ottaviano & Peri, 2012; Card, 2009). That σ_2 is lower in Latin America may reflect the larger socioeconomic gap in secondary school completion rates, as well as the barriers to entering the formal sector faced by workers without secondary. This implies that, in a developing country context, increased competition among workers without secondary will create relatively large wage consequences for these workers.

Finally, in Table 4 I present the estimated productivity parameters (averaged across metro areas). Migrants are relatively less productive than natives, reflected by the fact that α_{em} are all less than .5. This wage gap is most severe in high-skill occupations, where migrants have a relative productivity weight of .21, reflecting the larger wage penalty that migrants face in these occupations.[26] α_2 is .55 and α_3 is .71, reflecting the relatively higher productivity of high-skill labor. One benefit of shifting migrants from low- to high-skill occupations is an increase in total productivity, which will increase the total effective labor supply governed by equations (3) and (4).

In Section A2, I estimate an extended version of the model that separately adds a nest for age (over and under 30) or gender. Migrants are assigned to education-age or education-gender groups according to the demographic composition of their occupation, and

[26]It is important to note that these parameters also reflect other characteristics that effect the demand for migrant labor, such as discrimination.

Table 4: Productivity Parameters

α_{1m} (1)	α_{2m} (2)	α_{3m} (3)	α_2 (4)	α_3 (5)
0.439	0.373	0.210	0.547	0.709
(0.008)	(0.008)	(0.023)	(0.006)	(0.009)

See notes to previous table.

the remainder of the estimation procedure is identical. I also allow the education-specific migrant-native substitutability parameters to vary across demographic group and I find that they are similar within each of these subgroups. I estimate imperfect substitutability across age groups, and perfect substitutability across gender groups. Thus, while predicted wage effects are identical for men and women, the age-enhanced model generates interesting results regarding the distribution of wage effects across age groups that are discussed in the following section.

6.2 Robustness

In Table 5, I conduct various robustness checks for the migrant-native substitutability parameters. First, the estimates do not change when I restrict analysis to full-time workers. Second, one may be concerned that both historical migrants shares and economic trends are correlated with close proximity to the Venezuelan border, for example, due to changes in trade activity or daily commuting patterns from Venezuela.[27] I therefore show that results are robust to dropping the six metro areas within 100 km driving distance of the Venezuelan border, which also include all of the outliers with a migrant share greater than 15% that have the potential to disproportionately drive results.[28]

Third, I drop Bogotá, the largest city in the sample, to ensure it is not single-handedly driving any results.[29] Fourth, I drop all metro areas with an annual sample size of less than 1,000 observations, resulting in 27 major cities closely overlapping with the 23 official areas for which the GEIH is representative. This is to ensure that measurement error within small areas is not driving results. In both cases, the results remain stable.

[27] While Venezuela used to be a top trading partner of Colombia, its trade shares steadily declined during the 2000s such that it represented around 3.6% of exports and less than .01% of imports in 2010 (COMTRADE). However, trade may have persisted longer or had lagged economic effects closer to the border.

[28] Driving distance was calculated using Open Street Maps software, from the central municipality of the metro area to the closest crossing point along the Venezuelan border.

[29] I can also drop all metro areas one-by-one, and this is available upon request.

Table 5: Migrant-Native Substitutability Robustness

	$-\frac{1}{\sigma_{1m}}$ (1)	$-\frac{1}{\sigma_{2m}}$ (2)	$-\frac{1}{\sigma_{3m}}$ (3)
		Drop Part-Time Workers	
2SLS	14.76	12.23	5.97
	(2.66)	(1.85)	(1.25)
Mean K-P stat	63.1	73.4	96.7
		Drop <100km from Border	
2SLS	15.65	13.58	6.50
	(2.44)	(1.97)	(1.75)
Mean K-P stat	100.7	102.7	116.5
		Drop Bogota	
2SLS	15.67	14.74	8.31
	(2.20)	(2.19)	(1.51)
Mean K-P stat	57.5	67.2	73.6
		Drop Annual Sample <1,000	
2SLS	14.49	12.90	6.47
	(1.83)	(1.55)	(1.45)
Mean K-P stat	56.6	74.3	98.3
		Natives Assigned to Past Metro	
2SLS	14.63	12.87	6.54
	(2.00)	(1.62)	(1.64)
Mean K-P stat	60.0	71.8	89.7
		Include All Working-Age Natives	
2SLS	14.91	13.03	6.50
	(2.00)	(1.74)	(1.59)
Mean K-P stat	56.4	70.7	92.2
		Drop Wages of Recent Migrants	
2SLS	23.69	20.65	8.31
	(6.82)	(4.86)	(2.72)
Mean K-P stat	58.0	69.2	91.5

See notes to previous table, see paper for details.

Another concern already discussed is native internal migration. In Lebow (2021), I directly study native internal migration, and I find that Venezuelan arrivals caused small increases in out-migration among workers without post-secondary (a 1pp increase in the migrant share causes a .07pp and .1pp increase in native out-migration respectively for workers without and with completed secondary), and no significant changes in in-migration. It is not surprising that effects on internal migration are small given the short-run time period, but they nonetheless may have potential to bias results. In the current analysis I hold native labor supply fixed. However, if natives respond to Venezuelan arrivals by migrating internally, then this creates a compositional change in metro area populations that could bias the wage ratios. To account for this, I exploit the fact that I observe where natives lived 5 years before the survey, and I can assign all natives to their lagged metro area, to turn off any native compositional changes. The results remain unchanged, reflecting the small scale of the internal migration effects calculated in Lebow (2021).[30]

A sixth concern is native selection into employment, though I revealed in Section 4 that employment responses to the Venezuelan migration were small, both on average and within education group. To account for this, I show in panel 6 of Table 5 that results are identical when I use the total native working-age population instead of the working population to measure labor supply. Though this does not address any compositional changes resulting from endogenous native labor supply, it ensures that endogenous changes in L_{en} do not drive results.[31]

Finally, there may be a concern that changing composition of migrants arriving over this period is driving changes in the migrant-native wage ratio. I have already residualised wages from age and gender composition. To see if other unobserved characteristics are driving a compositional effect, I set the wages of migrants who arrived in the past year to missing, so that very recently arrived migrants impact the labor supply but not the migrant wage in each year. When I do this, the standard errors increase substantially, and the substitutability parameters increase, especially for workers in the lower two education groups. Therefore, absent compositional change, the true migrant-native substitutability is if anything larger than I estimate, and the gradient across education groups even steeper.

6.3 Total Wage Effects of Venezuelan Migration

In this section, I use the estimated parameters to calculate the total wage effects of immigration for natives at a national scale according to equations (8)-(9). To calculate changes in labor supply, using the appropriate survey weights, I calculate the number of Venezuelan migrants who arrived to Colombia between 2014-2019 and the 2014 native population by education group, and insert these into the CES equations for effective labor supply (3)-(5). As in the model estimation, migrants are assigned to education groups according to the educational distribution of natives in their observed occupation.

An important component of equations (8)-(9) is the term $(\zeta - 1)\,\Delta\ln L$, which captures the role of diminishing returns in the labor aggregate when capital is held fixed. In this

[30] If natives who would have experienced a migration-induced drop in wages leave their initial cities, then internal out-migration would also bias the substitutability parameters towards 0 (greater complementarity). However, this effect is likely to be small given the small scale of the internal migration response. Furthermore, the native education groups that responded with small levels of out-migration, groups 1 and 2, are those that have higher substitutability parameters in the estimation.

[31] I do not do this for migrants since non-working migrants do not have an occupation, which is necessary to assign them to education groups.

model, if capital is fully flexible, it is generally not possible for immigration to generate a non-trivial wage effect for the population average, because output becomes linear in the labor aggregate.[32] By holding capital fixed, I allow for negative average wage effects, which is what we observe in the 2SLS estimates of the total wage effect in Section 4. This can be motivated by the short-term nature of this immigration, since the majority of migrants did not arrive until 2018. I use the value of $\zeta = .49$, which is the preferred labor share for Colombia reported by the Penn World Tables (averaged across 2014-2019, though it is relatively stable over this period).[33] This parameter determines the size of the average wage effect, since capital constraints are the main ingredient of the model amplifying the average wage effect. The distributional effects, on the other hand, are driven by the differing magnitudes of within-group competition combined with the within- and across-group elasticity parameters, and are thus less sensitive to the choice of ζ.[34] Furthermore, unlike the average wage effect, these distributional effects persist in the long-run as capital adjusts to the labor supply increase. This can be seen by differencing the education-specific wage equations (8) and (9) and noticing that the $\left(\zeta - 1 + \frac{1}{\sigma_3}\right) d\ln L$ terms cancel.

Table 6 shows the total changes in observed labor supply, "effective" labor supply at each level of the nest, and predicted wage changes for natives by education group. The results show predicted wage effects of -3.98%, -3.35% and -0.29% for those without secondary, with secondary and with post-secondary respectively. These effects largely mirror the pattern of non-structural wage effects discussed in Section 4, though the magnitudes are a bit smaller than would be predicted by linearly multiplying the non-structural parameters by the national migrant share. This indicates that this model combined with the labor share $\zeta = .49$ does a reasonably good job of replicating the non-structural total wage effects.

I now consider a counterfactual in which I assign migrants to their observed education group rather than the one assigned according to their observed occupation. This can be thought of as a "no downgrading" scenario in which migrants compete with natives of the same education group, absent any frictions that force migrants into occupations below their skill level. Importantly, this assumes that Colombian and Venezuelan education are of comparable quality, which I presented evidence for in Section 3. It also assumes that, given their

[32] An extensive discussion of the role of capital flexibility in this framework can be found in (Ottaviano & Peri, 2008).

[33] The biggest challenge with measuring the labor share is that self-employment income includes a combination of labor and capital income. The method used here assumes that self-employed use labor and capital in same proportion as the rest of the economy. Methods that impute labor shares of self employed according to observable characteristics predict very similar values in Colombia over this period (Gomis, 2019).

[34] Various choices of ζ within plausible ranges of .45-.65 generate relatively similar magnitudes, without impacting the primary conclusions regarding the distributional benefits of no downgrading.

Table 6: Total Native Wage Effects of Immigration

	Observed Change in Labor Supply (1)	ΔL_e ($e \in \{1,2,3\}$) (2)	ΔL_e ($e \in \{m3,3\}$) (3)	ΔL (4)	$\Delta\ln(W_{en})$ (5)
Education Group 1	8.00	7.57	6.13	3.14	-3.98
Education Group 2	6.10	4.69	6.13	3.14	-3.35
Education Group 3	3.90	1.93	1.93	3.14	-0.29
Total	6.25			3.14	-1.07

Model predictions for change in native wages from the change in the population due to Venezuelan migration from 2014-2019. See paper for details.

Table 7: Counterfactual Native Wage Effects without Migrant Downgrading

	Observed Change in Labor Supply (1)	ΔL_e ($e \in \{1,2,3\}$) (2)	ΔL_e ($e \in \{m3,3\}$) (3)	ΔL (4)	$\Delta\ln(W_{en})$ (5)
Education Group 1	5.25	5.14	5.66	3.53	-3.09
Education Group 2	8.25	6.18	5.66	3.53	-3.25
Education Group 3	5.70	2.67	2.67	3.53	-0.66
Total	6.25			3.53	-1.24

Model predictions under counterfactual reallocation of migrants into observed education groups. See paper for details.

education, migrants who work in low-skill and high-skill occupations are exchangeable.

The results are presented in Table 7. Without downgrading, the labor supply increase shifts substantially towards education groups two and three - it decreases by 2.75pp in group one, and increases by 2.15pp and 1.8pp in groups two and three respectively. The wage decrease faced by natives in group one reduces from -3.98% to -3.09%. Interestingly, the wage effect for natives with completed high school remains almost unchanged, despite this group facing a substantial increase in competition absent occupational downgrading. There is a small amplification of the wage effect for post-secondary workers, from -.29% to -.66%, which in terms of percentage points is around 2/5 of the gains faced by workers without completed secondary.[35]

Why does the increase in within-group labor supply not substantially reduce wages for workers with secondary and post-secondary education? One contributing factor is that migrants and natives are more complementary in high-skill occupations, thus reducing the

[35]The average wage effect remains mostly unchanged, since it is highly sensitive to small changes in the wages of college workers, who make up a large fraction of the wage bill.

effects of within-group competition for more educated workers. Another is the increase in the total "effective" labor supplies from 3.14% to 3.53%, reflecting the higher productivity of workers in high-skill occupations, as well as the returns to spreading labor supply across skill groups in the CES framework. Because σ_3 is relatively small, there are large benefits from increasing the group 3 labor supply in terms of complementarities, working through the first term in equations (8) and (9). These productivity and complementarity effects enhance the benefits of reallocation for low-skill workers, and counteract the increase in within-group competition faced by medium- and high-skill workers.[36]

In Section A2, there is an enhanced version of the model that separately adds nests for age (over and under 30) and gender. I find perfect substitutability across genders, which implies that, despite migrants competing disproportionately with low-skill female workers, there are no effects of migration on the gender-wage gap either under downgrading or no downgrading. The age-enhanced model generates interesting predictions: younger migrants disproportionately compete with older low-skill workers, and there is imperfect substitutability across age groups. As a result, wage effects are slightly more negative for older workers without completed secondary, and this group benefits the most from undoing downgrading. The largest increases in competition under no downgrading are faced by younger workers with completed secondary or post-secondary. However, the total predicted wage effects by education group do not vary substantially relative to the initial model.

7 Conclusion

In this paper, I documented the occupational downgrading experienced by Venezuelan migrants in Colombia. I then showed that, using an instrumental variable strategy based on the historical share of Venezuelans living in a metro area, the effects of labor market competition on native Colombians were largest for less educated Colombians working in low-skill occupations. These workers suffered a wage loss without exiting the workforce or notable upgrading to higher skill occupations.

Then, using a nested CES production function framework that allows for imperfect substitutability between migrants and natives, I calculated the counterfactual wage effect

[36]While there is significant uncertainty in the estimates of σ_2 and σ_3, the general results of this exercise are similar using various values of these variables within the plausible ranges of 1-5 observed in the literature. Intuitively, as long as they are in this range of relatively low substitutability, then there are large benefits from complementarity across education groups, and large benefits to having labor supply balanced across groups.

of immigration on wages in the absence of occupational downgrading. The model predicts that, when migrants compete with workers of the same education level, the total wage effect of immigration between 2014-2019 on natives without a secondary degree reduces the increase in labor supply for this group from 8% to 5.3%, and reduces the wage impact on this group from -4% to -3.1%. At the same time, though it significantly increased competition faced by natives with secondary and post-secondary education, there was little change in the wage effects experienced by these groups. This reflects the increasing complementarity between migrants and natives in high-skill occupations, as well as the benefits that all workers experience from increased productivity and complementarities in a model that allows for imperfect substitutability.

The model also help to explain why the labor market effects of a large migration wave may be larger in the developing country context. I estimated high short-run migrant-native substitutability in low-skill occupations and showed that this is at least in part explained by the fact that these jobs are more likely to be informal. The relatively low substitutability between secondary and post-secondary workers in Latin America also inflates the the consequences of concentrating labor market competition among non-secondary workers.

These results highlight the importance of properly matching migrants to jobs given their education and experience. The Colombian example shows that even when migrants speak the same language and have de jure access to legal status, the problem of occupational downgrading persists. My analysis excludes various factors that may further increase the costs of occupational downgrading, such as reduced migrant spending and taxable income, skill-biased capital adjustments, and impacts on anti-immigrant sentiment. In this sense, I underestimate the total benefits of policies that seek to tackle the sources of occupational downgrading.

References

Acosta, Pablo, Cruces, Guillermo, Galiani, Sebastian, & Gasparini, Leonardo. 2019. Educational upgrading and returns to skills in Latin America: evidence from a supply–demand framework. *Latin American Economic Review*, **28**(1), 1–20.

Aksu, Ege, Erzan, Refik, & Kırdar, Murat Güray. 2018. *The impact of mass migration of syrians on the turkish labor market*. Tech. rept. Working Paper.

Alesina, Alberto, Miano, Armando, & Stantcheva, Stefanie. 2018. *Immigration and redistribution*. Tech. rept. National Bureau of Economic Research.

Altındağ, Onur, Bakis, Ozan, & Rozo, Sandra. 2020. Blessing or burden? Impacts of refugees on businesses and the informal economy. *Journal of Development Economics*, 102490.

Altonji, Joseph G, & Card, David. 1991. The effects of immigration on the labor market outcomes of less-skilled natives. *Pages 201–234 of: Immigration, trade, and the labor market*. University of Chicago Press.

Anderson, Michael L. 2008. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association*, **103**(484), 1481–1495.

Bahar, Dany, Özgüzel, Cem, Hauptmann, Andreas, & Rapoport, Hillel. 2019. Migration and post-conflict reconstruction: The effect of returning refugees on export performance in the former Yugoslavia.

Bahar, Dany, Choudhury, Prithwiraj, & Rapoport, Hillel. 2020. Migrant inventors and the technological advantage of nations. *Research Policy*, 103947.

Beerli, Andreas, Ruffner, Jan, Siegenthaler, Michael, & Peri, Giovanni. 2021. The abolition of immigration restrictions and the performance of firms and workers: evidence from Switzerland. *American Economic Review*, **111**(3), 976–1012.

Bonilla-Mejía, Leonardo, Morales-Zurita, Leonardo Fabio, Hermida, Didier, Flórez, Luz Adriana, Bonilla-Mejía, Leonardo, Morales, Leonardo Fabio, Hermida-Giraldo, Didier, & Flórez, Luz A. 2020. The Labor Market of Immigrants and Non-Immigrants Evidence from the Venezuelan Refugee Crisis. *Borradores de Economía; No. 1119*.

Borjas, George J. 1983. The substitutability of black, Hispanic, and white labor. *Economic Inquiry*, **21**(1), 93–106.

Borjas, George J. 2003. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, **118**(4), 1335–1374.

Borjas, George J, & Edo, Anthony. 2021. *Gender, Selection into Employment, and the Wage Impact of Immigration*. Tech. rept. National Bureau of Economic Research.

Borjas, George J, & Katz, Lawrence F. 2007. The evolution of the Mexican-born workforce in the United States. *Pages 13–56 of: Mexican immigration to the United States*. University of Chicago Press.

Borjas, George J, Freeman, Richard B, Katz, Lawrence F, DiNardo, John, & Abowd, John M. 1997. How much do immigration and trade affect labor market outcomes? *Brookings papers on economic activity*, **1997**(1), 1–90.

Borusyak, Kirill, & Jaravel, Xavier. 2021. Revisiting event study designs. *Unpublished working paper.*

Braun, Sebastian, & Mahmoud, Toman Omar. 2014. The employment effects of immigration: evidence from the mass arrival of German expellees in postwar Germany. *The Journal of Economic History*, 69–108.

Brell, Courtney, Dustmann, Christian, & Preston, Ian. 2020. The Labor Market Integration of Refugee Migrants in High-Income Countries.

Callaway, Brantly, & Sant'Anna, Pedro HC. 2020. Difference-in-differences with multiple time periods. *Journal of Econometrics.*

Card, David. 1990. The impact of the Mariel boatlift on the Miami labor market. *ILR Review*, **43**(2), 245–257.

Card, David. 2001. Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, **19**(1), 22–64.

Card, David. 2009. Immigration and inequality. *American Economic Review*, **99**(2), 1–21.

Card, David, & Lemieux, Thomas. 2001. Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, **116**(2), 705–746.

Card, David, Dustmann, Christian, & Preston, Ian. 2012. Immigration, wages, and compositional amenities. *Journal of the European Economic Association*, **10**(1), 78–119.

Caruso, German, Canon, Christian Gomez, & Mueller, Valerie. 2019. Spillover effects of the Venezuelan crisis: migration impacts in Colombia. *Oxford Economic Papers.*

Ceritoglu, Evren, Yunculer, H Burcu Gurcihan, Torun, Huzeyfe, & Tumen, Semih. 2017. The impact of Syrian refugees on natives' labor market outcomes in Turkey: evidence from a quasi-experimental design. *IZA Journal of Labor Policy*, **6**(1), 5.

Chiswick, Barry R, & Miller, Paul W. 2009. The international transferability of immigrants' human capital. *Economics of Education Review*, **28**(2), 162–169.

Clemens, Michael, Huang, Cindy, Graham, Jimmy, *et al.* . 2018. The economic and fiscal effects of granting refugees formal labor market access. *Center for Global Development and TENT. https://www. cgdev. org/publication/economic-and-fiscal-effects-granting-refugees-formal-labor-market-access-brief.*

Clemens, Michael A, & Hunt, Jennifer. 2019. The labor market effects of refugee waves: reconciling conflicting results. *ILR Review*, **72**(4), 818–857.

De Chaisemartin, Clément, & d'Haultfoeuille, Xavier. 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, **110**(9), 2964–96.

Del Carpio, Ximena V, & Wagner, Mathis C. 2015. The impact of Syrian refugees on the Turkish labor market. *World Bank policy research working paper*.

Delgado-Prieto, Lukas. 2021. Immigration, Wage Dynamics and Local Employment in Colombia.

Diewert, W Erwin. 1971. An application of the Shephard duality theorem: A generalized Leontief production function. *Journal of political economy*, **79**(3), 481–507.

Duranton, Gilles. 2015. A proposal to delineate metropolitan areas in Colombia. *Revista Desarrollo y Sociedad*, 223–264.

Dustmann, Christian, Frattini, Tommaso, & Preston, Ian P. 2013. The effect of immigration along the distribution of wages. *Review of Economic Studies*, **80**(1), 145–173.

Dustmann, Christian, Schönberg, Uta, & Stuhler, Jan. 2016. The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, **30**(4), 31–56.

Dustmann, Christian, Schönberg, Uta, & Stuhler, Jan. 2017. Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics*, **132**(1), 435–483.

Eckstein, Zvi, & Weiss, Yoram. 2004. On the wage growth of immigrants: Israel, 1990–2000. *Journal of the European Economic Association*, **2**(4), 665–695.

Edo, Anthony. 2020. The Impact of immigration on wage dynamics: Evidence from the Algerian independence war. *Journal of the European Economic Association*, **18**(6), 3210–3260.

Fernández, Manuel, & Messina, Julián. 2018. Skill premium, labor supply, and changes in the structure of wages in Latin America. *Journal of Development Economics*, **135**, 555–573.

Foged, Mette, & Peri, Giovanni. 2016. Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, **8**(2), 1–34.

Freitez, Anitza. 2011. La emigración desde Venezuela durante la última écada. *Revista temas de coyuntura*.

Friedberg, Rachel M. 2001. The impact of mass migration on the Israeli labor market. *The Quarterly Journal of Economics*, **116**(4), 1373–1408.

Goldin, Claudia, Katz, Lawrence F, *et al.* . 2020. Extending the race between education and technology. *Pages 347–51 of: AEA Papers and Proceedings*, vol. 110.

Goldin, Claudia Dale, & Katz, Lawrence F. 2009. *The race between education and technology*. harvard university press.

Goldsmith-Pinkham, Paul, Sorkin, Isaac, & Swift, Henry. 2020. Bartik instruments: What, when, why, and how. *American Economic Review*, **110**(8), 2586–2624.

Gomis, Roger. 2019. The Global Labour Income Share and Distribution: Methodological Description. *Geneva, Switzerland: International Labour Office. https://www. ilo. org/ilostatfiles/Documents/Labour% 20income% 20share% 20and% 20distribution. pdf.*

Goodman-Bacon, Andrew. 2021. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Graham, Jimmy, Guerrero Ble, Martha, Panayotatos, Daphne, & Leghtas, Izza. 2020. From Displacement to Development: How Colombia Can Transform Venezuelan Displacement into Shared Growth. *Center for Global Development*.

Guzmán, Sergio, & Marmolejo, Ivonne. 2021. *Colombia urgently needs help with the economic integration of Venezuelan immigrants*. Tech. rept.

Heckman, James J. 1979. Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.

Hunt, Jennifer. 1992. The impact of the 1962 repatriates from Algeria on the French labor market. *ILR Review*, **45**(3), 556–572.

IPUMS. 2019. Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.2 [dataset]. Minneapolis, MN.

Jaeger, David A, Ruist, Joakim, & Stuhler, Jan. 2018. *Shift-share instruments and the impact of immigration*. Tech. rept. National Bureau of Economic Research.

Katz, Lawrence F, & Murphy, Kevin M. 1992. Changes in relative wages, 1963–1987: supply and demand factors. *The quarterly journal of economics*, **107**(1), 35–78.

Katz, Lawrence F, *et al.* . 1999. Changes in the wage structure and earnings inequality. *Pages 1463–1555 of: Handbook of labor economics*, vol. 3. Elsevier.

Lebow, Jeremy. 2021. The Effect of Mass Migration on Native Labor Outcomes and Internal Migration in Colombia. *Mimeo*.

Lebow, Jeremy, Moreno Medina, Jonathan, & Coral, Horacio. 2021. Immigration and Trust: The Case of Venezuelans in Colombia. *SSRN Working Paper*.

Leuven, Edwin, & Oosterbeek, Hessel. 2011. Overeducation and mismatch in the labor market. *Handbook of the Economics of Education*, **4**, 283–326.

Lochmann, Alexia, Rapoport, Hillel, & Speciale, Biagio. 2019. The effect of language training on immigrants' economic integration: Empirical evidence from France. *European Economic Review*, **113**, 265–296.

Manacorda, Marco, Sánchez-Páramo, Carolina, & Schady, Norbert. 2010. Changes in returns to education in Latin America: The role of demand and supply of skills. *ILR Review*, **63**(2), 307–326.

Manacorda, Marco, Manning, Alan, & Wadsworth, Jonathan. 2012. The impact of immigration on the structure of wages: theory and evidence from Britain. *Journal of the European economic association*, **10**(1), 120–151.

Medina, Carlos, & Posso, Christian. 2010. Technical change and polarization of the labor market: Evidence for Colombia, Brazil and Mexico. *Borradores de Economía*, **606**.

Migración Colombia. 2019. *Total de Venezolanos en Colombia Corte a 31 de Marzo de 2019*. Tech. rept.

Migration Policy Institute. 2020. *An Uneven Welcome: Latin American and Caribbean Responses to Venezuelan and Nicaraguan Migration*. Tech. rept.

Monras, Joan. 2020. Immigration and wage dynamics: Evidence from the mexican peso crisis. *Journal of Political Economy*, **128**(8), 000–000.

Núñez, Jairo, & Sánchez, Fabio. 1998. Educación y salarios relativos en Colombia, 1976-1995. *Archivo de Macroeconomía*.

Ottaviano, Gianmarco IP, & Peri, Giovanni. 2008. *Immigration and national wages: Clarifying the theory and the empirics*. Tech. rept. National Bureau of Economic Research.

Ottaviano, Gianmarco IP, & Peri, Giovanni. 2012. Rethinking the effect of immigration on wages. *Journal of the European economic association*, **10**(1), 152–197.

Peri, Giovanni, & Sparber, Chad. 2009. Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, **1**(3), 135–69.

Peri, Giovanni, & Yasenov, Vasil. 2019. The labor market effects of a refugee wave synthetic control method meets the mariel boatlift. *Journal of Human Resources*, **54**(2), 267–309.

Peri, Giovanni, Rury, Derek, & Wiltshire, Justin. 2020. The Economic Impact of Migrants from Hurricane Maria.

Piracha, Matloob, Tani, Massimiliano, & Vadean, Florin. 2012. Immigrant over-and under-education: The role of home country labour market experience. *IZA Journal of Migration*, **1**(1), 1–21.

Poot, Jacques, & Stillman, Steven. 2010. The importance of heterogeneity when examining immigrant education-occupation mismatch: evidence from New Zealand.

Reina, Mauricio, Antonio Mesa, Carlos, & Ramírez Tobón, Tomás. 2018. *Elementos para una política pública frente a la crisis venezuela*. Tech. rept.

Rozo, Sandra, Ibáñez, Ana María, & Bahar, Dany. 2020. Give me Your Tired and Your Poor: Impact of a Large-Scale Amnesty Program for Undocumented Refugees. *Mimeo*.

Rozo, Sandra V, & Vargas, Juan F. 2021. Brothers or invaders? How crisis-driven migrants shape voting behavior. *Journal of Development Economics*, **150**, 102636.

Sanroma, Esteban, Ramos, Raúl, & Simón, Hipólito. 2008. The portability of human capital and immigrant assimilation: Evidence for Spain.

Santamaria, Julieth. 2020. 'When a Stranger Shall Sojourn with Thee': The Impact of the Venezuelan Exodus on Colombian Labor Markets.

Santamaría, Mauricio, *et al.* . 2004. *Income inequality, skills and trade: Evidence from Colombia during the 80s and 90s*. Universidad de los Andes Facultad de Economía, Colombia.

Tabellini, Marco. 2020. Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *The Review of Economic Studies*, **87**(1), 454–486.

Tribín-Uribe, Ana María, Achyuta, Adhvaryu, Anzola, Cesar, Ávila-Montealegre, Oscar, Bonilla-Mejía, Leonardo, Castro-Fernández, Juan Carlos, Flórez, Luz Adriana, Grajales-Olarte, Ánderson, Guarín-López, Alexander, Hamann-Salcedo, Franz, *et al.* . 2020. Migración desde Venezuela en Colombia: caracterización del fenómeno y análisis de los efectos macroeconómicos. *Revista Ensayos Sobre Política Económica; No. 97, octubre 2020. Pág.: 1-74.*

Tumen, Semih. 2016. The economic impact of Syrian refugees on host countries: Quasi-experimental evidence from Turkey. *American Economic Review*, **106**(5), 456–60.

UNHCR. 2019. *Global Trends: Forced Displacement in 2019.* Tech. rept.

Verme, Paolo, & Schuettler, Kirsten. 2021. The impact of forced displacement on host communities a review of the empirical literature in economics. *Journal of Development Economics*, 102606.

Walker, Maurice. 2011. PISA 2009 Plus Results: Performance of 15-year-olds in reading, mathematics and science for 10 additional participants.

Wilson Center. 2019. *Understanding the Venezuelan Refugee Crisis.* Tech. rept.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data.* MIT press.

A1 Additional Tables and Figures

Table A1: Characteristics of Migrants and Non-Migrants

	Non-Migrants	Migrants
Male (%)	48.5	49.6
	(50)	(50)
Age	36.5	31.5
	(13.9)	(11.3)
Labor Force Participation (%)	71.7	79.4
	(45)	(40.5)
Unemployment (%)	11.4	14.8
	(31.7)	(35.5)
Median Hrly Wage (2010 USD)	2.3	1.6
	(6.1)	(4.8)
Hrs per Week	45.2	49.6
	(15.9)	(17.4)
Own-Account (%)	25.3	32.1
	(43.5)	(46.7)
Informal (%)	56	88.3
	(49.6)	(32.1)
N	447264	21730

Means presented, standard deviations in parentheses. Restricted to urban residents age 15-64. Migrant defined as anyone in 2019 who was living in Venezuela 5 years ago. Population weights applied. Source: GEIH 2019

Table A2: Returns to Colombian and Venezuelan Schooling

	Colombian-Born	Venezuelan-Born
	Ln(Hrly Wage), In Colombia >5 Years	
Not Completed Secondary	(excluded)	(excluded)
Completed Secondary	0.34***	0.28***
	(0.00)	(0.05)
Any Post-Secondary	1.03***	1.15***
	(0.00)	(0.07)
N	1,689,727	2,575
	Ln(Income), US 2015	
Not Completed Secondary	(excluded)	(excluded)
Completed Secondary	0.31***	0.55**
	(0.09)	(0.21)
Any Post-Secondary	0.89***	1.17***
	(0.09)	(0.21)
N	3,389	1,316
	Ln(Income), Panama 2010	
Not Completed Secondary	(excluded)	(excluded)
Completed Secondary	0.36***	0.58**
	(0.04)	(0.28)
Any Post-Secondary	1.17***	1.07***
	(0.05)	(0.26)
N	1,356	209

Controlling age group and gender. Restricted to workers age 15-64. Population weights applied.
Source: Colombia 2014-2019 GEIH, US 2015 Census (IPUMS), Panama 2010 Census (IPUMS)

Table A3: Wage Returns to Schooling by Migrant Status and Occupation

	Non-Migrants	Migrants
	Occupation Mode: Not Completed Secondary	
Not Completed Secondary	(excluded)	(excluded)
Completed Secondary	15.80***	6.81***
	(0.68)	(2.40)
Any Post-Secondary	35.35***	14.55***
	(1.06)	(3.24)
	Occupation Mode: Completed Secondary	
Not Completed Secondary	(excluded)	(excluded)
Completed Secondary	36.70***	23.26***
	(1.43)	(4.71)
Any Post-Secondary	65.39***	24.15***
	(1.66)	(5.50)
	Occupation Mode: Any Post-Secondary	
Not Completed Secondary	(excluded)	(excluded)
Completed Secondary	35.35***	23.18
	(2.33)	(15.19)
Any Post-Secondary	96.42***	60.75***
	(2.23)	(14.73)

Outcome is 100 times log hourly wage. Controlling age group, gender, and city FE. Restricted to urban workers age 15-64. Population weights applied.
Source: 2019 GEIH

Figure A1: Relationship between 2005 and 2019 Migrant-Native Log-Ratios

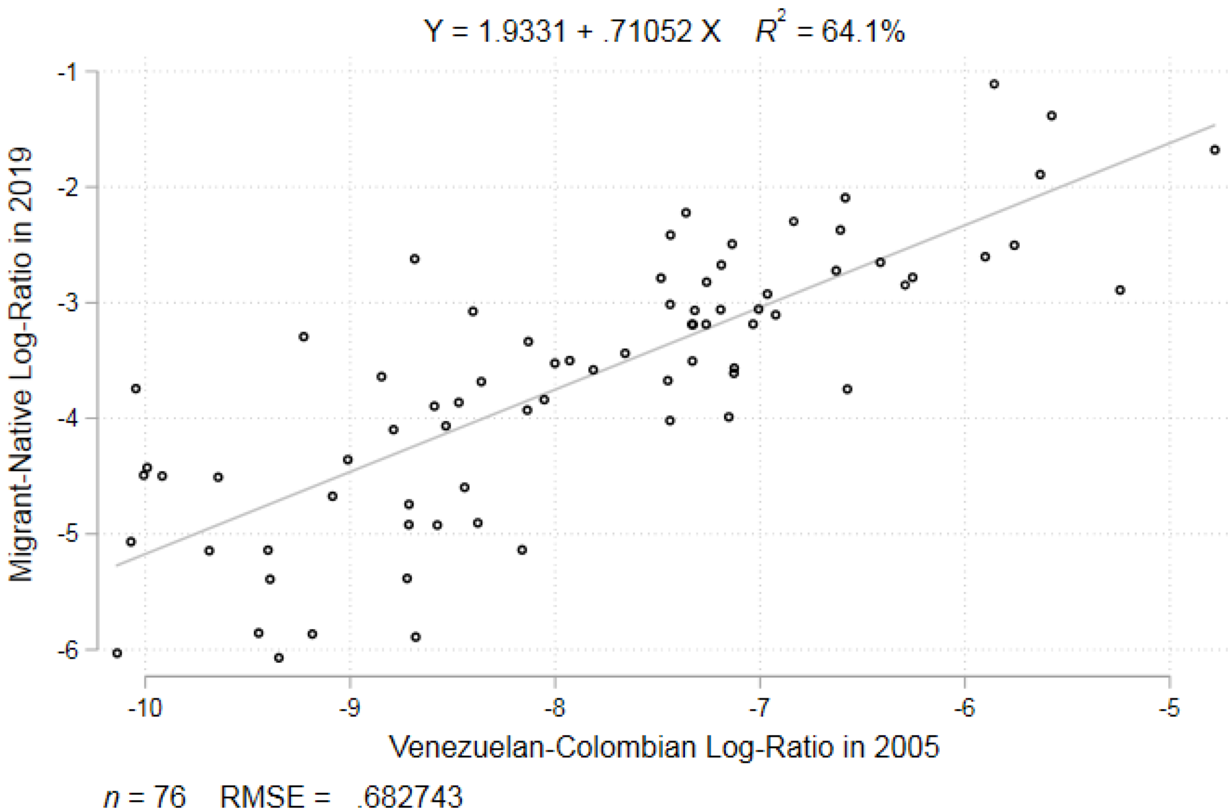
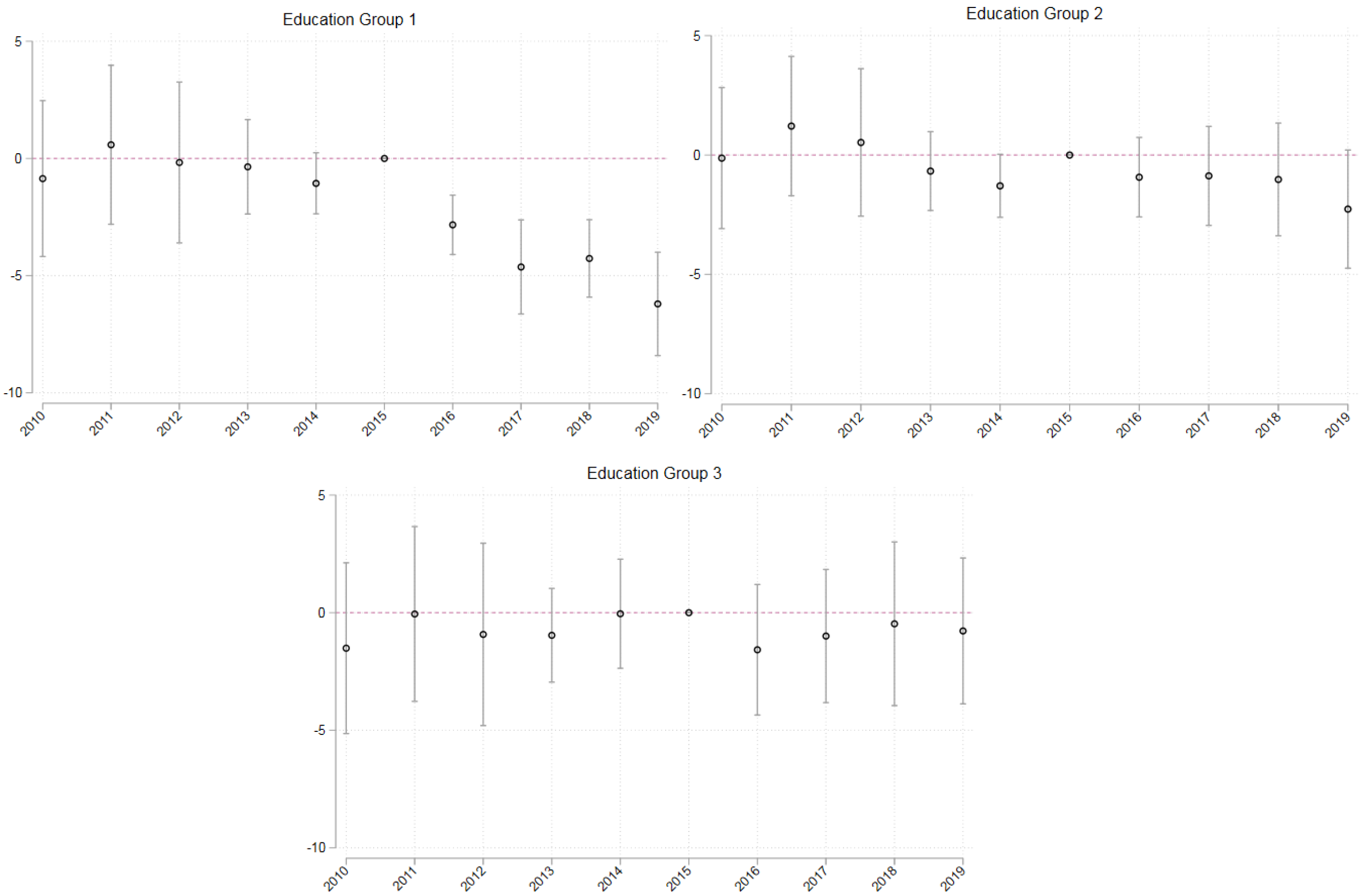


Table A4: Pre-Period Correlates of 2005 Venezuelan-Colombian Log-Ratio

2005 Venezuelan-Colombian Log-Ratio:	(1) < Completed Secondary	(2) Completed Secondary	(3) > Completed Secondary
2014 Metro Characteristic:			
Population (100,000)	0.001	0.003	0.004
	(0.004)	(0.004)	(0.004)
100*Ln(Hrly Wage)	-0.002	-0.008	-0.012**
	(0.006)	(0.005)	(0.005)
100*Ln(Hrs per Week)	0.021**	0.022*	0.015
	(0.010)	(0.011)	(0.012)
100*Unemployment Rate	-0.038	-0.063*	-0.059**
	(0.035)	(0.033)	(0.029)
100*LFP Rate	0.016	-0.018	-0.024
	(0.019)	(0.020)	(0.019)
N	79	79	79
R^2	.053	.16	.22

2005 log-ratio of the Venezuelan-born to Colombian-born population calculated by education group from the complete 2005 census and normalized to a standard deviation of 1. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Figure A2: Pre-Trends for Native Log-Wages



Outcome is native log-wage multiplied by 100. 2005 log-ratio of the Venezuelan- to Colombian-born population calculated by education from the complete 2005 censes and normalized to a standard deviation of 1. This is interacted with year (excluding 2015) and included with metro and year fixed effects. Metro-year observations weighted by working population. Cluster-robust 95% CIs presented.

Table A5: Migrant-Native Substitutability for Population Average with Year Fixed Effects

	OLS		2SLS	
	(1)	(2)	(3)	(4)
Migrant-Native Log-Ratio	-0.117	-0.092	-0.113	-0.391
	(0.019)	(0.052)	(0.016)	(0.380)
Kleibergen-Paap Wald stat			88.64	1.55
Metro FE	X	X	X	X
Year FE		X		X

Coefficient corresponds to $-\frac{1}{\sigma_m}$ estimated for the total population (not broken into education subgroups). Each metro-year cell is population-weighted. Metro-clustered standard errors in parenthesis.

Table A6: Migrant-Native Substitutability by Formality Status

	σ_{1m}	σ_{2m}	σ_{3m}
	(1)	(2)	(3)
		Formal	
2SLS	20.59	20.08	8.02
	(6.30)	(9.21)	(3.97)
Mean K-P Stat	31.3	45.8	54.7
		Informal	
2SLS	20.92	21.37	12.80
	(5.15)	(6.13)	(4.83)
Mean K-P Stat	55.8	63.7	80.1

Migrants and natives restricted to indicated formality status, defined according to compliance with mandatory pension and health insurance requirements. Clustered bootstrapped standard errors in parenthesis. See paper for estimation details.

A2 Adding Age and Gender to the Nesting Structure

Other demographic characteristics, such as age and gender, may be important dimensions along which differing levels of migrant competition impact wage inequality across groups. To investigate this, I adapt the model to separately incorporate an age and a gender division into the nesting structure. To mitigate small sample concerns, I apply the additional nest one at a time, and focusing on division into two groups: for gender, male and female, and for age, over-30 vs. under-30. I insert the nest between the education and nativity nests, and the new production function is characterized by the following set of equations:

$$Y = AK^{1-\zeta}L^{\zeta} \tag{A1}$$

$$L = \left(\alpha_3 L_3^{\frac{\sigma_3-1}{\sigma_3}} + \alpha_{-3} L_{-3}^{\frac{\sigma_3-1}{\sigma_3}}\right)^{\frac{\sigma_3}{\sigma_3-1}} \tag{A2}$$

$$L_{-3} = \left(\alpha_2 L_2^{\frac{\sigma_2-1}{\sigma_2}} + \alpha_1 L_1^{\frac{\sigma_2-1}{\sigma_2}}\right)^{\frac{\sigma_2}{\sigma_2-1}} \tag{A3}$$

$$L_e = \left(\alpha_{eD1} L_{eD1}^{\frac{\sigma_d-1}{\sigma_d}} + \alpha_{eD2} L_{eD2}^{\frac{\sigma_d-1}{\sigma_d}}\right)^{\frac{\sigma_d}{\sigma_d-1}} \tag{A4}$$

$$L_{ed} = \left(\alpha_{edm} L_{edm}^{\frac{\sigma_{edm}-1}{\sigma_{edm}}} + \alpha_{edn} L_{edn}^{\frac{\sigma_{edm}-1}{\sigma_{edm}}}\right)^{\frac{\sigma_{edm}}{\sigma_{edm}-1}} \tag{A5}$$

where $d\epsilon\{D1, D2\}$ notates the demographic group (either age or gender). The new equation in the system is equation (A4).The elasticity of substitutability between demographic groups, σ_d, is forced to be constant across education groups because of the difficulty in estimating this parameter precisely. However, I allow the migrant-native substitutability parameter, σ_{edm}, to vary within both education and age group. In this system, log-wages for $e \in \{1, 2\}$ are:

$$\begin{aligned} \ln W_{edj} = {} & \ln\left(AK^{1-\zeta}\zeta\right) + \ln\alpha_{-3} + \ln\alpha_e + \ln\alpha_{ed} + \ln\alpha_{edj} + \left(\zeta - 1 + \frac{1}{\sigma_3}\right)\ln L \\ & + \left(\frac{1}{\sigma_2} - \frac{1}{\sigma_3}\right)\ln L_{-3} + \left(\frac{1}{\sigma_d} - \frac{1}{\sigma_2}\right)\ln L_e + \left(\frac{1}{\sigma_{edm}} - \frac{1}{\sigma_d}\right)\ln L_{ed} - \frac{1}{\sigma_{edm}}\ln L_{edj} \end{aligned} \tag{A6}$$

and for $e = 3$ are:

$$\begin{aligned} \ln W_{3dj} = {} & \ln\left(AK^{1-\zeta}\zeta\right) + \ln\alpha_3 + \ln\alpha_{3d} + \ln\alpha_{3dj} + \left(\zeta - 1 + \frac{1}{\sigma_3}\right)\ln L \\ & + \left(\frac{1}{\sigma_d} - \frac{1}{\sigma_3}\right)\ln L_3 + \left(\frac{1}{\sigma_{3dm}} - \frac{1}{\sigma_d}\right)\ln L_{3d} - \frac{1}{\sigma_{3dm}}\ln L_{3dj} \end{aligned} \tag{A7}$$

The equation used to estimate the migrant-native substitutability parameters is analogous to equation (12) and derived in the same way:

$$\ln\left(\frac{W_m}{W_n}\right)_{edct} = \lambda_{edc} - \frac{1}{\sigma_{edm}}\ln\left(\frac{L_m}{L_n}\right)_{edct} + \epsilon_{edct} \tag{A8}$$

where both the wage-ratio and labor-ratio now vary by demographic group as well, and education-age-metro fixed effects are included to capture all fixed wage differences across demographic groups within each metro area. As before, these fixed effects are used to back out the productivity parameters, α_{edj}. Finally, analogous estimating equations are derived for the parameters higher up in the nesting structure:

$$\ln\left(\frac{W_{D2}}{W_{D1}}\right)_{ct} = \lambda_{1,c} + \lambda_{1,t} - \frac{1}{\sigma_d}\ln\left(\frac{L_{A2}}{L_{A1}}\right)_{ct} + \epsilon_{1,ct} \tag{A9}$$

$$\ln\left(\frac{W_2}{W_1}\right)_{ct} = \lambda_{2,c} + \lambda_{2,t} - \frac{1}{\sigma_2}\ln\left(\frac{L_2}{L_1}\right)_{ct} + \epsilon_{2,ct} \tag{A10}$$

$$\ln\left(\frac{W_3}{W_2}\right)_{ct} = \lambda_{3,c} + \lambda_{3,t} - \frac{1}{\sigma_3}\ln\left(\frac{L_3}{L_2}\right)_{ct} + \epsilon_{3,ct} \tag{A11}$$

The procedure for estimation is exactly as described in the paper, including the use of the instrument based on education-specific migrant enclaves presented in equation (15), as well the iterated random assignment of migrants into age- or gender-education groups according to the native distribution in their observed occupation. Thus, I estimate the substitutability between natives in each demographic-education bin and migrants who work in occupations competing with those workers, and predict the changes in the total wage effects if migrants instead compete with workers of their own demographic-education group.

I start by describing the results of the model with an added age nest, in the first Panel of Table A7. Migrant-native substitutability does not vary across age groups and is identical to the values estimated in the model without age. The elasticity of substitution across age groups, σ_d, is 3.4, indicating small imperfect substitutability across groups. This is consistent with the parameter values of around 5 that have been found in the US and the UK, and with Fernández & Messina (2018) in Latin America, which finds elasticities of substitution across experience groups of 3.6 among non-college graduates and 5.5 among college graduates.[37] Finally, $-\frac{1}{\sigma_3}$ is .31 and $-\frac{1}{\sigma_2}$ is .64, very close to the values estimated in the initial model.

Table A8 presents the changes in labor supply and predicted changes in wages under the scenarios of occupational downgrading (migrants assigned to demographic groups according to their observed occupation, as was done during estimation) and no downgrading (migrants assigned to their actual demographic group). Column 1 shows that migrants disproportionately work in jobs that tend to have more natives under age 30. Workers with less than high-school under age 30 experience over a 10% increase in labor supply. Thus, because of imperfect substitutability across age groups, the effects of wage competition in education group 1 are slightly larger for younger workers.

Under the no-downgrading counterfactual, migrants are shifted mostly from jobs with older, less-educated workers to jobs with younger, more educated workers. This interesting pattern highlights that the occupational downgrading is mostly increasing competition for older, less-educated workers, and they are the ones who benefit the most from undoing occupational downgrading. Likewise, it is the younger, better-educated workers who face the largest increases in competition and the largest resulting decrease in wages. The overall magnitude of the wage effects across education group both under downgrading and no downgrading is comparable to those based on the initial model without the age division.

I now turn to the results using a gender division rather than an age division are analogous. The education-specific substitutability between migrants and natives presented in the second panel of Table A7

[37] I do not allow the substitutability across experience groups to vary by education, but in practice differences in σ_d of this magnitude would not significantly impact total wage effects.

Table A7: Substitutability Parameters in Model with Demographic Nest

	$-\frac{1}{\sigma_{1dm}}$ (1)	$-\frac{1}{\sigma_{2dm}}$ (2)	$-\frac{1}{\sigma_{3dm}}$ (3)	$-\frac{1}{\sigma_{1dm}}$ (4)	$-\frac{1}{\sigma_{2dm}}$ (5)	$-\frac{1}{\sigma_{3dm}}$ (6)	$-\frac{1}{\sigma_{d}}$ (7)	$-\frac{1}{\sigma_{2}}$ (8)	$-\frac{1}{\sigma_{3}}$ (9)
Alternate Model 1: Added Age Nest									
		Age≤30			Age>30				
OLS	-0.07	-0.08	-0.14	-0.07	-0.07	-0.14	-0.29	-0.31	-0.64
	(0.01)	(0.01)	(0.03)	(0.02)	(0.01)	(0.04)	(0.14)	(0.23)	(0.41)
2SLS	-0.07	-0.08	-0.15	-0.07	-0.08	-0.17			
	(0.01)	(0.01)	(0.04)	(0.01)	(0.01)	(0.05)			
Mean K-P Stat	57.3	62.1	85.5	56.9	70.6	99.8			
Alternate Model 2: Added Gender Nest									
		Male			Female				
OLS	-0.06	-0.07	-0.13	-0.07	-0.08	-0.15	0.11	-0.29	-0.91
	(0.01)	(0.01)	(0.03)	(0.02)	(0.02)	(0.04)	(0.14)	(0.25)	(0.58)
2SLS	-0.07	-0.07	-0.15	-0.07	-0.08	-0.18			
	(0.01)	(0.01)	(0.05)	(0.01)	(0.02)	(0.05)			
Mean K-P Stat	57.6	72.2	88.9	50.4	63.5	97.2			

Means presenteed from 100 random assignments of migrants to demographic groups. Metro-clustered bootstrapped standard errors in parenthesis. Kleibergen-Paap Wald statistic is clustered at the metro level. See paper for estimation details.

again do not vary by gender, nor do the productivity parameters for migrants and natives (not shown). While I attain similar estimates for σ_2 and σ_3, the inverse elasticity of substitution across genders, $-\frac{1}{\sigma_d}$ is now in fact positive. This does not have a clear theoretical interpretation (it could be interpreted as "greater than perfect substitutability", or in other words, the relative wage of men increases as the relative supply of men increases). I simply interpret this as perfect substitutability across genders, and I set σ_d to 50 when calculating total wage effects. As a result, any differential competition across genders will not impact the distribution of wages across gender.

In principle, the fact that migrants in low-skill occupations tend to work in occupations with more women (causing an 9.3% increase in labor supply for women without completed secondary, relative to a 4.4% increase if only migrant women without completed secondary entered this category) could cause a disproportionate decrease in wages for less educated women under downgrading, and a disproportionate benefit for them under no downgrading. However, because of the perfect substitutability across genders, male and female wages are essentially equal under each scenario. Thus, the model predicts no effects of immigration or downgrading on the gender wage gap.

Table A8: Total Native Wage Effects in Model with Demographic Nest

	With Downgrading		Without Downgrading	
	%Δ Labor Supply	Predicted %Δ Wage	%Δ Labor Supply	Predicted %Δ Wage
	(1)	(2)	(3)	(4)
	Alternate Model 1: Added Age Nest			
Education Group 1				
Age $\leq$ 30	10.53	-3.78	10.51	-3.50
Age $>$ 30	7.32	-3.37	3.83	-2.38
Average	8.00	-3.47	5.25	-2.65
Education Group 2				
Age $\leq$ 30	7.03	-2.84	11.75	-3.31
Age $>$ 30	5.49	-2.65	5.96	-2.52
Average	6.10	-2.72	8.25	-2.83
Education Group 3				
Age $\leq$ 30	5.83	-0.73	9.51	-1.17
Age $>$ 30	2.97	-0.43	3.87	-0.67
Average	3.90	-0.50	5.70	-0.78
	Alternate Model 2: Added Gender Nest			
Education Group 1				
Male	7.20	-4.26	5.76	-3.21
Female	9.29	-4.11	4.43	-3.23
Average	8.00	-4.21	5.25	-3.21
Education Group 2				
Male	5.43	-3.60	8.72	-3.43
Female	7.02	-3.45	7.59	-3.37
Average	6.10	-3.54	8.25	-3.41
Education Group 3				
Male	3.84	-0.21	5.87	-0.60
Female	3.96	-0.17	5.55	-0.58
Average	3.90	-0.19	5.70	-0.59

Columns 1-2 use migrant demographic assignments according to the native distribution in their observed occupation (also used to estimate the model). Columns 3-4 are a counterfactual in which migrants are assigned to their actual demographic group. See paper for details.

A3 Generalized Production Function

A generalized production function would allow the elasticities of substitution to vary across every type of labor. With 6 groups (3 education groups and 2 nativity statuses), combined with a restriction that the cross-elasticities are symmetric, it is necessary to estimate 15 cross-elasticities and 6 own-elasticities. This is more demanding of the data relative to the nested CES model, which only required estimating 5 elasticities. However, it is still possible to estimate, and can help to motivate the choice of nesting structure in the nested CES framework. I consider a generalized Leontief production function, which, similar to a translog production function, can be thought of as a second-order approximation to an arbitrary production function. The advantage of the generalized Leontief is its empirically tractability, since it results in a linear-in-parameters wage equation (Diewert, 1971). Production takes the following form:

$$Y = \sum_j \sum_i \gamma_{ij} \left(L_i L_j\right)^{\frac{1}{2}} \tag{A12}$$

where the technology parameters are symmetric such that $\gamma_{ij} = \gamma_{ji}$ (by Young's Theorem, or the symmetry of second derivatives). By leaving capital out of the production function, I implicitly assume strong capital separability, motivated by both the difficulty of measuring capital and the short period of migration studied. Under perfect competition, wages are equated to the marginal product of labor:

$$w_i = \gamma_{ii} + \sum_{i \neq j} \gamma_{ij} \left(\frac{L_j}{L_i}\right)^{\frac{1}{2}} \tag{A13}$$

In this system, a positive (negative) γ_{ij} indicates complementarity (substitutability) between factors i and j. To convert these parameters into measures of the degree of substitution across parameters, I use the Hicks partial elasticity of complementarity, c_{ij}, which measures the effect of a relative change in the quantity of j on the price of i. This is derived for the generalized Leontief in Borjas (1983):

$$c_{ij} = \frac{\gamma_{ij}\bar{w}}{2w_i w_j \left(P_i P_j\right)^{\frac{1}{2}}} \quad \text{when } \; i \neq j \tag{A14a}$$

$$c_{ii} = \frac{\left(\gamma_{ii} - w_i\right)\bar{w}}{2p_i w_i^2} \quad \text{when } \; i = j \tag{A14b}$$

where $P_i = \frac{L_i}{\sum_k L_k}$ and $\bar{w} = \sum_k P_k w_k$.

Equation (A13) describes the system of equations that I simultaneously estimate using seemingly unrelated regression, imposing the constraint that $\gamma_{ij} = \gamma_{ji}$ and including metro and year fixed effects to absorb any elements of the error term that are fixed within these units. Like in the main text, I hold the native population fixed in 2014 and assign migrants to education groups according to their observed occupation. As before, I take the average of 100 iterations of this procedure and bootstrap the standard errors. The results are thus driven entirely by the arrival and occupational distribution of migrants across cities and years and not subject to the endogeneity bias of native labor supply responding to wage incentives. To simplify the procedure, I choose not to use the instrument for migrant arrival, motivated by the similarity of the OLS and 2SLS estimates in the nested CES framework.

The resulting cross-group coefficients (γ_{ij}) and elasticities of complementarity (c_{ij}) are presented in

Table A9. I use "NAT" to indicate natives and "MIG" to indicate Venezuelan migrants, and "E1", "E2" and "E3" to indicate education groups 1, 2 and 3 (less than secondary, secondary and post-secondary respectively). First, looking at the elasticities of complementarity in column 1, we see that natives in education groups 1-2 are substitutes, though t-statistic on this elasticity is only 1.42. In column 2, we see that natives in groups 1-3 or 2-3 are both closer to each other in magnitude and closer to 0, with confidence intervals consistent with them being complements or substitutes. These parameters are consistent with the results of the CES estimation, in which education groups 1-2 were substitutes and 3 and -3 were still substitutable but less so. However, the bootstrapped standard errors in the generalized Leontief are large such that it is hard to make informative conclusions from this analysis. This is also true of the standard errors on the elasticities for migrant groups, which are in many cases consistent with both large positive and large negative elasticities. One result that stands out is that migrants are more substitutable with other migrants than with natives. My primary take-away is that a more generalized model supports a CES nesting structure that first splits post-secondary and non-post-secondary, and then secondary into completed and not completed secondary (thus constraining the substitutability between groups 1-3 and 2-3 to be identical).

Table A9: Generalized Leontieff Estimation Results

	NAT E2	NAT E3	MIG E1	MIG E2	MIG E3
	(1)	(2)	(3)	(4)	(5)
			Coefficient Estimates		
NAT E1	-1.113	-0.130	0.039	0.031	-0.055
	(0.776)	(0.620)	(0.018)	(0.024)	(0.030)
NAT E2		0.064	-0.050	-0.029	0.036
		(0.411)	(0.018)	(0.012)	(0.042)
NAT E3			0.015	0.005	0.027
			(0.017)	(0.021)	(0.038)
MIG E1				-0.038	-0.049
				(0.022)	(0.048)
MIG E2					-0.061
					(0.047)
			Hicksian Elasticity of Complementarity		
NAT E1	-2.860	-0.170	0.663	0.681	-1.183
	(2.006)	(0.811)	(0.253)	(0.462)	(0.533)
NAT E2		0.076	-0.776	-0.595	0.703
		(0.494)	(0.220)	(0.221)	(0.668)
NAT E3			0.115	0.056	0.269
			(0.106)	(0.176)	(0.308)
MIG E1				-5.042	-6.297
				(1.988)	(4.305)
MIG E2					-10.271
					(5.585)

Means presenteed from 100 random assignments of migrants to education groups. Clustered bootstrapped standard errors in parenthesis.

A4 Migrant-Native Substitutability by Formality: Adjusting for Selection

To allow migrant-native elasticities of substitution to vary by formality status, I split the sample into formal and informal workers while estimating equation (12). Thus, I estimate this elasticity specifically for formal (informal) natives in an education group, and formal (informal) migrants assigned to the education group according to their occupation. However, this may suffer from a selection issue if natives sort into formal or informal jobs in response to the migrant arrival. In this section, I explore sensitivity to this issue using a Heckman correction for native sectoral choice.

To begin, we can write wages for individual i in the formal sector as a function of labor supply and an idiosyncratic error term, suppressing subscripts for education e, city c, and year t to simplify notation (this entire procedure can be analogously repeated for the informal sector):

$$\ln w_i^F = \beta_0 + \beta_1 \ln L^F + \epsilon_i \tag{A15}$$

Let D_i represent the event that person i is in the formal sector and wage w_i^F is observed. This occurs when the latent index $\omega_i^* = w_i^{F*} - w_i^{I*} > 0$, or when the individual's potential wage in the formal sector exceeds their potential wage in the informal sector. Express $D_i = 1$ as:

$$D_i = \gamma Z_i + \mu_i \tag{A16}$$

If we assume that that $(\epsilon, \mu) \sim N(0, 0, \sigma_\epsilon^2, \sigma_\mu^2 = 1, \rho)$, then under standard results from the sample selection literature (Heckman, 1979), expected wages can be written as

$$\begin{aligned} E(\ln w_i^F | D_i = 1) &= \beta_0 + \beta_1 \ln L^F + E(\epsilon_i | D_i = 1) \\ &= \beta_0 + \beta_1 \ln L^F + \sigma_\epsilon \rho \lambda(\gamma Z_i) \end{aligned} \tag{A17}$$

where $\lambda(\gamma Z_i)$ is the inverse mills ratio evaluated at the predicted probability of being in the formal sector. The sign of ρ determines the direction of selection bias, since this indicates whether individuals who select into the labor force are higher- or lower-wage. Equation (A17) can be estimated by controlling for the inverse mills ratio evaluated at the Probit estimate from the first stage equation (A16). Furthermore, the test that the coefficient $\sigma_\epsilon \rho = 0$ is a test for endogenous selection since $\sigma_\epsilon > 0$.

To estimate this with observations at the metro-year-education cell level, I follow the approach in Borjas & Edo (2021) and take the residual of native wages from a regression on the inverse mills ratio from the first stage Probit:

$$\ln w_{in}^F = \sigma_\epsilon \rho \lambda(\gamma Z_i) + \delta_i \tag{A18}$$

My new measure of native log-wages is $\ln W_n^F = \bar{\delta}_i$, or the mean value of the residual δ_i within each metro-year-education cell. I continue to assume that migrant job choice is exogenous, so $\ln W_m^F = \ln \bar{w}_{im}^F$. I then insert these measures into equation (12) to estimate the formality-specific substitutability parameter. I do this using the same estimation procedure, with random assignment of migrants to education groups according to their occupation, using the same instrument for the migrant share, and bootstrapping standard errors. Estimating 2SLS while controlling for the inverse mills ratio is consistent as long as the 2SLS instrumental variable, the enclave shift-share, is included in the first-stage Probit (Wooldridge, 2010).

For this method to be valid, I need an instrument Z_i that is a strong predictor of natives' formality status and is not correlated with wages via any other channels. The instrument that I use is the mean formality rate within the person's two-digit NAICS industry code, conditional on the mean hourly wage within that industry (in both the first-stage and the wage equation). This is a strong predictor of native formal sector participation and is assumed to be otherwise unrelated to wages controlling for the mean industry wage. Intuitively, a worker in an industry with a higher formality rate has a higher probability of being formal relative to a worker in an industry with the same average wages but a lower formality rate.

These results are presented in Table A10. The first thing to note is that the inverse mills ratio is strongly significant and indicates positive selection into formality for natives without secondary. However, it is less significant for workers with secondary or greater. Thus, the differences between this table and Table A6 are most notable for less educated workers. In particular, after adjusting for selection, migrant-native substitutability decreases in formal jobs. Overall, the results after the selection adjustment are comparable, and still support the conclusion that migrant-native substitutability is slightly higher in informal jobs, though these differences are not statistically significant.

Table A10: Migrant-Native Substitutability by Formality Status: Adjusted for Selection

	σ_{1m} (1)	σ_{2m} (2)	σ_{3m} (3)
		Formal	
2SLS	15.75	19.15	7.68
	(3.70)	(8.39)	(3.53)
Mean K-P Stat	31.3	45.8	54.7
IMR Mean Coef	0.307	0.013	-0.074
IMR Mean F-stat	213.7	1.0	5.1
		Informal	
2SLS	18.42	20.85	12.35
	(4.37)	(5.89)	(4.50)
Mean K-P Stat	55.8	63.7	80.1
IMR Mean Coef	-0.413	0.000	0.016
IMR Mean F-stat	125.1	0.0	0.5

Migrants and natives restricted to indicated formality status. Native log-wages residualized from regression on inverse mills ratio, using mean 2-digit industry formality rate conditional on industry mean wages as excluded variable in first-stage Probit. Clustered bootstrapped standard errors in parenthesis. See paper for estimation details.