Slowdown in Immigration, Labor Shortages, and Declining Skill Premia*

Federico S. Mandelman
Federal Reserve Bank of Atlanta

Yang Yu
Shanghai Jiao Tong University

Francesco Zanetti
University of Oxford and CEPR

Andrei Zlate
Federal Reserve Board

March 28, 2025

Abstract

We show that the steady decline in low-skilled immigration that began with the Great Recession in 2007 was associated with labor shortages in low-skilled service occupations and a decline in the skill premium. We develop and estimate a stochastic growth model with endogenous immigration and endogenous training to account for these empirical regularities. Lower immigration leads to higher wages for low-skilled workers and higher consumer prices. The decline in the skill premium discourages the training of native workers, which reduces aggregate productivity and welfare. Low-skill labor shortages lowered the effectiveness of fiscal policy stimulus during the COVID-19 pandemic.

JEL classification: F16, F22, F41.
Keywords: Immigration, skill premium, task upgrading, heterogeneous workers.

*Email: federico.mandelman@atl.frb.org (Federico Mandelman), yu.yang.econ@sjtu.edu.cn (Yang Yu), francesco.zanetti@economics.ox.ac.uk (Francesco Zanetti, correspondence), and andrei.zlate@frb.gov (Andrei Zlate). We thank our discussants Amelie Constant, Rosario Crino, Jerome Gonnot, Brad Hershbein, Pravin Krishna, Giovanni Peri, Fernando Rios-Avila, Henry Siu, and participants at several conferences and seminars for useful comments. We thank Nick Croteau and Li Guo for excellent research assistance. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of Atlanta, the Federal Reserve Board, or the Federal Reserve System.

1 Introduction

We document the protracted slowdown in low-skill immigration since the Great Recession in 2007, and outline four related facts based on U.S. regional data that link low-skill immigration, immigration policy, and the reduction in educational attainment of native workers to the fall in the skill-premium. We use these facts to construct a novel stochastic growth model with endogenous immigration, training, and offshoring choices that allows us to study the inter-linkages between the fall in the skill premium and the slowdown in low-skill immigration, along with the related changes in aggregate prices and broader macroeconomic conditions. The estimation of the structural model allows us to quantify the implications of the changes in labor market dynamics for macroeconomic performance and welfare.

We focus our study on the following questions: What is the interplay between the slowdown in low-skill immigration, the reduction in the skill premium, and the decision of native workers to invest in education? What are the implications of low-skill immigration and training for macroeconomic performance and welfare? Did low-skill labor shortages shape the effectiveness of policy stimulus such as the 2020 coronavirus stimulus package, and were they reversed by the 2022-2024 surge in low-skill immigration?

Motivation and evidence from the employment data. Immigration is central to understanding labor market dynamics in the United States. According to the U.S. Census Bureau, between 1980 and 2007, the foreign-born prime-age population (25-54 years old) increased by 16.3 million people, an average annual growth rate of 4.8 percent, and accounted for 29 percent of the increase in the total U.S. prime-age population. Since the Great Recession in 2007, however, the inflow of immigrants has declined sharply, and the average annual growth rate of the foreign-born population has more than halved to 1.9 percent for the prime-age group, while the total U.S. prime-age population has remained almost unchanged at about 126 million between 2007 and 2019. This decline in the foreign-born population has important implications for countries with aging populations like the U.S. We estimate that if the foreign-born population had continued to grow at its pre-Great Recession average rate, the total U.S. prime-age population would have been 11 percent higher in 2019.[1] These numbers are striking, yet they represent lower-bound estimates of the impact of the immigration slowdown on U.S. population growth since they likely understate the contribution of undocumented immigration, which is largely low-skill. By conservative estimates, undocumented immigration rose from negligible numbers in the early 1980s to about 12.2 million in 2007 but declined to 10.5 million in the following decade.[2]

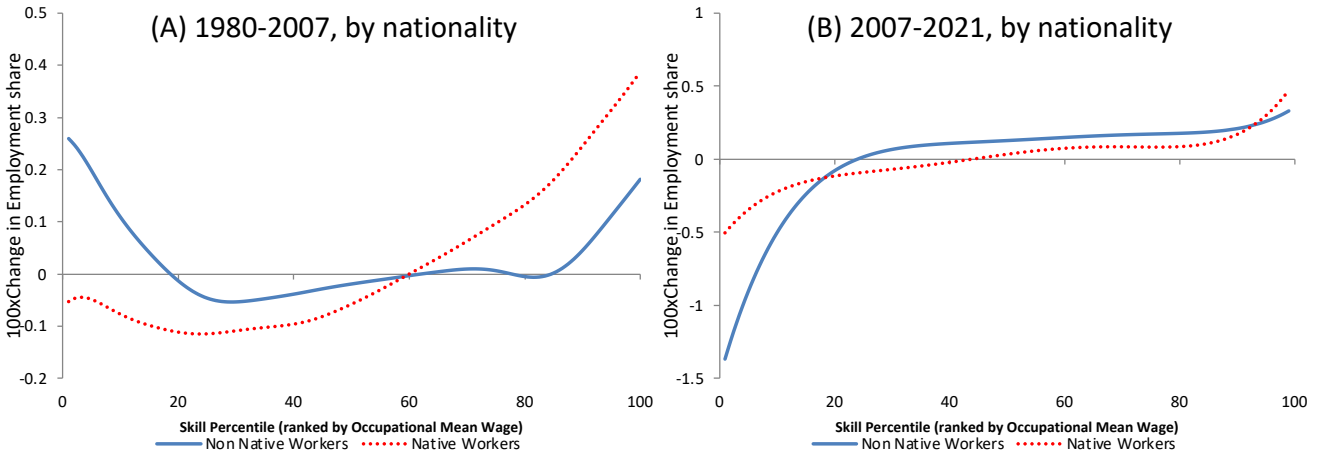
To contextualize our analysis, we first document marked differences in the evolution of native and non-native employment across the skill distribution in two distinct periods. Figure 1(A) shows employment growth across occupations grouped into percentiles and ranked by skill (x-axes) between 1980 and 2007 like in Autor and Dorn (2013), which we break down by native (dotted line) and foreign-born workers (solid line). For natives, net employment growth was mostly concentrated in high-skill occupations, with

[1] We compute these figures using data from the U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplement.

[2] Krogstad et al. (2019) documents that 71 percent of undocumented immigrants are in the prime age group, compared to 38 percent for the total U.S. population. Considering that most 16-24 year old undocumented immigrants also work, since they cannot pursue tertiary education in the U.S., around 85 percent of these immigrants are of working age.

negative growth in low-skill occupations. For foreigners, instead, the bulk of employment growth was at the bottom quartile of the skill distribution. Figure 1(B) shows the striking reversal of these employment dynamics in the 2007-2021 period that produced negative employment growth especially for foreign-born workers at the bottom quartile of the skill distribution, reflecting the decline in low-skill immigration. In contrast, there was no such reversal in high-skilled employment for both native and non-native workers. While the chart covers the pandemic years, the conclusions are the same for 2007-2019 only.[3] Employment in the bottom quartile of the skill distribution is concentrated in service occupations (such as child care workers, restaurant and hotel workers, domestic and office cleaners, beauticians, gardeners, health aides, etc.) and construction. Importantly for our analysis, these occupations require work to be performed in the same geographic location where the final consumption takes place, and thus cannot be offshored, making the hiring of immigrant workers the only viable alternative to meet increases in demand.

Figure 1: Smoothed changes in the growth of employment by skill percentile and nationality



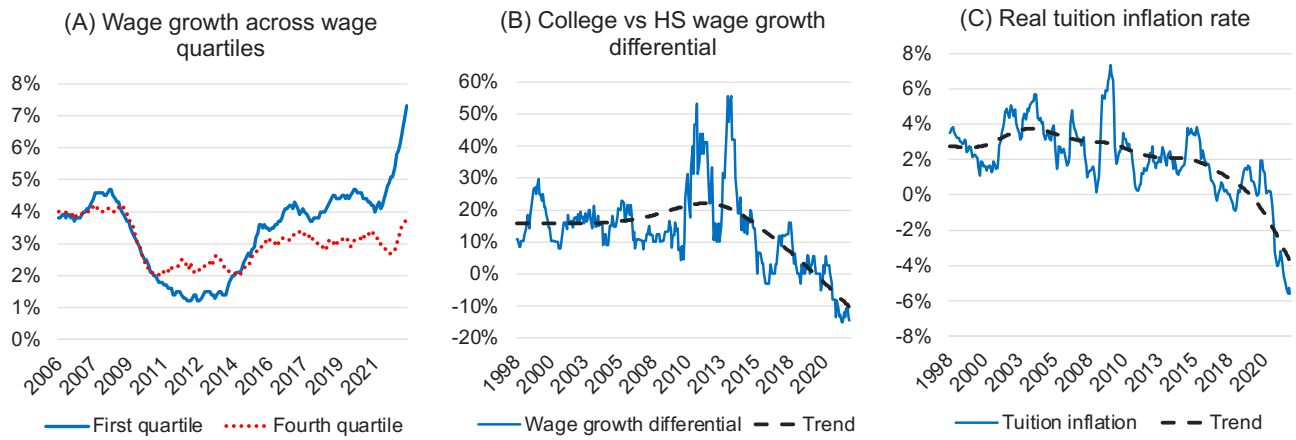
Note: We use U.S. Census/American Community Survey (ACS) data to compute changes in employment shares between 1980 and 2007 (left panel) and 2007-2021 (right panel). The occupations are sorted into 100 percentiles based on the mean occupational wages in the starting periods as a proxy for skill. The change in shares is obtained as the difference between the share of total U.S. employment in the starting and ending periods for each percentile. The smooth changes are obtained with a locally weighted polynomial regression between the change in employment shares and the corresponding percentiles.

In turn, the relative scarcity of low-skilled immigrant labor has contributed to the rapid growth of low-skilled wages over the past decade. Figure 2(A) shows nominal wage growth for jobs in the first quartile (red-dashed line) versus the fourth quartile (blue-solid line) of the earnings distribution. The faster wage growth for low-skill jobs coincided with a sharp decline in the college skill premium, shown in figure 2(B). The unprecedented decline in the skill premium reverses a four-decade upward trend.[4] Interestingly, the decline in the skill premium coincided with a decline in the inflation-adjusted cost of tuition, as shown in Figure 2(C), consistent with the lower return to investing in education over the period.

[3] See Appendix for details, where we also provide more details on the occupational sorting in these graphs

[4] See, for example, Krusell et al. (2000) and Acemoglu and Autor (2011) among many others

Figure 2: Wage growth across wage quartiles, college vs high school wage growth differential, and real tuition inflation



Note: Panel (A) illustrates the monthly year-on-year nominal wage growth for the first (red-dotted curve) and fourth (blue-solid curve) quartiles in the earnings distribution. Panel (B) illustrates the monthly year-on-year wage growth percentage differential between workers with college and high school (or less). Panel (C) illustrates the monthly year-on-year percentage change in tuition costs (tuition, other school fees, and childcare) deflated by the CPI index. The dashed curves in Panels (B) and (C) are the trend estimated with an HP filter with a smoothing parameter of 100,000. See Figure A13 in the Appendix for an alternative real tuition inflation rate series that captures tuition and fees only.

Evidence from regional data and the design of the theoretical model. We complement our novel aggregate-level evidence with some facts based on regional data from the U.S. Census' American Community Survey—which provides the characteristics of workers residing in different states, metropolitan areas and commuting zones—to shed light on the role of low-skill immigration in shaping U.S. labor market dynamics. In turn, we will use these facts to develop a theoretical model with realistic dynamics for immigration.

Fact I shows that the positive correlation between the skill premium and low-skill immigration is robust across states, metropolitan areas and commuting zones. Consistently, in our structural model, the arrival of low-skill immigrant workers boosts labor supply, lowers low-skill wages, and increases the skill premium in the destination country. Fact II shows that low-skill immigration across commuting zones is positively correlated with the initial low-skill wage at the destination, motivating that in the model, workers' migration decision is endogenous, and the key factor behind it is the expected wage in the non-tradable sector at the destination. Fact III shows that a more restrictive immigration policy at the state level is associated with lower immigrant inflows, motivating the presence of time-varying sunk migration costs in the model to capture the stance of immigration policy. Fact IV shows that educational attainment is positively correlated with the skill premium, justifying the key role of the skill premium in driving the household's endogenous training decision in the model.

Our model economy consists of two large economies (Home and Foreign) that trade with each other and a third small economy (South) that is the origin of low-skill immigrants, who in turn work in the (non-tradable) service sector in the Home region. We incorporate elements into the model that build on empirical and theoretical advances in the literature. As in Grossman and Rossi-Hansberg (2008), labor *tasks* rather than *goods* are endogenously traded between Home and Foreign countries, to which we

refer to as *offshoring* like in Mandelman and Zlate (2022). Due to remarkable declines in transportation and communication costs, international trade now increasingly consists on breaking down the production process of final goods into separate labor tasks that can be performed in different locations. We model this phenomenon with fixed and iceberg costs of task offshoring, as well as a stochastic shock to the iceberg costs that we estimate with data. We also model low-skill migration as an endogenous decision subject to a time-varying sunk migration cost. Following this approach, our model is able to disentangle the separate contributions of offshoring (international trade) and low-skill immigration to the dynamics of the skill premium. These developments, in turn, endogenously determine the native workers' choice of skill acquisition (i.e., training) subject to a time-varying sunk tuition costs.

We estimate our model using the Bayesian approach that establishes the numerical values of structural parameters from the best fit of the model to the data. The key advantage of our general equilibrium approach is that it addresses some of the identification issues in reduced-form analysis discussed below. The data used to estimate the stochastic growth model are not detrended, which allows for the estimation of long-run elasticities while accommodating the study of short-to medium-term dynamics. We show that the estimated model fits the data well across several dimensions. In particular, we find that the flows of low-skill immigrants and the time-varying cost of offshoring predicted by our model as latent variables are close to their empirical counterparts, which are proxied by the number of individuals apprehended at the U.S./Mexico border, and a trade-weighted measure of bilateral international trade costs we construct from the ESCAP-World Bank bilateral database–which is a comprehensive measure that goes beyond transportation costs.[5]

Key findings. Our estimated model yields three important results about the impact of low-skill immigration on labor market dynamics, macroeconomic performance, and welfare.

First, the recent decline in the skill premium is largely explained by the slowdown in low-skill immigration after the Great Recession of 2007. The historical quantitative analysis shows that before the Great Recession, high-skill workers benefited from the decline in barriers to international trade and offshoring, which reduced the earnings prospects of middle-skill workers (e.g., those with manufacturing jobs), and from robust low-skill immigration, which reduced consumer prices in the service sector. These effects later faded after the Great Recession, as stricter immigration policy reduced low-skill immigration and the supply of low-skilled workers, reversing the decades-long rise in the skill premium.

Second, the decline in low-skill immigration after 2007 took time to create labor shortages and affect the skill premium. The Great Recession, the prolonged downturn in the housing market, and the slow recovery in employment demand created substantial slack in labor markets. The effects of successive years of slow low-skill immigration only became visible as the labor market tightened in the mid-2010s.[6] During the COVID-19 pandemic, these acute labor shortages led to a spike in low-skilled wages in response to the CARES Act stimulus package, which sharply increased consumer prices, especially in the service sector, and undermined the effectiveness of the policy stimulus.

[5]The dataset factors in not only transportation costs and tariffs, but also language and communication barriers, among others. Data is available at: https://www.unescap.org/resources/escap-world-bank-trade-cost-database

[6]Consistent with our results, the unemployment to job openings ratio peaked at 6.9 in 2009 and declined very slowly to 1.1 in 2016 (U.S. Bureau of Labor Statistics).

Third, the decline in low-skill immigration since the mid-2000s reduced the welfare of the representative U.S. household—even under the extreme assumption of perfect substitution between equally skilled immigrants and natives.[7] Although reduced low-skilled immigration raises wages for low-skilled native workers, it also leads to more expensive non-tradable services that reduce purchasing power. Moreover, one important effect of reduced low-skill immigration is that the lower skill premium discourages investment in skill acquisition by native-born workers, which ultimately hampers labor productivity and reduces aggregate income and welfare.

We also examine the substantial surge in low-skill immigration that occurred in 2022-2024 and show that it helped partially, but not fully, to alleviate the labor shortage of low-skilled workers and mitigate the welfare losses that have accrued since 2007.

Related Literature. To the best of our knowledge, we are the first paper to: (a) study the empirical link between the fall in low-skill immigration, labor shortages, the unprecedented decline in the skill premium, and the associated impact on training choices of natives; and (b) assess the welfare implications through the lens of a structural model designed to account for this evidence. Hanson et al. (2017) and Mandelman and Zlate (2016) presented early evidence of the post-Great Recession immigration slowdown studied here. Our general result that immigration plays an important role in driving the employment and wage dynamics in the U.S. labor market is consistent with several studies. Borjas et al. (2012), and Friedberg and Hunt (1995) document a negative impact of migration on low-skill native employment and wages. Cortés (2008) finds that the inflow of immigrants into the United States lowers the price of services provided by low-skill workers. In turn, Autor and Dorn (2013) focus their analysis on employment at the left tail of the skill distribution, showing that employment growth in low-skill occupations through the mid-2000s was accounted for by the emergence of (non-tradable) service occupations. Burstein et al. (2020) highlight that the labor market adjustment to immigration differs between tradable and non-tradable occupations. Hunt (2017) and Jackson (2015) show empirically that a higher prevalence of low-skill immigration is associated with higher educational attainment among natives.

Our modeling of offshoring is based on the framework with *trade in tasks* developed by Grossman and Rossi-Hansberg (2008), which we extend to include a continuum of tasks performed by heterogeneous workers in a dynamic general equilibrium setting similar to Mandelman (2016).[8] Mandelman and Zlate (2022) assess the role of automation and offshoring using a model of trade in tasks. Unlike us, they focus on the pattern of labor market polarization prior to the Great Recession, which precedes the severe labor shortages that are the main focus of our analysis.[9]

[7] Several restrictive immigration policies have been enacted during this period. Major legislation passed by Congress includes: the Homeland Security Act of 2002, the Enhanced Border Security and Visa Entry Reform Act of 2002, the REAL ID Act of 2005, the Secure Fence Act of 2006, and the Jaime Zapata Border Enforcement Security Task Force Act of 2012.

[8] The modeling of worker heterogeneity across skills resembles the framework with firm heterogeneity across productivity levels proposed in Ghironi and Melitz (2005), which is also used to model offshoring through vertical FDI in Zlate (2016).

[9] Other important differences from the aforementioned study are the following. First, the previous paper used a two-country model in which labor immigration was exogenous. In the current paper, migration decisions are derived from the optimization problem of households in a third country (South) in response to changes to immigration policy and wage differentials. Second, the model in the above-mentioned study was deterministic and limited to the analysis of long-run transition dynamics. In contrast, the stochastic growth model in our paper allows us to study short-run dynamics following transitory shocks, which we need to identify the structural parameters that drive the recent reversal in the skill premium.

Di Giovanni et al. (2015), Kennan (2013), Klein and Ventura (2007), Klein and Ventura (2009), Bound et al. (2017), and Piyapromdee (2021) develop general equilibrium models of international labor migration, finding welfare gains from lower immigration barriers. Monràs et al. (2020), Bazzi et al. (2021) and East et al. (2023) study the labor market implications of the undocumented status of immigrants.

We are also related to studies on immigration, trade and offshoring by Ottaviano et al. (2013), Caliendo et al. (2021) and Mehra and Kim (2023). Like us, Albert and Monràs (2022), Allen et al. (2024), Cadena and Kovak (2016), and Monràs (2020), identify migration shocks that drive low-skill wage dynamics. Finally, Burstein and Vogel (2017) and Cravino and Sotelo (2019) study the effect of international trade on skill premium, and Autor et al. (2013, 2016) study the adjustment of the U.S. labor market to the emergence of China as a major player in global trade.

The rest of the paper is organized as follows. Section 2 presents facts based on regional data. Section 3 develops the model. Section 4 presents the results of estimating the model. Section 5 evaluates the fit of the model and studies the propagation of various shocks to the economy, interpreting historical episodes within the period 1983-2019. Section 6 discusses the welfare and implications of the post-pandemic immigration surge. Section 7 concludes.

2 Regional Evidence

In this section, we document empirical evidence based on regional data on how low-skill immigration shapes the U.S. labor market. The main dataset we use is the American Community Survey (ACS) data, which is administered annually by the U.S. Census Bureau, and covers more than 3.5 million households. The ACS provides individual information—such as place of birth, age, wage income, education, and occupation—of households across the country. We establish four empirical facts concerning the relationships between low-skill immigration, skill premium, immigration policy enforcement, and native educational attainment on three alternative regions concepts: state, metropolitan area, and commuting zone.[10] While these relationships have been documented in the literature, our analysis confirms that they hold in the updated sample (2021) of a unified comprehensive dataset, which is the largest available household survey (ACS). The reduced-form regional analysis presented below is not intended to settle causal inference, which would require deeper identification analysis. Our goal is to uncover a set of representative statistical relationships from such a unified dataset that will motivate key elements of our structural model specification in Section 3, which is the main contribution of this paper.

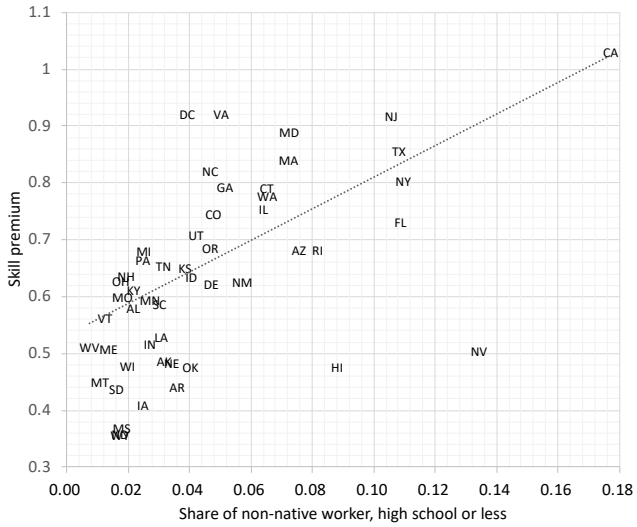
Third, our new study features endogenous training for the presence of a time-varying, educational sunk cost that we estimate using data on real tuition costs data. Fourth, we perform welfare analysis that is absent from the previous work.

[10]The metropolitan area is a conventional definition of a region in the literature on immigration (Card, 2009). Furthermore, commuting zone is a better-defined labor market than the previous two concepts because workers can easily commute within them, as shown in Autor et al. (2013).

2.1 Fact I: The Skill Premium is Positively Correlated with Low-skill Immigration

We measure region i's skill premium in year t as $sp_{i,t} = w^C_{i,t}/w^{HS}_{i,t} - 1$, where $w^C_{i,t}$ and $w^{HS}_{i,t}$ are the average hourly wage for *native* working-age full-time workers with college and high school or less, respectively.[11] We measure low-skill immigrant density, $im_{i,t}$, as the share of foreign-born low-skill (high school or less) population among the working-age population.

Figure 3: Skill premium is positively correlated with low-skill immigration



Note: This figure shows the scatter plot between skill premium and low-skill immigrant density in 2021 across states. Skill premium is the percentage hourly wage differential between native college graduates and high school or less for working-age full-time workers in the U.S. Low-skill immigrant density is the share of non-native low-skill (high school or less) population among the working-age population.

Figure 3 displays the scatter plot between skill premium ($sp_{i,t}$, y-axes) and the density of low-skill immigrants ($im_{i,t}$, x-axes) across the U.S. states in 2021, the ending year of our sample. States with a high density of low-skill immigrants, such as California and New Jersey, are associated with a high skill premium. More rigorously, we estimate the following cross-sectional regression using data in 2021 for three different definitions of region, respectively: $sp_{i,2021} = \alpha + \beta \text{Inflow}_{i,t0-2021} + \epsilon_i$, where $sp_{i,2021}$ is the skill premium in 2021, and $\text{Inflow}_{i,t0-2021}$ is the inflow rate of low-skill immigrants into region i between t0 and 2021.[12] We set $t0$ as 1970 for the state level and 2005 for the metropolitan area and commuting zone levels, respectively, which are the earliest periods when we can observe immigrants' years of immigration from most regions. For consistency, we use the inflow instead of the level of immigration as the independent variable, as we will instrument it with the inflow of immigration (predicted by the predetermined distribution of immigrants across regions).[13]

[11] We focus on the skill premium of native workers because it directly impacts their educational attainment decisions, which is a key focus of our paper. Our results remain robust when using the overall skill premium.

[12] Specifically, $\text{Inflow}_{i,t0-2021}$ is the ratio of the number of inflow of low-skill immigrants into region i between t0 and 2021 to the region's population in 2021.

[13] As shown in Table 1 of the Appendix, our results remain robust when using the level of immigration normalized by the region's population as the independent variable, instrumented with the predetermined inflow of immigration. This alternative specification would be consistent with Card (2009) who regresses dropout/high school graduate wage gap on the

Specifically, a classic endogeneity issue pointed out by Card (2001, 2009) is that skill premium and immigration are both affected by demand factors that are missing in our regression, making the OLS estimates potentially biased. Intuitively, higher demand for low-skill workers would increase low-skill immigration and decrease the skill premia. Hence, we follow Card (2009) by constructing the instrumental variable, $\widehat{\text{Inflow}}_{i,t0-2021}$, based on the historical distribution of immigrants from the same source country across regions. The intuition for the instrumental variable is that immigrants tend to settle in country-specific enclaves. For example, Mexican immigrants tend to cluster in Los Angeles and Chicago. The clustering entails a predetermined distribution of immigrants across regions that is uncorrelated with regional labor demand shock. Specifically, we define the instrument as $\widehat{\text{Inflow}}_{i,t0-2021} = \sum_k \lambda^k_{i,t0} \text{Inflow}^k_{t0-2021}$, where k indicates the source country of immigrants. The parameter $\lambda^k_{i,t0}$ is the fraction of low-skill immigrants from source country k who lived in region i in $t0$. The variable $\text{Inflow}^k_{t0-2021}$ is the inflow of low-skill immigrants from the source country k to *all* the United States between $t0$ and 2021. A natural prediction of the change in the number of low-skill immigrants from source country k between $t0$ and 2021 in region i is $\lambda^k_{i,t0}$ multiplied by $\text{Inflow}^k_{t0-2021}$. The key idea of the instrumental variable is that neither $\text{Inflow}^k_{t0-2021}$ or $\lambda^k_{i,t0}$ are endogenously determined by changes in labor demand specific to any of the regions between $t0$ and 2021. Columns (4)-(6) of Table 1 verify that $\widehat{\text{Inflow}}_{i,t0-2021}$ positively predicts low-skill immigrant density in 2021.

Table 1 shows the estimation results at the state, metropolitan area, and commuting zone levels, respectively, using $\widehat{\text{Inflow}}_{i,t0-2021}$ as an instrumental variable. Consistent with our hypothesis, the inflow of low-skill immigrants is positively correlated with skill premia, and the estimates are statistically significant. For example, the point estimates shown in Column (1) indicate that a state's skill premium would be 3.43 percentage points higher in a state with one percentage higher inflow rate of low-skill immigrants. Our findings are broadly consistent with the literature (e.g., Borjas et al. (2012), and Monràs (2020)) that find an adverse effect of low-skilled immigration on low-skilled wages, thus widening the skill premium. However, the size of the effect depends on the substitution elasticity between natives and immigrants.[14]

2.2 Fact II: Low-skill Immigration Positively Correlates with Initial Low-skill Wages

Next, we show that low-skill immigration is positively correlated with the initial level of low-skill wages. Intuitively, a region that experiences a positive shock to the demand for low-skill labor would see an increase in low-skill wages, and this would attract more immigrants of this skill group to live in the region. It is difficult to pin down these demand innovations with this reduced-form analysis, however. To address this identification issue, we resort to Autor et al. (2013), who show that the rapid surge in imports from China had an outsized impact on import-competing industries that were disproportionately located across the country. The regions most affected by China's trade shock, therefore, resulted in heterogeneous labor market outcomes for workers of different skills across regions. Column (2) of Table 2 verifies that higher import exposure to China's trade shock between 2000 and 2007 is associated with a lower low-skill wage in 2010 across commuting zones. Note that the sample size is smaller than in Columns (3)

ratio of labor hours by dropouts and high school graduates, instrumented with the predetermined inflow of immigrants.

[14]For instance, Ottaviano and Peri (2012) estimates a very low substitution elasticity. The dispersion in our estimation across alternative notions of the region may explain the different elasticity values.

Table 1: The skill premium is positively correlated with low-skill immigration

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Skill premium			Inflow of low-skill immigrants		
Definition of region	State	MET	CZ	State	MET	CZ
Inflow of low-skill immigrants	3.43**	3.53***	3.98***			
	(1.31)	(0.98)	(1.39)			
Predetermined immigration flow				0.44***	0.81***	0.25***
				(0.08)	(0.03)	(0.02)
Constant	0.44***	0.50***	0.37***	0.03***	0.01***	0.01***
	(0.07)	(0.02)	(0.02)	(0.002)	(0.001)	(0.001)
Adj R-squared	0.12	0.03	0.12	0.37	0.73	0.21
F stat	6.87	12.90	8.21	29.97	716.29	201.71
Observations	51	260	741	51	260	741

Note: The dependent variable is the skill premium in 2021. The independent variable is the low-skill immigrant density in 2021. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. MET is metro area and CZ is commuting zone.

and (6) of Table 1 because certain commuting zones lack trade shock measures. With this instrumental variable, we estimate the following cross-sectional regression: $\text{Inflow}_{i,2010-2021} = \alpha + \beta ln\left(w^{HS}_{i,2010}\right) + \epsilon_i$, where $\text{Inflow}_{i,2010-2021}$ is the inflow rate of low-skill immigrants into commuting zone i between 2010 and 2021.[15] The variable $w^{HS}_{i,2010}$ is the low-skill wage of commuting zone i in 2010, which is correlated with our instrumental variable given the persistent effect of trade shocks (Autor et al., 2016). Column (1) of Table 2 shows the estimation results. Namely, a higher low-skill wage is associated with larger inflows of low-skill immigrants.

Our findings are consistent with the literature that documents an active response of immigrants' immigration decisions to local labor market conditions, e.g., Cadena and Kovak (2016), Hauser and Seneca (2022), Albert and Monràs (2022), who all considered earlier sample periods than us.[16]

Table 2: Low-skill immigration inflow is positively correlated with low-skill wage

	(1)	(2)
Dependent variable	Inflow of low-skill immigrants	Low-skill wage
Low-skill wage	0.24**	
	(0.12)	
Change in import exposure		-0.002**
		(0.01)
Constant	-0.66***	2.74***
	(0.32)	(0.004)
Adj R-squared	0.04	0.01
F stat	4.51	4.06
Observations	660	660

Note: The dependent variables are the inflow rate of low-skill immigrants between 2010 and 2021. The independent variable is the log low-skill wage in 2010. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

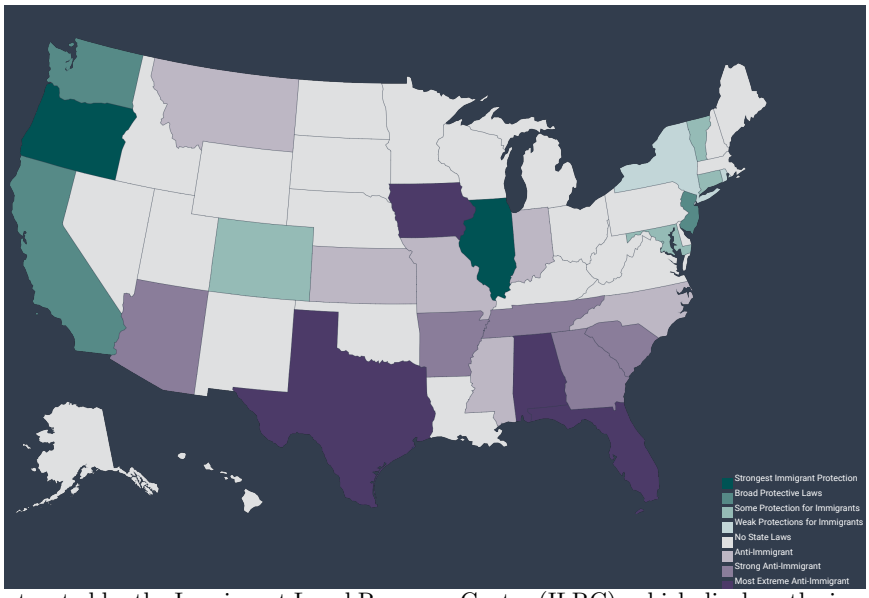
[15] Specifically, $\text{Inflow}_{i,2010-2021}$ is the ratio of the number of inflow of low-skill immigrants into commuting zone i between 2010 and 2021 to the commuting zone's population in 2021. We focus on commuting zones since the instrumental variable data is unavailable for state and metropolitan areas.

[16] Albert and Monràs (2022) find that low-skill immigrants prefer to move to expensive cities where salaries are higher so they can potentially remit more to their countries of origin.

2.3 Fact III: Immigration Policy Enforcement is Correlated with Low-skill Immigration

As we will show in section 4, the U.S. federal government has significantly increased immigration enforcement since 2007. The Immigrant Legal Resource Center (ILRC) measures how individual states contribute to immigration enforcement by examining recent state and local laws and legislation. ILRC constructs immigration enforcement measures that fall into four general categories: information and resource sharing with Immigration and Customs Enforcement (ICE), jail-to-ICE transfers, patrol officer cooperation with ICE, and contracts with ICE or U.S. Customs and Border Protection. These separate measures are aggregated into a composite index ranging from one to five, with a higher score indicating a more protective immigration policy. Figure 4 shows the scores for the U.S. states, excluding the District of Columbia which does not have a score. While we do not argue that the preference for enforcement levels is exogenous, this evidence suggests that it is not entirely endogenous to past immigration inflows. Policy preferences do seem to matter; California and Texas are border states and the two largest recipients of immigrants, but they have very different policy choices on this issue. Iowa is not only far from the border but also has a negligible low-skill foreign-born population, yet it has one of the most restrictive stances on immigration enforcement.

Figure 4: Immigration enforcement across U.S. states



Note: This map is constructed by the Immigrant Legal Resource Center (ILRC), which displays the immigration enforcement measures across U.S. states. A higher score (green-colored) indicates a more protective immigration policy. A lower score (purple-colored) indicates a stricter anti-sanctuary policy.

We estimate the following cross-sectional regression using data in 2021: $im_{i,2021} = \alpha + \beta ss_i + \epsilon_i$, where $im_{i,2021}$ is the low-skill immigrant density, and ss_i is the score of sanctuary policies constructed by ILRC. The estimation results are reported in Column (1) of Table 3: low-skill immigrants reside in states with more protective immigration policies. In turn, Column (2) reports the result when the dependent variable is the skill premium. The result suggests that sanctuary policies are associated with a higher skill

premium. An explanation is that sanctuary policies raise skill premia by increasing low-skill immigrant density (consistent with Fact I). Similarly, Allen et al. (2024) identifies various geographic extensions of the U.S.-Mexico border wall between 2007 and 2010 to show that they led to a decline in migration flows. Lessem (2018) finds similar results looking at individual migration decisions of Mexican Households.

Table 3: Sanctuary policy is positively correlated with low-skill immigration

	(1)	(2)
Definition of region	State (excluding the District of Columbia)	
Dependent variable	Inflow of low-skill immigrants	Skill premium
Sanctuary policy	0.02**	0.11***
	(0.01)	(0.04)
Constant	-0.01	0.34***
	(0.03)	(0.11)
Adj R-squared	0.08	0.13
Observations	50	50

Note: The dependent variables are low-skill immigrant density and skill premium in 2021 for Columns (1) and (2), respectively. The independent variable is the score of sanctuary policies constructed by ILRC. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

2.4 Fact IV: Natives' Educational Attainment is Positively Correlated with low-skill Immigration

Lastly, we examine the relation between low-skill immigration and natives' educational attainment. To this aim, we focus on the skill acquisition of the cohorts of individuals who reside in the same state where they were born and are between the ages of 22 and 26 years old in 2021. We denote with $s^C_{i,2021}$ the share of these people who had been to college in the year 2021. Our hypothesis is that the decision to pursue post-secondary education in the preceding five years is associated with the payoffs to education at the time of the choice (i.e., the skill premium in 2016), which is affected by low-skill immigration. Two caveats to our assumption. First, we focus on states only since the information for the place of birth is unavailable for metropolitan and commuting-zone areas. Second, since people with different educational attainment can move freely across regions to maximize their expected income, we remove from the sample the workers not living in the same state in which they were born.

We first estimate the following regression: $s^C_{i,2021} = \alpha + \beta im_{i,2016} + \epsilon_i$, where $s^C_{i,2021}$ is the share of local young people who had acquired post-secondary education by 2021. $im_{i,2016}$ is low-skill immigrant density. A potential concern is that all these variables are affected by demand factors that are missing in our regressions. Intuitively, higher demand for low-skill workers would increase low-skill immigration and decrease the skill premia, while depressing the demand for post-secondary education. As in Fact I, we instrument low-skilled immigration with the predetermined low-skill immigrant inflow.

Column (1) of Table 4 shows that a higher density of low-skill immigrants is associated with a higher acquisition of post-secondary education by local young people. A one-percentage point rise in immigrant inflows is associated with a percentage point higher share of college attainment, which is economically significant. Our preferred interpretation is that a pick-up in low-skill immigration leads to a higher skill premium, which encourages local young people to acquire post-secondary education. Indeed, Column

(3) of Table 4 shows that higher skill premia are associated with increased acquisition of post-secondary education by local young people, whereas Column (2) indicates that the skill premium is positively correlated with low-skill immigration in 2016, consistent with the finding established in Fact 1 for 2021.

Note that our estimates are based on individuals who reside in the same state where they were born. A potential selection issue arises because the inflow of low-skill immigration may prompt young local people, who choose not to pursue higher education, to move to other states. This could lead to an upward bias in Column (1).[17] To gauge the degree of this selection bias induced by the outflow of low-education young people, we construct a new variable $outflow^{local}_{i,2021}$, representing the share of young people born in state i who resided in different states and have not pursued higher education in 2021 out of all young people born in state i. We then regress $outflow^{local}_{i,2021}$ on $im_{i,2016}$, instrumented with the predetermined low-skilled immigrant inflow. Column (4) of Table 4 reports the estimation result: low-skill immigration is not significantly associated with the outflow of low-education young people. This insignificant result might be driven by the high moving costs faced by these individuals, who are presumably from lower-income families and experience higher information asymmetry. These findings suggest that the selection bias issue is likely to be negligible.

In sum, our findings are consistent with the literature. For example, using a similar approach to us (with different data sets and sample periods), Hunt (2017) and McHenry (2015) document a positive effect of immigration on natives' educational attainment. Llull (2018) uses a discrete choice model to estimate the heterogeneous (and overall positive) effect of immigration on natives' educational attainment.[18]

Table 4: Educational attainment is positively correlated with low-skill immigration and skill premium

	(1)	(2)	(3)	(4)
Dependent variable	Educ attainment	Skill premium	Educ attainment	Local outflow
Inflow of low-skill immigrants	1.00***	2.01***		0.05
	(0.35)	(0.48)		(0.08)
Skill premium			0.20**	
			(0.10)	
Constant	0.50***	0.46***	0.25***	0.01***
	(0.02)	(0.03)	(0.05)	(0.004)
Adj R-squared	0.03	0.25	0.15	-0.01
F stat	8.23	17.82	4.30	0.40
Observations	51	51	51	51

Note: The dependent variables are local young people's educational attainment in 2021, skill premium in 2016, and the outflow of local young people with low-education. The independent variables are skill premium and low-skill immigrant density in 2016. The instrumental variable is the predetermined low-skill immigrant inflow between 1970 and 2016. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

[17] For example, consider a group of 100 local young people. Without the inflow of low-skill immigration, 50 would pursue high education, resulting in an educational attainment of *Educ attainment*=50/100=0.5. With the inflow of low-skill immigration, 60 would pursue high education, raising the educational attainment to *Educ attainment*=60/100=0.6. However, if 20 of those who do not pursue higher education move to another state, our measured educational attainment rate would be *Educ attainment*=60/(100-20)=0.75, which is higher than the actual educational attainment rate of 0.6.

[18] Charles et al. (2018) use cross-city evidence to show that college enrollment responds strongly to labor market opportunities associated with booms and busts in the local housing market.

2.5 Taking Stock

Fact I suggests the potentially important role of low-skill immigration in determining the skill premium. Fact II reveals the endogenous nature of low-skill immigration driven by economic incentives. Fact III implies that low-skill immigration is strongly influenced by immigration costs, such as penalties from immigration enforcement. Fact IV shows that the skill premium is a critical determinant of natives' investment in education. Taken together, Facts I and IV suggest that low-skill immigration may have a key effect on skill upgrading. This section also illustrates the difficulty of making causal inferences. Households' optimization problem determines supply factors such as training, immigration, and trade that depend on, but also jointly affect, relative wages. To shed light on these pervasive identification issues, we take a structural general equilibrium approach that considers all these factors together. We develop a stochastic growth model and estimate it with non-detrended high-frequency aggregate time series data, which allows for the estimation of long-run elasticities while allowing for the study of short-to-medium term dynamics in the historical analysis. The structural setting also allows for a welfare analysis of counterfactual migration policies.

3 Model

Our model consists of two large economies (Home and Foreign), and a third small economy (South) that neighbors Home and is the source of low-skill immigrants. We mainly focus the discussion on the Home and South economies.[19] For Foreign, the equations are the same as those for Home, and the variables are marked with an asterisk.[20] The full derivation of the model is in the Appendix.[21]

3.1 Production in the Home Economy

The Home economy comprises two sectors. The first sector produces services, which are non-tradable by definition and require native and immigrant low-skilled labor. This service sector represents the service occupations that require either close contact with the final consumer or need to be executed where the final service is delivered (e.g., childcare or cleaning). The second sector produces a country-specific final good, which is obtained from the aggregation of a continuum of diverse labor tasks. These labor tasks can be either executed at Home or offshored to Foreign. Workers in this sector are heterogeneous in skill, which they acquire after undergoing training. In short, we will refer to this sector as the "tradable" sector. Notice, however, that the meaning of tradability is different from that typically found in the literature. Here, the labor tasks required to produce final goods are internationally traded, not the final goods themselves. It is also important to emphasize that some of these "tradable" tasks may be endogenously non-tradable if the costs of offshoring are high enough because of transportation, language, or even legal

[19] The Appendix describes the system of equations that characterize all the equilibrium conditions of the model as well as the auxiliary equations needed to take the model to the data.

[20] The model is symmetric for Home and Foreign, with the only exception being that Home receives immigrant low-skill labor from the South, whereas Foreign does not. See the Appendix for details.

[21] Since we focus on the labor market outcomes from offshoring and immigration, we abstract from capital, hence labor is the only factor of production.

barriers (e.g., financial services, project quality control, etc.).

Non-tradable sector. The first sector produces services that are non-tradable by definition. The labor input used in production, $L^A_{N,t}$, is a CES composite of aggregate units of low-skilled (untrained) native labor, $L_{N,t}$, and immigrant labor, $L^s_{\mathbf{i},t}$:

$$L^A_{N,t} = \left[\alpha \left(L_{N,t}\right)^{\frac{\sigma_N - 1}{\sigma_N}} + (1-\alpha)\left(L^s_{\mathbf{i},t}\right)^{\frac{\sigma_N - 1}{\sigma_N}}\right]^{\frac{\sigma_N}{\sigma_N - 1}}.$$

Output is a linear function of the labor input: $Y_{N,t} = \mathbb{X}_t L^A_{N,t}$. $\mathbb{X}_t$ is a stochastic, permanent world technology shock that affects all productive sectors in all countries. This global shock displays a unit-root which warrants a balanced-growth path for the economy. The price of this service good is $P_{N,t}$. The profit maximization problem implies the following wage equations for low-skill native and immigrant labor: $w_{\mathbf{u},t} = P_{N,t}\mathbb{X}_t \alpha \left(L^A_{N,t}/L_{N,t}\right)^{1/\sigma_N}$ and $w^s_{\mathbf{i},t} = P_{N,t}\mathbb{X}_t(1-\alpha)\left(L^A_{N,t}/L^s_{\mathbf{i},t}\right)^{1/\sigma_N}$.

Tradable sector. The tradable sector employs a continuum of skilled workers executing different labor tasks. In order to obtain the skill required for employment in the tradable sector, households invest in training every period. The cost of training involves an irreversible sunk cost, as will be specified later, and results in an idiosyncratic productivity level $\mathbf{z}$ for each worker. Workers draw this idiosyncratic productivity from a common distribution $\mathcal{F}(\mathbf{z})$ over the support interval $[1, \infty)$ upon completion of training. The untrained (raw) labor provided by each worker is augmented by idiosyncratic productivity $\mathbf{z}$ gained from training and expressed in efficiency units as follows: $l_{\mathbf{z},t} = \mathbf{z} l_t$, where l_t indicates units of raw labor. Idiosyncratic productivity $\mathbf{z}$ remains fixed thereafter until an exogenous skill destruction shock makes the skill obtained from training obsolete, transforming the efficiency units back into units of raw labor. The skill destruction shock is independent of the workers' idiosyncratic productivity level, so $\mathcal{F}(\mathbf{z})$ characterizes the efficiency distribution for all trained native workers at any point in time. The household's training decision is described in more detail further below.

The efficiency units of labor benefit from two technological innovations when used in production. One is the world productivity shock, $\mathbb{X}_t$, and the other is a temporary country-specific technology shock, ε^Z_t. The country-specifici technology shocks and all shocks introduced hereafter evolve as an AR(1) process. As a result, each efficiency unit of labor supplied is transformed in a production task, $n_t(\mathbf{z})$, as follows:

$$n_t(\mathbf{z}) = (\mathbb{X}_t \varepsilon^Z_t) l_{\mathbf{z},t} = (\mathbb{X}_t \varepsilon^Z_t)\mathbf{z} l_t. \tag{1}$$

Trained workers obtain skills and are employed in a variety of occupations, and each of these occupations allows them to execute a given set of tasks ξ, defined over a continuum of tasks Ξ (i.e., $\xi \in \Xi$), whose labor supply is denoted by $n_t(\mathbf{z}, \xi)$. At any given time, only a subset of these tasks Ξ_t ($\Xi_t \subset \Xi$) may be demanded by firms in the global labor market and effectively used in production.[22] The labor input of the tradable sector is obtained by aggregating over a continuum of tasks $n_t(\mathbf{z}, \xi)$ that are imperfect

[22] Although workers cannot migrate between Home and Foreign, the labor market of tasks is global as the tasks can be off-shored. In addition, the subset of tasks demanded by foreign companies is $\Xi^*_t \subset \Xi$, and may differ from Ξ_t

substitutes: $\mathbb{N}_t = \left[\int_{\xi\epsilon\Xi_t} n_t(\mathbf{z},\xi)^{\frac{\theta-1}{\theta}} d\xi\right]^{\frac{\theta}{\theta-1}}$, where $\theta > 1$ is the elasticity of substitution across tasks. The wage bill is $\mathbb{W}_t = \left[\int_{\xi\epsilon\Xi_t} w_t(\mathbf{z},\xi)^{1-\theta} d\xi\right]^{\frac{1}{1-\theta}}$, where $w_t(\mathbf{z},\xi)$ is the wage paid to each efficiency unit of labor. Importantly, some of these tasks may be executed in Foreign, as described in more detail below. With labor as the only input in production, the final tradable good is $Y_{T,t} = \mathbb{N}_t$, and the price of this final good is $P_{T,t} = \mathbb{W}_t$. The price of this tradable good is the numeraire, $P_{T,t} = \mathbb{W}_t \equiv 1$.

Trade in tasks and the skill income premium. In a symmetric equilibrium, the wage paid to each worker in the tradable sector is skill-specific but is symmetric across tasks. That is, $w_t(\mathbf{z},\xi) = w_t(\mathbf{z},.)$ for every task $\xi \in \Xi$. The skill premium $\pi_{D,t}$ in the domestic tradable sector is defined as the difference between the income obtained from a task executed for this sector and the income obtained by a raw unit of labor in the non-tradable sector:

$$\pi_{D,t}(\mathbf{z},.) = w_{D,t}(\mathbf{z},.)n_{D,t}(\mathbf{z},.) - w_{\mathbf{u},t}l_t, \tag{2}$$

where $n_{D,t}(\mathbf{z},.)$ denotes the task produced by one efficiency unit of labor in the tradable sector for the home market, and $w_{D,t}(\mathbf{z},.)$ is the associated wage. Some of the tasks embedded in the Home final good are executed in Foreign and imported (i.e., they are offshored by Home to Foreign). Conversely, Foreign demands some of the tasks executed in Home. Tasks executed in Home and delivered to Foreign, $n_{X,t}(\mathbf{z},.)$, are paid $w_{X,t}(\mathbf{z},.)$ and are subject to an iceberg offshoring cost $\tau \geqslant 1$ and also to a period-by-period fixed offshoring cost f_o, which is defined in terms of efficiency units of labor.[23] Changes in offshoring costs are reflected in shocks ε_t^{τ} to the level of the iceberg cost τ, so that $\tau_t = \varepsilon_t^{\tau}\tau$. The skill premium, $\pi_{X,t}$, for executing a task for Foreign is:

$$\pi_{X,t}(\mathbf{z},.) = \left(\frac{w_{X,t}(\mathbf{z},.)}{\tau_t} n_{X,t}(\mathbf{z},.) - f_{o,t}\right) - w_{\mathbf{u},t}l_t. \tag{3}$$

All Home workers have their tasks sold domestically. However, due to the iceberg trade cost and the fixed offshoring cost, only the most efficient Home workers engage in multinational production (i.e., execute tasks for both Foreign and Home).[24] A worker will take part in global production as long as the idiosyncratic productivity level $\mathbf{z}$ is above a threshold $\mathbf{z}_{X,t} = \inf\{\mathbf{z} : \pi_{X,t}(\mathbf{z},.) > 0\}$, to whom we refer as high-skill workers. Conversely, home workers with productivity below $\mathbf{z}_{X,t}$, to whom we refer as middle-skill workers, execute tasks for the domestic market only. To illustrate this idea with an example, consider a scenario in which U.S. multinationals hire professionals from the Silicon Valley area to research, develop, and design a state-of-the-art medical device to be sold worldwide. Other tasks may be offshored overseas to exploit comparative advantage (e.g., Chinese programmers develop software, and Thai technicians produce flash storage). However, some of the less productive skilled tasks will continue to be produced in-house for the domestic market only, given transportation and communication costs (e.g., customized

[23] The modelling of offshoring costs closely resembles the characterization of trade costs in Ghironi and Melitz (2005). For consistency with the economy-wide balanced growth path, this fixed cost is augmented by the world technology shock, then expressed in units of the Home numeraire as follows: $f_{o,t} = \frac{w_{\mathbf{u},t}}{(\mathbb{X}_t \varepsilon_t^Z)}(\mathbb{X}_t f_o)$.

[24] See Krishna et al. (2014) for evidence supporting this result.

assembly for U.S. hospitals or help desk services). In this context, a decrease in the offshoring costs allows multinationals to incrementally assign more tasks to foreign workers (e.g. the help desk may reallocate to India when telecommunication costs decrease). Trade integration enhances cross-country task specialization while displacing less skilled workers, and it is consistent with the evidence.[25] Shocks to aggregate productivity, demand, the iceberg trade cost will also result in changes to the threshold level $\mathbf{z}_{X,t}$.[26]

To solve the model with heterogeneous workers, it is useful to define average productivity levels for two representative groups, as in Ghironi and Melitz (2005). First, the average productivity of all workers is: $\tilde{\mathbf{z}}_{D,t} \equiv \left[\int_{\mathbf{1}}^{\infty} \mathbf{z}^{\theta-1} d\mathcal{F}(\mathbf{z})\right]^{\frac{1}{\theta-1}}$. Second, the average efficiency of the workers whose tasks are traded globally is: $\tilde{\mathbf{z}}_{X,t} \equiv \left[\frac{1}{1-\mathcal{F}(\mathbf{z}_{x,t})} \int_{\mathbf{z}_{x,t}}^{\infty} \mathbf{z}^{\theta-1} d\mathcal{F}(\mathbf{z})\right]^{\frac{1}{\theta-1}}$. To derive the average wage and skill premium, it is easier to consider an isomorphic setup where a mass of workers $N_{D,t}$ with average productivity $\tilde{\mathbf{z}}_{D,t}$ execute tasks for the domestic market. Within this group, a mass of high-skilled workers $N_{X,t}$ with average productivity $\tilde{\mathbf{z}}_{X,t}$ accomplish tasks for the foreign market in addition to the domestic market. The wages for each skill group are $\tilde{w}_{D,t} = w_{D,t}(\tilde{\mathbf{z}}_{D,t}, .)$ and $\tilde{w}_{X,t} = w_{X,t}(\tilde{\mathbf{z}}_{X,t}, .)$. Combining all these, the wage bill of the home tradable sector can be re-written as: $\mathbb{W}_t = \left[N_{D,t}\left(\tilde{w}_{D,t}\right)^{1-\theta} + N^*_{X,t}\left(\tilde{w}^*_{X,t}\right)^{1-\theta}\right]^{\frac{1}{1-\theta}}$, where $N^*_{X,t}$ denotes foreign workers executing tasks imported by Home, and $\tilde{w}^*_{X,t}$ is the corresponding wage expressed in units of the Home numeraire. The skill premia for each group are $\tilde{\pi}_{D,t} = \pi_{D,t}(\tilde{\mathbf{z}}_{D,t}, .)$ and $\tilde{\pi}_{X,t} = \pi_{X,t}(\tilde{\mathbf{z}}_{X,t}, .)$. Finally, the average skill premium is defined as: $\pi_t = (N_{D,t}\tilde{\pi}_{D,t} + N_{X,t}\tilde{\pi}_{X,t})/N_{D,t}$. By assumption, low-skilled labor is only used in non-tradable services.

3.2 Households in the Home Economy

Household members form an extended family and pool their labor income – obtained from working in the tradable and non-tradable sectors – and choose aggregate variables to maximize expected lifetime utility. As in Andolfatto (1996), the model assumes that household members perfectly insure each other against fluctuations in labor income resulting from changes in their employment status. This assumption eliminates any type of ex-post heterogeneity across workers at the household level.

Consumption. Household's consumption basket is:

$$C_t = \left[(\gamma_c)^{\frac{1}{\rho_c}} (C_{T,t})^{\frac{\rho_c-1}{\rho_c}} + (1-\gamma_c)^{\frac{1}{\rho_c}} (C_{N,t})^{\frac{\rho_c-1}{\rho_c}}\right]^{\frac{\rho_c}{\rho_c-1}},$$

[25] Inequality deepens in countries that lower their barriers to trade, irrespective of their degree of economic development. This implication contrasts with that of the traditional Heckscher-Ohlin/Stolper-Samuelson paradigm, which predicts a decrease in the skill premium in countries with abundant low-skilled labor. See Burstein and Vogel (2017) and Goldberg and Pavcnik (2007) for a related discussion.

[26] It is feasible to rationalize a scenario with two countries at different stages of economic development in this context. For instance, the distribution of idiosyncratic productivity in Home may stochastically dominate the one characterizing Foreign –i.e. $\mathcal{F}(\mathbf{z}) > \mathcal{F}^*(\mathbf{z})$. Therefore, workers at the top of the skill distribution in Foreign may have the same productivity as some of the workers in the middle of the skill distribution in Home. Notice, however, that same productivity across countries does not imply same wages in equilibrium. Consistent with the Balassa-Samuelson hypothesis, countries with higher average productivity in the tradable sector pay higher wages to low productivity workers in a sector that is either non-tradable or subject to trade costs. These wage differentials foster the offshoring of tasks despite low productivity in Foreign.

which includes amounts of the tradable good $C_{T,t}$ and the non-tradable services $C_{N,t}$. The consumer price index is: $P_t = \left[\gamma_c + (1-\gamma_c)\left(P_{N,t}\right)^{1-\rho_c}\right]$. The final good produced in the tradable sector in Home, $Y_{T,t}$, is a composite of domestic and foreign tasks. It is entirely used for consumption by the Home household, $C_{T,t}$, and also by the Southern immigrant workers established in Home, $C^s_{T,t}$, so that $Y_{T,t} = C_{T,t} + C^s_{T,t}$. The problem of the Southern household is described in Section 3.3.

Household problem. The household has standard additive separable utility over real consumption, C_t, and leisure, $1 - L_t$, where L_t is the aggregate supply of raw labor. By choosing consumption, labor, training, and bond holdings, the Home representative household maximizes a standard utility kernel, which is modified to be consistent with the balanced growth-path:

$$\mathbb{E}_t \sum_{s=t}^{\infty} \beta^{s-t} \varepsilon^b_t \left[\frac{1}{1-\gamma} C_t^{1-\gamma} - a_n \mathbb{X}_t^{1-\gamma} \frac{L_t^{1+\gamma_n}}{1+\gamma_n} \right], \tag{4}$$

where parameter $\beta \in (0,1)$ is the subjective discount factor, $\gamma > 0$ is the inter-temporal elasticity of substitution, $\gamma_n > 0$ is the inverse of the Frisch elasticity of labor supply, and $a_n > 0$ is the weight on the disutility from labor. Also, ε^b_t is a shock to the intertemporal rate of substitution, which may be interpreted as a consumption demand shock.

The period budget constraint expressed in units of the numeraire good is:

$$w_{\mathbf{u},t} L_t + \pi_t N_{D,t} = f_{j,t} N_{E,t} + P_t C_t + q_t B_t - B_{t-1} + \Phi(B_t). \tag{5}$$

Total income is captured by the two terms on the left-hand side. The first term, $w_{\mathbf{u},t} L_t$, captures the remuneration of all raw units of labor, which includes the income of low-skilled labor employed in the non-tradable service sector, as well as the potential income generated by the raw labor that undergoes training and works in the tradable sector. The second term captures the total skill income premium that results from training and selling tasks domestically, defined as the product between the total measure of skilled workers, $N_{D,t}$, and their average skill income premium, π_t as defined above.

On the right-hand side of (5), the first term represents the total investment in training, in which $N_{E,t}$ are the new skilled occupations created at time t, and $f_{j,t}$ is the sunk training cost required for each of these new skilled workers. Training costs are time-varying and are subject to cost-push shocks ε^{Tr}_t such that: $f_{j,t} = \left(\varepsilon^{Tr}_t f_j\right)^{\Theta_{f_j}}$, where Θ_{f_j} is defined as the elasticity of total training costs to observed tuition costs, which will be estimated below. Like offshoring costs, these costs also require a path consistent with balanced-growth.[27] International financial transactions are restricted to a one-period, risk free bond. The level of debt due every period is B_{t-1}, and the new debt contracted is B_t at price $q_t = 1/(1+r_t)$, with r_t representing the implicit interest rate. To induce model stationarity in balanced-growth, we introduce an arbitrarily small cost of debt, $\Phi(B_t) = \mathbb{X}_t \frac{\phi}{2} \left(\frac{B_t}{\mathbb{X}_t}\right)^2$.

[27] This sunk cost is expressed in units of the numeraire good as: $\tilde{f_{j,t}} = \frac{w_{\mathbf{u},t}}{(\mathbb{X}_t \varepsilon^Z_t)} (\mathbb{X}_t f_{j,t})$.

Measure of skilled workers. The newly-created skilled workers $N_{E,t}$ join the already-existing $N_{D,t}$ with a time-to-build lag, and together are subject to a skill destruction shock δ, that renders the skills obtained from training obsolete. Therefore, the resulting law of motion for the skilled occupations is: $N_{D,t} = (1-\delta)(N_{D,t-1} + N_{E,t-1})$, where we assume that new skilled workers start working at $t+1$.

Optimality conditions. The household maximizes utility subject to its budget constraint and the law of motion for skilled workers described above. The optimality conditions for labor effort and consumption/saving are conventional:

$$\hat{a}_n \left(L_t\right)^{\gamma_n} \left(C_t\right)^{\gamma} = \frac{w_{\mathbf{u},t}}{P_t}, \tag{6}$$

$$q_t = \beta E_t \left\{ \frac{\zeta_{t+1}}{\zeta_t} \right\} - \Phi'(B_t), \tag{7}$$

where $\hat{a}_n = a_n \mathbb{X}_t^{1-\gamma}$, and $\zeta_t = \varepsilon_t^b (C_t)^{-\gamma}/P_t$ characterizes the marginal utility of consumption. The optimality governing the choice of bonds for foreign households in conjunction with the Euler equation in (7) yields the following risk-sharing condition:

$$E_t \left\{ \frac{\zeta_{t+1}^*}{\zeta_t^*} \frac{\mathbb{Q}_t}{\mathbb{Q}_{t+1}} - \frac{\zeta_{t+1}}{\zeta_t} \right\} = -\frac{\Phi'(B_t)}{\beta}, \tag{8}$$

where $\mathbb{Q}_t$ is the factor-based real exchange rate (or terms of labor).[28] Finally, the optimality condition for training is pinned down by the following condition:

$$f_{j,t} = \mathbb{E}_t \sum_{s=t+1}^{\infty} \left[\beta\left(1-\delta\right)\right]^{s-t} \left(\frac{\zeta_s}{\zeta_t}\right) \tilde{\pi}_s, \tag{9}$$

This equation shows that in equilibrium the sunk training cost $f_{j,t}$ equals the present discounted value of the future skill premia resulting from the creation of new skilled occupations $\{\tilde{\pi}_s\}_{s=t+1}^{\infty}$ adjusted for the probability of skill destruction δ.

3.3 South Economy

The representative household in South provides raw labor without the possibility of training. This labor can either be employed in domestic production or emigrate to Home after incurring a sunk migration cost. The household members pool their total income, which is obtained from both domestic and emigrant labor, and choose aggregate variables to maximize lifetime utility. In the Appendix, we provide empirical evidence to justify the assumption of a unified household unit that includes immigrants and left-behind family members as well as their associated consumption choices.

[28] That is, $\mathbb{Q}_t = \frac{\varepsilon \mathbb{W}_t^*}{\mathbb{W}_t}$. Thus, the real exchange rate is expressed in units of the foreign numeraire per units of the home one, where ε is the nominal exchange rate.

Labor migration. The representative household supplies a total of $L^s_{\mathbf{u},t}$ units of raw labor every period. A portion of the household members $L^s_{\mathbf{i},t}$ resides and work as low-skill immigrant workers abroad (in Home). The remaining $L^s_{\mathbf{u},t} - L^s_{\mathbf{i},t}$ work in the country of origin (in South). The calibration ensures that the low-skill wage in Home is higher than the wage in South, so that the incentive to emigrate from South to Home exists every period. However, a fraction of total labor supply always remains in South ($0 < L^s_{\mathbf{i},t} < L^s_{\mathbf{u},t}$). The macroeconomic shocks are small enough for these conditions to hold in every period.

The household sends an amount $L^s_{\mathbf{e},t}$ of new emigrant labor to Home every period, where the stock of immigrant labor $L^s_{\mathbf{i},t}$ is built gradually over time. The time-to-build assumption in place implies that the new immigrants start working one period after arriving. They continue to work in all subsequent periods until a return-inducing exogenous shock, which hits with probability δ_l every period, forces them to return to South. This shock reflects issues such as termination of employment in the destination economy, deportation, or voluntary return to the country of origin, etc.[29] The resulting rule of motion for the stock of immigrant labor in Home is: $L^s_{\mathbf{i},t} = (1-\delta_l)(L^s_{\mathbf{i},t-1} + L^s_{\mathbf{e},t-1})$.

Household's decision problem. By choosing consumption, labor, and immigration, the South representative household maximizes lifetime utility:

$$\mathbb{E}_t \sum_{s=t}^{\infty} \beta^{s-t} \left[\frac{1}{1-\gamma} (C^s_t)^{1-\gamma} - a^s_n \mathbb{X}_t^{1-\gamma} \frac{(L^s_{\mathbf{u},t})^{1+\gamma_n}}{1+\gamma_n} \right], \tag{10}$$

subject to the law of motion for immigrant labor and the budget constraint:

$$w^s_{\mathbf{i},t} L^s_{\mathbf{i},t} + w^s_{\mathbf{u},t} \left(L^s_{\mathbf{u},t} - L^s_{\mathbf{i},t} \right) \geqslant f_{e,t} L^s_{\mathbf{e},t} + P^s_t C^s_t, \tag{11}$$

where $w^s_{\mathbf{i},t}$ is the immigrant wage earned in Home (the same as the low-skilled wage), so that the emigrant labor income is $w^s_{\mathbf{i},t} L^s_{\mathbf{i},t}$. Also, $w^s_{\mathbf{u},t}$ is the wage earned in South, so that $w^s_{\mathbf{u},t} \left(L^s_{\mathbf{u},t} - L^s_{\mathbf{i},t} \right)$ denotes the total income from hours worked by the non-emigrant labor. On the spending side, each new unit of emigrant labor sent to Home requires a sunk cost f_e, expressed in units of immigrant labor: $f_{e,t} = \frac{w_{\mathbf{i},t}}{(\mathbb{X}_t \varepsilon^Z_t)} (\varepsilon^{fe}_t \mathbb{X}_t f_e)$. Changes in labor migration policies (e.g. border enforcement) are reflected by shocks ε^{fe}_t to the level of the sunk emigration cost f_e. Consumption by the South household, C^s_t, is a CES composite of non-tradables produced in South, $C^s_{N,t}$, and the Home tradable composite $C^s_{T,t}$ which may account for immigrants' consumption in Home, as well as imports from Home to South.[30] P^s_t is the resulting consumer price index.

Optimality conditions. The optimization problem yields the typical conditions for consumption and labor supply. Using the law of motion for the stock of immigrant labor, the first order condition with

[29] Our endogenous emigration-exogenous return formulation is similar to the framework with firm entry and exit in Ghironi and Melitz (2005).

[30] We consolidate the current account for Home and Foreign and abstract from modeling migrants' remittances which, in principle, could be used to finance these imports.

respect to new emigrants $L^s_{\mathbf{e},t}$ implies:

$$f_{e,t} = \mathbb{E}_t \sum_{s=t+1}^{\infty} \left[\beta(1-\delta_l)\right]^{s-t} \left(\frac{\zeta^s_s}{\zeta^s_t}\right) \left(w^s_{\mathbf{i},t} - w^s_{\mathbf{u},t}\right). \tag{12}$$

In equilibrium, the sunk emigration cost, $f_{e,t}$, equals the benefit from emigration, with the latter given by the expected stream of future wage gains from working abroad *vis-à-vis* home (i.e. $w^s_{\mathbf{i},t} - w^s_{\mathbf{u},t}$) adjusted for the stochastic discount factor and the probability of return to the country of origin every period, δ_l.

Non-tradable sector. Southern output is non-tradable and obtained as a linear function of non-emigrant labor: $Y^s_{N,t} = (\varepsilon^s_t \mathbb{X}_t)\left(L^s_{\mathbf{u},t} - L^s_{\mathbf{i},t}\right)$. Here, $\mathbb{X}_t$ is the unit-root global technology shock and ε^s_t is a country-specific shock. The price of the non-tradable good is: $P^s_{N,t} = \frac{w^s_{\mathbf{u},t}}{\mathbb{X}_t \varepsilon^s_t}$. By definition, $Y^s_{N,t} = C^s_{N,t}$. For simplicity, we define a consolidated current account for Home and South and abstract from tradable production in the South.[31]

4 Data and Estimation

We estimate the model with the Bayesian approach that uses the full set of system equations, avoiding the potential identification issues of reduced-form models and allowing us to recover the latent variables (i.e., unobserved variables or shocks) that are critical to study the impact of changes in labor market dynamics and perform scenario analyses. The approach combines the prior distribution on the estimated parameters with the likelihood function of the solved DGSE model to obtain posterior distributions of the structural parameters.[32]

Data. We estimate the model using eight quarterly series for the sample period 1983:Q1 – 2018:Q4.[33] The sample period ends in 2018:Q4. Since our estimation method is not suitable for highly nonlinear

[31] Thus, the evolution of the net foreign asset position for this artificial economy is:

$$q_t B_t - B_{t-1} = N_{X,t} \left(\tilde{w}_{X,t}\right)^{1-\theta} N^*_t \mathbb{Q}_t - N^*_{X,t} \left(\tilde{w}^*_{X,t}\right)^{1-\theta} N_t, \tag{13}$$

where, on the right-hand side, the first term is the sum of all tasks executed by home skilled workers and exported to Foreign, and the second term represents the tasks executed by foreign skilled workers and imported to Home, expressed in units of the home numeraire. This trade in tasks is one of the key characteristics of this model. The Home and Foreign risk-free bonds are in zero net supply: $B_t + B^*_t = 0$. Our model is propagated by nine shocks: shocks to Global ($\mathbb{X}_t$), Home (ϵ^Z_t), Foreign ($\epsilon^{Z^*}_t$), and South technology (ϵ^s_t); shocks to demand in Home (ϵ^b_t) and Foreign ($\epsilon^{b^*}_t$); shocks to iceberg offshoring cost (τ_t), training cost (ϵ^{Tr}_t), and sunk emigration cost ($f_{e,t}$). The world technology shock has a unit root, as in Rabanal et al. (2011): $\log \mathbb{X}_t = \mathbb{X}_{t-1} + \eta^{\mathbb{X}}_t$. The other structural shocks in our model follow $AR(1)$ processes with i.i.d. normal error terms, $\log \varepsilon^{\hat{\imath}}_t = \rho^{\hat{\imath}} \log \varepsilon_{t-1} + \eta^{\hat{\imath}}_t$, in which the persistence parameter is $0 < \rho^{\hat{\imath}} < 1$. The error terms are $\eta \sim N(0, \sigma^{\hat{\imath}})$, and indexes $\hat{\imath} = \{\mathbb{X}, Z, Z^*, s, b, b^*, \tau, f_e, Tr\}$ denote the shocks. Country-specific shocks are independent.

[32] See the Appendix for detailed information on data sources and the estimation methodology. An and Schorfheide (2007) and Mandelman and Zanetti (2014) for an overview and implementation of the Bayesian estimation. In addition, the Appendix includes a description of the smoothing procedure implemented with the Kalman filter, the Monte Carlo Markov Chain convergence diagnostics, and the Bayes Factor used for model comparison.

[33] We use eight data series while the model has eight shocks, so the number of data series used in the estimation does not exceed the number of shocks, thus avoiding stochastic singularity.

events, our sample period is chosen to abstract from two unprecedented events: (a) the COVID-19 pandemic, which triggered an immediate collapse in employment and output due to lockdown measures, (b) the sharp increase in apprehensions following the *migrant caravans* in late 2018. Nevertheless, we use the estimated parameters to analyze the welfare implications of these forceful events in our policy application in Section 6.

The first three series are output for the Home, Foreign, and South economies. We use U.S. real GDP as a proxy for Home output; for Foreign we construct as a trade-weighted GDP aggregate of the U.S. major trade partners; and South output is proxied by Mexico's real GDP. The fourth observable is the total hours worked by U.S. border patrol agents assigned to patrol the U.S./Mexico border, shown in Figure 5, which serves as a proxy for the intensity of border enforcement and the restrictiveness of immigration policy more generally. An increase in border patrol hours is interpreted as an increase in the sunk migration cost, as in Mandelman and Zlate (2012).[34] The fifth to seventh observables are the per-capita employment for three skill groups (high-skill, middle-skill, and low-skill service occupations) in the U.S. which are measured with a similar approach to Acemoglu and Autor (2011) and Jaimovich and Siu (2020).[35] The final observable series consists of real tuition costs, defined as tuition, other school fees, and childcare deflated by the total consumer price index (CPI), which is the data counterpart for the sunk training cost in the model. For completeness, we provide some sensitivity analysis to this broad measure of tuition costs in the Appendix. Importantly, we allow for the sunk training cost to not necessarily be linked to its data counterpart in a linear fashion; instead, we let the data pin down the value of the elasticity of the sunk training cost to the real tuition cost series as defined in the model description. Finally, all series used for model estimation are seasonally adjusted and expressed in log-differences to obtain growth rates.[36]

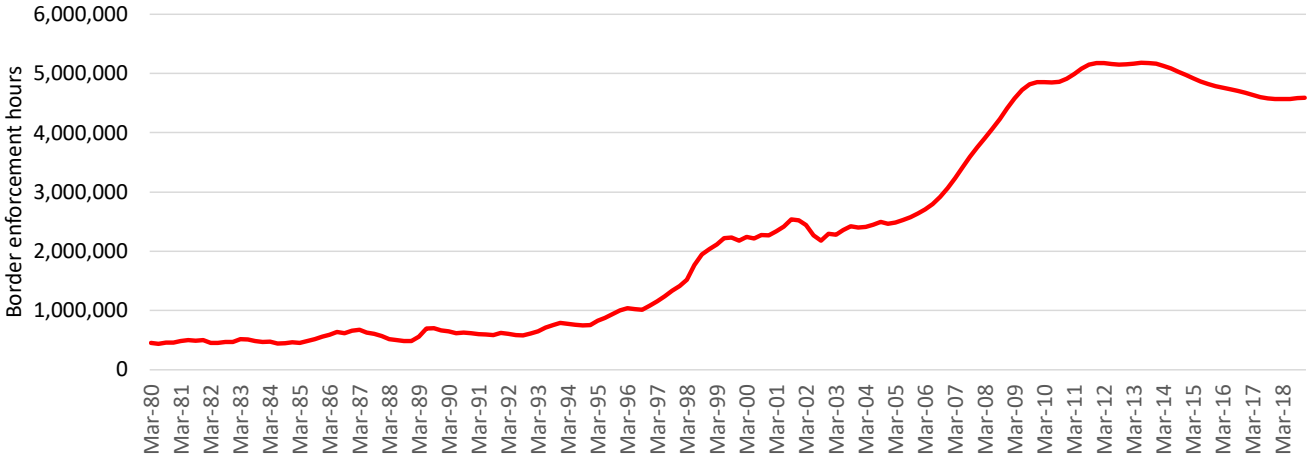
Two important variables are not used in the model estimation: (a) the inflows of low-skill migrant workers, and (b) the cost of offshoring. They do not enter the estimation for separate reasons. First, the number of individuals apprehended by U.S. patrol officers while attempting to illegally cross the U.S./Mexico border, shown in Figure 6, serves as a useful but noisy proxy for the inflow of low-skill immigrant workers. As pointed out in Hanson and Spilimbergo (1999), *ceteris paribus*, an increase in unautho-

[34]One caveat is that a border wall has been built over the past 40 years, and the model abstracts from that. Using construction data available for the period 2006-2010, Allen et al. (2024) show that this factor also enhanced border enforcement. Other factors omitted here are the role of increasing sanctions for those undocumented immigrants who are apprehended (Bazzi et al., 2021), and increasing interior enforcement (East et al., 2023).

[35]The U.S. Census employment data discussed in the introduction is decennial and thus not available on a high-frequency basis. In addition, it cannot be split easily into the three skill groups. We construct employment by skill group using data from the Current Population Survey (CPS). We consider three categories of employment based on the skill content of the tasks executed by each occupation in the Census data: Non-Routine Cognitive (high-skill), Routine Cognitive (middle-skill) and Non-Routine Manual (low-skill). Notice, that in the estimation we use total employment over population for each skill group while the introduction illustrates changes in employment shares in the Census data. In Jaimovich and Siu (2020) construction occupations are grouped with the middle-skill segment. We take a different approach for two reasons. First, construction is non-tradable by definition. Second, even though the earnings for the registered workers belong to the middle of the skill distribution. The underground economy is pervasive in this sector, and most low-skill laborers in this sector remain unregistered. See the Data Appendix for more details.

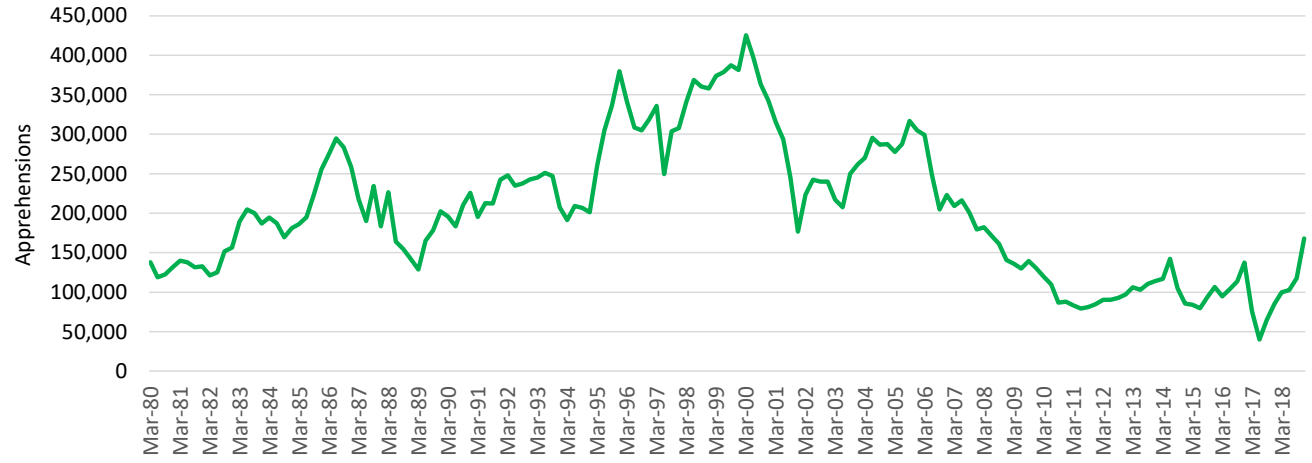
[36]The real GDP and tuition cost series are not detrended, while border enforcement is constructed as deviations from a linear trend. Along the balanced growth path in the model, border enforcement should grow at the same rate as output and tuition costs. Therefore, we follow Adolfson et al. (2007) and render border enforcement stationary by detrending its empirical counterpart around a linear trend.

Figure 5: U.S. border patrol enforcement at the U.S.-Mexico border



Note: This chart shows the number of border enforcement hours at the U.S.-Mexico border as a proxy for the stance of immigration policy and, in our model, the sunk migration cost. The data are from the U.S. Immigration and Naturalization Service (INS) for 1980-2004 (linewatch enforcement hours), used in Hanson and Spilimbergo (1999) and Hanson (2006), which we supplement with more recent data from the U.S. Customs and Border Protection (border patrol agent staffing at the Southern border).

Figure 6: Apprehensions at the U.S.-Mexico border



Note: This figure illustrates the number of apprehensions (arrests) at the U.S.-Mexico border as a proxy for the inflows of low-skill immigrant labor. The data are from the U.S. Customs and Border Protection for 2000-2018 (U.S. apprehensions at the Southwest border), which we supplement with earlier data from the U.S. Immigration and Naturalization Service (INS) for 1980-1999 used in Hanson and Spilimbergo (1999) and Hanson (2006) (linewatch apprehensions).

rized immigration would lead to more border apprehensions.[37] Nonetheless, the latter represent an imperfect indicator for immigration flows due to their complex relation with the intensity of border enforcement.

[37] We focus on the U.S.-Mexico border rather than the U.S.-Canada border because the former is more relevant for low-skilled immigration. Apprehensions (encounters) of migrants at the Canadian border were negligible until 2022 (covering the sample used for this manuscript). It is true that since then they have spiked, given the more lenient visitor policies recently implemented by the Trudeau administration. As a result, many migrants arrive on tourist visas and then illegally cross the border into the United States. While the surge after 2022 is notable, the numbers are still relatively small, with crossings from Canada representing less than 10 percent of those from Mexico in 2024.

Higher enforcement may discourage attempted unauthorized immigration, but for a given number of crossing attempts, higher enforcement can also result in more apprehensions. To address this issue, Hanson and Spilimbergo (1999) use instrumental variables methods to account for unauthorized immigration inflows, which result in the following reduced form specification: $ln(\text{Apprehensions}) - 0.8 \times ln(\text{OfficerHours})$. We adopt their approach and use this measure as a proxy for migration flows to assess the model fit, but not for the estimation. [38] Moreover, to achieve model parameter identification, our estimation strategy requires that an increase in border enforcement is assumed to be exogenous. However, the increase in enforcement in the mid-1990s may arguably be the result of many years of robust immigration growth since the 1980s. Unfortunately, the estimation strategy cannot capture this long-lagged response to low-frequency movements. However, as discussed in the regional analysis, predetermined policy preferences could be relevant.

Second, the cost of offshoring is affected not only by changes in trade costs but also by advances in telecommunications. These advances facilitate the disaggregation of the production process across locations by allowing workers in distant locations to interact and monitor each other in real time. Even though language and communication barriers are indirectly included in the ESCAP/World Bank measure of trade costs, it is not possible to directly quantify the impact of these technological advances on the actual cost of offshoring tasks. Therefore, these two variables are not directly included in the model estimation, but are used to assess the empirical adequacy of the model predictions by comparing them with the empirical proxies.

We abstract from skill-biased technological change (SBTC), which typically includes innovations in automation, computerization, and robotics. Skilled workers benefit from this type of innovation under the assumption of capital-skill complementarity–see Krusell et al. (2000). We abstract from their inclusion for several reasons. First, including SBTC innovations would require adding capital formation to the model, which would complicate the analysis considerably. Second, SBTC does not appear to be quantitatively relevant to the short- to medium-term dynamics examined here. In particular, it cannot explain the reversal of the skill premium in recent years which is the focus of this paper. See the Appendix for a discussion.

Prior distributions. We estimate the set of parameters shown in Table 5.[39] The prior probability density functions are centered at the values described below, and the size of the standard deviation is set to deliver a domain suitable to cover a wide range of empirically plausible parameter values. Shocks are harmonized with a very loose prior since we do not have much prior information about their magnitude. Nonetheless, some of these parameter values remain fixed through the estimation procedure, which can be interpreted as a prior that is extremely precise. This procedure is necessary to ensure identification because only a limited number of observables can be used to avoid stochastic singularity.

[38] The Appendix provides a robustness analysis to show that this high-frequency *flow* measure is consistent with other existing estimates of the *stock* of undocumented immigrants in the data.

[39] Model parameters are assumed to follow a `Gamma` distribution with a positive domain $[0,\infty)$. The auto-regressive parameters for the stationary shocks are assumed to follow a `Beta` distribution, which covers the range between 0 and 1. The standard deviation of all stochastic processes is assumed to follow an `InverseGamma` distribution that delivers a relatively large domain.

We use six empirical moments as targets to parameterize the prior values of six key model parameters affecting offshoring and labor migration. These moments include: (1) the ratio of high- to middle-skill jobs in the U.S. of 0.6; (2) the ratio between the high- and middle-skill labor income shares in the total U.S. labor income of 1.7;[40] (3) the share of low-skill workers in the native U.S. labor force of 0.2; (4) the ratio between U.S. low-skill wages and wages in Mexico of 2.2;[41] (5) the ratio of U.S. exports to GDP of 0.13; and (6) the ratio of U.S.-to-Mexico per-capita GDP of 5.4.

Regarding the associated six model parameters, we loosely center the prior distribution for sunk emigration cost at $f_e = 8.8$, while fixing the quarterly return rate of immigrant labor at $\delta_l = 0.05$, which is consistent with the data in Reyes (1997).[42] Following a similar strategy, we center the prior distribution for the iceberg offshoring costs at $\tau = \tau^* = 1.40$, consistent with Novy (2018), while fixing the cost of offshoring at $f_o = f_o^* = 0.0155$ to get identification while matching export ratios. The prior distribution for the elasticity of total training costs to observed tuition costs is centered at $\Theta_{f_j} = 0.35$, while the sunk training cost is normalized at $f_j = 1$ and the quarterly job-destruction rate is fixed at $\delta = 0.025$ as in Davis and Haltiwanger (1990). The response of labor supply to aggregate conditions is critical for the quantitative analysis, so we center the prior for γ_n and γ_n^s at 1.33, consistent with the micro estimates for the Frisch elasticity in Chetty (2012).[43]

Calibration. The remaining parameters are calibrated using standard values. Consistent with Borjas et al. (2012), the baseline specification assumes perfect substitution between native and low-skilled immigrant workers, so σ_N is set at an arbitrarily high value. This choice is disputed by other authors such as Ottaviano and Peri (2012) who find that these labor inputs tend to be complementary instead. We discuss the implications of this alternative scenario in Section 6. We set standard values for $\beta = 0.99$ and the inverse of the elasticity of inter-temporal substitution, $\gamma = 2$.[44] The idiosyncratic productivity of workers $\mathbf{z}$ follows the Pareto distribution $\mathcal{F}(\mathbf{z}) = 1 - \left(\frac{1}{\mathbf{z}}\right)^k$, as in Hamano and Zanetti (2017) .[45] The shape parameter is $k = 3.1$, and the elasticity of substitution across tasks in Home and Foreign is fixed

[40] The CPS survey reports the money income that includes wages and salaries, interest, dividends, rent, retirement income, as well as other transfers. There is one crucial caveat. Our parsimonious model abstracts from capital, so it is difficult to map each of these income sources to the skill groups defined in our setup. In addition, the CPS survey data are not suitable for studying high-income groups because of the small sample size and top coding of high incomes.

[41] BLS and INEGI are the data sources for the U.S. and Mexico, respectively. For the U.S., we consider median labor earnings for males with less than a high-school degree. For Mexico, we take the median wage for males.

[42] Reyes (1997) studies the return pattern of undocumented Mexican immigrants. They find that approximately only 50% remain in the U.S. after 2 years. Similarly, 35%, 25%, and 20% of them remain after 4, 10, and 15 years, respectively. We construct quarterly return rates based on these numbers. The resulting average is 0.05.

[43] The weights on the disutility from work are centered at $a_n = 3.9$ in Home and Foreign and $a_n^s = 8.6$ in the South, such that per-capita labor supply for each region is normalized to the data moments previously described along balanced growth.

[44] The cost of adjusting bond holdings is set at a very low value, $\phi = 0.0035$, which is sufficient to ensure stationarity. Per-capita labor supply is normalized in balanced-growth at $L_t = L_t^* = L_{\mathbf{u},t}^s = 0.5$. The share of tradable consumption is $\gamma_c = 0.75$ and the intra-temporal elasticity of substitution between the tradables and services is set at $\rho_c = 0.44$ as in Stockman and Tesar (1995). In the South, the share of the Home-produced tradable good γ_c^s in Household consumption is 0.2, the associated elasticity of substitution is $\rho_c^s = 1.5$.

[45] The shape parameter k is such that $k > \theta - 1$ so that $\mathbf{z}$ has a finite variance. When the parameter k is set at higher values, the dispersion of the productivity draws decreases and the idiosyncratic productivity becomes more concentrated toward the lower bound of the skill distribution.

at $\theta = 2.4$ to match the skewed U.S. income distribution.[46]

Table 5: Prior and Posterior distributions for estimated parameters

	Prior distribution				Posterior distribution			
Description	Name	Density	Mean	Std Dev	Mode	Mean	10%	90%
Sunk training cost elasticity	Θ_{f_j}	Beta	0.35	0.05	0.1927	0.1979	0.1639	0.2341
Sunk emigration cost	f_e	Gamma	8.8	0.1	7.2320	7.3579	6.4824	8.2679
Iceberg offshoring cost (H)	τ	Gamma	1.4	0.15	1.4141	1.3920	1.2399	1.5521
Iceberg offshoring cost (F)	τ^*	Gamma	1.4	0.15	1.4816	1.5014	1.3786	1.6268
Inv. elast. labor supply (H)	γ_n	Gamma	1.33	0.3	1.0768	1.1997	0.9604	1.4838
Weight disutility work (H)	a_n	Gamma	3.9	0.3	4.0379	4.0037	3.6298	4.3699
Inv. elast. labor supply (S)	γ_n^s	Gamma	1.33	0.3	1.0634	1.1439	0.8200	1.4873
Weight disutility work (S)	a_n^s	Gamma	8.6	1	8.5748	8.6473	7.4334	9.8875
Training cost shock	ρ_{f_j}	Beta	0.75	0.1	0.9988	0.9983	0.9974	0.9990
Migration cost shock	ρ_{f_e}	Beta	0.75	0.1	0.9802	0.9786	0.9676	0.9885
Offshoring cost shock	ρ_τ	Beta	0.75	0.1	0.9918	0.9894	0.9831	0.9948
Technology shock (H)	ρ_Z	Beta	0.75	0.1	0.9973	0.9966	0.9946	0.9983
Technology shock (F)	ρ_{Z^*}	Beta	0.75	0.1	0.7123	0.7057	0.6506	0.7582
Technology shock (S)	ρ_{Z^s}	Beta	0.75	0.1	0.9961	0.9944	0.9903	0.9977
Demand shock (H)	ρ_b	Beta	0.5	0.05	0.7806	0.7659	0.7410	0.7861
Demand shock (F)	ρ_{b^*}	Beta	0.5	0.05	0.4995	0.4989	0.4297	0.5657
Training cost shock	σ_{f_j}	Inv gamma	0.01	2*	0.0117	0.0119	0.0110	0.0128
Migration cost shock	σ_{f_e}	Inv gamma	0.01	2*	0.0285	0.0288	0.0266	0.0310
Offshoring cost shock	σ_τ	Inv gamma	0.01	2*	0.0059	0.0061	0.0055	0.0069
Technology shock (H)	σ_Z	Inv gamma	0.01	2*	0.0530	0.0532	0.0491	0.0575
Technology shock (F)	σ_{Z^*}	Inv gamma	0.01	2*	0.0287	0.0303	0.0261	0.0348
Technology shock (S)	σ_{Z^s}	Inv gamma	0.01	2*	0.0542	0.0538	0.0487	0.0591
Demand shock (H)	σ_b	Inv gamma	0.01	2*	0.0155	0.0158	0.0146	0.0172
Demand shock (F)	σ_{b^*}	Inv gamma	0.01	2*	0.0040	0.0047	0.0030	0.0068
Global technology shock	$\sigma_{\mathbb{X}}$	Inv gamma	0.01	2*	0.0290	0.0290	0.0268	0.0312

Note: The table shows the prior and posterior distributions of the model parameters subject to Bayesian estimation. (*) For the Inverse Gamma the degrees of freedom are reported.

Estimation results, posterior distributions. The last four columns of Table 5 report the posterior mean, mode, as well as the 10th and 90th percentiles of the posterior distribution of the parameters. Prior and posterior densities are displayed in the Appendix. The posterior mean of the sunk emigration cost, f_e, is equivalent to the immigrant labor income obtained over seven quarters in the destination economy. This value is only slightly higher than the estimate of five quarters found in Mandelman and Zlate (2012), which was based on a shorter time series for border enforcement (1983-2004). The elasticity of the sunk training cost to the observed real tuition costs, Θ_{fj}, is relatively low, which could reflect the possibility of substitution from private to public education or to alternative career paths with similar

[46] Notice, however, that the interpretation for some of these parameters may differ from the literature, since our framework features tradable *tasks* rather than tradable *goods*, and skill obsoletion rather than job destruction.

earning payoffs but lower education costs, which we do not model here. Productivity shocks are relatively more persistent than demand shocks, which is consistent with the literature (e.g. Smets and Wouters, 2007). Offshoring costs are very persistent and relatively less volatile than the stochastic innovations to border enforcement. Real tuition costs display near unit-root dynamics.

5 Model Fit and the Effect of Shocks

In this section, we study: (i) the fit of the model to the data, (ii) the propagation of shocks to a decline in migration, training, and offshoring costs, and (iii) we provide the historical contribution of the estimated shocks to key economic variables over the period 1983-2018.

5.1 Model Fit

We proceed with the posterior predictive analysis that compares the actual data with artificial time series generated with our estimated model. As already discussed, we do not use data series on immigrant flows or offshoring costs to estimate the model. Instead, we treat immigrant entry ($L_{e,t}$) and the iceberg offshoring cost (τ_t) as latent variables in the estimated model and we can use the comparison of them with data to assess the model fit.[47] Figure 7(A) shows model predictions for the flows of low-skill immigrant labor, expressed as deviations from balanced growth (dashed line) together with their empirical detrended proxy for migration adjusted for border enforcement like in Hanson and Spilimbergo (1999) and Hanson (2006) (solid line). The model predictions are broadly consistent with the data: Immigration falls sharply below trend growth at the onset of the Great Recession and, although it later rebounds, remains below trend growth until 2018, when the migrant *caravans* occurred (to be discussed later in the historical analysis). Of note, the final years of the sample data are remarkably noisy. This likely reflects changes in the seasonal pattern of unauthorized immigration after 2013 (see Gramlichv and Scheller, 2021), which are missed by the standard seasonal adjustment methods.[48]

Figure 7(B) shows that the model prediction for the cumulative stochastic innovations affecting the iceberg offshoring cost (dashed line) matches well the trade-weighted ESCAP/World Bank measure of trade costs (solid line). Following the Great Recession, the model predicts a transitory increase in trade costs from 2007 to 2010, while the data show a decline. This apparent discrepancy may be reconciled with additional information not reflected in the trade cost indicator, which fails to account for factors like the increase in trade protectionism during the crisis reflected in the increase in non-tariff barriers (see Georgiadis and Gräb, 2016), and the freeze in trade credit (see Coulibaly et al., 2013).

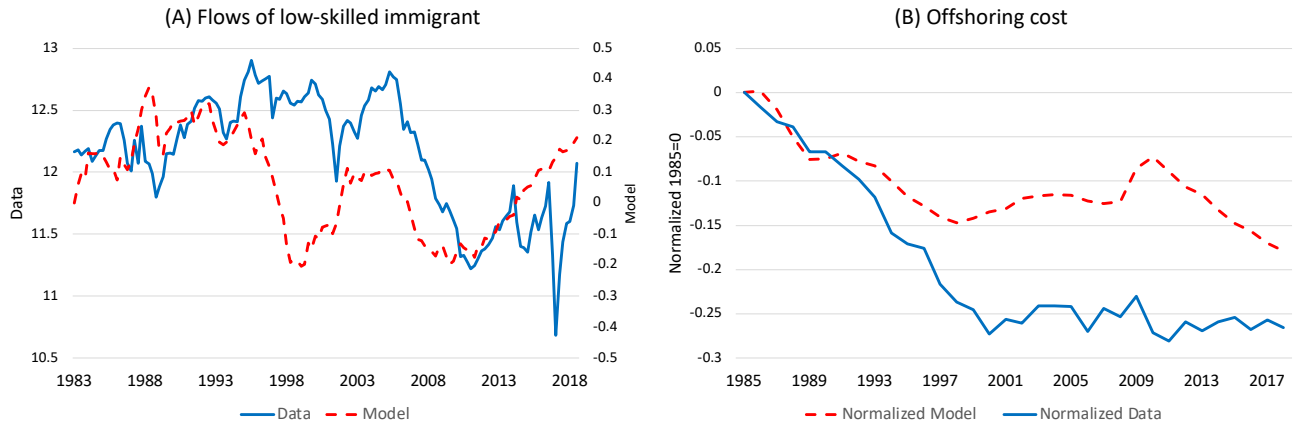
In sum, the model's predictions for the evolution of low-skilled immigration and offshoring costs appear to be broadly consistent with the data, which is noteworthy because we do not use data series on

[47] For this purpose, the Kalman filter provides smoothed estimated shocks to construct the predictions for the unobserved variables in each period, which allows for the reconstruction of the artificial historical series. See the Appendix for details on the smoothing procedure. One-sided estimates of the observed variables deliver a satisfactory in-sample fit.

[48] Traditionally, border apprehensions were highly seasonal and peaked in the spring before declining during the hot summer months. These long-lasting patterns vanished after 2013, and apprehensions spiked at different times during the year, which may reflect changes in entry strategies by *coyotes*, or alternatively, the increase in the proportion of migrants crossing the border to surrender to the authorities and seek refuge status.

labor migration or trade flows in estimating the model.

Figure 7: Flows of low-skill immigrant and offshoring cost predicted by the model



Note: This figures show the model predictions for low-skill immigration flows and the iceberg cost offshoring as latent variables as well as their data counterparts, i.e., border apprehensions adjusted for enforcement like in Hanson and Spilimbergo (1999) and Hanson (2006) and constructed trade costs from the ESCAP/World Bank database.

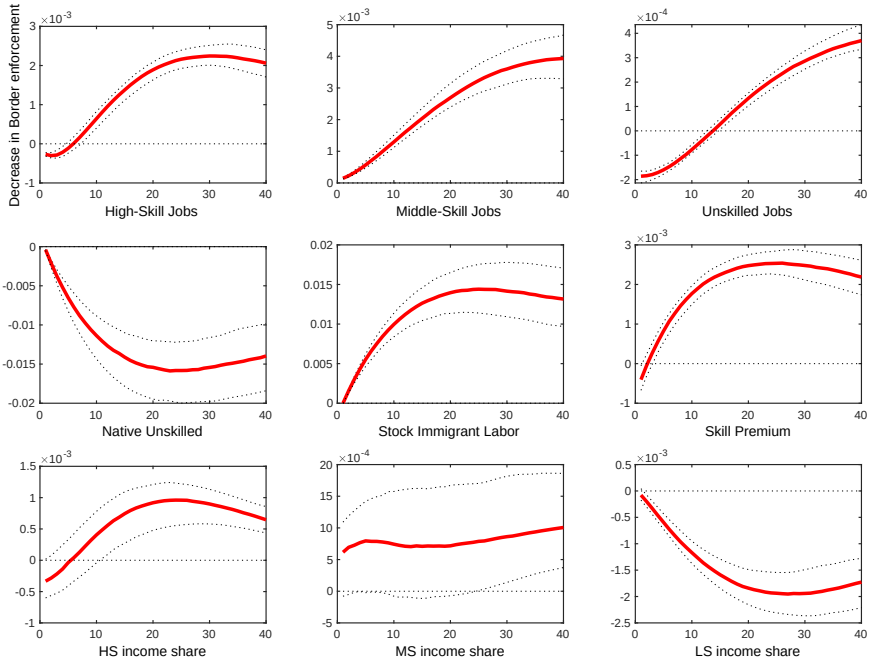
5.2 Impulse Response Functions

In this section, we describe the propagation of shocks to a decline in migration, training, and offshoring costs. The Appendix reports a discussion on the effect of technology (total factor productivity) and consumer demand shocks. We show the estimated median impulse responses (along the 10% and 90% posterior intervals) of key model variables to different stochastic innovations (one standard deviation for each). The historical analysis below provides an easier assessment of the quantitative impact of these innovation on macroeconomic variables.

Decline in the sunk migration costs. Figure 8 shows the response of key variables to a decline in migration costs, which can be interpreted as a decline in border enforcement. Immigration inflows rise on impact, but the stock of immigrant labor increases gradually–peaking only after five years (twenty quarters). The inertia in the increase of immigration explains its delayed impact on labor market variables, which is further discussed below for the historical analysis. The native household reacts to higher immigration by investing in training and reallocating labor away from low-skill service occupations and toward high- and middle-skill occupations (consistent with the job task upgrading in Ottaviano et al., 2013). As a result, the fraction of native low-skill jobs declines, while high- and middle-skill jobs rise slowly over time. The downward pressure of higher immigration on low-skill wages – along with the shift in native employment toward high- and middle-skill occupations – leads to an increase in the income shares of high- and middle-skill workers, but to a decline in the income share of low-skill occupations. The skill premium increases as a result.

Figure 8 reveals a quantitative relationship between the stock of low-skill immigrant labor and the skill premium (middle row, center and right panel) that is consistent with the literature (see Borjas (2014) and Mandelman and Zlate (2017)). Borjas (2014), employing a structural estimation approach with a CES

Figure 8: Impulse response to a decline in migration barriers



Note: This figure illustrates the impulse response of selected variables to the negative shock to immigration costs. The solid line is the median impulse response to one standard deviation of the estimated shock, the dotted lines are the 10 and 90 percent posterior.

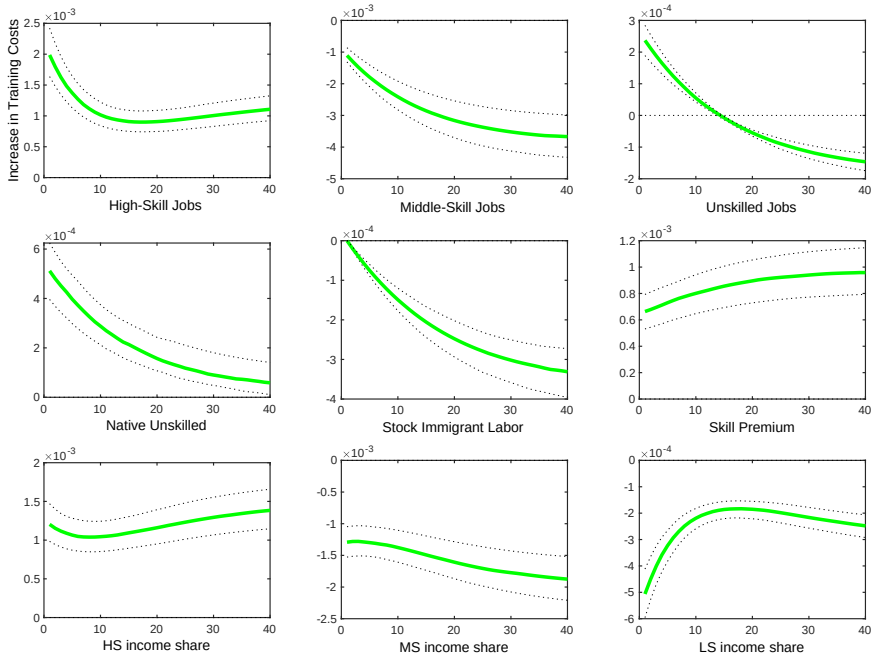
specification and perfect substitution, suggests that the 26% increase in the supply of high-school dropout workers due to immigration between 1990 and 2010 would have reduced wages for both immigrants and natives by 3.1%. For comparison, Figure 8 indicates that a 1.5% increase in the stock of immigrant labor coincides with a roughly 0.25% rise in the skill premium 25 quarters after the shock. Extrapolating from our model, a 26% increase in immigrant labor would correspond to a 4.3% rise in the skill premium, consistent with Borjas (2014)'s 3.1% decrease in low-skill wages. That said, in our model, the higher skill premium is not solely driven by the decline in low-skill wages——enhanced training and productivity among natives also contribute to this outcome.

To sum up, low-skill immigration encourages native workers to train and boosts the labor productivity and income for the trained workers.

Increase in training costs. Figure 9 shows the impulse responses to an increase in training costs. Due to the higher training costs, more workers choose to remain untrained, boosting the supply of native low-skill labor on impact. This puts downward pressure on the associated low-skill wages and the price of non-tradable services. Southerners have less incentive to migrate and the stock of immigrants declines gradually over time. This explains the contraction of the *aggregate* supply of low-skill labor at longer horizons despite a lower training choice by natives. In turn, the persistent decrease in training eventually dents labor productivity and further lowers real wages over time, further deterring immigrants to come.

A Harrod-Balassa-Samuelson effect takes place, in which lower productivity leads to lower non-tradable prices and, thus, a real exchange rate depreciation. These adjustments explain why, in the margin, there is some increase in high-skill employment. Native skilled workers become more competitive in the global marketplace due to the depreciation in the exchange rate, which allows for some upskilling among the most productive (higher $\mathbf{z}$) middle-skill workers.

Figure 9: Impulse response to an increase in training costs

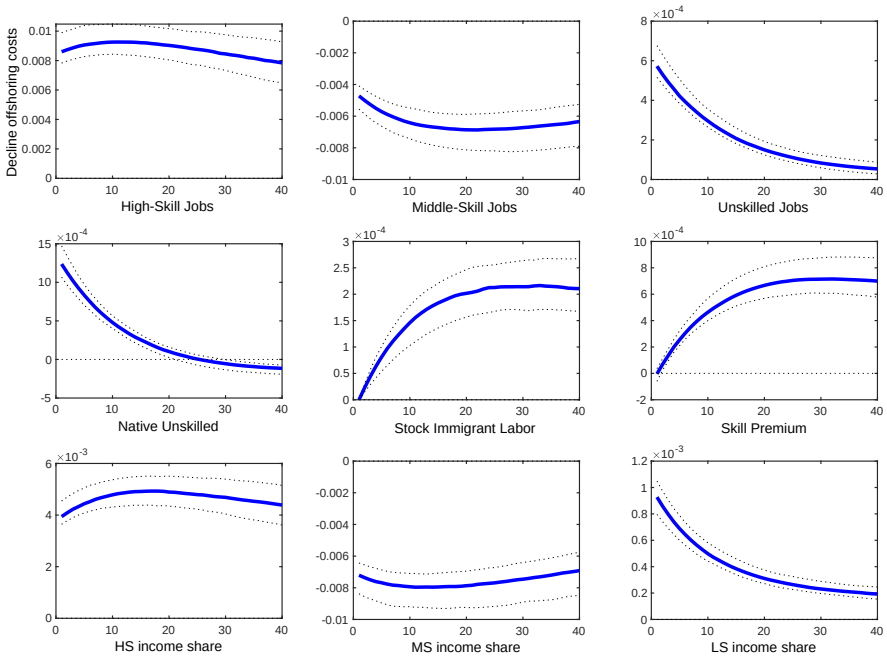


Note: This figure illustrates the impulse response of selected variables to a positive training cost shock. The solid line is the median impulse response to one standard deviation of the estimated shock, the dotted lines are the 10 and 90 percent posterior.

Decline in offshoring costs. Figure 10 shows the effect of a decline in offshoring costs. Since the effect of the shock is symmetric across countries, we show the impulse responses for Home only. Easier offshoring induces producers to expand the number of tasks executed abroad, which boosts the employment of high-skill workers who execute tasks for the global market, but displaces middle-skilled workers who face lower earnings resulting from the competition of offshore workers. Our result is consistent with the evidence in Oldenski (2014), who shows that middle-skill occupations in routine jobs (e.g., manufacturing) were the most affected by the globalization wave in the past decades. Efficiency gains from task specialization increase, however, which enhances aggregate labor productivity. As aggregate income increases, so does the demand for non-tradable services and low-skill employment, inducing an increase in low-skill immigration from the South (middle panel in the middle row). Thus, the impulse responses to the decline in offshoring costs are consistent with the evidence for the period 1980-2007 in Figure 1 in the introduction. A deeper integration of the U.S. labor markets into the global economy hollowed out

the middle-skill employment while inducing upskilling by natives towards more productive jobs, at the same time with an inflow of low-skill immigrant workers for non-tradable tasks that cannot be executed overseas (the top row). Workers at the upper and lower tails of the skill distribution not only enjoy better employment outcomes but also gain a higher share of income at the expense of those in the middle of the skill distribution (the bottom row). However, immigration dampens the increase in low-skill wages over time, boosting the skill premium at longer horizons.

Figure 10: Impulse response to a decline in offshoring costs



Note: This figure illustrates the impulse response of selected variables to a decline in offshoring costs in Home and Foreign. The solid line is the median impulse response to one standard deviation of the estimated shock, the dotted lines are the 10 and 90 percent posterior.

5.3 Historical Decomposition

Figures 11-16 show the historical contribution of the estimated shocks to key variables in the model for the period 1983:Q2-2018:Q4. The variables include employment and income shares for each skill group, and the labor migration-related indicators, expressed in deviations from balanced-growth (y-axes).

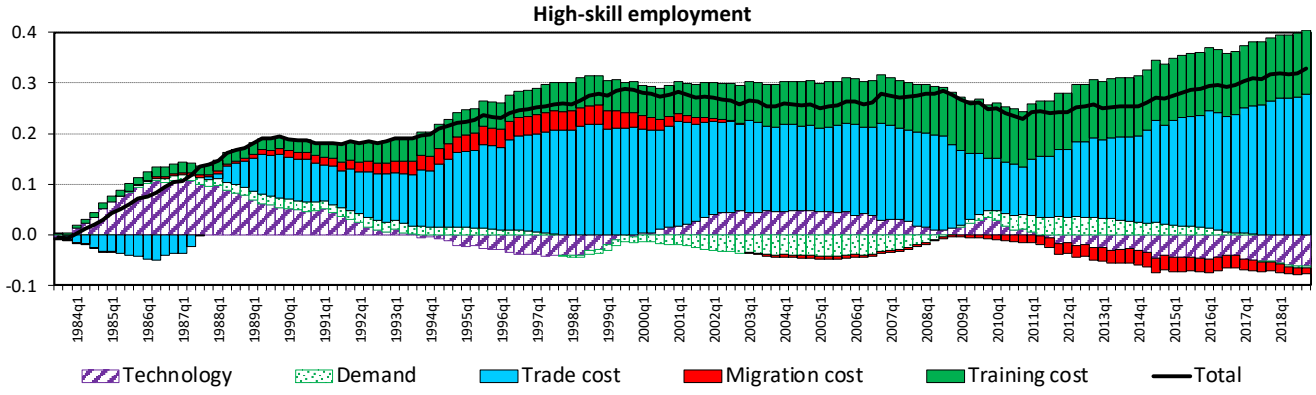
The historical contributions of the distinct shocks to the evolution of each variable are represented by the colored bars. The model predictions for the intensity of U.S.-Mexico border enforcement and for training costs reflect the actual data used in the estimation. The observed innovations to U.S.-Mexico border enforcement are fully reflected by fluctuations in migration cost shocks, which are exogenous to the model (ϵ_t^{fe}, dark red bars). The same applies to training cost shocks directly taken from the tuition data (ϵ_t^{Tr}, solid green bars). Figures A10-A12 in the Appendix display them. The remaining shocks include total factor productivity innovations grouped for all countries (purple), offshoring costs (blue),

and demand (white with green dots), which are not directly observed and are backed out with the Kalman filter. See Appendix for details.

Overall, the first half of the sample, comprising the period 1983-1996, shows a sustained relaxation in migration barriers associated with the significant increase in migration flows. Swings in border enforcement policy appear to be associated with the U.S. political cycle. The Immigration Reform and Control Act of 1986 provided amnesty for some of the workers who arrived before 1982, but also involved a short-lived increase in border enforcement in the late 1980s. More recently, the Unauthorized Immigration Reform Act under the Clinton Administration in 1996 was also accompanied by tightened enforcement, which this time showed to be stronger and more persistent. Border enforcement tightened further during the Great Recession, building over time the slowdown in immigration. Finally, training costs display a steady upward trend in tuition costs that persists through the sample period used for the estimation.

Accounting for historical events. The historical decomposition indicates that technological change (dashed-purple bars) played a central role in explaining the increase in inequality among employment skill groups in the 1980s. Moreover, the declining cost of offshoring (light-blue bars) became a dominating factor benefiting employment in high-skill occupations at the expense of middle-skill ones from the 1990s onwards, and thus driving the skill premium. These trends are notably consistent with the microeconomic evidence in Firpo et al. (2011). As shown in Figures 11 and 12, technology shocks lowered middle-skill employment during the three recorded recessions of 1990-91, 2001, and 2007-09, as documented in Jaimovich and Siu (2020). The decrease in migration costs contributed positively to the growth in both high- and middle-skill employment during the late 1980s and the 1990s, as immigration prompted native low-skill workers to undergo training and task upgrading. However, the increase in real tuition costs during the sample period was detrimental to skill upgrading, offsetting some of the increase in skilled labor. Due to the Harrod-Balassa-Samuelson effect explained above, higher tuition costs boosted the international competitiveness of the most productive skilled workers engaging in offshore production.

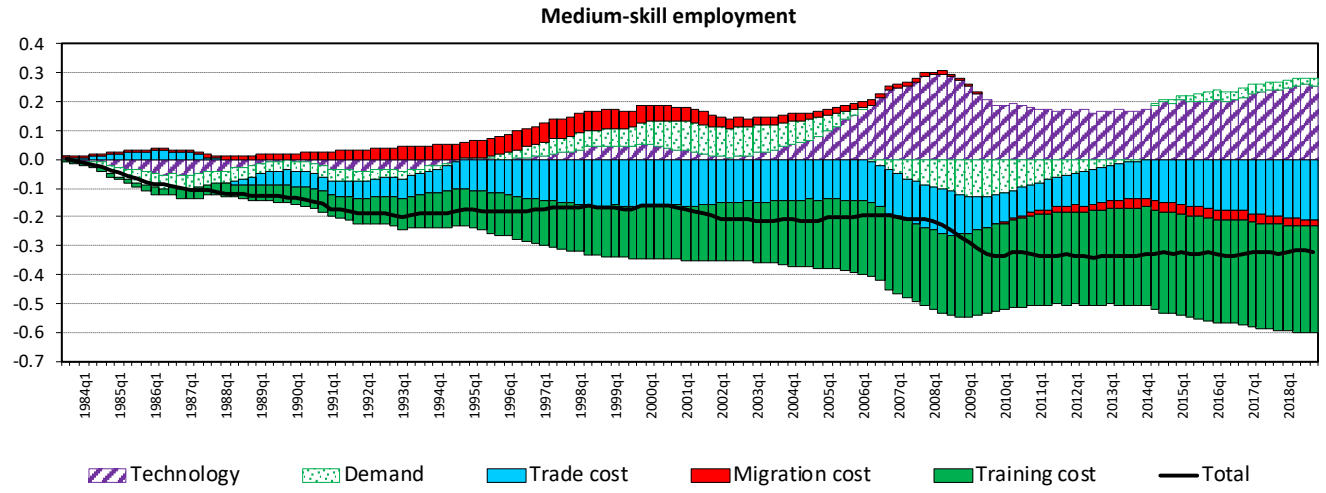
Figure 11: Historical decomposition, high-skilled employment



Note: This figure illustrates the historical decomposition of high-skilled employment between 1983 and 2019.

Figure 13 shows the evolution of aggregate low-skill employment, which includes both low-skill immigrants and natives. The decline in offshoring costs and immigration barriers positively contributed to

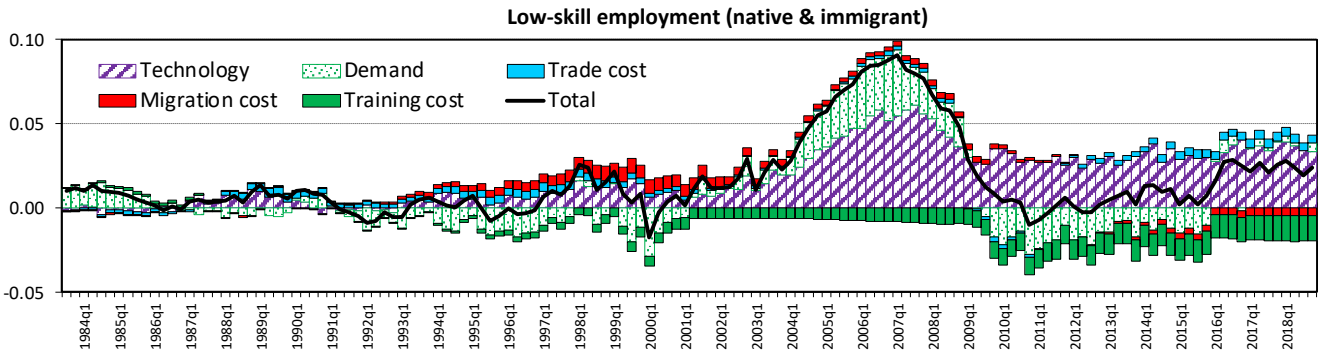
Figure 12: Historical decomposition, medium-skilled employment



Note: This figure illustrates the historical decomposition of medium-skilled employment between 1983 and 2019.

low-skill employment during the first half of the sample period. The relatively contained effect of the decline in immigration costs on the employment of low-skilled workers conceals sizable composition effects between natives and immigrants. As shown in Figures 14 and 15, a remarkable decline in low-skill native employment was offset a steady increase in immigration inflows in line with the ACS data discussed in the introduction.

Figure 13: Historical decomposition, low-skill employment

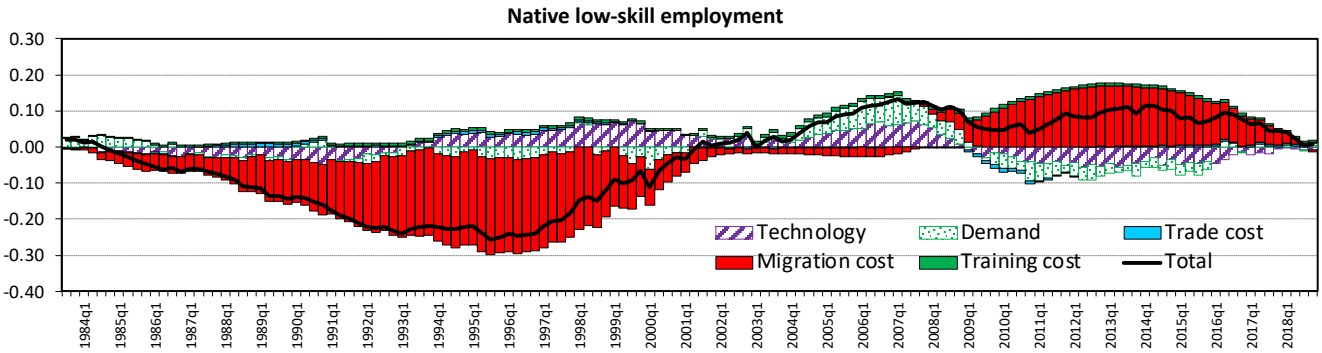


Note: This figure illustrates the historical decomposition of low-skill employment between 1983 and 2019.

The positive shocks to productivity and consumption demand (dotted-green bars) that are linked to the housing bubble also contributed to the sizable increase in aggregate low-skill employment in the early 2000s. Conversely, the reversal of these transitory shocks explained the decline in low-skill employment in the aftermath of the Great Recession in 2007.[49]

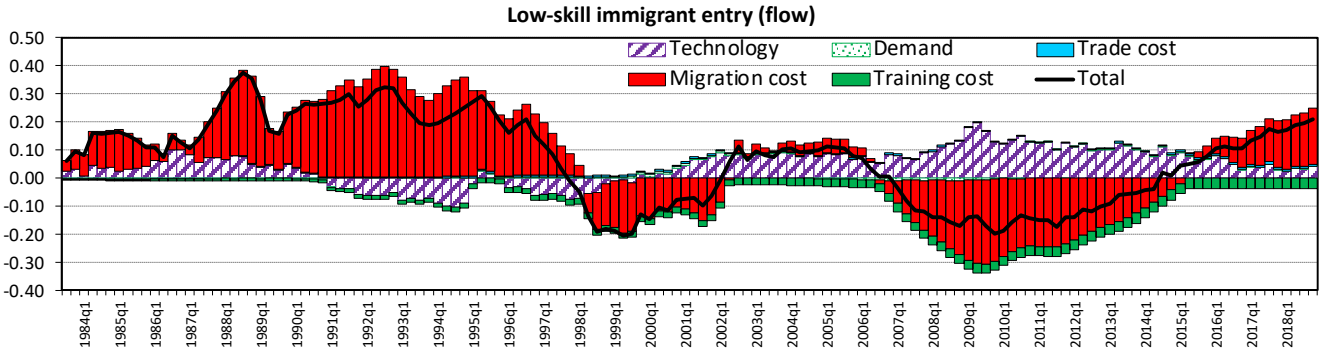
[49]The intuition for the consumption demand shocks, shown in the Appendix, is straightforward: the demand shock enhances the demand for both non-tradable and tradable tasks in Home due to their complementary in consumption. As a result, Home relied on Foreign to provide more tradable tasks (leading to an increasing trade deficit) and instead devoted more of its labor to producing non-tradable tasks that couldn't be substituted with imports from Foreign.

Figure 14: Historical decomposition, low-skill native employment



Note: This figure illustrates the historical decomposition of employment of low-skill natives between 1983 and 2019.

Figure 15: Historical decomposition, inflow of low-skill immigrants



Note: This figure illustrates the historical decomposition of the inflow of low-skill immigrants between 1983 and 2019.

In our model, labor markets and trade deficits are tightly interconnected. A consumption boom driven by wealth effects can be satisfied by more imported tasks (boosting the current account deficit), while the increase in demand for non-tradable requires a rise in immigration since those services cannot be offshored.[50] The boom-bust in low-skill (non-tradable) employment during the housing bubble coincided with a sizable increase in the U.S. current account deficit and enhanced immigration, followed by a remarkable correction after the Great Recession.[51]

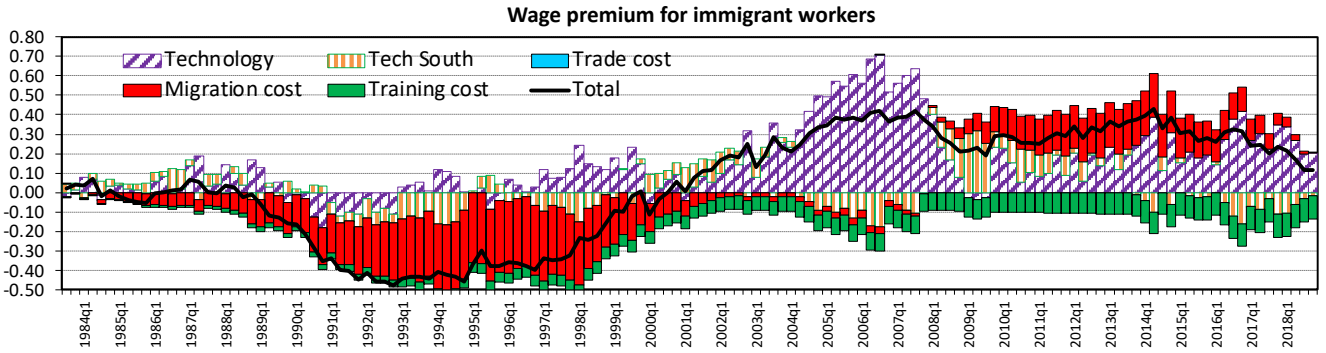
Figure 15 shows that the entry of low-skill immigrants was driven by the sustained lowering of immigration barriers in the 1980s. This policy stance, briefly interrupted by the 1986 immigration reform, continued until the mid-1990s. The 1996 reform, however, ushered in a sustained increase in enforcement that decisively turned the tide of migration for the next two decades. This prolonged slump in immigration was temporarily interrupted by a surge in low-skilled immigration from the housing boom of the early 2000s. After the bubble burst, this immigration slump resumed after 2007 and persisted through 2018, largely explaining the low-skilled labor shortages that motivates our analysis.

[50] Conversely, negative demand shocks in Foreign may reflect the increase in the supply of foreign savings documented during those years –referred to as the global savings glut– that is not directly accounted by our model. See Kehoe et al. (2018) for a discussion on those forces for macroeconomic performance.

[51] The current account deficit decreased from 6.2% of GDP in 2006:3 to 2.5% in 2009:2.

Finally, Figure 16 shows the historical decomposition of the wage premium for immigrant workers, calculated as the difference between the wages of Southern workers (and low-skilled natives) in Home and those residing in South. Unlike in previous figures, the purple dashed bars group together the productivity shock and demand in Home, while the orange bars show the measured shocks in South. The decrease in immigration barriers in the 1980s generated a delayed decrease in the immigration wage premium in the 1990s given the inertia previously discussed for Figure 8.[52] Low-skill wages at Home surged temporarily with the housing boom in the mid-2000s and after 2016 as labor markets tightened. The slowdown in immigration, which began in 2006, had its largest impact in the mid-2010s as the economic recovery matured. This delayed increase in the immigrant wage premium boosted overall low-skill wages, allowing the model to quantitatively match the cumulative decline in the average U.S. skill premium over the last decade of the sample (5.06 % in the model vs. 5.00 % in the data).

Figure 16: Historical decomposition, wage premium for immigrant workers



Note: This figure illustrates the historical decomposition of the wage premium for immigrant workers between 1983 and 2019.

6 Welfare Analysis

Using the second-order approximation around the balanced-growth path of the system, we derive welfare gains (losses) as the percent of the expected stream of consumption that one should add (or subtract) that make the representative households as well-off in each counterfactual scenario as in the estimated model, *ceteris paribus*. We cannot distinguish the welfare effects for workers with different skills, given the model assumptions about perfect risk sharing within the representative household in each country. Implicitly, these aggregate welfare results assume that an appropriate set of taxes/transfers across households in each country can address idiosyncratic inequality.[53] Our measure of welfare is the present discounted value of the flow of utility.[54] Table 6 shows the welfare outcomes from alternative counterfactual scenarios

[52] Such inertia explains why disruptive, but transitory, economic events in Mexico like the 1980's debt crisis, Peso crisis, or the acute contraction in 2009 following the disruption of trade links have limited impact on the wage premium of immigrants.

[53] The approximation is based on the approach developed in Woodford (2003).

[54] The representative household in Home only accounts for the native workers while the Southern household accounts for the welfare of migrant workers. Implicitly, we assume that migrants in Home use remittances to transfer funds to their country of origin to equalize utility across household members in different locations (see Mandelman and Zlate, 2012, for details). Remittances are netted out in the consolidated current account for Home and South.

for immigration and offshoring policies, and alternative trajectories in training costs, which we discuss below.

Table 6: Welfare implications of counterfactual changes in immigration, skill acquisition, and trade

	Home	Foreign	South
(1) Decline in Immigration	−2.54%	−0.19%	−9.06%
(2) Decline in Immigration (with shocks)	−2.52%	−0.29%	−11.00%
(3) Decrease in Training Costs	5.28%	−0.31%	2.07%
(4) Decrease in Training Costs (with shocks)	5.28%	−0.31%	2.35%
(5) Decrease in Offshoring Costs	2.99%	4.19%	1.30%
(6) Decrease in Offshoring Costs (with shocks)	3.86%	4.31%	1.24%

Note: The table shows welfare gains (losses) from three different scenarios—slowdown in immigration, lower training costs, and lower offshoring costs—with or without estimated shocks for each scenario. The gains (losses) are computed relative to alternative scenarios in which low-skill immigration during 2007-2018 would have matched its higher average rate from pre-2007; real tuition costs would have remained at their relatively low levels from 1983; and offshoring costs would have remained at their high levels from 1983, respectively.

Rows (1) and (2) in Table 6 show the welfare losses generated by the slowdown in low-skill immigration to Home for households in each economy during 2007-2018. To measure this, we compute welfare differences between a counterfactual scenario in which the average annual growth of low-skilled immigration for the period 2007-2018 would be identical to the one observed during 1983-2006 and the observed one with the immigration slowdown.[55] Row (1) in the table considers the scenario which abstracts from the estimated transitory AR(1) shocks. Results indicate that the representative U.S. native household loses 2.54% of its average consumption in every period as a result of the downshift in immigration from the average growth observed during the 1983-2006 period towards those observed during 2007-2018. Not surprisingly, the losses are much larger in the South (-9.06%), for which immigrant work represents a larger source of income, and smaller for Foreign (-0. 19%), which is only marginally affected through indirect channels affecting the real exchange rate.

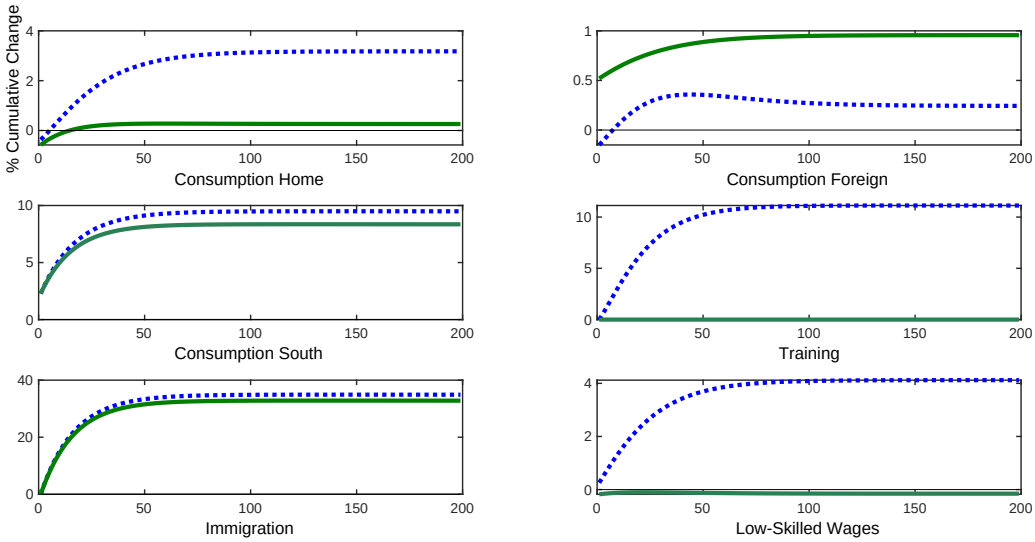
In Figure 17, we explore the mechanism behind the welfare loss experienced by native workers in the Home country due to slower immigration. The reduction in training (skill acquisition) emerges as a key driver. The scenario depicted in the figure assumes that immigration barriers decline immediately, leading to an increase in immigration levels. Specifically, we replicate the annual growth rate of immigration observed from 1983 to 2006 for the period from 2007 to 2018. As a result, the stock of immigrant labor in the new balanced-growth path becomes 35% higher than estimated in the model (indicated by the blue-dotted line). However, when households are unable to optimally adjust their training choices in response to lower migration barriers (as shown by the green-solid line), the welfare gain (reflected by consumption in Home) is greatly diminished, which illustrates the role of training.

While the baseline delivers inter-temporal consumption gains of around 3% in the new equilibrium, the inter-temporal gains practically disappear with a fixed training choice.[56] That is, if households cannot

[55] Our structural approach builds on the short-run elasticities of substitution between natives and immigrants estimated in the literature to account for the dynamic response to skill acquisition (training) that follows. Taken together this explains why our results show a much larger displacement of low-skilled native workers at longer horizons.

[56] The 2.54% gains mentioned above are computed as stable stream (i.e., like an annuity). In reality, consumption increases

Figure 17: Welfare analysis: transitional dynamics



Note: Blue-dotted line: Immigration returns to pre-2007 trend. Green-Solid: Same case with a fixed-training margin.

invest in training, the gains from immigration are practically muted, with the potential gains from lower consumer prices for non-tradables being largely offset by lower wages for low-skilled natives. A central result of our analysis is that the native workers' decision to invest in skills can deliver sizable gains from low-skill immigration even under the extreme assumption of perfect substitutability between low-skill native and immigrant labor (and no capital accumulation). Allowing for some degree of complementarity in production between natives and immigrants would deliver even larger gains from immigration.[57]

Row (2) in Table 6 includes the estimated transitory shocks to the analysis. The inclusion of shocks allows us to account for the welfare effects of transitory innovations perturbing the model dynamics along balanced growth under different policy scenarios. The results are different from those in row (1), since the model is approximated to a second order and entails non-linearities. Thus, the welfare losses from lower migration for the South are significantly higher (11%) in the presence of shocks. This finding highlights the role of labor migration as a possible insurance mechanism for the Southern household. For instance, if the South is hit by a negative shock (e.g., the 1994 Mexican Peso crisis), earnings are much lower for those residing in the South, but not for the immigrants working in Home who can support their families back at home. Consistent with our model, empirical evidence shows that migrants send more remittances in times of distress to support consumption in their native countries.[58]

slowly as training boosts labor productivity over time.

[57] If we introduce a moderate degree of complementarity along the lines of the findings in Ottaviano and Peri (2012), immigration is slightly less harmful to natives and they choose to acquire slightly less education when there is some complementarity. However, none of the main results and implications change. The decline in the skill premium is similar, as the smaller negative impact on low-skilled wages is offset by less education and a more abundant supply of low-skilled labor. Results are available upon request.

[58] For Home, instead, the losses are practically the same with or without shocks, which is due to high immigration barriers causing a lock-in effect whereby the stock of established immigrants responds slowly to shocks. For instance, migrants do not return to South during a recession in Home, as they have to pay a hefty sunk migration cost to re-enter after the Home economy recovers. Similarly, immigrants would be slow to arrive if there is an increase in consumer demand in Home, which

Rows (3) and (4) in Table 6 show the counterfactuals with real training costs remaining at the 1983 values, which are substantially lower than those observed in 2018.[59] As expected, the benefits are sizable for the Home economy, whose consumption stream rises by 5.28% due to increases in labor productivity. More training by Home natives is also beneficial for the South, as the need to fulfill low-skill tasks through enhanced immigration is greater. The effect on Foreign welfare is negligible. On one hand, Foreign benefits from a greater variety of labor tasks provided by a relatively more skilled labor force in Home. On the other hand, more tasks from Home also cause some displacement of skilled workers in Foreign.[60]

Rows (5) and (6) in Table 6 account for the potential welfare gains derived from the decrease in trade/offshoring costs observed during the sample period. The reduction in the iceberg offshoring costs in Home is welfare-improving for all the three economies. Home becomes specialized in its most productive labor tasks while Foreign benefits from the increasing availability of complementary Home tasks, which also enhances task specialization.[61] Comparing row (5) to row (6), economies more open to trade/offshoring are also better hedged from domestic shocks, given the larger welfare gains on row (6), which accounts for estimated transitory shocks, than on row (5), which abstracts from such shocks.[62]

COVID-19 policy responses: the CARES Act and labor shortages. We finally apply the estimated model to study the impact of labor shortages during the COVID-19 pandemic.[63] One notable development amid the labor shortages resulting from the pandemic was the unprecedented fiscal stimulus payments under the CARES Act enacted in March 2020. The stimulus bill unleashed approximately 5 trillion dollars (about 25 percent of GDP at the onset of the crisis) in direct transfers (see U.S. Treasury Report 2022), which was accompanied by a large-scale monetary policy easing by the Federal Reserve. On the supply side, the COVID-19 pandemic significantly restrained labor force mobility and halted immigration due to travel restrictions.

While our stylized model cannot accommodate interventions in fiscal and monetary policy, we can approximate them using a large positive shock to consumer demand (i.e., a negative shock to the discount rate), as standard in the macro literature.[64] The blue-dotted line in Figure 18 illustrates the case with the positive consumer demand shock. A demand shock boosts the demand for consumption but also for

is only transient in nature (as we will discuss in the next section).

[59] Real tuition and other school fees tripled during the sample period used in the estimation (1983-2018).

[60] The results are essentially identical when we include transitory shocks, as the training choice involves a costly and irreversible sunk cost and is unlikely to be affected by transitory shocks.

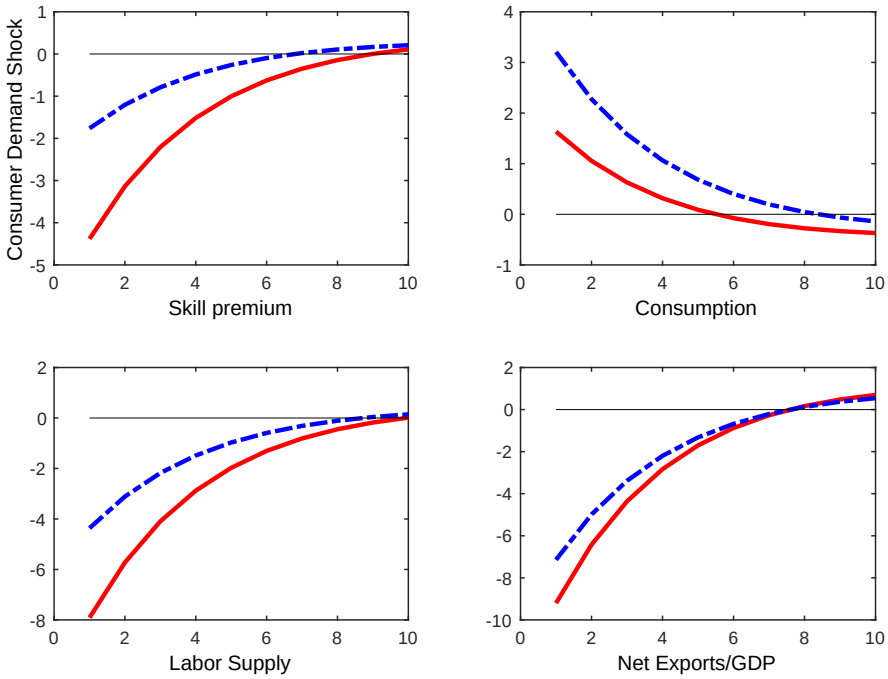
[61] The price (wage) impact on the terms-of-trade (labor) resulting from increasing availability of Home tasks in Foreign explains why the gains are relatively larger for Foreign. In addition, some of the Home's welfare gains are transferred to South through the income of immigrant workers.

[62] While Kim and Kim (2003) shows that international bond trading helps to hedge minor and highly transitory shocks, our results indicate that trade and offshoring enhances the scope of risk-sharing in face of multiple highly persistent shocks (some close to be near-root).

[63] The shock processes estimated for the period (1983-2018) and used in welfare computations rule out the possibility of a shock of the magnitude of the COVID-19 pandemic.

[64] We must generate a sizable decline in the discount rate to match the large jump in household consumption observed following successive rounds of fiscal stimulus. To make the simulation relevant for our context, we suppress the skill acquisition channel in this counterfactual scenario. Without this assumption, households would borrow against their future income to invest heavily in training, which would be counterfactual. See Ikeda et al. (2023) for an application with large shocks to consumer demand achieved with shocks to the discount factor.

Figure 18: COVID-19 simulation: IRF to the CARES demand shock



Note: Blue-dotted line: CARES demand shock, Red-solid line: Same scenario with negative shock to labor supply.

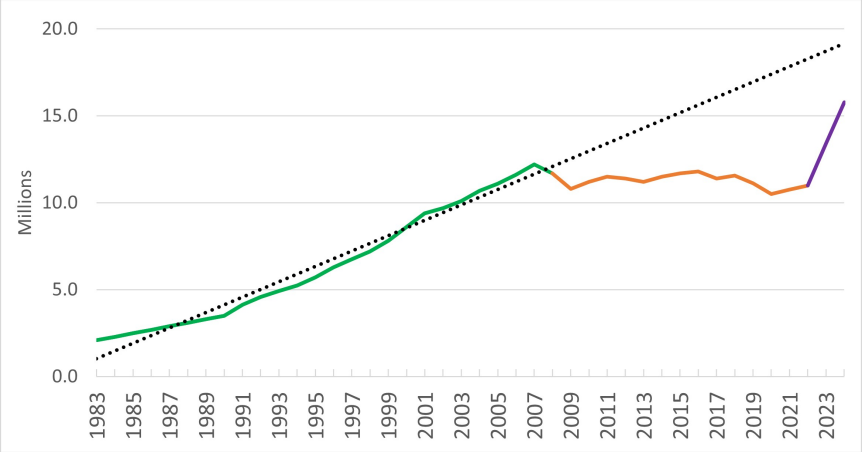
leisure, leading to a decline in the aggregate labor supply. The model's implication is consistent with a phenomenon popularly known as the *Great Resignation* that took place amid generous transfers from the government.[65] The boost to consumption was mostly accommodated by an increase in imported goods, evinced by a sharp decline in the net exports to GDP ratio. The COVID-19 recession was atypical in that it was associated with a notable increase in the trade deficit (from 2.5 percent to 5 percent of GDP) rather than the current account reversal typically linked with economic downturns. The supply of low-skill services was, however, greatly constrained by pervasive shortages of low-skill immigrant labor. Therefore, the positive demand shock increases the price of the non-tradable services and the wages of low-skill workers producing these services. Low-skill immigration does not respond significantly to the higher low-skill wages in Home, however, given the transitory nature of this shock. As a result, the change in the entry of immigrants is limited.

The sluggish recovery in the labor force in the aftermath of the COVID-19 pandemic cannot be solely explained by a higher preference for leisure. Mobility restrictions and fear of contagion of the disease, as new variants of the COVID-19 virus emerged, also had a substantial adverse impact on labor force participation. The fall in participation was particularly acute in the case of low-skill service occupations that require physical contact with consumers and cannot be offshored or performed with remote working (i.e., via telecommunication). The red-solid line in Figure 18 illustrates a scenario in which we simulate

[65] Evidence suggests that overly generous stimulus payments and unemployment benefits triggered an increase in voluntary separations; see Arbogast and Dupor (2022).

a negative labor supply shock to match the observed decrease in the participation rate at the time the stimulus payments. The restricted labor participation, combined with a positive demand shock, generates a spike in low-skill wages and the price of low-skill services, a larger drop in the skill premium, which closely matches the quantitative evidence shown in the introduction–See Figure 2(A)– along with a smaller increase in consumption.

Figure 19: Stock of Undocumented Immigrants in the United States: Current Levels vs. pre-2007 Trend



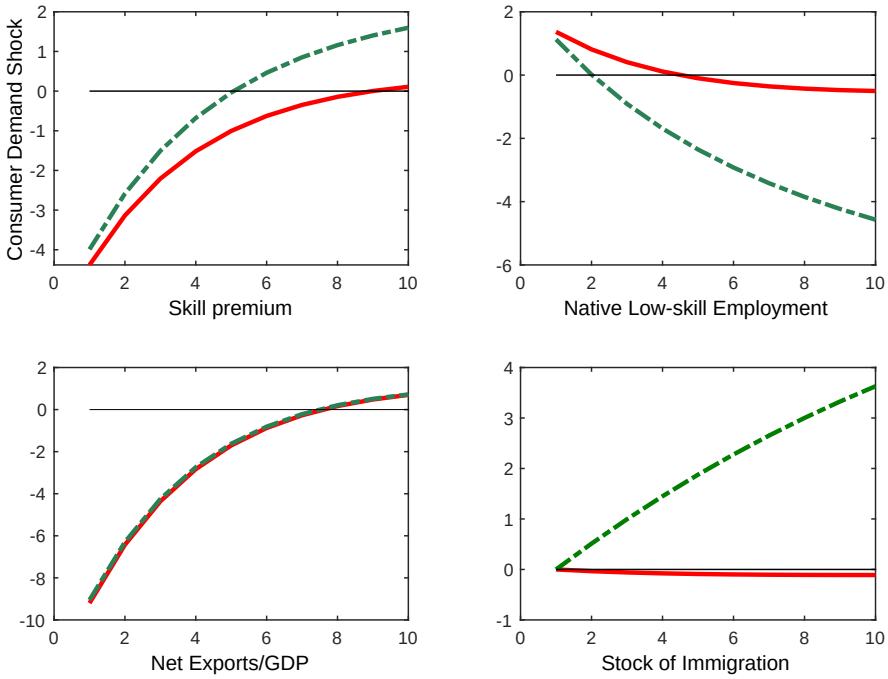
Note: The figure shows estimates for the stock of undocumented immigrants in the United States, in millions at the annual frequency, obtained from PEW Research (1983-2008, in green); the U.S. Department of Homeland Security and the Citizenship and Immigration Services (2009-2022, in orange); and the CBO (2024) (for 2023-2024, in purple), as well as the linear pre-2007 trend (in dotted black).

The 2022-24 surge in immigration. Figure 19 shows the stock of undocumented immigrants in the United States as a proxy for the stock of unskilled immigrant labor updated through 2024 using estimates from the CBO (2024). The stock of undocumented immigrants flattened out around the Global Financial Crisis in 2008, consistent with the slowdown in arrivals recorded by our apprehensions proxy series in Figure 6.[66] The gap between the stock of undocumented immigrants in 2018 and the level implied by the pre-2007 linear trend explains the welfare losses for the Home and South economies reported in rows (1) and (2) of Table 6. More recently, using the CBO (2024) estimates for the net migration of other foreign nationals to extrapolate the stock of undocumented immigrants to 2022-2024 (Fig. 19, in purple), the gap between the actual stock and the pre-2007 trend has narrowed significantly, indicating that the recent surge in low-skill immigration has mitigated partially—but not entirely—the welfare loss estimates in Table 6.

From a normative perspective, a potential policy solution for an acute shortage of low-skill labor is to relax immigration policy and encourage low-skill immigration. Figure 20 replicates the previous scenario (positive demand shock with restricted labor supply) along with a counterfactual increase in low-skill

[66]In the Appendix, we show that our apprehensions series as a proxy for undocumented immigrant arrivals (obtained from the U.S. Customs and Border Patrol) is consistent with the stock of undocumented immigrant labor adjusted for exits (obtained from PEW Research, the U.S. Department of Homeland Security and the Citizenship and Immigration Service, see Figure A14); it is also consistent with the net migration flow estimates for other foreign nationals from the CBO (2024), including for the surge in 2022-2024 (see Figure A15).

Figure 20: COVID-19 simulation: The impact of an increase in immigration



Note: Red-solid line: see Figure 18, Green-dotted line: Same scenario with a policy-induced increase in immigration.

immigrant inflows (i.e., calibrated after the spike in apprehensions during 2021-2022). The red-solid line is the same as the red line in Figure 18. The green-dotted line shows the counterfactual case in which immigration barriers are relaxed, which features a long-term increase in the skill premium and decline in native low-skill employment as natives undergo training. For robustness, in the Appendix we use industry level data to support our modeling assumptions. In sum, the evidence indicates that industries that had acute labor shortages in 2021 experienced in the subsequent years: (a) a larger increase in the employment of immigrants, (b) a faster decline in the number of unfilled vacancies relative to other industries with less severe shortages.

7 Conclusion

We document a sustained decline in immigration and a severe shortage of low-skill workers that began with the Great Recession of 2007. We show that the slowdown in low-skill immigration led to higher low-skill wages and a decline in the skill premium, reversing its four-decade upward trend. The lower skill premium coincided with a decline in the educational attainment of the native-born population. To account for this evidence and study the implications for macroeconomic performance and welfare, we develop and estimate a novel stochastic growth model with endogenous immigration, training, and offshoring decisions. The historical decomposition of the model shows that the initial increase in the skill premium that began in the 1980s can be explained primarily by the decline in immigration and trade

barriers that favored higher-skill workers at the expense of lower-skilled workers. This upward trend in the skill premium was reversed when the shortage of low-skilled immigrant workers, which began with the 2007 immigration slowdown, became severe as labor markets tightened during the mid-2010s.

The estimated structural model allows us to examine the welfare effects from the slowdown in low-skill immigration incurred by the representative household in the destination economy. The main welfare impact is related to low-skill immigrants' occupations in non-tradable services that cannot be offshored, meaning that labor shortages translate directly into higher consumer prices when demand increases. In the quantitative analysis, the lower skill premium and diminishing return to education discourage skill acquisition, which is detrimental to welfare by persistently reducing labor productivity and amplifying the price increase in response to demand expansion.

Our study suggests several promising lines of research on the interaction between immigration and macroeconomic performance. Since our focus is primarily on the interplay between low-skill immigration, the compression of the skill premium, and labor market outcomes, we have abstracted from significant changes in the demographic structure of the labor force. Extending the analysis to examine the effects of declining labor force participation due to population aging would be an important extension for future research. Moreover, the demographic shift toward lower fertility rates in Latin America that began in the 1980s may have altered the dynamics of low-skilled immigration in recent years, with important implications for the U.S. economy. After all, the slowdown in low-skilled immigration documented in our study was abruptly reversed at the end of 2021, with a surge of about six million new immigrants arriving in 2023 and 2024 alone (see CBO, 2024). However, the origins and characteristics of migrant workers are changing, with Latin America's importance as a supplier of low-skill migrant workers to the United States declining sharply over the past four years, replaced by greater internationalization and a broader mix of skill sets among immigrants. The macroeconomic implications of these changes are unclear, but they will certainly play an important role in the U.S. labor market dynamics and living standards in the future.

References

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, volume 4, pages 1043–1171. Elsevier.

Adolfson, M., Laséen, S., Lindé, J., and Villani, M. (2007). Bayesian estimation of an open economy dsge model with incomplete pass-through. *Journal of International Economics*, 72(2):481–511.

Albert, C. and Monràs, J. (2022). Immigration and Spatial Equilibrium: The Role of Expenditures in the Country of Origin. *American Economic Review*, 112(11):3763–3802.

Allen, T., de Castro Dobbin, C., and Morten, M. (2024). Border Walls.

An, S. and Schorfheide, F. (2007). Bayesian Analysis of DSGE Models. *Econometric Reviews*, 26(2-4):113–172.

Andolfatto, D. (1996). Business cycles and labor-market search. *American Economic Review*, pages 112–132.

Arbogast, I. and Dupor, B. (2022). The jobs effect of ending pandemic unemployment benefits: A state-level analysis. *Federal Reserve Bank of St. Louis, Working Papers*.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, 103(5):1553–1597.

Autor, D. H., Dorn, D., and Hanson, G. H. (2013). The China Syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, 103(6):2121–2168.

Autor, D. H., Dorn, D., and Hanson, G. H. (2016). The China Shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics*, 8:205–240.

Bazzi, S., Hanson, G., John, S., Roberts, B., and Whitley, J. (2021). Deterring Illegal Entry: Migrant Sanctions and Recidivism in Border Apprehensions. *American Economic Journal: Economic Policy*, 13(3):1–27.

Borjas, G. J. (2014). *Immigration economics*. Harvard University Press.

Borjas, G. J., Grogger, J., and Hanson, G. H. (2012). Comment: On Estimating Elasticities Of Substition. *Journal of the European Economic Association*, 10(1):198–210.

Bound, J., Khanna, G., and Morales, N. (2017). Understanding the Economic Impact of the H-1B Program on the U.S. *National Bureau of Economic Research, Working Papers*.

Burstein, A., Hanson, G., Tian, L., and Vogel, J. (2020). Tradability and the Labor-Market Impact of Immigration: Theory and Evidence From the United States. *Econometrica*, 88(3):1071–1112.

Burstein, A. and Vogel, J. (2017). International trade, technology, and the skill premium. *Journal of Political Economy*, 125(5):1356–1412.

Cadena, B. C. and Kovak, B. K. (2016). Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession. *American Economic Journal: Applied Economics*, 8(1):257–290.

Caliendo, L., Opromolla, L. D., Parro, F., and Sforza, A. (2021). Goods and factor market integration: A quantitative assessment of the EU enlargement. *Journal of Political Economy*, 129(12):3491–3545.

Card, D. (2001). Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration. *Journal of Labor Economics*, 19(1):22–64.

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

CBO (2024). The Demographic Outlook: 2024 to 2054. *CBO report*.

Charles, K. K., Hurst, E., and Notowidigdo, M. J. (2018). Housing Booms and Busts, Labor Market Opportunities, and College Attendance. *American Economic Review*, 108(10):2947–2994.

Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, 80(3):969–1018.

Cortés, P. (2008). The effect of low-skilled immigration on US prices: evidence from CPI data. *Journal of Political Economy*, 116(3):381–422.

Coulibaly, B., Sapriza, H., and Zlate, A. (2013). Financial frictions, trade credit, and the 2008–09 global financial crisis. *International Review of Economics & Finance*, 26:25–38.

Cravino, J. and Sotelo, S. (2019). Trade-induced structural change and the skill premium. *American Economic Journal: Macroeconomics*, 11:289–326.

Davis, S. J. and Haltiwanger, J. (1990). Gross job creation and destruction: Microeconomic evidence and macroeconomic implications. *NBER Macroeconomics Annual*, 5:123–168.

Di Giovanni, J., Levchenko, A. A., and Ortega, F. (2015). A global view of cross-border migration. *Journal of the European Economic Association*, 13(1):168–202.

East, C. N., Hines, A. L., Luck, P., Mansour, H., and Velásquez, A. (2023). The Labor Market Effects of Immigration Enforcement. *Journal of Labor Economics*, 41(4):957–996.

Firpo, S., Fortin, N. M., and Lemieux, T. (2011). Occupational tasks and changes in the wage structure.

Friedberg, R. M. and Hunt, J. (1995). The impact of immigrants on host country wages, employment and growth. *Journal of Economic Perspectives*, 9(2):23–44.

Georgiadis, G. and Gräb, J. (2016). Growth, real exchange rates and trade protectionism since the financial crisis. *Review of International Economics*, 24(5):1050–1080.

Ghironi, F. and Melitz, M. J. (2005). International trade and macroeconomic dynamics with heterogeneous firms. *The Quarterly Journal of Economics*, 120(3):865–915.

Goldberg, P. K. and Pavcnik, N. (2007). Distributional effects of globalization in developing countries. *Journal of Economic Literature*, 45(1):39–82.

Gramlichv, J. and Scheller, A. (2021). What's happening at the U.S.-Mexico border in 7 charts. *Pew Research Center*.

Grossman, G. M. and Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring. *American Economic Review*, 98(5):1978–1997.

Hamano, M. and Zanetti, F. (2017). Endogenous Turnover and Macroeconomic Dynamics. *Review of Economic Dynamics*, 26:263–279.

Hanson, G. (2006). Illegal Migration from Mexico to the United States. *Journal of Economic Literature*, XLIV(December 2006):869–924.

Hanson, G., Liu, C., and McIntosh, C. (2017). The Rise and Fall of U.S. Low-Skilled Immigration. NBER Working Papers 23753, National Bureau of Economic Research, Inc.

Hanson, G. H. and Spilimbergo, A. (1999). Illegal immigration, border enforcement, and relative wages: Evidence from apprehensions at the US-Mexico border. *American Economic review*, 89(5):1337–1357.

Hauser, D. and Seneca, M. (2022). Labor mobility in a monetary union. *Journal of International Economics*, 137(C).

Hunt, J. (2017). The impact of immigration on the educational attainment of natives. *Journal of Human Resources*, 52(4):1060–1118.

Ikeda, D., Li, S., Mavroeidis, S., and Zanetti, F. (2023). Testing the effectiveness of unconventional monetary policy in Japan and the United States. *American Economic Journal: Macroeconomics*, forthcoming.

Jackson, O. (2015). Does immigration crowd natives into or out of higher education? *Working Paper*.

Jaimovich, N. and Siu, H. E. (2020). Job polarization and jobless recoveries. *Review of Economics and Statistics*, 102(1):129–147.

Kehoe, T. J., Ruhl, K. J., and Steinberg, J. B. (2018). Global imbalances and structural change in the United States. *Journal of Political Economy*, 126(2):761–796.

Kennan, J. (2013). Open borders. *Review of Economic Dynamics*, 16(2):L1–L13.

Kim, J. and Kim, S. H. (2003). Spurious welfare reversals in international business cycle models. *Journal of International Economics*, 60(2):471–500.

Klein, P. and Ventura, G. (2007). Tfp differences and the aggregate effects of labor mobility in the long run. *The B.E. Journal of Macroeconomics*, 7(1):1–38.

Klein, P. and Ventura, G. (2009). Productivity differences and the dynamic effects of labor movements. *Journal of Monetary Economics*, 56(8):1059–1073.

Krishna, P., Poole, J. P., and Senses, M. Z. (2014). Wage effects of trade reform with endogenous worker mobility. *Journal of International Economics*, 93(2):239–252.

Krogstad, J. M., Passel, J. S., and Cohn, D. (2019). 5 facts about illegal immigration in the U.S. *Pew Research Center*.

Krusell, P., Ohanian, L. E., Rios-Rull, J.-V., and Violante, G. L. (2000). Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*, 68(5):1029–1054.

Lessem, R. (2018). Mexico–U.S. Immigration: Effects of Wages and Border Enforcement. *The Review of Economic Studies*, 85(4):2353–2388.

Llull, J. (2018). Immigration, Wages, and Education: A Labour Market Equilibrium Structural Model. *The Review of Economic Studies*, 85(3):1852–1896.

Mandelman, F. S. (2016). Labor market polarization and international macroeconomic dynamics. *Journal of Monetary Economics*, 79:1–16.

Mandelman, F. S. and Zanetti, F. (2014). Flexible prices, labor market frictions and the response of employment to technology shocks. *Labour Economics*, 26(C):94–102.

Mandelman, F. S. and Zlate, A. (2012). Immigration, remittances and business cycles. *Journal of Monetary Economics*, 59(2):196–213.

Mandelman, F. S. and Zlate, A. (2016). The slump in undocumented immigration to the united states. *Macroblog-Federal Rerserve Bank of Atlanta*.

Mandelman, F. S. and Zlate, A. (2017). Economic aspects of international labour migration.

Mandelman, F. S. and Zlate, A. (2022). Offshoring, automation, low-skilled immigration, and labor market polarization. *American Economic Journal: Macroeconomics*, 14(1):355–89.

McHenry, P. (2015). Immigration and the Human Capital of Natives. *Journal of Human Resources*, 50(1):34–71.

Mehra, M. and Kim, D. (2023). Skilled immigration, offshoring, and trade. *Journal of Economic Dynamics and Control*, 146.

Monràs, J. (2020). Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis. *Journal of Political Economy*, 128(8):3017–3089.

Monràs, J., Vázquez-Grenno, J., and Elias, F. (2020). Understanding the Effects of Granting Work Permits to Undocumented Immigrants. *Working Paper*.

Novy, D. (2018). Is the iceberg melting less quickly? international trade costs after world war ii. *Working Paper*.

Oldenski, L. (2014). Offshoring and the Polarization of the U.S. Labor Market. *ILR Review*, 67(3):734–761.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, 10(1):152–197.

Ottaviano, G. I. P., Peri, G., and Wright, G. C. (2013). Immigration, offshoring, and american jobs. *American Economic Review*, 103(5):1925–1959.

Piyapromdee, S. (2021). The Impact of Immigration on Wages, Internal Migration, and Welfare. *Review of Economic Studies*, 88(1):406–453.

Rabanal, P., Rubio-Ramirez, J. F., and Tuesta, V. (2011). Cointegrated tfp processes and international business cycles. *Journal of Monetary Economics*, 58(2):156–171.

Reyes, B. I. (1997). *Dynamics of immigration: Return migration to Western Mexico*. Public Policy Institute of CA.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606.

Stockman, A. C. and Tesar, L. L. (1995). Tastes and technology in a two-country model of the business cycle: Explaining international comovements. *American Economic Review*, 85(1):168–185.

Woodford, M. (2003). *Interest and prices*. Princeton Univ. Press, Princeton, NJ [u.a.].

Zlate, A. (2016). Offshore production and business cycle dynamics with heterogeneous firms. *Journal of International Economics*, 100:34–49.

Slowdown on Immigration, Labor Shortages, and Declining Skill Premia.

TECHNICAL APPENDIX

Federico S. Mandelman, Yang Yu, Francesco Zanetti, Andrei Zlate[1]

This section presents additional materials and results. It includes:

1. The system equations characterizing the equilibrium conditions of the model, where real variables are rescaled to account for the unit-root technology process.
2. Data sources and Bayesian estimation: description of the data sources, the estimation methodology, and the Kalman smoothing procedure.
3. Additional estimation results for the baseline model: the prior and posterior densities of model parameters, Markov Chain Monte Carlo (MCMC) multivariate convergence diagnostics, impulse responses and variance decomposition.
4. Additional charts for the series of shocks to the iceberg trade cost, the sunk migration cost, and the sunk training costs relevant for the historical decomposition in Section 5.3; for an alternative series of the real tuition inflation rate; and charts showing the consistency of our apprehensions series as a proxy for undocumented immigrant arrivals with alternative measures, such as the stock of undocumented immigrant labor from the U.S. Citizenship and Immigration Service or the net migration flow estimates for other foreign nationals from the Congressional Budget Office.
5. Additional results for regional regressions.
6. Industry Evidence: Immigration and job vacancies during 2021-2024.
7. The Role of skill-biased technological change.
8. Southern household modeling assumptions.

[1]The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Banks of Atlanta, The Board of Governors or the Federal Reserve System.

1 Normalized Model Equations

The presence of a unit-root global technology shock makes the model non-stationary. Therefore, some of the real variables–i.e., mostly those expressing quantities such as output, consumption, and bond holdings–are re-scaled by the world productivity $\mathbb{X}_t$ to become stationary. In the equations below, the variables marked with hats are subject to such a normalization. For instance, the normalized total value added in Home is $\hat{Y}_t = \frac{Y_t}{\mathbb{X}_t}$. A similar normalization holds for the remaining variables in Home, Foreign, and the South. Notice that employment and prices are stationary, so they are not rescaled. In what follows, we provide the equations for Home and South. For Foreign, the equations and variables are similar to those in Home, except that there is no immigration from South into Foreign. Variables for Foreign and South are marked with an asterisk and a s superscript, respectively.

Equations for Home The average real wages in the middle-skill and high-skill occupations:

$$\tilde{w}_{D,t} = w_{D,t}(\tilde{\mathbf{z}}_{D,t},.) = \frac{\theta}{\theta-1}\frac{\hat{w}_{\mathbf{u},t}}{\varepsilon_t^z v}, \text{ where } v = \left[\frac{k}{k-(\theta-1)}\right]^{\frac{1}{(\theta-1)}} \tag{A1}$$

$$\tilde{w}_{X,t} = w_{X,t}(\tilde{\mathbf{z}}_{X,t},.) = \frac{1}{\mathbb{Q}_t}\left(\tau\varepsilon_t^{\tau}\right)\frac{\theta}{\theta-1}\frac{\hat{w}_{\mathbf{u},t}}{\varepsilon_t^z \tilde{\mathbf{z}}_{X,t}} \tag{A2}$$

The average skill income premia for the tasks executed and delivered domestically ($\tilde{\pi}_{D,t}$, which includes both middle-skill and high-skill occupations) and for tasks executed domestically and delivered to Foreign ($\tilde{\pi}_{X,t}$, which includes only high-skill occupations):

$$\tilde{\pi}_{D,t} = \hat{\pi}_{D,t}(\tilde{\mathbf{z}}_{D,t},.) = \frac{1}{\theta}\left(\tilde{w}_{D,t}\right)^{1-\theta}\hat{\mathbb{N}}_t \tag{A3}$$

$$\tilde{\pi}_{X,t} = \hat{\pi}_{X,t}(\tilde{\mathbf{z}}_{X,t},.) = \mathbb{Q}_t\frac{1}{\theta}\left(\tilde{w}_{X,t}\right)^{1-\theta}\hat{\mathbb{N}}_t^* - \hat{f}_{o,t} \tag{A4}$$

where $\hat{f}_{o,t} = \frac{\hat{w}_{\mathbf{u},t} f_o}{\varepsilon_t^z}$ is the fixed cost of offshoring, and $\hat{\mathbb{N}}_t$ and $\hat{\mathbb{N}}_t^*$ are the demand for the composite of tradable tasks in Home and Foreign.

The average skill income premium in middle-skill and high-skill occupations:

$$\tilde{\pi}_t = \frac{\left(N_{D,t}\tilde{\pi}_{D,t} + N_{X,t}\tilde{\pi}_{X,t}\right)}{N_{D,t}} \tag{A5}$$

The wage bill:

$$\mathbb{W}_t = \left[N_{D,t}\left(\tilde{w}_{D,t}\right)^{1-\theta} + N_{X,t}^*\left(\tilde{w}_{X,t}^*\right)^{1-\theta}\right]^{\frac{1}{1-\theta}}, \text{ where } \mathbb{W}_t \equiv 1 \text{ is set as the numeraire.} \tag{A6}$$

The sunk training cost in units of the numeraire:

$$\hat{f}_{j,t} = \frac{\hat{w}_{\mathbf{u},t}\left(f_j\varepsilon_t^{Tr}\right)^{\Theta_{fj}}}{\varepsilon_t^z} \tag{A7}$$

The share of offshoring occupations and the offshoring equilibrium condition:

$$\frac{N_{X,t}}{N_{D,t}} = \left(\frac{1}{\tilde{\mathbf{z}}_{X,t}}\right)^k v^k \tag{A8}$$

$$\tilde{\pi}_{X,t} = \hat{f}_{o,t}\frac{\theta-1}{k-(\theta-1)} \tag{A9}$$

where the average productivity in high-skill occupations $\tilde{\mathbf{z}}_{X,t} = \mathbf{z}_{X,t}v$ is a function of the offshoring productivity cutoff $\mathbf{z}_{X,t}$.

The law of motion for the number of trained occupations (job turnover):

$$N_{D,t} = (1-\delta)(N_{D,t-1} + N_{E,t-1}) \tag{A10}$$

Household optimality conditions (i.e., marginal utility of consumption, labor supply, training decision, and the Euler equation for bonds):

$$\hat{\zeta}_t = \frac{\varepsilon_t^b \left(\hat{C}_t\right)^{-\gamma}}{P_t}, \text{ where } \hat{\zeta}_t = \zeta_t(\mathbb{X}_t)^{\gamma} \tag{A11}$$

$$a_n \left(L_t\right)^{\gamma_n} \left(\hat{C}_t\right)^{\gamma} = \frac{\hat{w}_{\mathbf{u},t}}{P_t} \tag{A12}$$

$$\hat{f}_{j,t} = \beta \left(1-\delta\right) \mathbb{E}_t \left[\frac{\hat{\zeta}_{t+1}}{\hat{\zeta}_t}(\hat{f}_{j,t+1} + \widetilde{\pi}_{t+1}) \left(\frac{\mathbb{X}_{t+1}}{\mathbb{X}_t}\right)^{1-\gamma}\right] \tag{A13}$$

$$q_t = \beta \mathbb{E}_t \left[\left(\frac{\mathbb{X}_{t+1}}{\mathbb{X}_t}\right)^{-\gamma} \frac{\hat{\zeta}_{t+1}}{\hat{\zeta}_t}\right] - \phi \hat{B}_t \tag{A14}$$

Uncovered interest rate parity condition:

$$\mathbb{E}_t \left[\frac{\hat{\zeta}^*_{t+1}}{\hat{\zeta}^*_t} \left(\frac{\mathbb{X}_{t+1}}{\mathbb{X}_t}\right)^{-\gamma} \frac{\mathbb{Q}_t}{\mathbb{Q}_{t+1}}\right] = \mathbb{E}_t \left[\frac{\hat{\zeta}_{t+1}}{\hat{\zeta}_t} \left(\frac{\mathbb{X}_{t+1}}{\mathbb{X}_t}\right)^{-\gamma}\right] - \phi \frac{\hat{B}_t}{\beta} \tag{A14a}$$

Aggregate accounting and current account:

$$q_t \hat{B}_t = \left(\frac{\mathbb{X}_{t-1}}{\mathbb{X}_t}\right) \hat{B}_{t-1} + \hat{w}_{\mathbf{u},t} L_t + N_{D,t} \widetilde{\pi}_t - P_t \hat{C}_t - \hat{f}_{e,t} N_{E,t} - \frac{\phi}{2} \hat{B}_t^2 \tag{A15}$$

$$q_t \hat{B}_t = \left(\frac{\mathbb{X}_{t-1}}{\mathbb{X}_t}\right) \hat{B}_{t-1} + \mathbb{Q}_t N_{X,t} \left(\tilde{w}_{X,t}\right)^{1-\theta} \hat{\mathbb{N}}^*_t - N^*_{X,t} \left(\tilde{w}^*_{X,t}\right)^{1-\theta} \hat{\mathbb{N}}_t - \frac{\phi}{2} \hat{B}_t^2 \tag{A16}$$

Income-based GDP in units of the consumption good:

$$\hat{Y}_t = \frac{(N_{D,t} \widetilde{\pi}_t + \hat{w}_{\mathbf{u},t} L_t + \hat{w}_{\mathbf{i},t} L^s_{\mathbf{i},t})}{P_t}. \tag{A17}$$

The production for the traded aggregate and the non-traded goods, as well as the relative demand for low-skill native and immigrant labor:

$$\hat{Y}_{T,t} = \hat{\mathbb{N}}_t \tag{A18}$$

$$\hat{Y}_{N,t} = L^A_{N,t} = \left[\alpha \left(L_{N,t}\right)^{\frac{\sigma_N - 1}{\sigma_N}} + (1-\alpha) \left(L^s_{\mathbf{i},t}\right)^{\frac{\sigma_N - 1}{\sigma_N}}\right]^{\frac{\sigma_N}{\sigma_N - 1}} \tag{A19}$$

$$\frac{\hat{w}_{\mathbf{u},t}}{P_{N,t}} = \alpha \left(\frac{L_{N,t}}{\hat{Y}_{N,t}}\right)^{-\frac{1}{\sigma_N}} \tag{A20}$$

$$\frac{\hat{w}_{\mathbf{i},t}}{P_{N,t}} = (1-\alpha) \left(\frac{L^s_{\mathbf{i},t}}{\hat{Y}_{N,t}}\right)^{-\frac{1}{\sigma_N}} \tag{A21}$$

Since there is no immigration into Foreign, $\hat{Y}^*_{N,t} = L^*_{N,t}$ and $P^*_{N,t} = \hat{w}^*_{\mathbf{u},t}$.
The The Consumer Price Index (CPI), consumption basket, and relative demand:

$$P_t = \left[\gamma_c \left(P_{T,t}\right)^{1-\rho_c} + (1-\gamma_c) \left(P_{N,t}\right)^{1-\rho_c}\right]^{\frac{1}{1-\rho_c}}, \text{ where } P_{T,t} = \mathbb{W}_t \equiv 1 \text{ is the numeraire.} \tag{A22}$$

$$\hat{C}_t = \left[\gamma_c{}^{\frac{1}{\rho_c}} \left(\hat{C}_{T,t}\right)^{\frac{\rho_c - 1}{\rho_c}} + (1-\gamma_c)^{\frac{1}{\rho_c}} \left(\hat{Y}_{N,t}\right)^{\frac{\rho_c - 1}{\rho_c}}\right]^{\frac{\rho_c}{\rho_c - 1}} \tag{A23}$$

$$\frac{\hat{C}_{T,t}}{\hat{Y}_{N,t}} = \frac{\gamma_c}{1-\gamma_c} \left(\frac{1}{P_{N,t}}\right)^{-\rho_c} \tag{A24}$$

Equations for the Southern economy Free entry condition and the sunk emigration cost:

$$\hat{f}_{e,t} = \frac{\hat{w}_{\mathbf{i},t}\left(f_e \varepsilon_t^{f_e}\right)}{\varepsilon_t^z} \tag{A25}$$

Law of motion of the stock of immigrant labor:

$$L^s_{\mathbf{i},t} = (1-\delta_l)(L^s_{\mathbf{i},t-1} + L^s_{\mathbf{e},t}) \tag{A26}$$

Household budget constraint:

$$\hat{w}_{\mathbf{i},t} L^s_{\mathbf{i},t} + \hat{w}^s_{\mathbf{u},t}\left(L^s_{\mathbf{u},t} - L^s_{\mathbf{i},t}\right) = \hat{f}_{e,t} L^s_{\mathbf{e},t} + P^s_t \hat{C}^s_t, \tag{A27}$$

Household optimality conditions (i.e., marginal utility of consumption, labor supply, and the emigration decision) :

$$\hat{\zeta}^s_t = \frac{\left(\hat{C}^s_t\right)^{-\gamma}}{P^s_t}, \text{ where } \hat{\zeta}^s_t = \zeta^s_t(\mathbb{X})^\gamma \tag{A28}$$

$$a^s_n \left(L^s_{\mathbf{u},t}\right)^{\gamma^s_n} \left(\hat{C}^s_t\right)^\gamma = \frac{\hat{w}^s_{\mathbf{u},t}}{P^s_t} \tag{A29}$$

$$\hat{f}_{e,t} = \beta(1-\delta)\,\mathbb{E}_t\left[\frac{\hat{\zeta}^s_{t+1}}{\hat{\zeta}^s_t}\left(\hat{f}_{e,t+1} + (\hat{w}_{\mathbf{i},t+1} - \hat{w}^s_{\mathbf{u},t+1})\right)\left(\frac{\mathbb{X}_{t+1}}{\mathbb{X}_t}\right)^{1-\gamma}\right] \tag{A30}$$

Production of the non-traded good:

$$\hat{C}^s_{N,t} = \varepsilon^s_t\left(L^s_{\mathbf{u},t} - L^s_{\mathbf{i},t}\right) \tag{A31}$$

$$P^s_t = \frac{\hat{w}^s_{\mathbf{u},t}}{\varepsilon^s_t} \tag{A32}$$

Consumption of the traded aggregate good from Home:

$$\hat{C}^s_{T,t} = \hat{Y}_{T,t} - \hat{C}_{T,t}$$

Aggregate consumption:

$$\hat{C}^s_t = \left[(\gamma^s_c)^{\frac{1}{\rho^s_c}}\left(\hat{C}^s_{T,t}\right)^{\frac{\rho^s_c-1}{\rho_c}} + (1-\gamma^s_c)^{\frac{1}{\rho^s_c}}\left(\hat{C}^s_{N,t}\right)^{\frac{\rho^s_c-1}{\rho_c}}\right]^{\frac{\rho^s_c}{\rho^s_c-1}} \tag{A33}$$

$$P^s_t = \left[\gamma^s_c + (1-\gamma^s_c)\left(P^s_{N,t}\right)^{1-\rho^s_c}\right]^{\frac{1}{1-\rho^s_c}}, \text{ since } P_{T,t} \equiv 1 \tag{A34}$$

$$\frac{\hat{C}^s_{T,t}}{\hat{C}^s_{N,t}} = \frac{\gamma^s_c}{1-\gamma^s_c}\left(\frac{1}{P^s_{N,t}}\right)^{-\rho^s_c} \tag{A35}$$

Employment and income shares by skill group in Home The number of high-skill occupations is $N_{X,t}$; the number of middle-skill occupations is $N_{M,t} = N_{D,t} - N_{X,t}$. Aggregate hours for low-skill employment (natives and immigrants) are:

$$L^{Aggr}_{N,t} = L_{N,t} + L^s_{\mathbf{i},t} \tag{A36}$$

Income share of high-skill labor:

$$Share_{H,t} = \frac{N_{X,t}(\widetilde{\pi}_{X,t} + \widetilde{\pi}_{DX,t}) + L_t^{\tilde{\mathbf{z}}_{X,t}} \hat{w}_{\mathbf{u},t}}{\hat{Y}^{\mathbf{W}}_t}, \tag{A37}$$

where $\widetilde{\pi}_{X,t}$ (defined above) is the skill income premium for tasks executed in high-skill occupations that are actually offshored; $\widetilde{\pi}_{DX,t}$ is the skill income premium for the tasks executed in high-skill occupations that are suitable to be offshored (i.e. with productivity above the threshold value $\mathbf{z}_{X,t}$) but are executed for the domestic market; $L_t^{\tilde{\mathbf{z}}_{X,t}}$ is the total units of raw labor embodied in tasks executed in high-skill occupations and sold to Foreign; and Y_t^{W} is the income-based GDP net of training costs expressed in terms of the numeraire. All these are defined below:

$$\widetilde{\pi}_{DX,t} = \frac{1}{\theta}\left(\frac{\theta}{\theta-1}\frac{\hat{w}_{\mathbf{u},t}}{\varepsilon_t^z \tilde{\mathbf{z}}_{X,t}}\right)^{1-\theta}\hat{\mathbb{N}}_t \tag{A38}$$

$$L_t^{\tilde{\mathbf{z}}_{X,t}} = N_{X,t}\left[(\theta-1)\frac{\widetilde{\pi}_{DX,t}}{\hat{w}_{\mathbf{u},t}} + (\theta-1)\left(\frac{\widetilde{\pi}_{X,t}}{\hat{w}_{\mathbf{u},t}} + \frac{f_o}{\varepsilon_t^z}\right) + \frac{f_o}{\varepsilon_t^z}\right] \tag{A39}$$

$$\hat{Y}_t^{\mathrm{W}} = P_t\hat{Y}_t - \hat{f}_{j,t}N_{E,t} \tag{A40}$$

Income share for the low-skill labor input, including both natives and immigrants:

$$Share_{L,t} = \frac{\hat{w}_{\mathbf{u},t}L_{N,t} + \hat{w}_{\mathbf{i},t}L_{\mathbf{i},t}^s}{\hat{Y}_t^{\mathrm{W}}} \tag{A41}$$

Consequently, the income share of middle-skill labor is:

$$Share_{M,t} = 1 - Share_{H,t} - Share_{L,t}. \tag{A42}$$

Additional definitions for Home The CPI-based real exchange rate is:

$$RER_t = \frac{P_t^*\mathbb{Q}_t}{P_t} \tag{A43}$$

Net exports-to-GDP ratio:

$$\frac{NX_t}{GDP_t} = \frac{\mathbb{Q}_t N_{X,t}\left(\tilde{w}_{X,t}\right)^{1-\theta}\hat{\mathbb{N}}_t^* - N_{X,t}^*\left(\tilde{w}_{X,t}^*\right)^{1-\theta}\hat{\mathbb{N}}_t}{P_t\hat{Y}_t} \tag{A44}$$

Exports-to-GDP ratio:

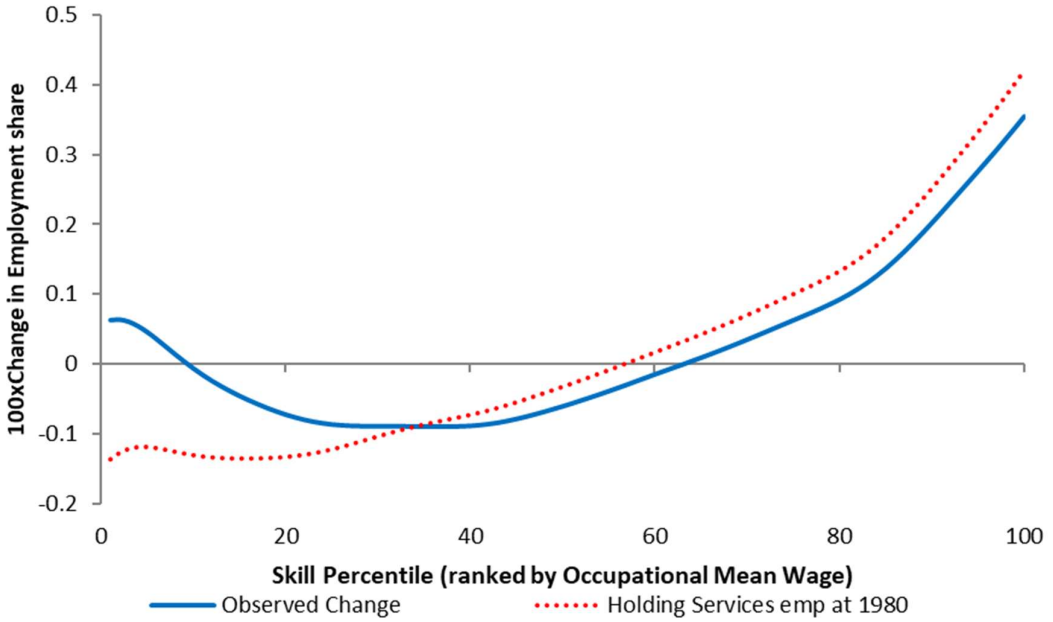
$$\frac{EX_t}{GDP_t} = \frac{\mathbb{Q}_t N_{X,t}\left(\tilde{w}_{X,t}\right)^{1-\theta}\hat{\mathbb{N}}_t^*}{P_t\hat{Y}_t} \tag{A45}$$

2 Data Sources and Bayesian Estimation

Data sources for Fig. 1 The panels in Fig. 1 are constructed following the methodology in Autor and Dorn (2013). We use data from the American Community Survey (which includes 1% of the population) and the IPUMS census data (5% of the population) for the years 2021, 2007, 1980, respectively. Occupations are sorted into 100 percentiles based on the mean occupational wages in 1980.[2] The employment shares are computed for each occupation and then are aggregated at the percentile level. The change in employment shares is obtained as the simple difference between the share of employment for each year considered for the occupations in each percentile. For completeness, we consider a counterfactual scenario in which employment changes are calculated assuming that employment in all *service occupations* remains at the level of 1980. Mimicking the methodological approach in Autor and Dorn (2013), this counterfactual is constructed by pooling ACS data from 2007 with census data from 1980. This approach consists of estimating a weighted logit model for the odds, from which an observation is drawn from the 1980 census sample (relative to the actual sampling year), using as predictors a service occupation dummy and an intercept. Weights used are the product of census sampling weights and annual hours of labor supply. Observations in 2007 are re-weighted using the estimated odds multiplied by the hours-weighted census sampling weight, weighting downward the frequency of service occupations in 2007 to their 1980 level. Given the absence of other covariates in the model, the extra probability mass is implicitly allocated uniformly over the remainder of the distribution.

[2] As discussed in Acemoglu and Autor (2011), the ordering does not change significantly if a different base year is used.

Figure A1: Changes in Employment: The Role of Non-tradable Service Occupations



Note: Changes in employment by skill percentile. Red-dotted line: counterfactual in which *Service Occupations* remain constant at its 1980 level

Data sources for model estimation To estimate the model, we use data series for the interval from 1983:Q1 to 2018:Q4. First, we use the U.S. real GDP as a proxy for Home GDP; real GDP in the rest of the world as a proxy for Foreign GDP, which is constructed as a trade-weighted aggregate of the U.S. major trade partners; and Mexico's real GDP as a proxy for the South GDP.[3] The U.S. Census employment data used in Fig. 1 are decennial and thus not available on a high-frequency basis. In addition, the census data cannot be split easily into the three skill groups needed for the system estimation. Therefore, we closely follow Jaimovich and Siu (2020), and use the Current Population Survey from the Bureau of Labor Statistics to construct quarterly time series of employment by skill group. We consider three categories of employment based on the skill content of the tasks executed by each occupation in the Census data: Non-Routine Cognitive (high-skill), Routine (middle-skill), and Non-Routine Manual (low-skill).[4] This classification is based on the categorization of occupations in the 2000 Standard Occupational classification system. *Non-routine cognitive* workers are in "management, business, and financial operations occupations" and "professional and related occupations." *Routine cognitive* workers are those in "sales and related occupations" and "office and administrative support occupations." *Routine manual* occupations are "production occupations", "transportation and material moving occupations," and "installation, maintenance, and repair occupations." *Non-routine manual* occupations are "service occupations" and "construction and extraction occupations." As explained in Jaimovich and Siu (2020) and Firpo et al. (2011), this group classification corresponds to rankings in the occupational income distribution: non-routine cognitive occupations tend to be high-skill occupations whereas non-routine manual occupations tend to be low-skill. Routine occupations both cognitive and manual are middle-skill occupations. The data are seasonally adjusted with the X-12 ARIMA method from the U.S. Census Bureau.

The categorization of occupations in our paper is slightly different than that in Jaimovich and Siu (2020). Specifically, construction occupations are grouped among those providing low-skill/non-tradable tasks, for two reasons. First, construction jobs are intrinsically non-tradable and thus not subject to offshorability. Second, even though the average hourly earnings of construction workers belong to the middle (and not the bottom) of earnings distributions in the CPS classification, some important caveats exist. The underground economy is particularly pervasive

[3]The U.S. trade partners included are: among the advanced economies, Australia, Canada, the euro area (Germany, France, Italy, Netherlands, Belgium, Spain, Ireland, Austria, Finland, Portugal, Greece), Japan, Sweden, Switzerland and the U.K.; among the emerging markets, China, India, Hong Kong, Taiwan, Korea, Singapore, Indonesia, Malaysia, Philippines, Thailand, Mexico, Brazil, Argentina, Venezuela, Chile, Colombia, Israel, Russia and Saudi Arabia. The data are collected from Haver Analytics.

[4]Jaimovich and Siu (2012) show that their classification in three groups is consistent with the analysis in Autor and Dorn (2012), which provides a more comprehensive definition of six categories based on an occupation's degree of intensity in abstract, routine, and manual tasks, respectively.

in this sector. Construction is densely populated by low-skill laborers who execute non-routine manual tasks that hardly can be mechanized. Many contractors are unregistered workers, and many of the registered ones subcontract by hiring hourly low-wage laborers without keeping records.[5] Having said this, the model implications are somewhat similar when construction occupations are included within the middle-skill segment.

To evaluate the model fit, we build and use two series that serve as proxies for (i) the inflows of low-skill migrant workers and (ii) the cost of offshoring. The series of apprehensions at the U.S.-Mexico border are constructed as follows: For January 1980 to September 2004, we use monthly data on apprehensions at the U.S.-Mexico border provided by the U.S. Immigration and Naturalization Service and made available on Gordon Hanson's website ("border linewatch apprehensions"). For October 1998 to September 2019, we use monthly data on apprehensions provided by the U.S. Border Patrol. We seasonally-adjust the monthly series and convert them to quarterly values using a cubic spline.

Estimation Methodology This section briefly explains the estimation approach used in this paper. A more detailed description of the method can be found in Justiniano and Preston (2010), among others. Let's define Θ as the parameter space of the DSGE model, and $z^T = \{z_t\}_{t=1}^T$ as the data series used in the estimation. Their joint probability distribution, $P(z^T, \Theta)$, results in a relationship between the marginal, $P(\Theta)$, and the conditional distribution $P(z^T|\Theta)$, which is known as the Bayes theorem: $P(\Theta|z^T) \propto P(z^T|\Theta)P(\Theta)$. The method updates the a prior distribution using the likelihood to obtain the conditional posterior distribution of the structural parameters in the data. The resulting posterior density $P(\Theta|z^T)$, is used to draw statistical inference on the parameter space, Θ. Combining the state-form representation implied by the solution for the linear rational expectation model and the Kalman filter, we can compute the likelihood function. The likelihood and the prior permit a computation of the posterior that can be the starting value of the random walk version of the Metropolis-Hastings (MH) algorithm, which is a Monte Carlo method that generates draws from the posterior distribution of the parameters. In this case, the results reported are based on 100,000 draws from this algorithm. We choose a normal jump distribution with covariance matrix equal to the Hessian of the posterior density evaluated at the maximum. The scale factor is chosen to deliver an acceptance rate between 35% and 50% depending on the run of the algorithm. Measures of uncertainty follow from the percentiles of the draws.

Smoothing The DSGE model can be written in a state-space representation as $\xi_{t+1} = F\xi_t + v_{t+1}$ and $z_t = H'\xi_t + w_t$, in which ξ_t is the vector of unobserved variables at date t, and z_t is the vector of observables; shocks v_t and w_t are uncorrelated, normally distributed, white noise vectors. The first expression is the *state* equation, and the second is the *observed* equation.

Smoothing involves the estimation of $\xi^T = \{\xi_t\}_{t=1}^T$, conditional on the full data set, z^T, used in the estimation. The smoothed estimates are denoted as $\xi_{t|T} = E(\xi_t|z^T)$ and, as shown in Bauer et al. (2003), can be written as:

$$\xi_{t|T} = \xi_{t|t} + P_{t|t}F'P_{t+1|t}^{-1}\left[\xi_{t+1|T} - \xi_{t+1|t}\right], \tag{A46}$$

in which $P_{t+1|t} = E(\xi_{t+1} - \xi_{t+1|t})(\xi_{t+1} - \xi_{t+1|t})'$ is the mean squared forecasting error associated with the projection of ξ_{t+1} on z^t and a constant, projection which is denoted as $\xi_{t+1|t} = E(\xi_{t+1}|z^t)$. Using the Kalman filter to calculate, $\{\xi_t\}_{t=1}^T$, $\left\{\xi_{t+1|t}\right\}_{t=0}^{T-1}$, $\left\{P_{t|t}\right\}_{t=1}^T$, and $\left\{P_{t+1|t}\right\}_{t=0}^{T-1}$, the sequence of smooth estimates, $\left\{\xi_{t|T}\right\}_{t=1}^T$, is determined from equation (A46).

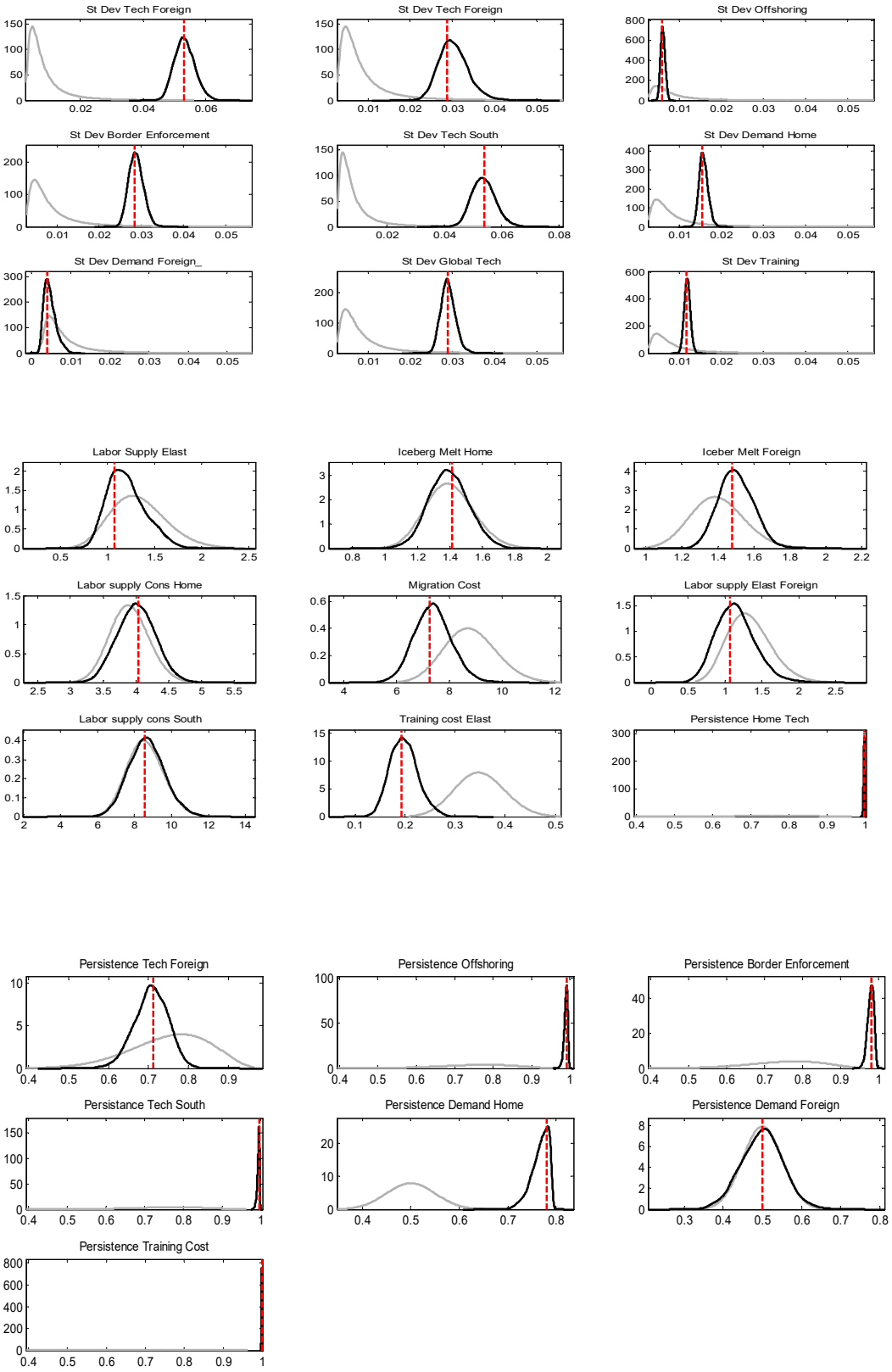
3 Additional Estimation Results for Baseline Model

Prior and posterior density Fig. A2 shows the prior density (grey line), posterior density (black line), and the mode (red line) from the numerical optimization of the posterior kernel for the benchmark model. These results complement those reported in Table 1 of the paper.

Convergence diagnostics Fig. A3 shows the convergence of iterative simulations with the multivariate diagnostic methods described in Brooks and Gelman (1998). The empirical 80% interval for any given parameter,

[5] For instance, a FPI report (2007) shows that despite the residential construction boom of the early 2000s in the New York City metropolitan area in which construction permits more than doubled, there was negligible increase in the official count of the New York City residential construction workers (which contradicts the evidence). In a related paper, Hotchkiss and Quispe-Agnoli (2012) find that the construction industry is, proportionally, the largest employer of undocumented immigrants. See also Cebula and Feiger (2011).

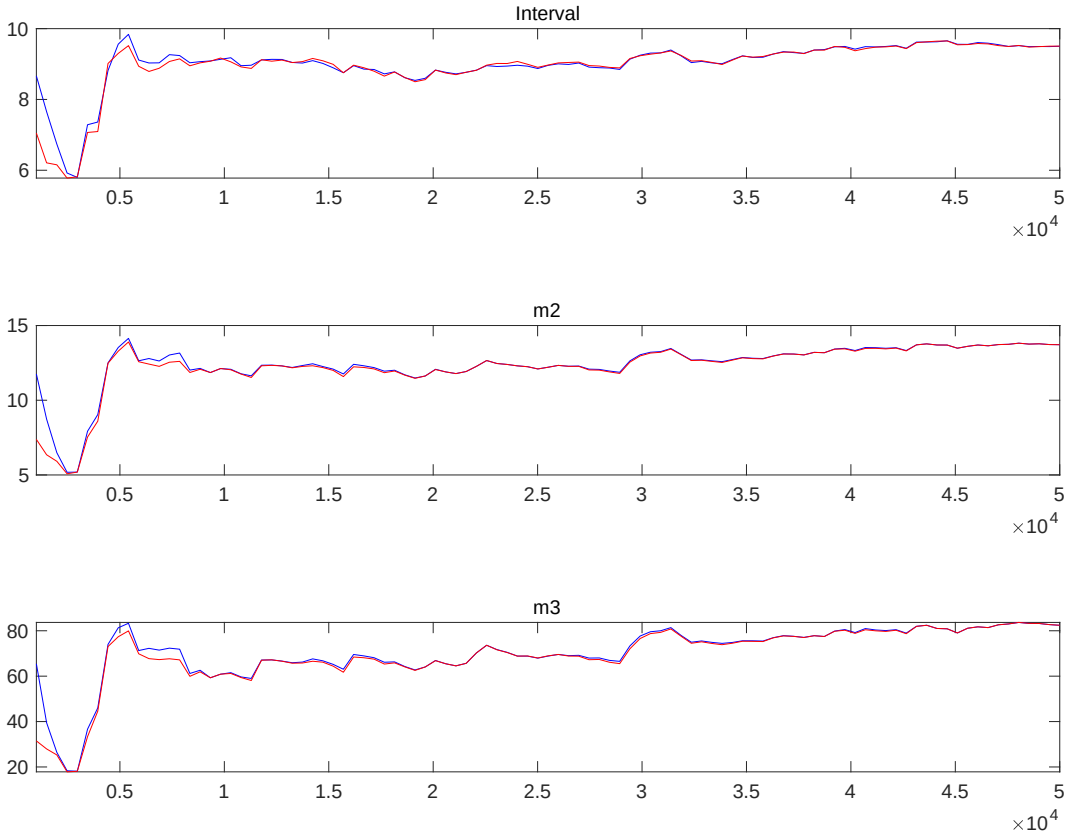
Figure A2: Prior and posterior distributions



Note: This figure shows the prior density (grey line), posterior density (black line), and the mode (red line) from the numerical optimization of the posterior kernel for the benchmark model.

ϱ, is first taken from each individual chain. The interval is described by the 10% and 90% of the n simulated draws. In this multivariate approach, ϱ is defined as a vector parameter based upon observations, $\varrho_{jt}^{(i)}$, denoting the i_{th} element of the parameter vector in chain j at time t. The direct analogue of the univariate approach in higher dimensions is to estimate the posterior variance-covariance matrix as: $\hat{V} = \frac{n-1}{n}W + (1 + \frac{1}{m})B/n$, where $W = \frac{1}{m(n-1)}\sum_{j=1}^{m}\sum_{t=1}^{n}(\varrho_{jt} - \bar{\varrho}_{j}.)(\varrho_{jt} - \bar{\varrho}_{j}.)'$ and $B/n = \frac{1}{m-1}\sum_{j=1}^{m}(\bar{\varrho}_{j.} - \bar{\varrho}_{..})(\bar{\varrho}_{j.} - \bar{\varrho}_{..})'$. It is possible to summarize the distance between $\hat{V}$ and W with a scalar measure that should approach 1 (from above) as convergence is achieved, given suitably over-dispersed starting points. We can monitor both $\hat{V}$ and W, determining convergence when any rotationally invariant distance measure between the two matrices indicates that they are sufficiently close. Fig. A4 reports measures of this aggregate.[6] Convergence is achieved before 100,000 iterations. General univariate diagnostics are not displayed, but they are available on request.

Figure A3: Convergence diagnostics



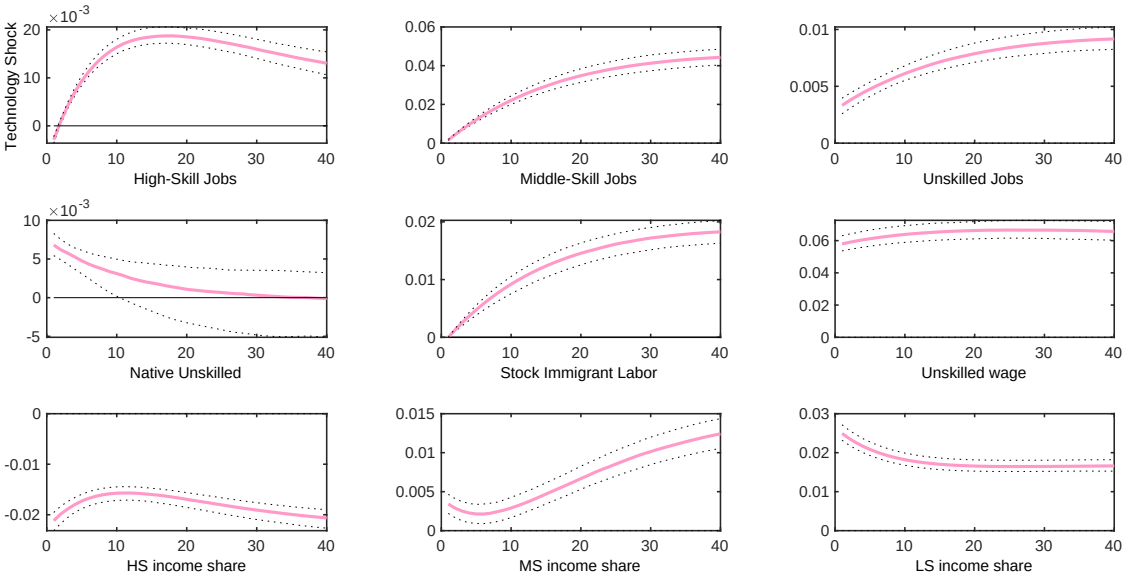
Note: This figure shows the convergence of iterative simulations with the multivariate diagnostic methods described in Brooks and Gelman (1998).

Impulse responses Figures A4-A8 show additional impulse responses for technology and demand shocks. They are consistent with the model implications discussed in the paper. In Fig. A4, a positive technology shock in the Home "tradable" sector boosts the number of high-skill and middle-skill occupations and encourages task/skill upgrading among natives. As low-skilled native employment declines, immigration from the South is enhanced. In Fig. A5, a positive technology shock in foreign tradables tasks leads to a decrease in the number high-skill occupations in Home, which are substituted with relative more productive high- and middle-skilled foreign tasks. In contrast, the number of middle-skill occupations in Home which provide tasks only domestically increases. Since Home consumption receives a boost from the higher productivity in Foreign, the existing complementarity between goods and services prompts an increase in the low-skill native employment and immigration into Home. In Fig. A6, a positive technology shock in the Southern economy, where the immigrant labor originates, leads to a decrease

[6]Note that, for instance, the interval-based diagnostic in the univariate case becomes now a comparison of volumes of total and within-chain convex hulls. Brooks and Gelman (1998) propose to calculate for each chain the volume within 80%, say, of the points in the sample and compare the mean of these with the volume from 80% of the observations from all samples together.

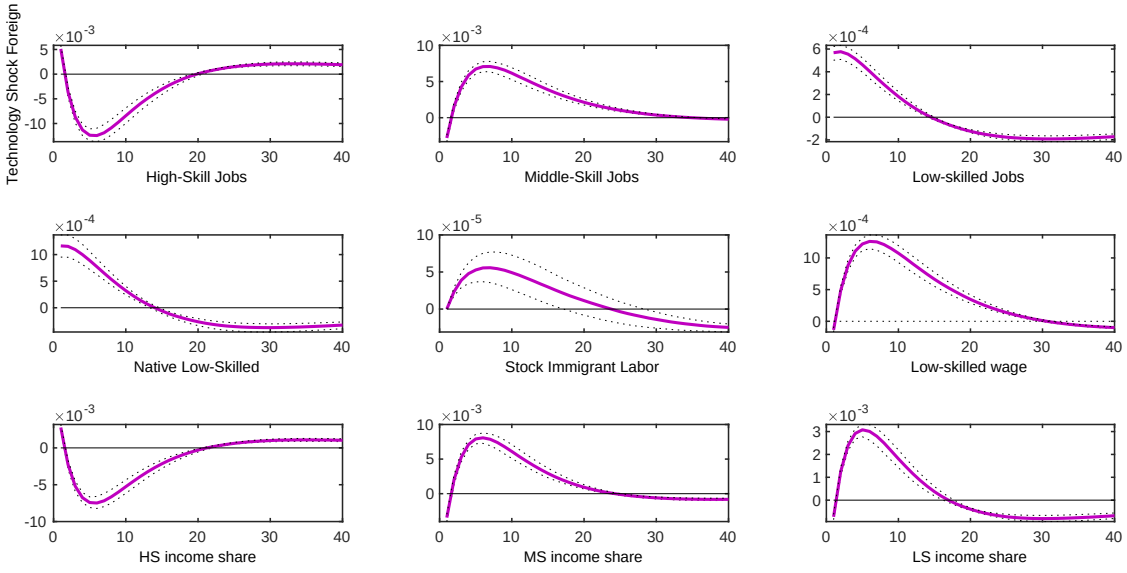
in the stock of immigrant labor in Home. The lower supply of immigrant labor also causes "task downgrading" in Home, i.e., the native workers reduce training, which leads to a decrease in the number of high-skill and middle-skill occupations, as well as to an increase in the low-skilled native employment in Home. In Fig. A7, a positive demand shock in Home, which encourages current consumption at the expense of future consumption, leads to an increase the number of high-skill occupations in Foreign (not shown), and–due to complementarity between goods and services–also to an increase in the low-skill native employment in Home. Similarly, in Fig. A8, a positive demand shock in Foreign leads to an increase in the number of high-skill occupations in Home and an decrease in low-skilled native employment in Home.

Figure A4: Impulse response to Home technology shock



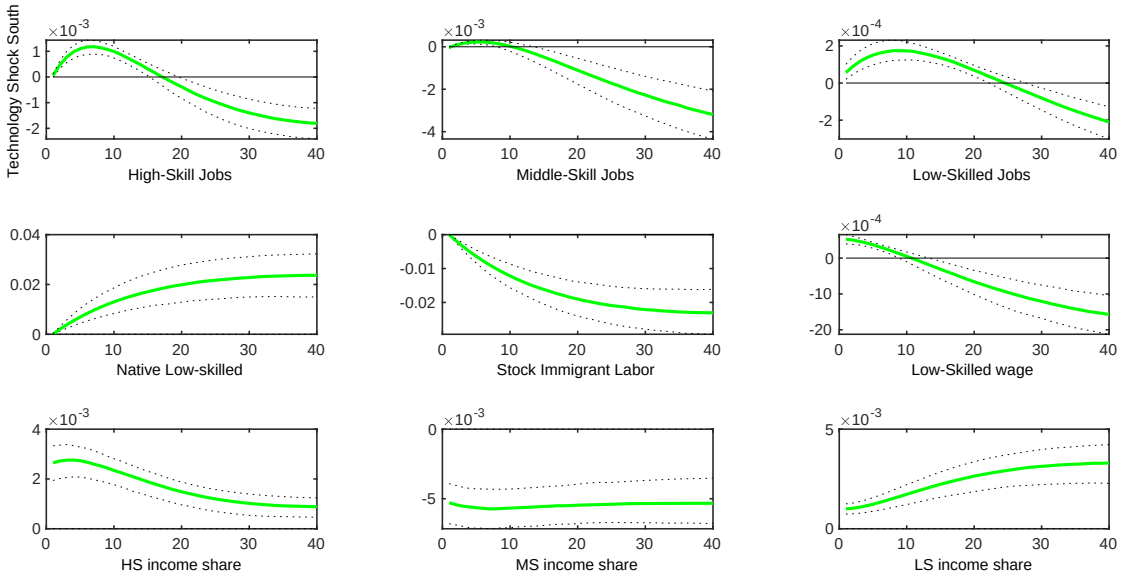
Note: This figure illustrates the impulse response of selected variables to the technology shock in Home. The solid line is the median impulse response to one standard deviation of the estimated shock, the dotted lines are the 10 and 90 percentiles.

Figure A5: Impulse response to Foreign technology shock



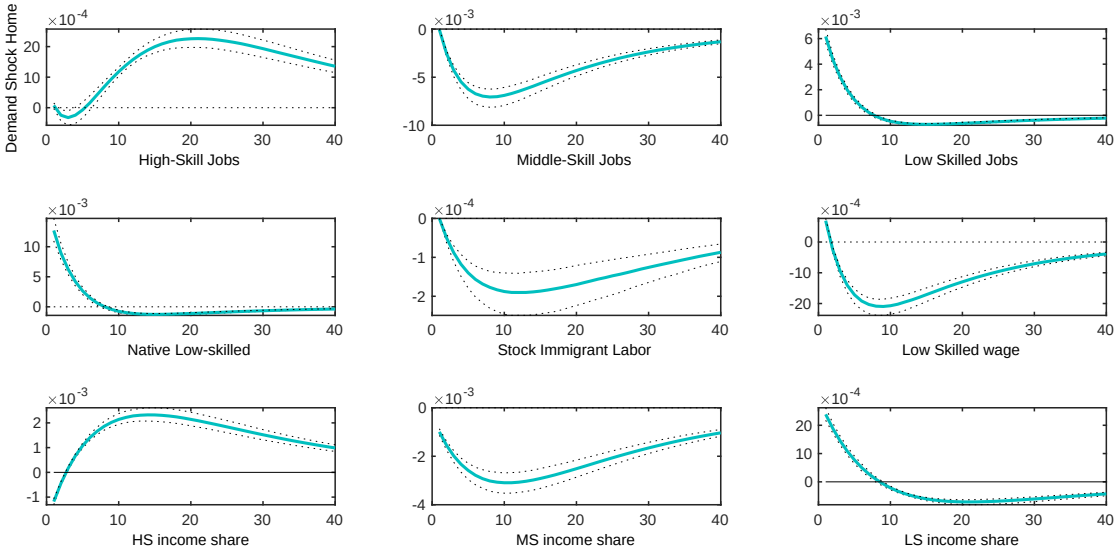
Note: This figure illustrates the impulse response of selected variables to the technology shock in Foreign. The solid line is the median impulse response to one standard deviation of the estimated shock, the dotted lines are the 10 and 90 percentiles.

Figure A6: Impulse response to South technology shock



Note: This figure illustrates the impulse response of selected variables to the technology shock in South. The solid line is the median impulse response to one standard deviation of the estimated shock, the dotted lines are the 10 and 90 percentiles.

Figure A7: Impulse response to Home demand shock

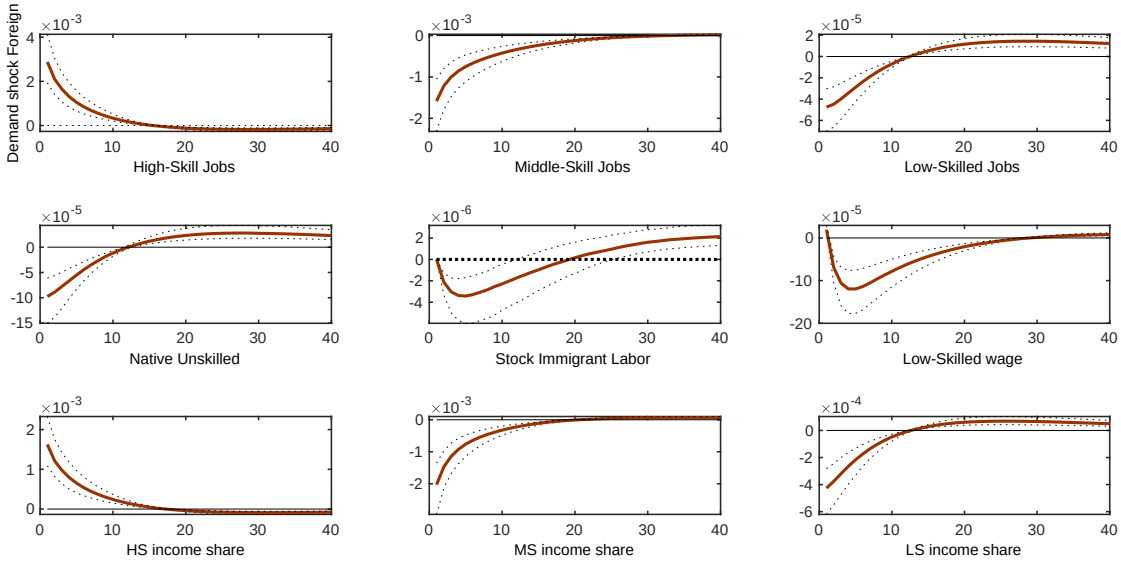


Note: This figure illustrates the impulse response of selected variables to the demand shock in Home. The solid line is the median impulse response to one standard deviation of the estimated shock, the dotted lines are the 10 and 90 percentiles.

Variance decomposition Fig. A9 displays the forecast error variance decomposition of key economic variables at various quarterly horizons (Q1, Q4, Q16, and Q40), based on the benchmark posterior estimation. As discussed, the model identifies shocks affecting the iceberg offshoring cost (cross-country symmetric), the sunk migration cost, training costs, technology, and consumption demand.

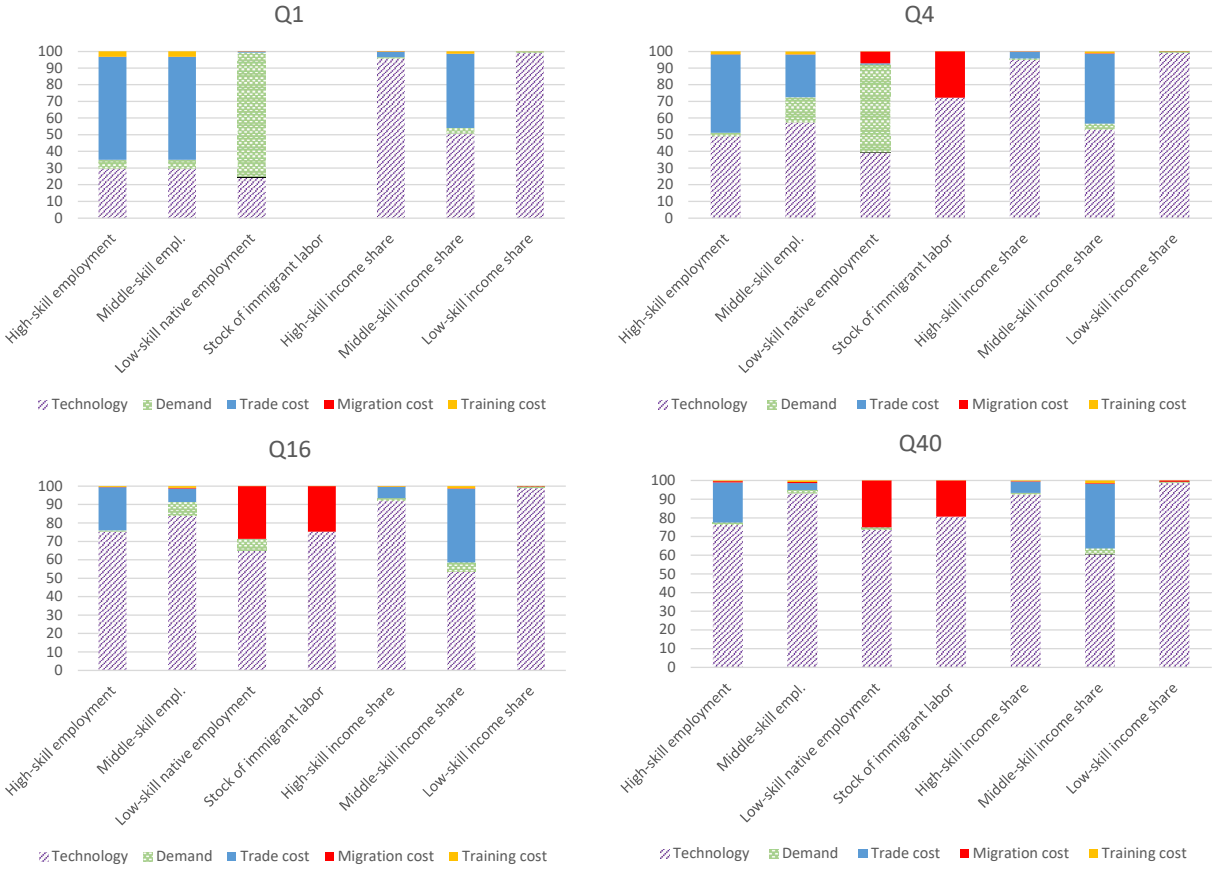
In the model, the high-skill and middle-skill employment in the tradable sector are rendered as state variables by the sunk training cost. Therefore, the estimated technology and demand shocks have small effects on these variables at very short horizons, while the shock to the iceberg trade cost have sizable effects on the offshoring margin in the short term. In contrast to high- and medium-skill employment, the low-skill employment (reflecting demand for services) reacts on impact to demand shocks. However, the impact of shocks to the iceberg trade cost and demand on the three types of employment tends to decline over time, in favor of shocks to productivity that become increasingly important.

Figure A8: Impulse response to Foreign demand shock



Note: This figure illustrates the impulse response of selected variables to the demand shock in Foreign. The solid line is the median impulse response to one standard deviation of the estimated shock, the dotted lines are the 10 and 90 percentiles.

Figure A9: Forecast error variance decomposition



Note: This figure shows the forecast error variance decomposition of key economic variables at various quarterly horizons (Q1, Q4, Q16, and Q40), based on the benchmark posterior estimation.

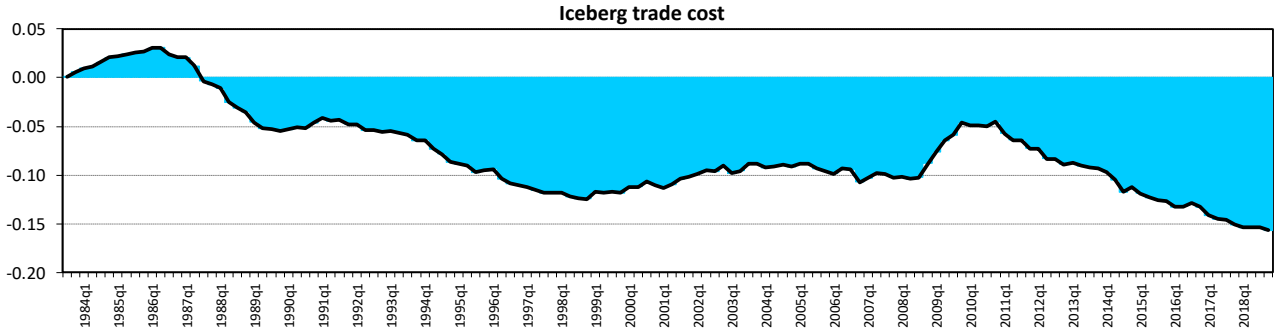
The stock of immigrant labor does not react to shocks on impact, but reacts to technology and border enforcement shocks at both medium- and long-term horizons. Due to the substitutability between low-skill native

and immigrant employment, the shock to border enforcement similarly affects the low-skill native employment at medium- and long-term horizons.

Finally, the income shares of high-skill and middle-skill labor are affected by shocks to technology and the iceberg trade cost, just like their corresponding employment groups. The shocks to border enforcement have little effect on the low-skill income share, since the latter includes both native and immigrant labor.

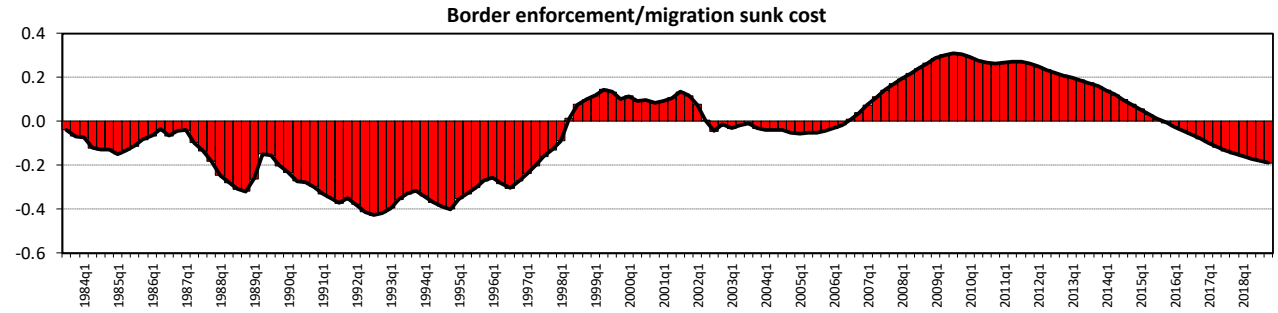
4 Additional figures

Figure A10: Shock to iceberg trade costs



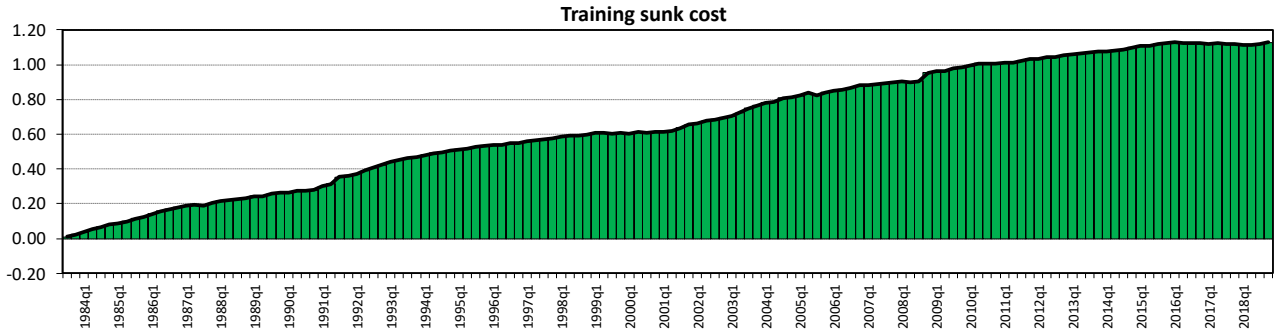
Note: This figure shows the shock to iceberg trade cost, which is disciplined with the ESCPAP/WB data.

Figure A11: Shock to sunk migration costs



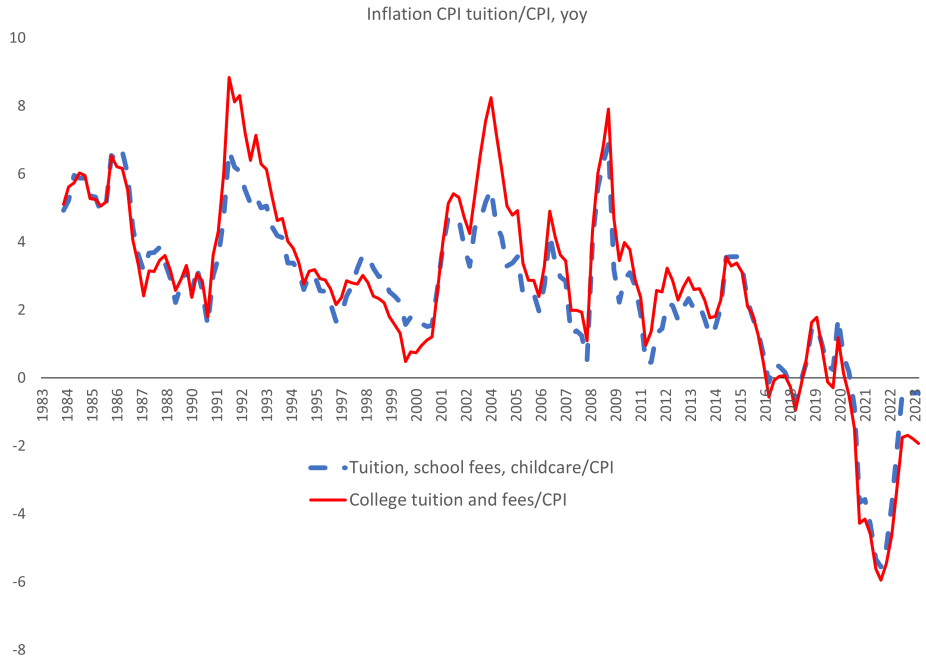
Note: This figure shows the shock to sunk migration costs, which is disciplined with Border enforcement data.

Figure A12: Shock to training costs



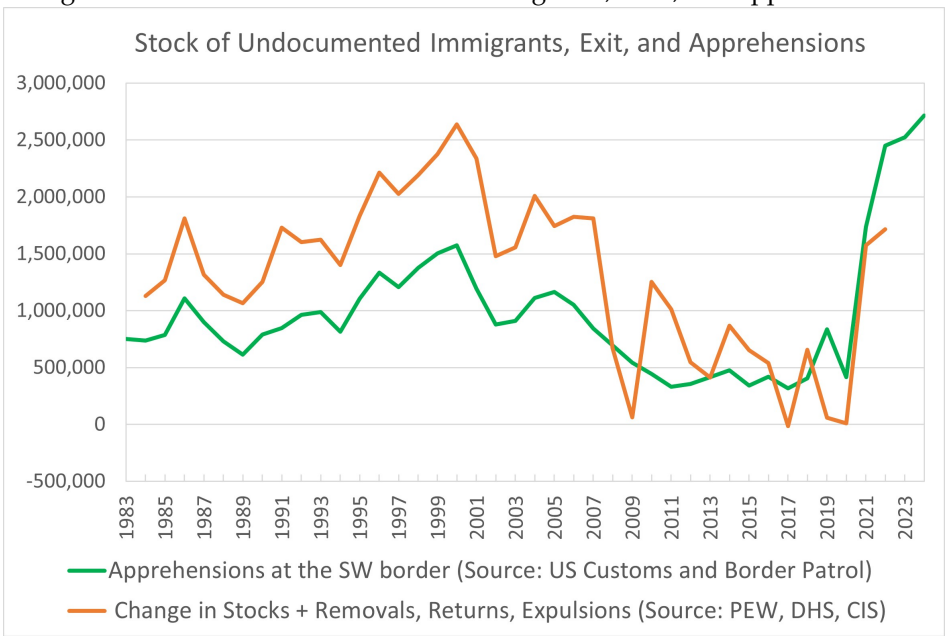
Note: This figure shows the shock to training costs disciplined with Tuition CPI data.

Figure A13: Real tuition inflation rate abstracting from childcare



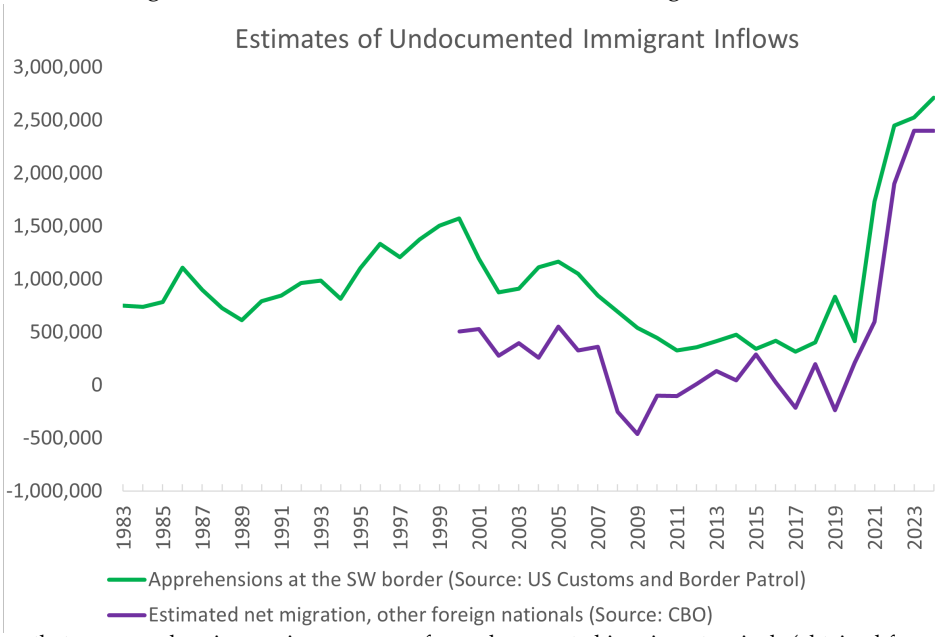
Note: This figure compares the real tuition inflation rate from Figure 2 panel c, which uses CUSR0000SEEB "Consumer Price Index for All Urban Consumers: Tuition, Other School Fees, and Childcare in U.S. City Average, Index 1982-1984=100, Quarterly, Seasonally Adjusted", with an alternative series that abstracts from childcare costs, CUSR0000SEEB01 "College tuition and fees in U.S. city average, all urban consumers, 1982-84=100, quarterly, seasonally adjusted". Both are normalized by CPIAUCSL, "Consumer Price Index for All Urban Consumers: All Items in U.S. City Average, Index 1982-1984=100, Quarterly, Seasonally Adjusted". The source is the Bureau of Labor Statistics.

Figure A14: Stock of undocumented immigrants, exits, and apprehensions



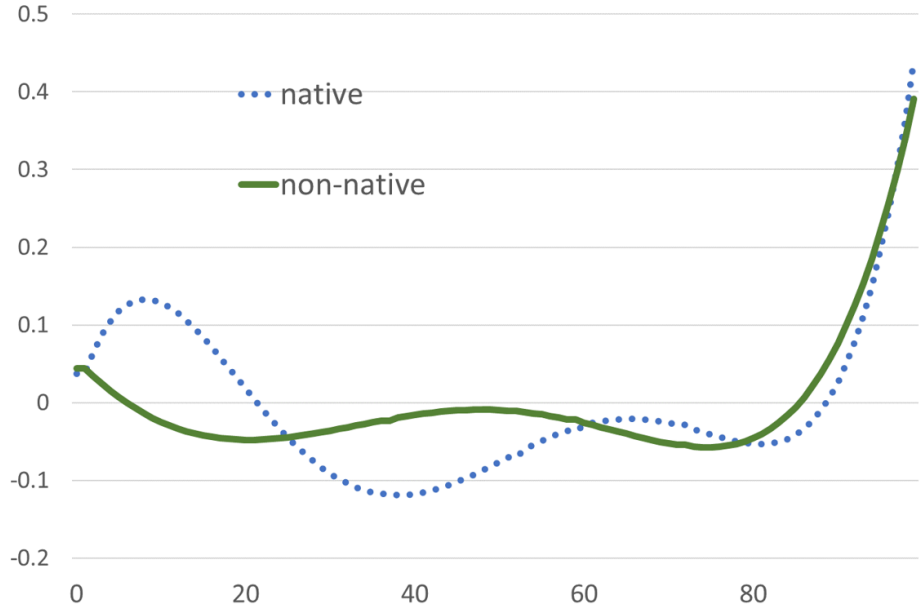
Note: This figure shows that our apprehensions series as a proxy for undocumented immigrant arrivals (obtained from the U.S. Customs and Border Patrol) is consistent with the stock of undocumented immigrant labor adjusted for exits (obtained from PEW Research, the U.S. Department of Homeland Security and the Citizenship and Immigration Service).

Figure A15: Estimates of undocumented immigrant inflows



Note: This figure shows that our apprehensions series as a proxy for undocumented immigrant arrivals (obtained from the U.S. Customs and Border Patrol) is consistent with the net migration flow estimates for other foreign nationals from the Congressional Budget Office, including for the surge in 2023-2024.

Figure A16: Changes in Employment: Robustness to sample period.



Note: This figure replicates Figure 1(B) in the main text, but restricts the sample period to 2007-2019 to exclude the years of the COVID-19 pandemic. The figure highlights that the shortage of low-skilled workers precedes the pandemic by showing muted job gains for foreign-born workers in the bottom 25% of the skill distribution, in sharp contrast to the preceding period. The low job creation in the bottom quartile is now explained by natives.

5 Additional regional regression results

Table 1: Skill premium is positively correlated with low-skill immigration

	(1)	(2)	(3)	(4)	(5)	(6)
Definition of region	State		Metropolitan area		Commuting zone	
Estimation method	OLS	GMM	OLS	GMM	OLS	GMM
Low-skill immigrant density	3.36***	4.05***	1.26***	3.26***	1.72***	2.27***
	(0.62)	(1.67)	(0.26)	(0.61)	(0.17)	(0.54)
Constant	0.47***	0.43***	0.55***	0.43***	0.36***	0.34***
	(0.04)	(0.08)	(0.02)	(0.03)	(0.01)	(0.02)
Adj R-squared	0.36	-	0.08	-	0.12	-
Observations	51	51	260	260	741	741

The dependent variable is the 2021 skill premium, while the independent variable is low-skill immigrant density in the same year.

6 Industry Evidence: Immigration and job vacancies during 2021-2024

We provide additional evidence on the cross-sectional correlation between low-skilled immigration and labor shortages across U.S. industries. If our model assumptions are correct, immigrants who arrived in large numbers after 2021 should have been attracted to industries that had a high excess of unfilled vacancies at that time (i.e., high labor shortages). Therefore, we should see that industries that had acute labor shortages in 2021 would experience in subsequent years: (a) a larger increase in the employment of immigrants, (b) a faster decline in the number of unfilled vacancies relative to other industries with less severe shortages.

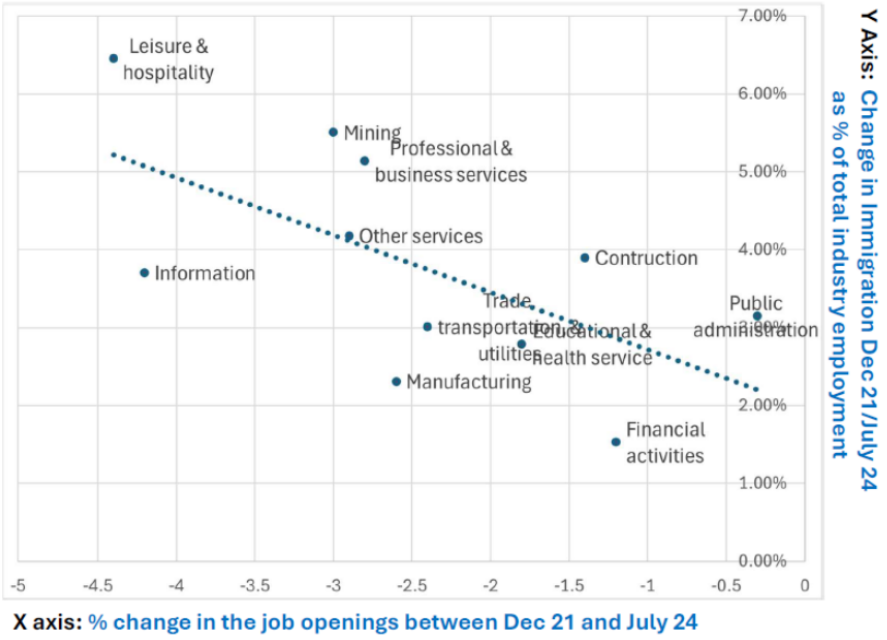
This is indeed the case. Figure A17 shows a negative relationship between (a) the change in immigration as a percentage of total industry employment over the period December 2021 to July 2024 (y-axes) and (b) the change in job openings (vacancies) over the same period (x-axes). Of particular note is the case of leisure and hospitality, which has historically been the largest employer of low-skilled immigrants in the service sector. Immigrants hired in 2021-2024 alone account for 6.4 percent of the total employment level in this industry in 2024. In turn, the vacancy rate, which represents the number of unfilled jobs in the industry, fell from a record 11 percent to 6.6 percent (a 4.4 percent decline) in those three years alone.

7 The Role of Skill Biased Technological Change

In the model, we include and estimate the following key structural innovations that drive the U.S. skill premium: trade/offshoring costs (making it easier, for example, to offshore middle-skill jobs such as blue-collar manufacturing jobs), education costs (capturing changes in college tuition costs), and crucially, the impact of low-skilled immigrants on the labor force. In our quantitative analysis, we show that, of all the factors considered, only the shortage of low-skilled labor associated with the fall in low-skilled immigration can explain the steady decline in the skill premium after the mid-2010s.

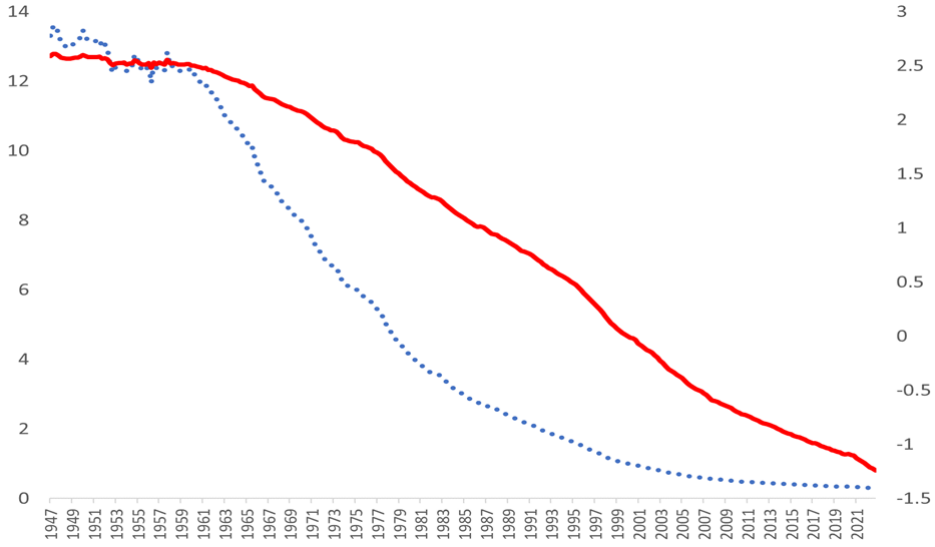
However we abstract from skill-biased technological change (SBTC), which typically includes innovations in automation, computerization, and robotics. As noted in the seminal work of Krusell et al. (2000), skilled workers benefit from this type of innovation under the assumption of capital-skill complementarity. We abstract from their inclusion for two reasons. First, including SBTC would require adding capital formation to the model, which would complicate the analysis considerably. Second, we don't believe that SBTC innovations are quantitatively relevant to the discussion here. To see this, consider first Figure A18, which shows the evolution of the relative price of quality-adjusted information and communication technology (ICT) capital (blue dotted line). The literature uses these series to pin down skill-biased technological shocks. The figure shows a persistent secular decline in ICT prices since the 1970s, contributing to the steady rise in the skill premium observed over most of this period. To better see this, the red solid line in this figure displays this price in *logs* reflecting percentage changes. Crucially for our results, this trend did not reverse in the final years of the sample, implying that SBTC cannot explain the recent reversal in the skill premium since the mid-2010s. More generally, this variable shows a steady secular decline over

Figure A17: Cross-Industry evidence: Immigration and Vacancy filling rate



decades, with no notable temporary fluctuations. As such, it is not particularly informative for our study, which is designed to capture short- to medium-term fluctuations.

Figure A18: ICT price



8 Southern Household Modeling Assumptions

For the sake of model tractability, we assume that immigrants' consumption decisions are made at the level of the household residing in the South (Mexico). In each period, the extended representative household optimally decides how much of its labor supply to devote to domestic production and how much of that labor to emigrate to Home (U.S.) after paying a sunk migration cost. Household members in both countries pool their income, which implicitly

implies that migrants can channel funds across the border via remittances.[7] Southern households, in turn, allocate consumption between non-tradable Southern goods and the tradable Home composite. In this appendix, we provide empirical justification for two modeling assumptions: (1) Migrants' consumption decisions are made at the household level in the South. (2) Immigrants do not consume non-tradables produced at Home. The first assumption implicitly assumes that migrants arrived in the United States on their own, while maintaining strong ties with household members in the country of origin (who use migrant remittances to pay their living expenses in Mexico). In principle, however, Southern workers could migrate with their own families or start a new family in Home (i.e., the U.S.), in which case they would integrate more firmly and make consumption decisions independently of the Southern family. All of these alternatives are possible, of course, but we believe our assumption is better suited to capturing the evidence. The first thing to show is whether immigrants crossing the border come alone. Figure A19 breaks down border apprehensions into *single adults*, *family unit members* and *unaccompanied children*. As can be seen in the graph, until mid-2018, the vast majority of apprehensions (encounters) consisted of *single adults*. This is the time frame we used to estimate our model. Family units become somewhat more relevant after 2019 with the emergence of the migrant *caravans* and the post-COVID migrant surge (which also has a significant number of unaccompanied children).

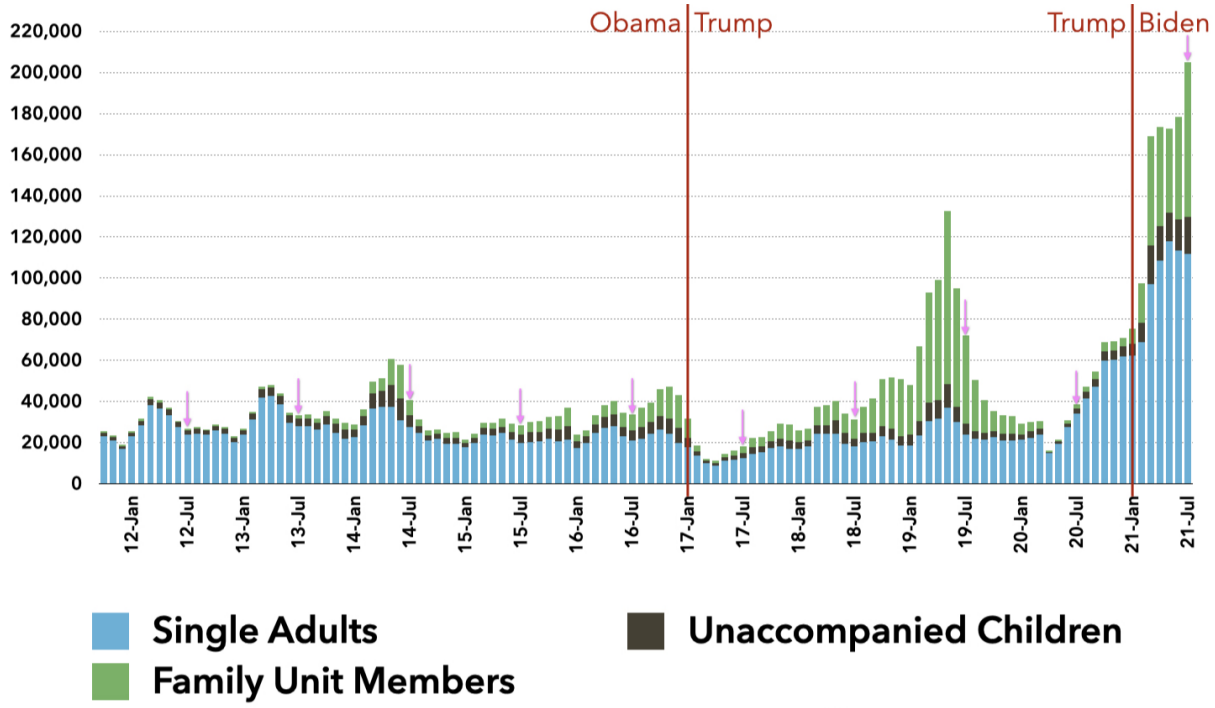


Figure A19: Border Encounters: Single adults, Family Units, and unaccompanied children (Source: Washington Office of Latin America)

The next concern is whether *single* migrants typically arrive alone to settle and raise their own families, cutting ties with those left behind. The evidence suggests that this is not a typical scenario either. Mota (2023) uses the Mexican Family Life Survey (MXFLS) to show that about 5 percent of Mexican children (17 years and younger) have a parent living abroad. Consistently, 5.2 percent of Mexico's working-age population resides in the United States, with remittances accounting for 4.5 percent of Mexico's GDP (official flows only). Görlach (2023) also finds that a husband is seventeen times more likely to migrate abroad than a wife. Still, our model may implicitly allow for the possibility that immigrants marry natives and become naturalized U.S. citizens. Moreover, we calibrate the "exit rate" in the law of motion of the immigrant stock at 5 percent per quarter following the empirical evidence, which is sizable, consistent with a higher rate of return to the country of origin (i.e., about 50 percent of the immigrants returned to Mexico within two years and 70 percent returned within 10 years; see Reyes, 1997). In each case, the evidence supports our modeling assumption of established immigrants maintaining strong ties and making

[7]For simplicity, we define a consolidated current account for Home and South in footnote 31. Thus, we do not need to specify the level of remittances, which is implicitly determined by the consumption choices of goods produced in Home and South-given the hand-to-mouth assumption of Southern households. In other words, migrants send remittances, which households residing in the South use to import goods produced in the Home. Since we use a consolidated current account, we do not need to explicitly distinguish these cross-border flows.

consumption decisions jointly with household members in the country of origin.

The second question is why we assume that immigrants do not consume non-tradable services produced in Home. This assumption is made purely for analytical convenience. However, we believe that it is also reasonable in this context. Critically, in our paper, non-tradable consumption includes *only* those personal services provided by low-skill workers at the very bottom of the wage distribution (e.g., child care providers, restaurant workers, gardeners, cleaners, beauticians). These represent a different type of of non-tradable services than those in the standard international trade model.[8] Low-skilled immigrants do not typically purchase most of these personal services in the market, but rather provide them for themselves through home production, given their very low income levels. This assumption is not critical to our results, however.

References

[1] **Bauer, A., Haltom, N., Rubio-Ramírez, J. 2003.** Using the Kalman filter to smooth the shocks of a dynamic stochastic general equilibrium model. FRB Atlanta WP 332.

[2] **von Below, D., Vézina, P. 2016.** The Trade Consequences of Pricey Oil. *IMF Economic Review*, 64(2): 303–318.

[3] **Brooks, Stephen and Andrew Gelman. 1998.** General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics,* 7(4): 434–455.

[4] **Cebula, R., Feige, E., 2011.** America's Underground Economy: Measuring the Size, Growth and Determinants of Income Tax Evasion in the U.S. Munich Personal RePEc Archive 29672.

[5] **Görlach, J., 2023.** Borrowing Constraints and the Dynamics of Return and Repeat Migration. *Journal of Labor Economics*, 41(1), 205-243.

[6] **Hotchkiss, J., Quispe-Agnoli, M., 2012.** The labor market experience and impact of undocumented workers. Federal Reserve Bank of Atlanta, mimeo.

[7] **Justiniano, A., Preston, B., 2010.** Monetary policy and uncertainty in an empirical small open-economy model. *Journal of Applied Econometrics* 25: 93–128.

[8] **Krusell, P., L. Ohanian, J. Rios-Rull, and G. Violante, 2000.** Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*, vol. 68(5),1029-1054.

[9] **Mota, J. 2023.** International Migrants and the Human Capital Formation of Their Left-Behind Children. University of Houston, manuscript.

[10] **Rabanal, P., Rubio-Ramirez, J. 2005.** Comparing New Keynesian models of the business cycle: A Bayesian approach. *Journal of Monetary Economics* 52(6): 1151–1166.

[11] **Reyes, B.I. 1997.** Dynamics of immigration: Return migration to Western Mexico. *Public Policy Institute of CA.*.

[8] We do not treat financial services, for example, as non-tradables because our model is a model of trade in tasks rather than goods. See the model description for details.