Immigration and Spatial Equilibrium:
The Role of Expenditures in the Country of Origin

Christoph Albert[1] and Joan Monras*[2]

[1] *CEMFI*

[2] *Universitat Pompeu Fabra, CREI, Barcelona GSE, and CEPR*

September 8, 2020

Abstract

We show that immigrants in the US concentrate in expensive cities, the earnings gap between natives and immigrants is larger in these cities, and these patterns are stronger when prices in the country of origin are lower. To rationalize this empirical evidence, we propose a quantitative spatial equilibrium model in which immigrants spend a fraction of their income in their countries of origin. Our model serves two purposes. First, to develop a new instrument for immigrant shocks that we use to test the model's predictions on native internal relocation responses. Second, to evaluate the consequences of immigration for aggregate productivity.

JEL Categories: F22, J31, J61, R11.

Keywords: Immigration, location choices, spatial equilibrium.

*Correspondence: jm3364@gmail.com. We are grateful to David Albouy, George Borjas, Paula Bustos, Guillermo Caruana, Donald Davis, Albrecht Glitz, Nezih Guner, Stephan Heblich, Emeric Henry, Ethan Lewis, Melanie Morten, Florian Oswald, Fernando Parro, Diego Puga, and Jorge de la Roca for their insightful comments and to the audiences at the NBER SI ITI and URB/RE, CEPR-CURE and a number of seminars and conferences for their useful questions, discussions, and encouragement. We also thank the insightful research assistance of Micole De Vera and Ana Moreno. Monras thankfully acknowledges financial support from the Fundación Ramon Areces and from the Spanish Ministry of Science and Innovation, through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S). All errors are ours. This paper subsumes the paper "Immigrants' Residential Choices and their Consequences".

1 Introduction

There are fundamentally two theories explaining immigrants' location choices. First, it is well known that immigrants tend to move to cities or regions that are thriving. One implication of this fact is that immigrants may be particularly important for "greasing the wheels" of the labor market since they arbitrage away excess demand across locations – explored theoretically in Borjas (2001) and empirically in Cadena and Kovak (2016). This observation has had a profound influence on the literature that estimates the effects of immigration on labor market outcomes. Many papers have compared high- to low-immigration regions to try to infer the causal effect of immigration on outcomes of interest. However, if immigrants' primary motive for choosing particular cities or regions is that they are thriving, there is a spurious correlation between labor market outcomes and immigrant settlement patterns.

Second, and in part as a solution to this endogeneity concern, many authors emphasize that immigrants tend to move where previous immigrants settled (Munshi 2003). Local immigrant communities help new immigrants find jobs and suitable neighborhoods for their stay in the host country. This idea has been the basis for the networks instrument, the most widely used instrument in the migration literature (Altonji and Card 1991). Simply stated, the networks instrument is based on the fact that past stocks of immigrants are good predictors of future flows, which, in the absence of serially correlated outcomes, provides regional variation in immigrant "shocks".[1] The widespread use of this instrument – despite numerous criticisms (Borjas et al. 1996; Jaeger et al. 2018; Goldsmith-Pinkham et al. 2020; Adao et al. 2019; Borusyak et al. 2019) – highlights the importance of understanding immigrant location choices within host economies.

In this paper, we provide a novel theory that explains immigrants' location choices, based on the following idea. Immigrants tend to spend large fractions of their income in their home countries. Many send remittances to family members left behind, plan on returning, or simply spend their leisure time at home. This means that they care not only about the prices in the location where they live but also about the prices in their home countries. We argue that immigrants' expenditure in the country of origin significantly shapes their residential choices and wages, which, in turn, affects the distribution of economic activity across locations and the general equilibrium in the host economy. We use this theory, which we argue is strongly supported by the data, to propose a novel instrument for immigrant shocks.

In the first part of the paper, we use different data sets to document a number of empirical regularities.[2] First, we report that immigrants concentrate disproportionately in metropolitan areas with high living costs, where, as it is well known in the urban economics literature, nominal wages and productivity tend to be higher (Combes and Gobillon 2014; Glaeser 2008). Second, the gap in earnings between natives and immigrants is *larger* in these cities, even when we consider wage gaps within narrowly defined skill cells and when we control for native-immigrant imperfect substitutability, following Ottaviano and Peri (2012) and Card (2009), and other observable characteristics. Third, we show that there is a strong heterogeneity across immigrant groups both in location choices and relative wages. We use cross-origin and arguably exogenous within-origin variation in real exchange rates to document that when exchange rates are lower, immigrants' concentration in expensive cities is stronger and the gap in wages between natives and immigrants is higher. We then use data on state-to-state migration flows from Mexico to the United States to show that Mexican immigrants from poorer states, where presumably

[1] Immigrant shocks are typically measured as the inflow of immigrants over a certain period divided by the size of the market of interest.

[2] In particular we use data from the US Census, the American Community Survey (ACS), the Current Population Survey (CPS), the Consumer Expenditure Survey, the Matricula Consular, and the World Bank's International Comparison Program database.

price levels are lower, tend to disproportionately migrate to the richest and most expensive states in the United States. Using variation in years since arrival, we also show that the patterns are weaker for immigrants that have been longer in the US and, hence, might be less attached to their home countries. Finally, we provide evidence that immigrants consume less locally than natives, by comparing housing and total expenditures of otherwise comparable native and immigrant households within locations.

In the second part of the paper, we explain these empirical regularities with a spatial equilibrium model in which immigrants' location choices are influenced by price differences between host and origin countries. When part of immigrants' consumption is related to their country of origin, immigrants have strong incentives to settle in locations with high nominal prices and wages, whereas natives only care about real wages. Thus, all else being equal, a native is indifferent between one location and an alternative one that is twice as expensive as long as wages are also twice as high. However, immigrants' smaller share of local consumption implies that they prefer the high-wage, high-price city. As a consequence, immigrants concentrate in expensive cities and, if wages partly reflect the value of living in a city – which is the case in non-competitive labor markets – the native-immigrant wage gap is higher in these locations.[3] Some degree of substitutability between home and local goods makes this mechanism stronger for immigrants coming from cheaper countries, which is in line with the data both when we compare location and wage patterns across countries of origin and when we relate them to fluctuations in exchange rates.

It is worth noting that these patterns cannot be explained by native-immigrant imperfect substitutability unless one makes the – most likely too strong – assumption that there are as many different types of labor as countries of origin. The patterns are also unlikely to be driven by differences in productivity levels across origin countries, since in some of the specifications we use variation over time within countries of origin. Finally, both in the empirical part and in our model, we also account for immigrant networks, which, according to the data, act as an amenity shifter that helps to explain immigrant location patterns within our framework.

We estimate the model by matching the relative distribution of immigrants and their wages relative to those of natives in 1990, using variation across metropolitan areas and across countries of origin conditional on the relative size of local immigrant networks. The estimation identifies two key parameters. First, our baseline estimates imply that immigrants' weight of total expenditure in the home country is around 17%, and, second, the elasticity of substitution between consuming locally and in the country of origin is 1.86.[4] We validate our estimation by showing that the model performs well at predicting aggregate population and wage data and two additional non-targeted moments: the variation in the consumption of housing across origins and the immigrant inflows and native population growth patterns across cities over the 1990s observed in the data.

In the final part of the paper, we use the model for two purposes. First, the insights of the model allow us to construct a new instrumental variable strategy for immigrant shocks as an alternative to the networks IV. In particular, we allocate immigrant flows for each country of origin during a decade based on the *relative* share of immigrants across locations, predicted by the exchange rate fluctuations and its interaction with an exogenous determinant of housing prices like the housing supply elasticity estimates of Saiz (2010). This acts as the "share"

[3]In order to obtain this result, wage differences between workers cannot be competed away. This means that we depart from standard perfectly competitive models of the labor market and instead consider wage bargaining. See Becker (1957) and Black (1995). A recent paper studying immigration in non-competitive frameworks is Amior and Manning (2020) where they introduce monopsony power in an aggregate model of the labor market based on Borjas (2003). As Amior and Manning (2020) also emphasize, the literature on immigration is mostly (and perhaps surprisingly) centered in perfect competition frameworks.

[4]We also analyze the sensitivity of these estimates. Using a large number of alternatives, we obtain a range of estimates for the average share of consumption in the origin country between 15% and 25%. The range of estimates that we obtain for the elasticity of substitution between consuming locally and in the country of origin goes from 1.5 to 2.5.

in a shift-share-type instrument.

We then compare this new instrument with the networks instrument and Bartik labor demand shocks, the two existing theories that have been used to explain immigrants' location choices. We show that our new instrument is the main predictor of immigrant settlement patterns in cities with low housing supply elasticities (and hence high house prices) and high productivity. It retains predictive power for less expensive, less productive locations, where the networks instrument and Bartik shocks become also relevant.

Using the new instrument, we test the predictions of the quantitative model further. In the estimated model, immigrant inflows have a causal effect on native relocation patterns. Natives are predicted to relocate from high-productivity cities where immigrants enter to low-productivity ones. Indeed, we show that on average around 0.8 natives relocate for each immigrant arrival. This estimate, as we show, reflects the native internal migration responses in highly productive cities with low housing supply elasticities, such as San Francisco, Miami, or New York City. We give further credibility to our IV estimates by using the networks instrument to perform over-identification tests. The null hypothesis of the Sargan-Hansen test for the validity of all instruments cannot be rejected, and hence, we cannot rule out that our instrument is indeed exogenous. Moreover, when we estimate the effect of lagged immigrant shocks on current native internal migration responses, we obtain tightly estimated zeros, ruling out the possibility that our results are driven by pre-trends.

In the final exercise of the paper, we use our model to estimate the effect of immigration on aggregate productivity. In a seminal contribution, Hsieh and Moretti (2019) argue that the most productive cities in the US have imposed constraints in their housing supply, which limit the extent to which labor can relocate and hence take advantage of these higher levels of productivity. They label this phenomenon as spatial misallocation of labor. Our model and empirical evidence suggests that immigration reduces, at least to some extent, this spatial misallocation. As a result of immigration, a higher share of aggregate production in the host economy takes place in more productive locations. We estimate that over the 1990s aggregate productivity increased by around 0.40% thanks to the roughly 8 million immigrants that arrived to the US during the decade.

Overall this paper has three main contributions. First, it provides a new theory for immigrants' location choices within host economies that is well supported by a large set of empirical facts – many of which were not documented in previous research. Other theories on immigrant location choices can at best explain a subset of these facts, but, to the best of our knowledge, none is able to explain all these facts in a unified framework. Second, we use our model to provide a new instrument for immigrant shocks which can be used in future research, beyond testing the quantitative implications of our model as we do in this paper. Third, we use our model to better understand how immigration is shaping economic activity across locations and in the aggregate, something which typically escapes empirical accounts of immigration based on difference-in-difference type of comparisons.

Our work is related to several strands of the literature. A range of papers have investigated the effect of immigration on host economies. Following the pioneering work of Altonji and Card (1991), many have evaluated the effect of immigration on host economies using spatial variation, summarized in the review article by Dustmann et al. (2016). Our paper contributes to this literature by analyzing how immigrants decide to locate within host economies based on a new mechanism. This allows us to introduce a novel instrument that can be used together with the commonly used networks IV.

The model we propose is related to a large body of recent work on quantitative spatial equilibrium models, summarized in Redding and Rossi-Hansberg (2017). In this literature, only Burstein et al. (2020), Caliendo et al. (2018), Piyapromdee (Forthcoming), and Monras (2020) use spatial equilibrium models to study immigration. In these papers, immigrants are not characterized by how they consume but, rather, by differences in observable characteristics that may be important for labor market outcomes but that are silent on many of the empirical facts uncovered in our paper.

In this paper, we take the view that immigrants are characterized by the fact that they have an extra good in their utility function, which can only be consumed in their country of origin. This extra good may represent remittances, future consumption, or income spent during periodical stays in the origin country. Only a relatively small number of papers have effectively seen immigration in this way. This is the case, for example, in studies of temporary and return migration. In most of this work, authors think of migration decisions as a way to accumulate human capital or savings for the eventual return to the home country (see a review of the literature in Dustmann and Gorlach 2016). This literature, however, has not studied the effects of immigration on the spatial equilibrium of the host economy.

Some studies have investigated how changes in relative prices between host and destination country due to nominal exchange rate fluctuations affect immigrants' behavior in terms of the amount of remittances sent and their use in the home country (Yang 2006, 2008). The findings in these studies are in line with this paper. In contrast to these studies, it is worth emphasizing that we explore the role of nominal disparities for consumption and location choices of immigrants in the host country rather than in the sending economy.

In what follows, we first describe our data in Section 2. Section 3 documents a number of stylized facts on immigrants' location choices, wages, and consumption patterns. We present a quantitative spatial equilibrium model consistent with these facts in Section 4. In Section 5, we use our model to develop a new instrument and to quantify the aggregate effects of immigration. Section 6 concludes the paper.

2 Data

2.1 Census, American Community Survey, and Current Population Survey

Our main data sources are the US Census of the years 1980, 1990, and 2000, and the American Community Survey 2009-2011 downloaded from IPUMS (Ruggles et al. 2016). We use information on the metropolitan statistical area (MSA) of residence of surveyed individuals, the wage they received in the preceding year, the number of weeks that they worked in the preceding year, and their country of birth. We keep all individuals aged between 18 and 65, excluding military personnel and persons living in group quarters. Further, we drop the population living outside MSAs and in MSAs that are not identified across all years, whose boundaries are not consistent over time and for which the housing supply elasticity from Saiz (2010) is not available. This leaves us with a total of 185 MSAs. We define as immigrants all individuals born outside the United States who were not US citizens at birth. To obtain aggregate wage levels, we only consider salary workers aged between 25 and 59 who are not in school and report positive weeks and hours worked. To control for the potential heterogeneity in demographic characteristics of workers across cities and migrant status, we construct composition-adjusted wages. That is, we use the residuals of a regression of the logarithm of the weekly wage on dummies for sex, education, experience, race, and marital status.[5]

We also use data from the Census and ACS to compute local price indices. We thereby follow Moretti (2013)

[5] The education categories are high school dropout, high school graduate, some college (less than three years), and college graduate. The categories for experience are eight five-year intervals, starting with experince years 1-5 and ending with experience years 36-40. We follow Ottaviano and Peri (2012) in assuming that high school dropouts enter the labor market at age 17, high school graduates at age 19, workers with some college at age 21, and college graduates at age 23. The racial categories are Black, American Indian or Alaska Native, Chinese, Japanese, other Asian and other race. The categories for marital status are married, separated, divorced, widowed, and single.

and apply his code to our sample including the pooled 2009-2011 ACS data. From this, we obtain a local price index for each of the MSAs, which is based on the variation in local housing cost.[6]

To explore whether our results hold for data collected more frequently, we rely on the Current Population Survey (CPS). In particular, we use the CPS March supplement to generate yearly cross-sectional data for individuals, including information on their demographic characteristics, labor market variables, and location of residence. Sample selection and variable definitions are the same as for the Census data. As information on the birthplace of respondents is only available after 1994, we only use CPS data for the period 1994-2011.

To give a sense of the characteristics of the MSAs in our sample, Table 1 reports the MSAs with the highest immigrant shares in the United States in 2000, together with some of the main economic variables used in the analysis. Many of the MSAs with high levels of immigration are also large, expensive, and pay high wages. The gap in earnings between natives and immigrants is also larger in these locations. There are a few notable outliers, which are mostly MSAs in California and Texas close to the border between the US and Mexico.

2.2 Consumer Expenditure Survey

To document consumption patterns, we employ two different data sets. First, we use the Consumer Expenditure Survey (CEX), which is maintained by the Bureau of Labor Statistics and has been widely employed to document consumption behavior in the United States. It is a representative sample of US households and contains detailed information on consumption expenditure and household characteristics. Unfortunately, it contains no information on birthplace or citizen status, making it unfeasible to directly identify immigrants. Instead, we rely on one of the Hispanic categories that identifies households of Mexican origin in the years 2003 to 2015.[7] The data set contains around 30,000 households per year, of which around 7% are of Mexican origin.

We also document consumption patterns using the Census and the ACS. Both the Census and the ACS record the rent paid by households who are renting. This is the most important local expenditure for many households and typically accounts for roughly 25% of (gross) household income.

2.3 World Bank data

An important aspect of our empirical analysis is to document the heterogeneity by price level of immigrants' origin country. To measure these, we use real exchange rate (RER) data, which are provided by the World Bank with respect to the United States for a large number of countries in its International Comparison Program database.[8] In Table E1 in online Appendix E, we provide a list of the top 10 and bottom 10 countries in terms of the average real exchange rate with respect to the United States for the years 1990, 2000, and 2010. While price levels in countries like Norway or Japan are around 35% higher, the number of countries with real exchange rates larger than 1 is relatively low. Australia, ranked 10th in the table, is only 7% more expensive than the United States. At the other end, countries like Vietnam (with large immigrant communities in the United States) have prices that are only 20% of those in the United States.

[6] We use the version of Moretti's price index that is calculated as the weighted sum of local housing cost and the cost of non-housing consumption, which is assumed to be the same across areas. Local housing costs are measured as the average of the monthly cost of renting a two- or three-bedroom apartment in an MSA.

[7] Monras (2020) shows that the overlap between individuals identified as Hispanics of Mexican origin and Mexican-born individuals is around 85% in Census data. This gives us confidence that, by using the Hispanic variable in Consumption Expenditure data, we are capturing a large number of Mexican-born individuals.

[8] The data series is titled "Price level ratio of PPP conversion factor (GDP) to market exchange rate."

Further, we use the bilateral Remittance Matrix provided by the World Bank for 2010, which is the earliest year available, to compute the total amount of remittances sent from immigrants in the US to their origin countries.

2.4 Matricula Consular data

Mexican immigrants in the United States are encouraged to register in the local consulates, which issue a card called the Consular ID. In order to obtain this card, they need to show their birth certificate or passport. In principle, both immigrants legally admitted to the US and those that are undocumented can obtain this card. The card is useful, among other things, to open bank accounts in a number of financial institutions, which gives many Mexicans sufficient incentive to register. Among the information recorded with this registration process is the address of destination in the US and the municipality of origin within Mexico. This data allows to compute the bilateral flows of Mexicans in any given year. In theory, these can be computed at the municipality level, although only state-to-state flows are publicly available. Caballero et al. (2018) and Allen et al. (2019) show that these data match well representative data sets on stocks. In this paper, we use the state-to-state migration flows for the year 2016.

3 Motivating facts

3.1 Immigration and cities

In this section, we document two facts on the cross-sectional relationship between immigration and cities. First, we show that immigrants concentrate much more than natives in expensive cities. Second, we document that the gap in wages between immigrants and natives is larger (i.e. more negative) in these cities.

Estimation equations

To document the first fact, that immigrants concentrate more than natives in expensive cities, we define the *immigrant concentration* as the relative number of immigrants living in city c and regress this measure (in logs) on the price level of city c. As we show later, this regression is a reduced-form version of the relationship between the relative distribution of immigrants and city price levels implied by the model. More specifically, we run the following regression:

$$\ln\left(\frac{\mathrm{Imm}_{c,t}}{\mathrm{Imm}_t}\Big/\frac{\mathrm{Nat}_{c,t}}{\mathrm{Nat}_t}\right) = \beta_Q \ln P_{c,t} + \psi_c + \psi_t + \varepsilon_{c,t}, \tag{3.1}$$

where $\mathrm{Imm}_{c,t}$ is the number of immigrants, $\mathrm{Nat}_{c,t}$ the number of natives, and $P_{c,t}$ the corresponding price index in city c at time t. Imm_t and Nat_t are the overall numbers of immigrants and natives in our sample, respectively. An estimate $\beta_Q > 0$ implies that immigrants concentrate more in expensive MSAs than natives.

The second fact that we document is that the gap in wages between natives and immigrants is strongly related to city prices. If wages reflect, at least in part, the value of living in a location – which is a natural result in non-competitive labor market models (see Section 4) – we should expect the relationship between wage gaps and city prices to be the mirror image of relative location choices. We obtain the wage gaps at the city level by calculating the difference in the average composition-adjusted (log) wage of natives and immigrants in each MSA.

When workers are heterogeneous and can be divided into different factor types, there are three potential problems for interpreting city-specific wage gaps as reflecting differences in the utilities of living in a city. First, it may be that immigrants and natives within narrowly defined skill groups are imperfect substitutes. Hence, in locations where they are more concentrated, immigrants might either earn relatively less, if their higher concentration is supply-driven (e.g. because of larger networks), or relatively more, if it is demand-driven (e.g. because the demand for labor services usually offered by immigrants is higher in more expensive cities). Second, the skill sorting of natives and immigrants across cities might be different and could partly drive local wage gaps. Third, it could be that the gap in earnings between natives and immigrants varies with education. For instance, it might be larger among high-skilled workers than among low-skilled workers. In this case, the higher concentration of high-skilled workers in larger cities (i.e the sorting of skills across cities) could generate higher average wage gaps in these locations.

To control for these factors, we combine the empirical approaches of Card (2009) and Ottaviano and Peri (2012). In particular, we estimate a model in which we relate the gap in wages between natives and immigrants to their relative labor supplies in an MSA, following Card (2009). Moreover, we group workers in skill cells based on education and experience and calculate wage and employment ratios within those cells, as in Ottaviano and Peri (2012).[9] The inclusion of skill cell fixed effects absorbs any variation in wage gaps across cities due to different sorting along the education and experience dimensions. More concretely, we estimate the following regression:

$$\hat{w}_{I,k,c,t} - \hat{w}_{N,k,c,t} = \phi_k + \phi_{c,t} + \gamma \ln(\frac{L_{I,k,c,t}}{L_{N,k,c,t}}) + \varepsilon_{i,t}, \tag{3.2}$$

where $\hat{w}_{I,k,c,t}$ and $\hat{w}_{N,k,c,t}$ are the average composition-adjusted log wages and $L_{I,k,c,t}$ and $L_{N,k,c,t}$ the total hours worked in skill cell k and city c at time t for immigrants and natives, respectively. Since, compared to Ottaviano and Peri (2012), our data is further disaggregated at the MSA level, we opt for larger skill cells in order to have enough observations in each cell. In particular, cells are defined by two education groups (high school or less and at least some college) and four 10-year experience intervals.[10]

To account for the potential endogeneity problem that arises because the relative labor supply of immigrants might be driven by a higher relative demand for immigrant labor at the local level, we instrument their labor supply by the typical shift-share networks instrument. To do so, we allocate the national immigrant inflows to locations according to the distribution of the stock of immigrants from the same origin 10 years before. This strategy allows us to extract the city-time-specific component of the wage gaps as the city-time fixed effects $\phi_{c,t}$, which are adjusted for any effects due to spatial sorting or imperfect substitution between natives and immigrants within education-experience cells.

The last column of Table E2 in online Appendix E presents the estimate of the coefficient γ, which gives the negative inverse elasticity of substitution between natives and immigrants. For comparison, we show the estimates obtained with alternative specifications used in the literature in the first three columns.[11] In column 4, we estimate regression (3.2) without MSA or MSA-year fixed effects. Thus, γ is identified using variation within skill cells over time and across MSAs. With this specification, we obtain a coefficient that implies an elasticity of substitution of around 20, which is the actual consensus estimate in the literature. However, after including MSA fixed effects in column 5, the estimate becomes much smaller, implying an elasticity of 74. In our final

[9] Our main estimates use the mean of the log composition-adjusted wages as in Card (2009), although we replicate the original Ottaviano and Peri (2012) results and its sensitivity as discussed in Borjas et al. (2012).

[10] Results are robust to using the same cell definition as Ottaviano and Peri (2012).

[11] In the first column, we replicate the specification of Ottaviano and Peri (2012), which relies on national variation within skill cells across time (in particular, our sample selection and specification corresponds to *Pooled Men and Women* in column 2 of table 2 in their paper. The fact that our coefficient is slightly lower might be driven by not including the year 1970 in our sample.) In column 2, we follow Borjas et al. (2012) by using the mean of log wages instead of the log of mean wages as Ottaviano and Peri (2012). In column 3, we replicate the specification of Card (2009), Table 6, which relies on variation across the 124 largest MSAs in the year 2000 and uses log composition adjusted wages.

specification with MSA-year fixed effects, the coefficient essentially becomes zero. Thus, once we account for MSA-specific time trends, we find no indication for imperfect substitutability between natives and immigrants at the local level (see also Borjas et al. 2012; Ruist 2013).[12]

In the next step, we relate these MSA-specific adjusted wage gaps, identified through $\hat{\phi}_{c,t}$, to the city price level using the following regression:

$$\hat{\phi}_{c,t} = \beta_P \ln P_{c,t} + \psi_c + \psi_t + \varepsilon_{c,t}. \tag{3.3}$$

An estimate of $\beta_P < 0$ indicates that the adjusted log difference in wages between immigrants and natives is greater (i.e. more negative) in more expensive cities.

Results

We report the results of estimating equations (3.1) and (3.3) in Table 2, and provide a large number of robustness checks in Table E3 in online Appendix E. Before turning to the estimates, Figure 1 illustrates the relationships using cross-sectional data from the 2000 Census. In Panel A, we observe a tight positive relationship between the distribution of immigrants and city price indices (marker sizes reflect city population levels). The few outliers far from the linear fit are mostly smaller cities along the US-Mexico border. In online Appendix A.1, we show that this relationship also holds when using commuting zones instead of MSAs.[13] This evidence is in line with recent work by Albouy et al. (2018), who argue that immigrants live in relatively high-wage, high-amenity locations. Panel B of Figure 1 shows that the relationship between wage gaps and city price levels is the mirror image of the relationship in Panel A: a higher price level is associated with a larger wage gap (i.e. lower earnings of immigrants relative to natives).

Table 2 shows the regression coefficients pooling all years and including year fixed effects in all specifications. A potential endogeneity problem arises if unobserved city-specific economic shocks affect internal immigrant flows relatively more than native flows and, at the same time, affect price levels. To address this issue, we use an IV strategy. It is well documented that part of the variation in house prices across cities is driven by geographic constraints (see, for example, Gyourko and Saiz 2006). Based on this idea, Saiz (2010) builds estimates of housing supply elasticities that take into account the share of developable land around each city center. The housing supply elasticity is a good predictor for local housing cost. Hence, we use these (time-invariant) elasticities to predict the local price indices in the first-stage regression. The exclusion restriction is that the geographic constraints underlying the housing supply elasticity estimates affect the immigrant concentration or wage gaps only through their effect on price levels.[14]

Both the OLS and IV regressions confirm that a higher price level is associated with a higher concentration of immigrants (columns 1 and 2) and a larger wage gap (columns 4 and 5). In columns 3 and 6, we include MSA fixed effects, and hence β_Q and β_P are identified by within-city variation over time only. While still significant at the 5% level, coefficients are much smaller in this specification. Given that the city price ranking is highly persistent, internal geographic mobility is imperfect, and wages may be sticky, it is not surprising that the pure time variation in prices has a more limited effect for both outcome variables. Our focus in the remainder of this

[12] This result is robust to only restricting the sample to large MSAs or alternately constructing the IV by allocating national inflows always based on the 1980 distribution of immigrants instead of using the preceding decade.

[13] Commuting zones are a partition of the US territory and can be divided between urban and rural ares. Urban commuting zones are equivalent to MSAs, whereas rural commuting zones are not captured by MSA information. Saiz (2010) does not provide housing supply elasticities for non-urban commuting zones. This is why we focus attention on metropolitan areas.

[14] In the regression with the immigrant concentration as dependent variable, the restrictions might be violated, if "entry" cities with ports such as New York have a higher immigrant share and are also more geographically constrained. We have therefore confirmed that the results also hold when dropping port cities.

paper, in particular for the estimation of the model, will therefore be on the cross-sectional variation.

There are different potential explanations for the results reported in Table 2. In particular, they could be driven by particular subsamples of immigrants. We explore this in Table E3 in online Appendix E, in which we report the regressions shown in columns 1, 2, 4, and 5 of Table 2 for various subsamples. Panels A and C present the OLS results on the immigrant concentration and wage gaps, respectively, while panels B and D present the IV specification based on housing supply elasticities.

Columns 1 and 2 of Table E3 show that when we restrict the sample of immigrants to those who are likely documented or to those who are likely undocumented, we obtain similar results. The results are somewhat stronger for undocumented immigrants, potentially reflecting (as we argue in subsection 3.2) that undocumented immigrants are less attached to the host country.[15] Columns 3 to 6 show the results restricting the sample of immigrants and natives to a particular education level. It is well documented that immigrants are concentrated in the bottom and the top of the education distribution (Borjas 2003; Ottaviano and Peri 2012). If very high- and very low-skilled workers had higher incentives to live in large cities, it could be that the results in Table 2 are driven by this. However, we obtain strongly significant coefficients for each education subsample. In the final column, we show that the results are also robust to excluding immigrants from Latin America.

In the following section, we disaggregate the sample of immigrants to investigate the potential heterogeneity of our findings by origin country. This disaggregation allows us to further control for origin-specific immigrant networks at the city level, which might be an important driver of immigrants' location choices and wage gaps.

3.2 Immigrant heterogeneity

Heterogeneity across and within origin countries

If there is some degree of substitutability between consuming locally or consuming in the country of origin, we would expect the patterns documented above to be stronger for immigrants from countries with lower relative price levels, or, in other words, lower real exchange rates with respect to the US. To test this hypothesis, we split immigrants into groups defined by their country of birth.

To explore the heterogeneity by country of origin, we estimate equation (3.1) at the city-origin level (i.e. we replace Imm by the immigrant population from the respective origin). This further disaggregation also enables us to control for the presence of immigrant networks, a factor we neglected thus far. Because the contemporaneous network, which we measure as the fraction of immigrants from origin o among the total city population at time t, is mechanically related to the immigrant concentration, we instead use the network predicted by the distribution of immigrants in the previous Census year.[16] To first inspect the variation across origins in a nonparametric way, we estimate origin-specific coefficients β_o. We then estimate equation (3.2) at the skill-city-origin-time level and replace the city-time with city-origin-time fixed effects. In the second step, we estimate origin-specific elasticities using the following regression:

$$\hat{\phi}_{c,o,t} = \beta_o \ln P_{c,t} + \beta_o^{Net} Net_{c,o,t} + \phi_t + \phi_c + \varepsilon_{c,t}, \tag{3.4}$$

where $Net_{c,o,t}$ denotes the immigrants network constructed as described above.

[15] We code likely documented or likely undocumented immigrants following Borjas (2017), who uses, among others, variables indicating participation in welfare programs that require being legally in the country, to identify documented immigrants.

[16] In particular, we allocate the total immigrant population from a certain origin according to its distribution across MSAs 10 years before and then divide this predicted immigrant stock by current city population (natives plus immigrants) to get the predicted network. Results are virtually identical when always allocating based on the 1980 distribution.

Figure 2 plots the city price elasticities of the immigrant concentration against real exchange rates in Panel A and the elasticities of the wage gap in Panel B. Although the elasticity estimates are somewhat noisy, both plots show statistically significant relationships that go in the expected direction. Panel A shows that the city price elasticity of the concentration of immigrants from origins with *lower* real exchange rates is *higher*. Again, the corresponding plot for the wage gap in Panel B is the mirror image of that for the immigrant concentration. The elasticity of the wage gap is higher for immigrants from more expensive countries, meaning that the difference in wages between natives and immigrants does not increase as steeply for them. Overall, these graphs show that there is substantial variation across countries of origin, which is well aligned with the role that differences in relative prices should play in shaping the importance of home country expenditures for immigrants from different countries of origin.

To document these relationships more systematically using, arguably, exogenous exchange rate variation within countries of origin, we expand equations (3.1) and (3.3) by interacting the city variable directly with the real exchange rate (RER) and pooling the Census years 1990 and 2000 and the combined ACS data for 2009-2011 (real exchange rates are not available for 1980 in the World Bank data). In particular, we estimate the following regressions:

$$\ln\left(\frac{\text{Imm}_{c,o,t}}{\text{Imm}_{o,t}} / \frac{\text{Nat}_{c,t}}{\text{Nat}_t}\right) = \alpha_1 \ln P_{c,o,t} + \alpha_2 \ln RER_{o,t} + \alpha_3 \ln P_{c,t} \times \ln RER_{o,t} + \alpha_4 Net_{c,o,t} + \delta_t + \delta_o + \delta_c + \varepsilon_{c,o,t} \tag{3.5}$$

$$\hat{\phi}_{c,o,t} = \beta_1 \ln P_{c,t} + \beta_2 \ln RER_{o,t} + \beta_3 \ln P_{c,t} \times \ln RER_{o,t} + \beta_4 Net_{c,o,t} + \zeta_t + \zeta_o + \zeta_c + \varepsilon_{c,o,t}, \tag{3.6}$$

where, as before, $\ln P_{c,t}$ denotes the price level of city c, and $RER_{o,t}$ is the real exchange rate of origin country o with respect to the United States at time t. We estimate the relative share equation using a Poisson pseudo maximum likelihood (PPML) regression model in order to deal with the incidence of zeros (Santos Silva and Tenreyro 2006).[17]

The estimates of interest are α_3 and β_3. A negative estimate of α_3 indicates that immigrants from cheaper countries tend to concentrate more in expensive cities. In line with this, a positive estimate of β_3 indicates that the wage gap of immigrants from these countries of origin is larger in these cities. With origin fixed effects, the identifying variation comes from RER fluctuations across decades for each country of origin.

Table 3 presents the results. In the first/fourth columns, we estimate the model excluding the immigrant network and MSA fixed effects. In the second/fifth columns, we add the immigrant network, and in the third/sixth columns we add MSA fixed effects. All columns include origin country fixed effects.

The coefficient of the interaction between price levels and the RER is significantly negative with the immigrant concentration as dependent variable, and positive with the wage gap as dependent variable. As expected, the predicted network of immigrants from the same origin has strong predictive power for the immigrant concentration. However, adding this control or the MSA fixed effects to the regression leaves the coefficients of the interaction term virtually unchanged. Interestingly, the effect of immigrant networks on the wage gap is negative. This can be interpreted as evidence of a compensating differential in wages for immigrants living in locations with larger communities of compatriots, consistent with immigrant networks acting as a positive amenity shifter. Moreover, we obtain a significantly positive coefficient β_2, which indicates that a 10% appreciation of the real exchange rate

[17] These results are based on the top 68 sending countries. In particular, we drop origin countries with less than 100 observations in any of the Census years, small island countries and Eastern European countries that did not exist before 1990. The resulting 68 origins account for around 90% of all immigrants in the sample.

with respect to the origin country reduces the wage gap by around 0.4 percentage points.[18] This result supports the notion that local wages (at least partly) compensate workers for the *individual* price index of the goods they purchase with their income, which in the case of remittance-sending immigrants is affected by real exchange rates.

Heterogeneity across Mexican immigrants

Given the large number of immigrants from Mexico, which is by far the top sending country, we can explore in more detail the heterogeneity among Mexican immigrants in the US to check whether we obtain results consistent with those documented above.

One limitation of the results shown in Table 3 is that we only observe outcomes at 10-year intervals, implying that the time variation in exchange rates might be confounded with long-run economic trends in sending countries, which could affect, for example, the composition of entering immigrants. Thus, we also exploit fluctuations in real exchange rates between the United States and Mexico at a higher frequency (yearly). Exchange rate fluctuations are common and difficult to anticipate over short time horizons, offering arguably exogenous variation.[19] More concretely, we focus our analysis on Mexican immigrants who move to the United States from abroad or who move within the United States across states in a given year and relate their new locations and the wages they receive there to the real exchange rate, whose variation at a yearly frequency is primarily driven by the *nominal* exchange rate. As the CPS data does not include the necessary information to compute city price levels with Moretti (2013)'s method, we proxy prices by population levels for this exercise.

To obtain yearly elasticities of the immigrant concentration with respect to city size, we estimate equation (3.1) separately for each year.[20] As the CPS contains a much lower number of observations than the Census, we are not able to calculate wages of immigrants at the skill-MSA-time level to estimate equation (3.2) (i.e. we cannot adjust for potential imperfect substitutability between natives and immigrants). Instead, we account for the potential underlying individual heterogeneity by estimating for each year standard Mincerian wage regressions using individual data, including an immigrant indicator, a dummy variable for each skill cell (again defined by experience and education groups), and additional controls for race, marital status, and 82 occupation categories. The dependent variable is the weekly wage, and the city size elasticity of the wage gap is identified by an interaction between the immigrant dummy and the MSA population.

Figure 3 plots the estimated elasticities against the yearly average real exchange rate between Mexico and the US. The two plots show a linear fit that goes in the expected direction. The lower the prices in Mexico relative to US prices, the more positive the elasticity of the relative share of Mexican immigrants and the more negative the elasticity of the wage gap with respect to city size. In other words, Mexicans concentrate more in large cities and are paid lower wages relative to natives when the real exchange rate is lower.

We next exploit variation across regions *within* Mexico to relate origin characteristics with location choices in the US. For this, we use data on bilateral flows of Mexican migrants between particular states of origin in Mexico and states of destination in the US. As before, if consumption in Mexico (e.g. through remittances or circular migration, plays an important role, immigrants from cheaper origin states should disproportionately move towards more expensive destinations in the United States). Unfortunately, we do not have data on price levels for Mexican states. Instead, we use gross domestic product (GDP) per capita at the state level to proxy for local prices.

[18] Due to the inclusion of the interaction between the RER and local price level, this effect varies between around 0.3 percentage points in the cheapest cities and 0.5 in the most expensive cities.

[19] In this paper, we complement the evidence shown in Nekoei (2013). Exchange rate fluctuations affect not only the intensive margin of immigrant labor supply decisions (i.e., hours worked) but also, and very importantly, the extensive margin (i.e., location choices).

[20] To avoid losing MSAs with no Mexican immigrant movers, we do not take the log of the left-hand side of equation (3.1).

More specifically, to investigate the heterogeneity in Mexican migration flows by state of origin, we use the following estimation equation:

$$\ln F_{ij} = \beta_1 \ln P_i + \beta_2 \ln P_j + \beta_3 \ln P_i * \ln P_j + (\delta_i + \delta_j) + \varepsilon_{ij}, \tag{3.7}$$

where F_{ij} is the fraction of all migrants from Mexican state of origin i that move to US state of destination j, P_i is the GDP per capita in i, and P_j is the GDP per capita in j. In some specifications we also control for the level of population at origin and destination, or include origin and destination fixed effects. Our main hypothesis is that β_3 is negative, in which case there are relatively less flows to high GDP per capita destinations from high GDP per capita origin states.

Table 4 shows the estimation results. Mexicans of all origins tend to disproportionately move to US states with high GDP per capita. This can be read as further corroborating that immigrants prefer to move to high-wage locations. The positive coefficient of the GDP per capita in the origin state indicates that, overall, wealthier states send more migrants. Selection or credit constraints when migrating can explain this result (Angelucci 2015). The interaction between origin and destination GDP per capita is negative, indicating that immigrants from low GDP per capita states disproportionately locate in high GDP per capita capita destinations. Column 2 expands the regression in column 1 by including population controls. Again, the estimate of β_3 is negative and significant. Column 3 includes origin and destination fixed affects. Column 4 excludes California, the largest receiving state. All these specifications confirm the main finding.

Heterogeneity by attachment to the United States

On top of price variation, there are other factors that may generate heterogeneity in the importance of the host country for an immigrant. In this section, we explore heterogeneity along one such factor, namely the number of years that immigrants spend in the US. In online Appendix A.3 we further explore heterogeneity by legal status using the Immigration Reform and Control Act (IRCA) of 1986 as exogenous shock.

According to Dustmann and Mestres (2010), immigrants who do not intend to return to their countries of origin remit a smaller share of their income. They are also less likely to spend time back home and, thus, are more similar to natives. Hence, if immigrants who have been longer in the US care less about consuming in the home country, we should find that immigrants concentrate more and the wage gap is larger in more expensive cities for immigrants who arrived more recently.

To investigate this idea, we use the year of immigration available in the Census data and divide immigrants into groups of five-year intervals depending on their time spent in the United States. We then estimate equation (3.4), but instead of the origin o, we disaggregate the immigrant concentration and the wage gaps by the groups defined by time in the US and estimate β for each of these groups. We plot the different city price-immigrant concentration and city price-wage gap elasticities in Panels A and B of Figure 4. As before, Panel B is the mirror image of Panel A. The negative and strongly linear fit in Panel A indicates that the relationship between immigrant concentration and city prices decreases with the length of immigrants' stay in the United States. This means that immigrants concentrate less in expensive US cities as they spend more and more time in the US, although the variation of the coefficient is not large, probably indicating stickiness in location choices. Similarly, as shown in Panel B, the differences in the wage gap in expensive relative to cheap locations is less pronounced as immigrants stay longer in the US.

3.3 Immigrants' consumption patterns

We argue that the previous results on relative wages and the geographic distribution of immigrants are at least partly driven by the fact that immigrants spend part of their incomes in their countries of origin. Related to this mechanism, Yang (2006) provides evidence that exchange rate fluctuations affect consumption in the origin countries using data for households in the Philippines and migrants in the US. In particular, an appreciation of the US dollar leads to a higher probability of vehicle ownership and entrepreneurial income in households with members that currently reside in the US. Moreover, in line with this evidence, Yang (2008) shows that an appreciation increases remittances sent home by Filipino migrants.

Dustmann and Mestres (2010) report that immigrants in Germany remit around 10% of their income. While data of similar quality as that used in their study do not exist for the US, we can compute the overall flow of outgoing remittances based on the Bilateral Remittance Matrix provided by the World Bank.[21] The earliest available year is 2010, which is also the latest year of ACS data we use in this paper. Hence, we base our calculation of the remittance share of income on the total remittances leaving the US during that year. This figure amounts to around \$110 billion. To obtain the total disposable income of immigrants residing in the US, we sum up their total wage income and multiply it with the average US tax burden of around 32%, which results in \$540 billion. Thus, the remittance share of wage income is a very sizeable 20%.[22]

We now turn to the analysis of the local consumption of immigrants in the host economy relative to natives. As described in Attanasio and Pistaferri (2016), measuring consumption is not an easy task with difficulties arising from measurement issues and from the treatment of durable goods. Our assumption is that total observed consumption can be decomposed into local consumption (part of which can be separately observed), savings, and remittances (or income spent in another country). Hence, if we find that immigrants spend less than natives locally (conditional on income), this must imply that they either save more (possibly to spend the savings after return to their home country), or spend a part of their income on remittances. Hence, by looking at local consumption we can infer something about whether immigrants are likely to spend an important share of their income in their countries of origin, as hypothesized.

More concretely, we investigate whether immigrants consume less locally than natives, using two alternative data sets: the Consumption Expenditure Survey (CEX) and the Census. The former allows us to investigate overall consumption but only identifies Mexican immigrants, as explained in more detail in Section 2. The latter allows us to identify immigrants from multiple countries of origin but contains only expenditures on housing, which represent around 25% of total expenditures (Davis and Ortalo-Magne 2011).

We start by analyzing the overall local consumption with CEX data using the following regression:

$$\ln \text{Total Expenditure}_{i,s,t} = \alpha + \beta \text{Mexican}_{i,s,t} + \sum_j \gamma_j \text{HH Income category } j_{i,s,t} + \eta X_{i,s,t} + \delta_s + \delta_t + \varepsilon_{i,s,t} \quad (3.8)$$

where the dependent variable is the quarterly total expenditure at the household level.[23] CEX data identifies income only by category; hence we use income bracket dummies indexed by j. *Mexican* is a dummy variable identifying households of Mexican origin, and δ_s and δ_t are state and time dummies, respectively. The coefficient of interest (β) measures the difference in total expenditures between Mexicans and non-Mexicans conditional on

[21] https://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data

[22] Considering also non-wage income, the remittance share would be around 18%. For this exercise, we include all immigrants aged 18 to 65.

[23] More specifically, we use the variable "totexpcq" from the Consumer Expenditure Survey. This variable combines expenditures on all items.

observable characteristics, most importantly income categories. The location fixed effects (δ_s) ensure that the identification of β comes from within-location comparisons.[24]

The results, reported in Panel A of Table 5, suggest that, unconditionally, Mexican households consume on average 33% less than non-Mexican households, and as much as 38% when we force within-state comparisons (column 2). In column 3, we include income controls, which reduce the estimate to 15%. When we add all controls in column 4, which include age, family size, race, and marital status, we find that Mexican households consume around 22% less locally than non-Mexican households. With close to 0 saving rates (in the early 2000s the savings rate was around 2%), this number also represents the share of income that is potentially devoted to consuming in the home country and aligns well with the average income share of remittances.[25]

An important part of local expenditures are housing costs, for which we have information in both the CEX and Census data. In Panel B of Table 5, we repeat the regressions of Panel A but use rent expenses as the dependent variable. The patterns are similar to those we observed in Panel A, showing that an important part of the difference between how immigrants and natives consume is likely driven by housing expenses. In particular, columns 3 and 4 suggest that Mexicans consume around 10% to 20% less than similar looking non-Mexicans.

To explore in more detail how immigrants and natives consume housing, in Panel C we turn to Census data, in which we can observe both ownership status and the rental expenses for renters. In Table E5 in online Appendix E, we show that immigrants are less likely to own the place where they live. Here, we concentrate on renters by running the same regression as in equation 3.8 but using a continuous measure of income and MSA instead of state fixed effects.[26] When including the MSA fixed effect in the second column of Panel C, we find that immigrant households spend around 14% less on rents than observationally similar natives, consistent with the results in Panel B. Additionally controlling for household income shrinks the coefficient to a bit less than 10%.

In column 4, we restrict the sample to immigrant households and include the real exchange rate as a predictor to check whether immigrants from more expensive countries spend more on housing, conditional on income, household characteristics, and MSA fixed effects. Consistent with the model that we present in Section 4, immigrants from more expensive origins spend relatively more on housing in the host economy. One way to assess the magnitude of the estimates is to note that if immigrants spend between 10% to 20% less (as estimated in Panel A), and housing rents represent around 25% of income, the share of income devoted to housing should be around 3% to 5% lower for immigrant households (0.15×0.25), broadly consistent with the direct estimates of this number provided in Panel B of Table E6 in online Appendix E.[27]

Overall, these results suggest that immigrants spend less than natives on local goods, with only small differences across MSAs, and that immigrants coming from more expensive countries spend more income on housing in the host country. We return to these results in Section 4.4 when we compare the heterogeneity in consumption patterns as a function of the real exchange rate predicted by the model and found in the data.

[24] In these data, the lowest level of geographic disaggregation is the state.

[25] As can be seen in data from the St Louis Fed (`https://fred.stlouisfed.org/series/PSAVERT`), the aggregate personal savings rate has fluctuated between 2% and 12% since 1980. See also Dynan et al. (2004) for a discussion on how savings are lower for low-income households.

[26] Controls include a vector of individual characteristics, including dummies for the number of persons present in the household, sex, marital status, and age. An alternative approach that yields similar results is to use the share of income devoted to paying housing rents as dependent variable (either in logs or levels), which we show in Table E6 in online Appendix E.

[27] A direct comparison is not possible given that Consumer Survey data identifies Mexicans, while US Census data identifies all immigrants.

4 A spatial equilibrium model with immigration

In this section, we introduce a quantitative spatial equilibrium model with immigrants that consume part of their income in their countries of origin. This model rationalizes the empirical evidence presented so far.

4.1 Model setup

Utility and location choices

The utility of individual i from country of origin j in location c is given by:

$$\ln U_{ijc} = \rho + \ln A_{jc} + \alpha_t \ln C_T + (1-\alpha_t)\frac{\sigma}{\sigma-1}\ln\left(\frac{\alpha_l}{\alpha_l+\alpha_f}(C_H)^{\frac{\sigma-1}{\sigma}} + \frac{\alpha_f}{\alpha_l+\alpha_f}(C_F)^{\frac{\sigma-1}{\sigma}}\right) + \ln\varepsilon_{ijc},$$

where α_t denotes the expenditure share devoted to tradable goods C_T, and α_l and α_f denote the expenditure weights of local non-tradable goods C_H and foreign goods C_F, respectively. The elasticity of substitution between local non-tradable and foreign goods is denoted by σ.[28] A_{jc} denotes the utility derived from local amenities, which consist of a component A_c that yields the same utility to both natives and immigrants independent of their origin, and a component $\tilde{A}_{jc}$ that is specific to immigrants from origin j. Thus, $\ln A_{jc} = \ln A_c + \ln \tilde{A}_{jc}$. Finally, ε_{ijc} is a Frechét distributed idiosyncratic taste shock for living in location c and ρ is a constant that ensures there is no constant term in the indirect utility function to be derived in what follows.

Note that there are alternative interpretations for what foreign consumption C_F represents. It could include consumption of non-tradables in the home country, remittances sent to relatives, or future consumption in the home country. Rather than attempting to model the specificities of each of these channels explicitly, we opt for a simple formulation that encapsulates all of them.

Note that C_H represents the consumption of housing and other non-tradable goods which need to be consumed in location c. For simplicity, we will henceforth refer to C_H as housing.

Individuals maximize their utility subject to a standard budget constraint, given by:

$$C_T + p_c C_H + p_j C_F \le w_{jc},$$

where p_c is the price of housing and p_j is the price of foreign goods denominated in the home currency. The price of tradable goods is the numeraire and therefore equal to 1.

We assume $\alpha_t + \alpha_l + \alpha_f = 1$ and define the auxiliary parameters: $\bar{\alpha}_l = \frac{\alpha_l}{\alpha_l+\alpha_f}$ and $\bar{\alpha}_f = \frac{\alpha_f}{\alpha_l+\alpha_f}$. By utility maximization, we then obtain the following indirect utility of living in each location (derivation in Appendix B.1):

$$\ln V_{ijc} = \ln V_{jc} + \ln\varepsilon_{ijc} = \ln A_{jc} + \ln w_{jc} - (1-\alpha_t)\ln \bar{p}_{jc}(\bar{\alpha}_l, \bar{\alpha}_f) + \ln\varepsilon_{ijc}, \tag{4.1}$$

where

$$\bar{p}_{jc}(\bar{\alpha}_l, \bar{\alpha}_f) = (\bar{\alpha}_l^{\sigma} p_c^{1-\sigma} + \bar{\alpha}_f^{\sigma} p_j^{1-\sigma})^{\frac{1}{1-\sigma}}.$$

Note that V_{jc} captures the value of living in c for individuals from j, net of the idiosyncratic component.

[28] We opt for allowing only non-tradable goods to be substitutable with foreign goods for three main reasons. First, tradable goods to a large extent contain periodically purchased non-durable goods like food products that cannot be easily substituted with origin consumption. Second, due to their very nature, many durable tradable goods have similar prices across countries, which eliminates any motive for substitution. Third, the empirical evidence suggests that immigrants primarily save on housing expenses relative to natives. Allowing for a more flexible substitution across goods would yield similar insights as this model.

Given this indirect utility, workers decide where to live by selecting the location that delivers the highest level of indirect utility given the realization of the taste shock. Denoting the inverse shape parameter of the distribution of ε_{ijc} by λ, which governs the variance of the idiosyncratic taste shocks, the outcome of this maximization yields:[29]

$$\pi_{jc} = \frac{V_{jc}^{1/\lambda}}{\sum_k V_{jk}^{1/\lambda}} = \left(\frac{V_{jc}}{V_j}\right)^{1/\lambda}, \tag{4.2}$$

where $V_j = (\sum_k V_{jk}^{1/\lambda})^{\lambda}$ is the expected value for living in this economy for workers from country of origin j.

π_{jc} denotes the share of workers from j that decide to live in city c, which depends on the indirect utility in location c relative to all other locations. Note that the indirect utility increases in wages and local amenities and decreases in local prices. Thus, locations with higher wages, higher amenity levels, and lower price indices will attract more people.

Production of tradable goods

Each worker inelastically supplies one unit of labor to local firms, which produce tradable goods with a linear production function using labor as the only input. The productivity of firms varies at the local level, so that the total output of tradable goods in city c is given by:

$$Q_c^T = B_c L_c, \tag{4.3}$$

where $L_c = \sum_j L_{jc} = \sum_j \pi_{jc} L_j$ is the sum of workers from all origins who live in c. Note that, based on the evidence presented in the empirical part of this paper, we assume that natives and immigrants are perfect substitutes at the city level.[30]

The marginal revenue of hiring an extra worker is given by B_c. The cost of hiring an additional worker, possibly from origin j, is the wage that the worker receives. Thus, the firm profit generated by hiring an additional worker is given by $B_c - w_{jc}$. The average cost of hiring workers across all the cities is given by $\bar{w}$. We normalize the values of the productivity and amenity parameters B_c and A_c to be equal to 1 on average. As long as these values do not deviate too far from their average, their logarithmic values can thus be interpreted as percentage deviations from the average. We can then approximate firm profits by $(B_c - w_{jc}) \approx B_c - 1 - \ln w_{jc} = \ln B_c - \ln w_{jc} = S_{jc}^{Firm}$. This expression gives the marginal value of a hired worker for a firm.

Labor market

We assume that there is imperfect competition in labor markets and that wages are an outcome of a negotiation between firms and workers. Most of the literature on immigration has assumed that labor markets are perfectly competitive. However, there is ample evidence that a number of phenomena are better understood with non-competitive labor markets. This includes, for example, the long-lasting consequences of job loss or year of entry into the labor market (Oreopoulos et al. 2012), the literature on discrimination in the labor market (Becker 1957; Black 1995), and most of the literature investigating unemployment.[31] In this paper, we argue that the facts shown

[29] In principle λ could be different for natives and immigrants. However, Amior (2020a) provides strong evidence for assuming the same λ for natives and immigrants.

[30] It is straightforward to introduce imperfect substitutability between natives and immigrants by assuming that $L_c = ((\sum_{j \neq N} L_{jc})^{\rho_I} + L_{Nc}^{\rho_I})^{1/\rho_I}$, where L_{Nc} indicates natives in location c. Introducing imperfect native-immigrant substitutability cannot explain the heterogeneity of the empirical patterns documented in subsection 3.2 and complicates the algebra. Hence, we abstract from it.

[31] In the migration literature, Amior and Manning (2020) study the effects of immigration in a monopsonistic framework without taking into account local labor markets and the spatial equilibrium of the host economy.

in our empirical section are explained quite naturally by assuming non-competitive labor markets, although this assumption is not key for our predictions on relative location patterns, which also hold in a setting where wages and marginal products are equalized. However, we need to assume frictional labor markets to make the model consistent with the evidence that immigrant-native wage gaps, even after controlling for skill heterogeneity or imperfect substitutability, are related to local price levels and real exchange rates.[32]

Firms and workers meet, negotiate over wages, and split the total surplus of the match. A worker's surplus in matching with a firm is given by $S_{jc}^{Worker} = \ln V_{jc}$. Hence, since workers' taste shocks are unobservable to firms, we assume that the surplus relevant for bargaining is the utility in location c given a zero taste shock. The worker's outside option is receiving the average utility across all cities, which is normalized to one. That is, the surplus of accepting a job in city c is the local indirect utility obtained in that city relative to the average city.

The outcome of the negotiation between workers and firms is determined by Nash bargaining. Workers' weight in the negotiation is given by a constant β that is independent of endogenous variables. Thus, a (constant) share β of the total surplus generated by a match accrues to workers. Using this assumption, we can determine the wage levels of a worker from country of origin j living in location c:

$$\ln w_{jc} = \beta \ln B_c + (1-\beta)[(1-\alpha_t)\ln \bar{p}_{jc} - \ln A_{jc}] \tag{4.4}$$

This equation shows standard results from the spatial economics literature. Higher wages in a city reflect lower amenity levels, higher local productivity, or higher local price indices.

The firm profits are given by $\sum_c B_c L_c - \sum_c \sum_j w_{jc} L_{cj}$ and are positive as long as B_c is sufficiently large because of the non-competitive nature of the labor markets. We assume that they are distributed to absentee firm owners, who only consume tradable goods.[33]

Housing market

The supply of housing is provided by combining land, which is a fixed factor, and final tradable good as inputs. Denoting by $\tilde{\eta}_c$ the weight of land in the production of housing, we have that the supply of housing is proportional to housing prices.[34] This can be represented by

$$Q_c^H = \left(\frac{p_c}{p_{0,c}}\right)^{\frac{1-\tilde{\eta}_c}{\tilde{\eta}_c}}. \tag{4.5}$$

Total demand for housing is given by the sum of the demands of natives and immigrants. The demand for housing differs between natives and immigrants because of differences in wages and weights of housing in total expenditures. Local housing prices are defined by market clearing in each city:

[32] Since immigrants disproportionately locate in more expensive cities, one could also generate a relationship between wage gaps and city prices as found in the data by assuming perfectly competitive labor markets and imperfect substitutability of immigrants and natives in production. With the appropriate calibration, such a model would yield similar predictions. However, such a model would *not* be able to account for the documented heterogeneity in the price elasticity of the wage gap across immigrant groups, unless we made the perhaps not very realistic assumption that immigrants from different countries of origin are different factors of production.

[33] Alternatively, we could assume that firm profits are distributed to workers who each hold a representative portfolio of the firms in the economy. While this would have implications for welfare, it would not affect the main predictions of the model in terms of location choices, wages, and local prices.

[34] Profit maximization of the construction of housing with the production function $Q_c^H = H(Q^T, F_c) = (Q^T)^{1-\tilde{\eta}_c} F_c^{\tilde{\eta}_c}$ leads to a total production of housing equal to $Q_c^H = (1-\tilde{\eta}_c)^{\frac{1-\tilde{\eta}_c}{\tilde{\eta}_c}} (p_c)^{\frac{1-\tilde{\eta}_c}{\tilde{\eta}_c}} F_c$, where land is denoted by F_c and owned by absentee landlords. A number of papers in this literature make this assumption. See, as an example, Eeckhout et al. (2014). Alternatively, the return on land, $r_c F$ – where r_c represents land prices – can be distributed to workers, who each hold a representative portfolio of the land in the economy.

$$(\frac{p_c}{p_{0,c}})^{\eta_c} = (1-\alpha_t)(\sum_{j\neq N}((\frac{\bar{\alpha}_l}{p_c})^{\sigma}\bar{p}_{jc}^{\sigma-1}L_{jc}w_{jc}) + \frac{1}{p_c}L_{Nc}w_{Nc}), \tag{4.6}$$

where $\eta_c = \frac{1-\tilde{\eta}_c}{\tilde{\eta}_c}$ is the elasticity of housing supply. Note that when land is more important in production, the elasticity of housing supply is lower.

4.2 Properties

Given these primitives of the model, in this subsection we derive a number of properties, which are the basis for the structural estimation described in Section 4.3. The difference between natives and immigrants is the weight they give to local and foreign price indices.

Assumption. *Natives only care about local price indices so that $\alpha_f = 0$ and $\alpha_l = \alpha$. Immigrants care about local and foreign price indices so that $\alpha_f \neq 0$ and $\alpha_l + \alpha_f = \alpha$.*

Proposition 1. *There is a gap in wages between natives and immigrants that is given by the following expression, which increases in the local price level p_c. The increase becomes steeper with lower origin prices p_j, iff $\sigma > 1$.*

$$\ln\frac{w_{Nc}}{w_{jc}} = (1-\beta)[(1-\alpha_t)\ln\frac{p_c}{\bar{p}_{jc}} + \ln\tilde{A}_{jc}] \tag{4.7}$$

Proof. Online Appendix B.2

□

This expression highlights that, if $\beta = 1$, there is no wage gap between natives and immigrants, implying that workers always receive a wage equal to productivity. Thus, the utility derived from both local amenities and prices of non-tradables is not reflected in wages. This would be the case with perfectly competitive labor markets. In case $\beta < 1$, the size of the wage gap depends on the consumption share of non-tradables, relative prices, and origin-specific amenities.

It is also worth noting that differences in the price index of the country of origin do not play a direct role *only* in the special case of $\sigma = 1$. In this case, the share of expenditure in home country goods is always the same, irrespective of the prices, as income and substitution effects cancel each other out. If, instead, $\sigma > 1$, the substitution effect dominates, and origin consumption increases when origin goods become relatively cheaper.

In Section 3, we show empirically that immigrants concentrate in higher proportions in more expensive cities. This is captured by the following proposition.

Proposition 2. *The spatial distribution of immigrants relative to natives is given by the following expression, which increases in the local price level p_c. The increase becomes steeper with lower origin prices p_j, iff $\sigma > 1$.*

$$\ln\frac{\pi_{jc}}{\pi_{Nc}} = \frac{\beta}{\lambda}[(1-\alpha_t)\ln\frac{p_c}{\bar{p}_{jc}} + \ln\tilde{A}_{jc}] + \ln\frac{\sum_k\left(A_kB_k/p_k^{1-\alpha_t}\right)^{\frac{\beta}{\lambda}}}{\sum_k\left(A_{jk}B_k/\bar{p}_{jk}^{1-\alpha_t}\right)^{\frac{\beta}{\lambda}}} \tag{4.8}$$

Proof. Online Appendix B.2 □

In the limiting case of $\beta = 1$, immigrants benefit most from locating in more expensive cities as they enjoy the same wages as natives despite being less affected by the local price levels. Therefore, differences in price levels and origin-specific amenities are reflected exclusively in the relative probabilities of choosing a specific city. On the other hand, if firms have all the bargaining power (i.e. when $\beta = 0$), the wages offered to immigrants fully take into account the prices they face, which implies that they do not benefit more than natives from living in expensive cities. Hence, the spatial distribution of natives and immigrants would be the same.

Propositions 1 and 2 are directly linked to the facts that we report in Section 3. They show that the concentration of immigrants is higher and that immigrants receive lower wages than natives in expensive cities. We can use the allocation of workers across locations to obtain the equilibrium size of the city, given the total native and immigrant populations (L_N and L_j for each country of origin j).

Proposition 3. *The equilibrium size of the city is given by the following equation:*

$$L_c = \sum_j \frac{(A_{jc}B_c/\bar{p}_{jc}^{1-\alpha_t})^{\frac{\beta}{\lambda}}}{\sum_k (A_{jk}B_k/\bar{p}_{jk}^{(1-\alpha_t)})^{\frac{\beta}{\lambda}}} L_j + \frac{(A_c B_c/p_c^{1-\alpha_t})^{\frac{\beta}{\lambda}}}{\sum_k (A_k B_k/p_k^{1-\alpha_t})^{\frac{\beta}{\lambda}}} L_N \tag{4.9}$$

Proof. Online Appendix B.2 □

Note that this proposition also implies that immigrants make large cities even larger, as they care less than natives about the higher living costs in large cities (i.e., congestion). Moreover, it shows that cities are large because they are either productive (B_c) or pleasant to live in (A_{jc}). Thus, conditional on amenity levels, immigration tilts the population distribution towards more productive cities.

To see the aggregate effect of immigration on total output via immigrants' location choices, we can obtain an expression of total output per capita depending on immigrant shares, if we divide (4.9) by overall population and take the sum across cities, weighted by productivity B_c:

$$q = \frac{1}{L} \sum_c \left[\sum_j \frac{(A_{jc}B_c^{\frac{\beta+\lambda}{\beta}}/\bar{p}_{jc}^{1-\alpha_t})^{\frac{\beta}{\lambda}}}{\sum_k (A_{jk}B_k/\bar{p}_{jk}^{(1-\alpha_t)})^{\frac{\beta}{\lambda}}} L_j + \frac{(A_c B_c^{\frac{\beta+\lambda}{\beta}}/p_c^{1-\alpha_t})^{\frac{\beta}{\lambda}}}{\sum_k (A_k B_k/p_k^{1-\alpha_t})^{\frac{\beta}{\lambda}}} L_N \right] \tag{4.10}$$

A final property of the model is that immigrants have a heterogeneous effect on the demand for housing across locations. Although they add to overall housing demand, they spend a smaller fraction of their income on housing as natives. As a result, immigration increases the demand for housing to a lesser extent than it would, if immigrants consumed like natives. Hence, immigrants tend to reduce the per capita demand for housing at the local level. However, as immigrants are drawn towards expensive cities more so than natives, the increase in the local aggregate demand for housing is larger in these cities, exacerbating house price disparities across locations.

4.3 Estimation

We estimate the model by multivariate feasible generalized least squares based on equations (4.7) and (4.8). Conveniently, these equations give us expressions for the wage gaps and immigrant concentrations at the city-origin level without the need to determine the common amenity values A_c and productivity levels B_c. We only need the origin-specific amenity levels as well as local and origin prices to compute $p_c/\bar{p}_{jc}$ as functions of the parameters α_f and σ. We base the estimation on the 1990 sample of 185 MSAs and 68 origin countries used in our empirical section and measure local prices as the rent component of the city-specific price index, computed following Moretti (2013), and origin prices as real exchange rates. The wage gaps are those that we use as the dependent variable in regression equation (3.4). The immigrant concentration, computed as before, is the share of immigrants from origin j in city c among all immigrants from that origin divided by the share of natives in city c among all natives.

There are four key parameters in the model affecting consumption, relative wages, and location choices of immigrants. First, α_f governs the (average) income share of expenditures in the origin country. Second, σ

measures the sensitivity of consumption shares to relative prices. Third, λ and β jointly govern how much of the differences in local utility levels are reflected in wages versus location choices. As we cannot identify α_t and β separately from the four parameters stated above with the data we use in the estimation, we set the former equal to 0.4, which is close to the 0.38 used in Moretti (2013) (see further details justifying the choice of this value in online Appendix C.1) and the latter equal to 0.5.[35]

In order to jointly estimate λ, α_f, σ and $\tilde{A}_{jc}$, we impose the following structure on $\tilde{A}_{jc}$ to reduce the dimensionality of the parameter vector:

$$\tilde{A}_{jc} \equiv \exp(\phi \; Net_{jc} + \nu_{jc})$$

Thus, we assume that the log origin-specific amenity value of city c is a linear function of the existing immigrant network Net_{jc} plus an i.i.d. error term capturing unobserved heterogeneity. We measure Net_{jc} as the 10-year lagged share of immigrants from origin j among the total population in city c. From this, we obtain the estimating equation

$$\begin{bmatrix} \hat{w}_{jc} \\ \hat{\pi}_{jc} \end{bmatrix} = 0.3 \begin{bmatrix} 1 \\ 1/\lambda \end{bmatrix} \hat{p}_{jc}(p_c, p_j, \sigma, \alpha_f) + \phi \begin{bmatrix} Net_{jc} \\ Net_{jc} \end{bmatrix} + \begin{bmatrix} 0 \\ \Psi_j \end{bmatrix} + \begin{bmatrix} \nu^w_{jc} \\ \nu^\pi_{jc} \end{bmatrix},$$

where $\hat{w}_{jc}$ and $\hat{\pi}_{jc}$ are the (inverse) immigrant-native wage gap and the immigrant concentration from the data, respectively, $\hat{p}_{jc} \equiv \ln(p_c/\bar{p}_{jc})$, and Ψ_j is the last term on the right-hand side in equation (4.7). The residuals ν^w_{jc} and ν^π_{jc} are linear combinations of the unobserved heterogeneity ν_{jc} and measurement error in wage gaps and immigrant concentrations, respectively. We derive a consistent estimator of $\theta \equiv (\sigma, \alpha_f, \lambda, \phi)$ by defining the residual vector $\bar{\nu}_{jc} \equiv [\nu^w_{jc}, \nu^\pi_{jc}]'$ and imposing the orthogonality condition $E[\bar{\nu}_{jc}|\hat{p}_{jc}, Net_{jc}] = 0$. Hence, the identifying assumption is that the unobserved residuals are exogenous to immigrant price indices and networks.

As both α_f and σ affect immigrants' expenditure on origin goods (see the expression for C_F in online Appendix B.1), we add an additional restriction in order to separately identify the two parameters: we set the economy-wide average of the expenditure share equal to its data equivalent of 0.2.

Hence, we define the estimator as

$$\hat{\theta} \equiv \arg\min_{\theta} \quad \sum_{jc} \left[\bar{\nu}_{jc}(\theta)\right]' W \left[\bar{\nu}_{jc}(\theta)\right] \quad \text{s.t.} \quad \frac{\sum_j \sum_c \omega_{jc} p_j C_{F,jc}}{\sum_j \sum_c \omega_{jc} w_{jc}} = 0.2, \tag{4.11}$$

where the weights ω_{jc} are the shares of immigrants from origin j living in city c in the total immigrant population across all cities. We obtain our final estimate $\hat{\theta}$ using a two-step procedure. First, we estimate the parameter vector $\tilde{\theta}$ that minimizes (4.11) using the identity matrix as weighting matrix W. We then repeat the minimization substituting it with the estimated covariance matrix $\hat{W}(\tilde{\theta}) = \left[\frac{1}{JC}\sum_{jc} \hat{\nu}_{jc}\hat{\nu}'_{jc}\right]^{-1}$ in order to obtain $\hat{\theta}$ in the second step.

Table 6 presents the values of the estimated parameters. Our estimate of λ is 0.043. The implied (internal) migration elasticity (which is $1/\lambda$) is somewhat higher than other estimates in the literature (see Monras 2015; Caliendo et al. 2015). However, the comparison is not straightforward since the variation used in previous literature is quite different from the one used in this paper and previous papers did not have non-competitive labor markets. The elasticity of substitution between local and origin goods is 1.86, α_f is 0.17, and the sensitivity of the amenity

[35] In online Appendix C, we report the estimates of α_f and σ for alternative assumptions on α_t and β. Within an interval of 0.2 to 0.6 for the former and 0.2 to 0.8 for the latter, we find that the estimates of α_f vary between 0.13 and 0.19 and those of σ between 1.5 and 2.5. The model predictions remain virtually unchanged for any values of α_t and β within these intervals.

value to the immigrant network ϕ is 1.49.

As a final step, we use the calibrated and estimated parameters as well as natives' wages from the data to compute the productivity levels B_c and common amenity values A_c implied by the model. These are needed in our counterfactual simulations, in which we vary the aggregate native and immigrant population levels, in order to compute the full general equilibrium. We solve for the productivity parameters using equation (4.4).[36] The indirect utility of natives is given by equation (4.1) after substituting A_{jc} with A_c and $\bar{p}_{jc}(\bar{\alpha}_l, \bar{\alpha}_f)$ with p_c. Thus, with our estimate of λ, we can determine the common component of city amenity levels A_c by matching the observed native population distribution with its counterpart predicted by the model through equation (4.2). Further, when conducting counterfactual simulations, we make use of the housing market equations (4.5) and (4.6) to calculate the new equilibrium house prices resulting from changes in housing demand combined with the housing supply elasticities from Saiz (2010). For this, we back out the parameter $\bar{p}_{0,c} = p_c^*(Q_c^{H*})^{-1/\eta_c}$, where p_c^* is observed prices in the baseline year 1990 and Q_c^{H*} is the housing supply that equals demand obtained from equation (4.6).

4.4 Comparison of the model versus the data

Goodness of fit

With the estimated parameters at hand, we can inspect the goodness of fit of the model by comparing predicted aggregate outcomes with their data equivalents. In particular, based on the predicted distributions of immigrants and natives across cities and their wages in each location for each country of origin, we compute the overall immigrant population shares and wage gaps (aggregating across origins) for each city and plot it along with the data for 1990 in Figure 5. Because ultimately we are interested in how the origin consumption channel built into our model affects cities depending on their productivity, we sort MSAs by their productivity level along the x-axis. For the purpose of illustrating the predictive power of this channel, Panel A and B show the model predictions when shutting down the possibility of consumption at origin. That is, $\alpha_f = 0$ for all workers, and therefore the only difference between natives and immigrants is that the latter derive utility from network amenities. Panels C and D show the predictions in the full model.

As Panel A of Figure 5 suggests, allowing immigrants' relative location choices to be driven exclusively by network amenities results in an immigrant population share very close to the economy-wide average of 12.2% in the majority of cities. This reduced version of the model is only able to predict higher immigrant shares in cities with large networks of Mexican immigrants, among them three cities at the lower end of the productivity distribution, which are located in Texas directly at the US-Mexican border (Brownsville, McAllen, and El Paso), several cities with intermediate productivity mostly in California (e.g., Salinas, Fresno, and San Diego), and Los Angeles at the high end of the productivity distribution. Since the relative wages of natives and immigrants are also affected by network amenities, wage gaps are around 0 except for the aforementioned group of cities with large networks.

When allowing for consumption heterogeneity between natives and immigrants, the model is able to replicate well both the general increase of the immigrant share with city-level productivity, driven by the higher price levels in more productive cities, and the higher shares in the three border cities with low productivity but a large share of (Mexican) immigrants. Naturally, there is more unexplained variation in the wage gap, especially for smaller

[36] Native wages are measured as average residuals in city c obtained from a Mincerian regression, as described in Section 2.1.

cities, as the low number of immigrant observations in the Census data leads to very volatile wage gaps. However, the model is able to replicate the decrease of immigrants' relative wages with productivity.

Non-targeted moments I: Heterogeneity in local expenditures by origin

Since immigrants spend a part of their income in their origin countries, more so if the real exchange rate is favorable, we can use the model to predict the expenditure on local non-tradable goods by origin. It is not straightforward to find a data equivalent for this expenditure, which is why we do not directly use it as a targeted variable in the model estimation. However, we can check whether the model predictions are consistent with the data by correlating them with the income share of rent expenses in the data, which we have already used in Section 3.3. To calculate these shares, we rely on rent expenditure from the Census, as the CEX data does not allow us to identify immigrants by origin. Specifically, we use the rent expenses divided by household income as the dependent variable in the same specification shown in Table 5, Panel C, column 3, replacing the immigrant dummy with origin country fixed effects. We then plot these fixed effects against the deviation of the share of local non-tradable expenditure from its average in the model in Figure 6. We find a tight significant relationship between data and model predictions (the regression coefficient is 1.73 with standard error 0.268). Naturally, the share of rent expenses in the data shows a lower variation (between -6% and +6% deviation from the mean) than the share of local non-tradable expenses in the model (between -10% and +10% deviation from the mean according to the linear fit), as the latter contains more goods than just housing.

Non-targeted moments II: Immigration over the 1990s

As a second check on non-targeted moments, we use the model estimated using 1990 cross-sectional data to predict the change in the spatial distribution of natives and immigrants following the immigrant inflow that took place during the 1990s, the decade with the strongest increase in the total population share of immigrants in the US since the mid-19th century.

To do so, we change the native and immigrant population levels (aggregated across cities) from their values in 1990 to those in 2000 and compare the equilibrium population distribution across cities implied by the model before and after the change. Thus, we predict city population growth driven by the channels present in our model, abstracting from potential changes in relative amenity values and productivity levels across cities in the 1990s. During this decade, the overall urban immigrant population in the sample increased from 10.7 million to 18.8 million, while the urban native population increased from 87.5 million to 98.9 million.[37] Thus, the total population share of immigrants rose from 12% to 19%. More than a third of this increase was due to inflows from Mexico, which amounted to around 3.35 million, followed by India with 0.45 million and Vietnam with 0.39 million.

In our simulations, house price changes depend on the increase in housing demand and the elasticity of housing supply in each city (see equations (4.5) and (4.6)). These price changes together with the immigrant network and the origin consumption channel will be key in driving city-specific population growth. Based on the change in the spatial distribution of workers across cities, we can then compute the change in aggregate TFP between 1990 and 2000, which we discuss in Section 5.3. In order to separately quantify the contributions of the different channels operating in the model, we conduct three counterfactual simulations. In the first one, there is no immigration and

[37] Note that these numbers only refer to individuals aged 18-65 living in one of the 185 MSAs included in the sample.

thus population growth is only driven by natives.[38] In the second one, immigration takes place but immigrants are identical to natives (i.e. $\alpha_f = \phi = 0$). Finally, in the third counterfactual, we simulate the full model with immigrants' deriving utility from network amenities and goods consumed in their origin countries purchased with local income.

Figure 7 shows the immigrant and native growth rates for each city, predicted by the three counterfactual simulations (panels A to C) and the actual growth rates observed between 1990 and 2000 in the data (panel D). The rates are defined as the change in the respective population group (immigrants represented by blue dots, natives by red crosses) divided by the overall population of the group in each city.

By keeping the aggregate immigrant population constant, Panel A shows how the population distribution would have changed just from the fact that the native population residing in MSAs increased over the 1990s. The variation in native population growth rates across cities is purely driven by the heterogeneity in housing supply elasticities: natives locate disproportionately in cities with a higher elasticity. If housing supply elasticities were constant across cities, the native population growth rates predicted by the model would be constant and equal to the aggregate growth in native population. As can be seen in the plot, the elasticity tends to be higher in less productive cities. Hence, low-productivity cities can attract relatively more native workers than more productive cities, where housing is more constrained.

The simulation depicted in Panel B adds the immigrant population that entered the US during the 1990s to the overall immigrant population, but assumes that these immigrants are identical to natives. Hence, they do not derive utility from existing immigrant networks or origin consumption. As a result of the higher size of the overall population, natives (and immigrants) concentrate even more strongly in those cities that have elastic housing supplies. Note that while the incoming immigrants distribute identically across cities as natives, the shape of their inflow rates is different, because we compute the immigrant inflows relative to the immigrant population in the 1990s and immigrants in 1990 are differently distributed than natives.[39]

In Panel C of Figure 7, we finally simulate our baseline model incorporating both the network and origin consumption channels. Now, the distribution of the inflow rates looks drastically different. In particular, immigrant inflows increase with city productivity, with the exception of the predicted inflows above 20% for the three border cities with large Mexican communities.[40] While native inflows resemble those in Panel B at the lower end of the productivity distribution, natives get "pushed out" of some of the cities at the high end of the productivity distribution. These are, in particular, Los Angeles, San Francisco, and San Jose, which are the cities that have the strongest house price increases due their very inelastic housing supplies. Panel D of the figure plots the actual data of native and immigrant inflow rates. While the model cannot replicate the large variation in these rates for low and mid productivity, which are also driven by local amenity and/or productivity shocks that the model does not capture, the model is able to replicate the general pattern that immigrants disproportionately locate in high-productivity cities and those with large Mexican communities, while the opposite holds for natives. The model performs especially well with respect to the immigrant inflows for the most productive cities, predicting larger inflows for San Jose, San Francisco, and Washington DC and somewhat lower ones for Chicago, Philadelphia, and Detroit. On the other hand, the model over-predicts the inflows considerably for Los Angeles and somewhat for Boston. In sum, the origin consumption channel goes a long way in explaining the observed heterogeneity in

[38] Note that this can happen both through rural-urban migration of natives and through a fertility rate that is higher than the replacement fertility rate.

[39] To avoid that the relocation of the immigrant population already present in 1990 drives the differences between Panel A and Panel B, we assume that these immigrants always behave as in the full model with network amenities and origin consumption in both counterfactuals. Thus, only the behavior of incoming immigrants differs between the counterfactuals.

[40] In these cities, the share of Mexican immigrants among the working-age population (18-64) was around 23%-24% in 1980.

immigrant and native population growth rates across US cities during the 1990s.

5 New instrument, native relocation, and productivity

Our model can be used for a number of applications. We investigate two of them in this section. First, we use the main insights of the model to propose a new instrument for immigrant shocks. In particular, we can use fluctuations in exchange rates over time and their interaction with an exogenous driver of local rents, like the housing supply elasticity estimated in Saiz (2010), to predict the number of immigrants from the various countries of origin entering each metropolitan area in a given period of time. We examine this point in subsection 5.1, where we argue that our instrument outperforms both the networks instrument and labor demand shocks in predicting immigration flows. Our quantitative model predicts that immigration displaces or crowds out natives from the most expensive locations. We test this prediction of the model in subsection 5.2, using our new instrument, and we show that it is strongly supported in the data.

Second, the model can help us understand the effect of immigration on aggregate productivity. The movement of economic activity towards more productive locations has aggregate implications. We analyze these aggregate outcomes in subsection 5.3.

5.1 Instrumental variables for immigrant shocks

Our framework captures three reasons why immigrants may be drawn to particular locations. Variation that is uncorrelated with the outcome of interest and predicts immigrant locations is the basis of any instrument for immigrant shocks. According to the model, the fraction of immigrants from country j that locate in city c is given by:

$$\ln \pi_{jc} = \frac{1}{\lambda}(\ln V_{jc} - \ln V_j) = \frac{1}{\lambda}(\underbrace{\ln A_{jc}}_{\text{Networks}} + \underbrace{\ln w_{jc}}_{\text{Bartik}} - \underbrace{(1-\alpha_t)\ln \bar{p}_{jc}(\bar{\alpha}_l, \bar{\alpha}_f)}_{\text{Exchange rate}\times\text{HS elasticity}} + \ln V_j). \tag{5.1}$$

This equation highlights ways in which researchers can obtain variation in immigrant shocks. First, immigrants may be driven to particular locations because of higher wages. One way to generate variation in wages is to construct Bartik shocks (Bartik 1991). Second, when researchers want to investigate the effects of immigration on wages and, hence, cannot use Bartik shocks, they can rely on variation coming from the immigrant networks instrument. Finally, our model suggests that variation in the price index relevant to the various immigrant groups, which is an interaction between the local price index and the home country price index, can also be used as a shifter of immigrant location patterns.

The state of the art instrument for immigrant shocks is the networks instrument, which is based on the pioneering work of Altonji and Card (1991) and refined in subsequent literature. This shift-share instrument can be constructed as follows. The first step uses the equation

$$\Delta\widehat{\text{Imm}}_{cot}^{IV-networks} = \Delta\text{Imm}_{ot} \times \frac{\text{Imm}_{co,t-1}}{\text{Imm}_{o,t-1}}.$$

This implies that the predicted number of immigrants from each country of origin o to each location c can be obtained from the aggregate inflow of immigrants from each country (the *shift* part of the instrument), denoted by ΔImm_{ot}, and the relative number of immigrants from country o that were living in c relative to all immigrants from that country at time $t-1$ (the *share* part of the instrument). This instrument is based on the idea that networks drive, at least in part, the location choices of immigrants. A natural justification of this instrument is presented in the model, as discussed in Section 4. The size of the network acts as a shifter of local amenities that drives immigrants to particular locations, which is orthogonal to labor demand shocks if they are not serially correlated. Viewing the instrument in this way is consistent with the data, as shown in Table 3.

Mimicking this logic, we can build an instrument using variation in home country prices and our theory on how immigrants decide where to locate, as described in Section 4. For this, we need to take two steps. The first step builds on previously estimated equation (3.5), which is a linearized version of equation (4.8) of the model.

$$\ln\left(\frac{\text{Imm}_{cot}}{\text{Imm}_{ot}} / \frac{\text{Nat}_{ct}}{\text{Nat}_t}\right) = \beta \text{ HS elasticity}_c + \gamma \ln \text{RER}_{ot} + \eta \ln \text{RER}_{ot} \times \text{ HS elasticity}_c + [\delta_c] + \delta_t + \delta_o + \varepsilon_{cot}. \quad (5.2)$$

The only difference between this equation and equation (3.5) is that we substitute the local price index by estimates of the housing supply elasticity (denoted by "HS elasticity") provided in Saiz (2010).[41] The results of estimating equation (5.2) are shown in Table E7 in online Appendix E. They are in line with Table 3.[42] Note that when we include city fixed effects (δ_c) we are allowing locations to have fixed characteristics that make them attractive to immigrants, like being close to the US-Mexican border.

With the estimated equation (5.2), we can predict the relative distribution of immigrants in each decade that is driven exclusively by exchange rate fluctuations and the (time-unvarying) housing supply elasticity of each location. To obtain an estimate of the predicted inflow of immigrants from each country of origin in each metropolitan area, we can then take the predicted relative distribution of immigrants as estimated in equation (5.2) and multiply it with the aggregate change in the number of immigrants from each country of origin ΔImm_{ot} and the relative distribution of natives at the beginning of the period $\frac{\text{Nat}_{c,t-1}}{\text{Nat}_{t-1}}$:

$$\Delta\widehat{\text{Imm}}_{cot}^{IV-AM} = \Delta\text{Imm}_{ot} \times \widehat{\left(\frac{\text{Imm}_{cot}}{\text{Imm}_{ot}} / \frac{\text{Nat}_{ct}}{\text{Nat}_t}\right)} \times \frac{\text{Nat}_{c,t-1}}{\text{Nat}_{t-1}}. \quad (5.3)$$

This expression highlights both the similarities and the differences between the standard networks instrument and the new instrument that we propose in this paper, which we label by "AM". In both cases, we need to assign the aggregate inflows from each country of origin across locations. Our new instrument assigns these according to the relative distribution of immigrants and natives based on exchange rate fluctuations interacted with the housing supply elasticity and the distribution of natives (not immigrants) across locations. This is different from the networks instrument, which assigns these aggregate immigrant inflows by country of origin according to the initial immigrant distribution across locations. In other words, the difference between this new instrument and the networks instrument is that the we use a different *share* part. To the extent that identification comes from the shares, the two instruments are different.

Using the predicted inflow of immigrants from each country of origin, we obtain an overall predicted immigrant inflow in a location by aggregating over all countries of origin:

$$\Delta\widehat{\text{Imm}}_{ct}^{IV-k} = \sum_o \widehat{\text{Imm}}_{cot}^{IV-k},$$

where k indicates the strategies to obtain this predicted immigrant shock at the origin-MSA level. With this, we can build a predicted relative immigrant inflow or, in short, "immigrant shock" at the local level as:

[41] It is worth noting that when estimates of the housing supply elasticity are not available one can use lagged housing prices or (native) population levels as a proxy in equation (5.2).

[42] The first column of Table E7 does not allow for country of origin variation, while columns 2 and 3 both use this variation. The only difference between columns 2 and 3 is the exclusion or inclusion of metropolitan area fixed effects. The coefficient of interest is the interaction between the housing supply elasticity and the exchange rate. A positive point estimate means that when the exchange rate is low, immigrants concentrate more in low housing supply elasticity cities like San Francisco.

$$Z^k_{ct} = \frac{\widehat{\Delta \text{Imm}}^{IV-k}_{ct}}{\text{Nat}_{c,t-1}}.$$

With this we can formulate the first-stage regression as:

$$\frac{\Delta \text{Imm}_c}{\text{Nat}_{c,t-1}} = \beta^{FS} Z^k_{ct} + \delta_t + \delta_r + \mu_{ct}, \tag{5.4}$$

where δ_t and δ_r indicate time and Census region fixed effects.

This instrument identifies the causal effect of interest, if two conditions are satisfied. First, the instrument needs to be relevant. Second, the error term μ_{ct} needs to be orthogonal to the error term of the second-stage equation. Intuitively, the identifying variation when using both the networks instrument and our new instrument comes from predicting the importance of immigrant shocks across locations at different points in time. In the case of the networks instrument, the identification assumption is that past migration does not directly influence current changes in labor market outcomes. In the case of our new instrument, the identification assumption is that exchange rate fluctuations and the underlying (time-unvarying) housing market elasticity does not directly impact changes in the outcomes of interest.

Our proposed instrument captures, for example, that during periods when the Mexican peso is particularly low, Mexicans are more likely to concentrate in places like Los Angeles or San Francisco; whereas when the peso is high, they are more likely to choose a cheaper city like Sacramento. This variation determines the relative distribution of Mexicans across locations within states or census regions (depending on the fixed effects used) for reasons that are unrelated to changes in the outcome of interest varying at the local labor market level. To obtain the local area shock, we aggregate this country by country variation by summing over all countries of origin.

Table 7 shows the results of predicting immigrant flows based on the three components highlighted in equation (5.1). Depending on the second-stage regression, each of these three predictors of immigrant location choices may be a valid instrument. Table 7 reflects the relevance of each element for a potential IV strategy.

In the first three columns of the table, we show that all three elements of equation (5.1) are, when used alone, a good predictor of the actual immigrant shock. The point estimates are strongly significant and, together with the fixed effects, explain a large fraction of the variation in the data. Indeed, immigrant workers are drawn to strong labor markets, measured by the Bartik shock (column 1), to locations with large immigrant communities (column 2), and to expensive locations (column 3).

In columns 4 to 6 we compare the ability of the Bartik shock, the networks instrument, and our instrument to predict immigrant shocks in different types of cities in terms of housing supply elasticities. It is apparent from columns 4 and 5 that our new instrument has more predictive power than both the networks instrument and the Bartik shock for metropolitan areas with low housing supply elasticities. In column 6, we see that for locations with high elasticities ($HS > 2$), the three variables contribute to predicting immigrant shocks.

In columns 7 to 9 we repeat the exercise but investigate heterogeneity with respect to productivity. In the most productive places (locations that are estimated to be as productive as Miami or more), Bartik shocks and especially our instrument are good predictors of immigrant shocks, while the immigrant networks instrument is not statistically significant and irrelevant (the estimate is close to 0), as can be seen in column 7. As we move to cities with lower productivities in columns 8 and 9, we see that while our instrument is still highly predictive of immigrant locations, the immigrant networks instrument starts to gain relevance. In column 10 we put the three variables together. This column shows that both Bartik shocks and our new instrument remain relevant when all three variables are put together.

5.2 Native reallocation

As we explain in Section 4, the model predicts that immigrant shocks lead to a relocation of natives from expensive, highly productive locations to cheaper and less productive locations. In this subsection, we test this prediction of the model. To do so, we use the following regression:

$$\frac{\Delta \text{Nat}_c}{\text{Nat}_{c,t-1}} = \beta \frac{\Delta \text{Imm}_c}{\text{Nat}_{c,t-1}} + \delta_t + \delta_r + \varepsilon_{ct}. \tag{5.5}$$

With random variation in immigrant shocks, this equation identifies how many natives on net relocate as a function of immigrants arrivals. An estimate of $\beta = 0$ means that natives are unresponsive to immigrant shocks. A coefficient $\beta < 0$ (respectively $\beta > 0$) means that natives move away from (respectively into) locations where immigrants enter. δ_t denotes time fixed effects and δ_r denotes region fixed effects that control for regional trends.[43]

The results of estimating equation (5.5) are reported in Table 8. In Panel A, we show the results using our new instrument, either alone or in combination with each of the two other potential IV strategies. In the first column, we report the OLS estimate, which, like in prior literature, results in a small estimate that is not distinguishable from 0. The different IV specifications used in columns 2 to 6 indicate that immigrant shocks lead to a net decline in the native population. The point estimate fluctuates around -0.8, suggesting that each immigrant arrival crowds out 0.8 natives. Our preferred estimate is shown in column 8. In this column we use our instrument together with the networks instrument and we control for the Bartik shock, its interaction with a dummy for whether the MSA's housing supply elasticity is below 1, and the dummy itself. These controls account for the divergence in labor demand shocks across MSAs documented in Moretti (2011), and the specification uses variation for identification that is not serially correlated, as Bartik shocks tend to be (Amior and Manning 2018).

Several aspects give credibility to this empirical strategy. First, having many instruments is useful because we can check with the Sargan-Hansen test whether we can reject the null hypothesis that our instruments are valid. We cannot reject this null hypothesis in any of our specifications. Second, in online Appendix E, Table E8, we check whether our results are driven by differential pre-trends. When using the lagged immigrant shocks as the dependent variable, we see that our IV (and also our IV combined with the immigrant networks instrument), shows no signs that pre-trends are driving the result. This is not true when using Bartik shocks, which is in line with the finding in Amior and Manning (2018) that local labor demand shocks are very persistent.

In Panel B of Table 8, we explore the heterogeneity of this result by splitting MSAs in three groups, along either the housing supply elasticity or productivity dimensions. The first three columns show the split as a function of the housing supply elasticity. These columns suggest that the relative decline in native population concentrates in areas with housing supply elasticities below 1, consistent with our model. As can be seen in columns 4 to 6, native relative population declines are also stronger in the more productive locations.

The results reported in Panel A of Table 8 are in line with our model, but contradict prior literature. In a seminal contribution, Card and DiNardo (2000) conclude that natives do not relocate after immigrant shocks. This result is confirmed in later work by Peri and Sparber (2011) using the same specification that we used here. The difference in the results is explained by the fact that prior literature did not take into account the heterogeneity in the data. In an unweighted regression using all MSAs, it is possible to obtain an estimate that is not different to 0 when the true effect is -1 for a small number of expensive, populous, and productive cities,

[43]Results are the same when not controlling for regional trends and when controlling for state-specific trends. However, including regional trends and Bartik shocks as controls is important for ensuring that results are not driven by pre-trends. Standard errors are clustered at the state level.

and a tightly estimated 0 for all other locations.[44]

5.3 The impact of immigration on aggregate productivity

The final application of our framework is to study aggregate outcomes through the lens of our model. As argued by Hsieh and Moretti (2019), high-productivity cities like New York City typically have higher housing constraints. These constraints on housing lead to spatial misallocation: workers would be more productive in these locations, but they cannot afford the cost of living there. Our framework suggests that immigrants contribute to alleviating this labor misallocation, which we quantify in this subsection.

To highlight the different mechanisms operating in the model in detail, we conduct several counterfactual simulations of the inflow of immigrants and the increase in native population across cities over the 1990s. In particular, together with using the full model, we also simulate the model under two alternative assumptions: first, we set the housing supply elasticities at a constant value of 100 across all cities (thus, there is no scope for labor misallocation anymore) and, second, we keep the native population fixed at 1990 levels (thus effectively assuming that there is no *additional* misallocation due to the growth in the native urban population that we see over the 1990s).

We combine each of these simulations with the following scenarios on immigrant behavior: 1) no heterogeneity in the utility function of natives and immigrants; 2) only heterogeneity with respect to network amenities of immigrants; and 3) simulation of the full model. This results in nine counterfactual simulations, for each of which we present the percentage change difference in TFP between 1990 and 2000 in Table 9.[45] We define the immigrant contribution to TFP growth, shown in the last row, as the difference in the model predicted aggregate TFP change between assuming that immigrants behave like natives and assuming that they behave as posited in our model, i.e. taking into account both local networks as amenities and country of origin-specific price indices.

When simulating the model without any heterogeneity across workers by origin and with very high and uniform housing supply elasticities, we obtain a slight decrease in TFP of 0.10%. This is driven by the fact that immigrants already in the economy in 1990 (who are located according to the full model) are disproportionately located in high-productivity cities, and hence, when new immigrants and natives enter the economy, the population share of pre-1990 immigrants falls, which shifts overall population to less productive cities. When we allow the elasticities to vary but keep native population fixed at the 1990 level, TFP falls by 0.24%. This is the misallocation effect due to immigration predicted by a "naive" model in which immigrants are assumed to be identical to natives. This percentage reflects that entrants into the economy are pushed to low-productivity locations. Finally, in the scenario with varying elasticities and native urban population growth, TFP falls by 0.46%. This percentage indicates that when natives and immigrants are the same they are both pushed equally to low-productivity cities.

In the second set of counterfactuals, displayed in the second row of Table 9, we repeat the exercise assuming that immigrants value networks. In this case the results are similar to the first set of counterfactuals. If anything, the spatial misallocation is slightly exacerbated, most likely driven by Mexicans locating in the low-productivity border cities.

In the third set of counterfactuals, where we add the role of consumption heterogeneity, we find that immigration alleviates spatial misallocation. The possibility of substituting local consumption by consumption at origin, implies that incoming immigrants disproportionally locate in more expensive, more productive cities. When abstracting from native population growth, we obtain that immigration led to an increase in TFP of around 0.15%.

[44] More recent accounts also obtain estimates that suggest that immigration crowds out native population (Borjas 2006; Amior 2020b).
[45] TFP is computed as the sum of output over the number of workers, according to equation (4.10).

Since the native urban population has grown during the 1990s, which counteracts the positive effects of immigrants, we find that accounting for both immigrant inflows and native population growth over the 1990s results in a decline in TFP of around 0.08%. This decline is 0.38 percentage points lower than that predicted without the immigrant consumption channel.

This exercise suggests that immigration has a positive effect on overall TFP and that this is entirely driven by the higher probability to locate in expensive cities. Regarding the size of the effect, there are several things to keep in mind. First, our model abstracts from city-level changes in TFP. If productivity levels increase in locations with higher population growth (e.g. because of agglomeration forces), the positive effects of immigration would accordingly be larger. Second, our predictions only refer to the immigrants that entered the US during the 1990s, who make up just a fraction of the total immigration population present today (which is around 21% in our sample of cities). Third, in our analysis, we abstract from rural areas, since Saiz (2010) does not provide housing supply elasticity estimates for non-urban commuting zones. If immigrants are more likely than natives to locate in urban areas, which is likely given that rural areas tend to be cheaper and less attractive to migrants in terms of networks, and rural areas are less productive, this will have an additional positive effect on aggregate productivity that is not accounted for.

Altogether, our model highlights the importance of taking into account that immigrants spend a substantial fraction of their income in their countries of origin. The model-based counterfactuals presented in this section show that acknowledging this defining feature of immigrants changes our understanding of how immigration affects the host economy, not only across locations, but also in the aggregate.

6 Conclusion

In the first part of this paper, we document that immigrants concentrate in more expensive cities and that their earnings relative to natives are lower there. We show that these patterns are stronger for immigrants coming from countries with low price levels and that immigrant households consume less locally than similar-looking native households.

Taking all this evidence together, we posit that these patterns emerge because a share of immigrants' consumption is not affected by local price indices but, rather, by prices in their country of origin. That is, given that immigrants send remittances home and are more likely to spend time and consume in their countries of origin, they have a greater incentive than natives to live in high-nominal-income locations.

We build a quantitative spatial equilibrium model with labor market frictions to quantify the importance of this mechanism. We estimate the model and show that the differential location choices of immigrants relative to natives have important consequences, as economic activity moves from low-productivity to high-productivity cities. Model simulations suggest that through this mechanism, immigration over the 1990s might have increased labor productivity by around 0.4%.

Using the model insights, we build a new instrument that relies on the interaction between exchange rate fluctuations and the housing supply elasticity to assign immigrants to locations. Our new instrument performs well at predicting immigrant location patterns. For cities with low housing supply elasticities, our instrument outperforms those based on the traditional theories for immigrants' location choices (labor demand shocks and immigrant networks), while for cities with high elasticities, all three theories have predictive power.

References

Adao, R., M. Kolesar, and E. Morales, "Shift-Share Designs: Theory and Inference," *Querterly Journal of Economics*, 2019.

Albert, Christoph, "The Labor Market Impact of Undocumented Immigrants: Job Creation vs. Job Competition," *American Economic Journal: Macroeconomics*, Forthcoming.

Albouy, D., H. Cho, and M. Shappo, "Immigration and the Pursuit of Amenities," *mimeo*, 2018.

Allen, T., K. Dobbin, and M. Morten, "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *mimeo*, 2019.

Altonji, J. and D. Card, *The Effects of Immigration on the Labor Market Outcomes of Less-Skilled Natives*, in John Abowd and Richard Freeman (eds.), Immigration, Trade, and the Labor Market, University of Chicago Press, 1991.

Amior, M., "The Contribution of Immigration to Local Labor Market Adjustment," *mimeo*, 2020.

_ , "Immigration, Local Crowd-Out and Undercoverage Bias," *mimeo*, 2020.

_ **and A. Manning**, "The Persistence of Local Joblessness," *American Economic Review*, 2018.

_ **and** _ , "Monopsony and the Wage Effects of Migration," *mimeo*, 2020.

Angelucci, M., "Migration and financial constraints: evidence from Mexico," *The Review of Economics and Statistics*, 2015, *97* (1).

Attanasio, O. and L. Pistaferri, "Consumption Inequality," *Journal of Economic Perspectives*, 2016, *30* (2).

Bartik, T. J., "Who Benefits from State and Local Economic Development Policies?," *Books from Upjohn Press. W.E. Upjohn Institute for Employment Research*, 1991.

Becker, G., *The Economics of Discrimination* 1957.

Black, D., "Discrimination in an equilibrium search model," *Journal of Labor Economics*, 1995.

Borjas, G., "Does Immigration Grease the Wheels of the Labor Market?," *Brookings Papers on Economic Activity*, 2001.

_ , "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics*, 2003, pp. 1335–1374.

_ , "Native Internal Migration and the Labor Market Impact of Immigration," *Journal of Human Resources*, 2006, *41(2)*, 221–258.

_ , "The Labor Supply of Undocumented Immigrants," *Labour Economics*, 2017.

_ **, J. Grogger, and G. Hanson**, "On Estimating Elasticities of Substitution," *Journal of European Economic Association*, 2012.

_ **, R. Freeman, and L. Katz**, "Searching for the Effect of Immigration on the Labor Market," *American Economic Review Papers and Proceedings*, 1996.

Borusyak, K., P. Hull, and X Jaravel, "Quasi-experimental Shift-share Designs," *mimeo*, 2019.

Burstein, A., G. Hanson, L. Tian, and J. Vogel, "Tradability and the Labor-Market Impact of Immigration: Theory and Evidence from the U.S.," *Econometrica*, 2020.

Caballero, E., B. Cadena, and B. Kovak, "Measuring Geographic Migration Patterns Using Matrículas Consulares," *Demography*, 2018.

Cadena, B. and B. Kovak, "Immigrants Equilibrate Local Labor Markets: Evidence from the Great Recession," *American Economic Journal: Applied Economics*, 2016.

Caliendo, L., L. Opromolla, and Sforza A. Parro F. and, "Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement," *mimeo*, 2018.

_ , M. Dvorkin, and F. Parro, "Trade and Labor Market Dynamics," *NBER Working Paper No. 21149*, 2015.

Card, D., "Immigration and Inequality," *American Economic Review Papers and Proceedings*, 2009, *99(2)*, 1–21.

_ and J. DiNardo, "Do immigrant Inflows lead to Native Outflows," *American Economic Review*, 2000.

Combes, P-P. and L. Gobillon, "The Empirics of Agglomeration Economies," *Handbook of Regional and Urban Economics*, 2014.

Davis, M. and F. Ortalo-Magne, "Household Expenditures, Wages, Rents," *Review of Economic Dynamics*, 2011.

Dustmann, C. and J. Mestres, "Remittances and Temporary Migration," *Journal of Development Economics*, 2010, *92(1)*, 70–62.

_ and J.-S. Gorlach, "The economics of temporary migrations," *Journal of Economic Literature*, 2016, *54(1)*, 98–136.

_ , U. Schonberg, and J. Stuhler, "The Impact of Immigration: Why Do Studies Reach Such Different Results?," *Journal of Economic Perspectives*, 2016.

Dynan, K., J. Skinner, and S. Zeldes, "Do the Rich Save More?," *Journal of Political Economy*, 2004, *112* (2).

Eeckhout, J., R. Pinheiro, and K. Schmidheiny, "Spatial Sorting," *Journal of Political Economy*, 2014, *122(3)*, 554–620.

Glaeser, E., *Cities, Agglomeration and Spatial Equilibrium*, Oxford University Press, 2008.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift, "Bartik Instruments: What, When, Why, and How," *American Economic Review*, 2020.

Gyourko, J. and A. Saiz, "Construction Costs and the Supply o Housing Structure," *Journal of Regional Science*, 2006, *46(4)*, 661–680.

Handbury, J. and D. Weinstein, "Goods Prices and Availability in Cities," *Review of Economic Studies*, 2015, *82 (1)*, 258–296.

Hsieh, C.-T. and E. Moretti, "Housing Constraints and Spatial Misallocation," *American Economic Journal: Macroeconomics*, 2019.

Jaeger, D., J. Ruist, and J. Stuhler, "Shift-share Instruments and the Impact of Immigration," *NBER working paper*, 2018.

Monras, J., "Economic Shocks and Internal Migration," *IZA Discussion Paper No. 8840*, 2015.

— , "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis," *Journal of Political Economy*, 2020, *128(8).*

Moretti, E., "Local Labor Markets," *Handbook of Labor Economics*, 2011.

— , "Real Wage Inequality," *American Economic Journal: Applied Economics*, 2013.

Munshi, K., "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics*, 2003.

Nekoei, A., "Immigrants' Labor Supply and Exchange Rate Volatility," *American Economic Journal: Applied Economics*, 2013, *5* (4), 144–164.

Oreopoulos, P., T. von Wachter, and A. Heisz, "Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, 2012.

Ottaviano, G. and G. Peri, "Rethinking the Effect of Immigration on Wages," *Journal of the European Economic Association*, 2012, *10(1)*, 152–197.

Peri, G. and C. Sparber, "Assessing Inherent Model Bias: An Application to Native Displacement in Response to Immigration," *Journal of Urban Economics*, 2011, *69(1)*, 82–91.

Piyapromdee, S., "The Impact of Immigration on Wages, Internal Migration and Welfare," *Review of Economic Studies*, Forthcoming.

Redding, S. and E. Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017.

Ruggles, S., M. Sobek, T. Alexander, C.A. Fitch, R. Goeken, PK Hall, M. King, and C. Ronnander, "Integrated Public Use Microdata Series: Version 4.0 [Machine-readable database].," *Minneapolis, MN: Minnesota Population Center [producer and distributor]*, 2016.

Ruist, J., "Immigrant-Native Wage Gaps in Time Series: Complementarities or Composition Effects?," *Economic Letters*, 2013.

Saiz, A., "The Geographic Determinants of Housing Supply," *Quartely Journal of Economics*, 2010, *125(3)*, 1253–1296.

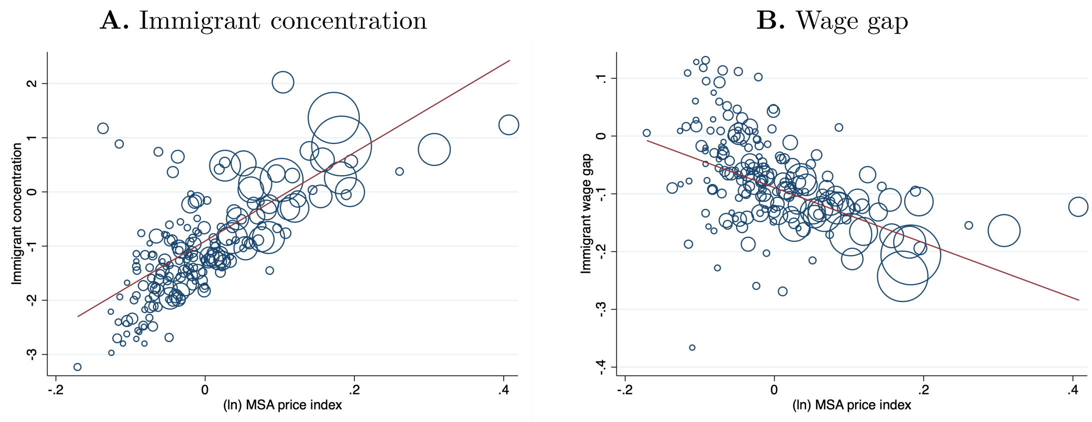
Silva, J.M.R. Santos and S. Tenreyro, "The log of Gravity," *Review of Economics and Statistics*, 2006, *88(4)*, 641–658.

Yang, D., "Why Do Migrants Return to Poor Countries? Evidence from Philippine Migrants' Responses to Exchange Rate Shocks," *Review of Economics and Statistics*, 2006, *88(4)*, 715–735.

— , "International Migration, Remittances, and Household Investment: Evidence from Philippine Migrants'Exchange Rate Shocks," *Economic Journal*, 2008, *118*, 591–630.

7 Figures

Figure 1: Immigration and cities

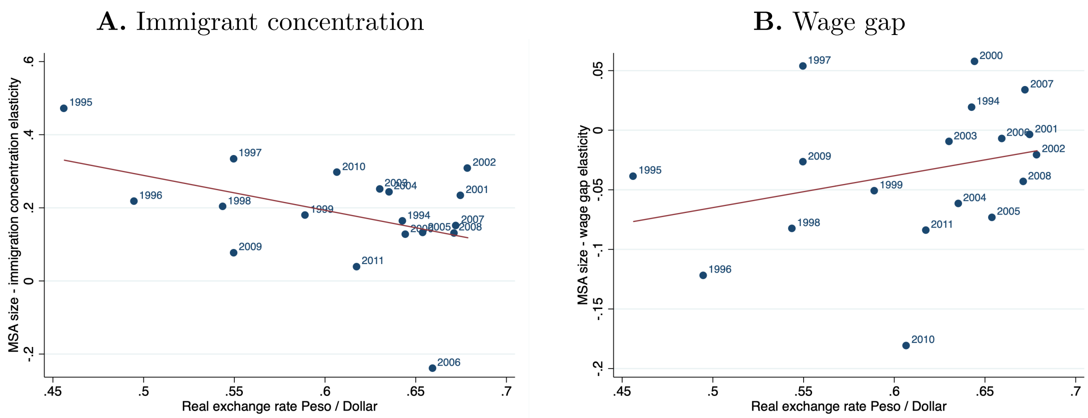


Notes: Panel A shows the relationship between the immigrant concentration in an MSA and the MSA price index. The relative share of immigrants is measured as the number of immigrants in the MSA relative to all immigrants in the United States divided by the number of natives in the MSA relative to all natives in the United States. Panel B plots the MSA-specific component of the wage gap ($\hat{\phi}_c$), estimated by equation (3.2), against the MSA price index, which is computed following Moretti (2013). Marker sizes reflect MSA population levels. Data are from the 2000 US Census.

Figure 2: City price elasticity of immigrant concentration and wage gap by real exchange rate

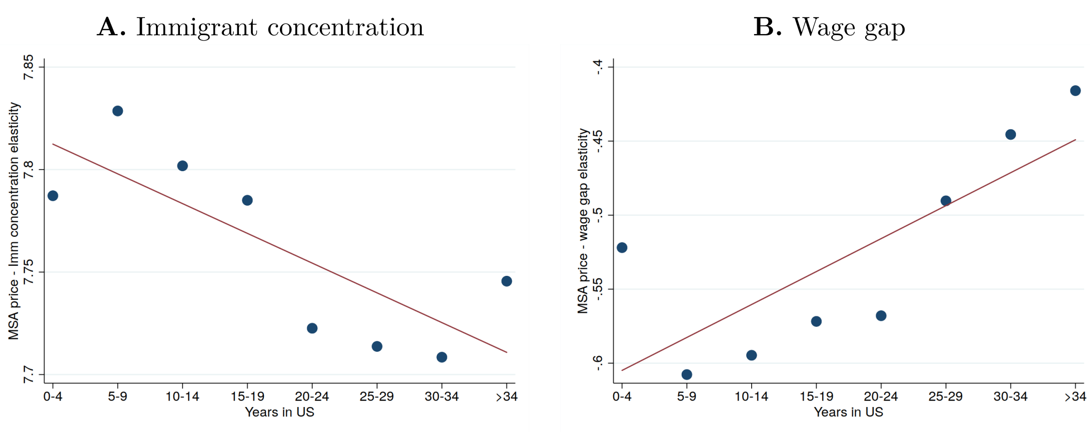
Notes: This figure uses data from the 2000 Census to show the heterogeneity in the elasticity between the immigrant concentration (Panel A) and the native-immigrant wage gaps (Panel B) with respect to the MSA price, for each of the different countries of origin as a function of the country of origin price level. The dots represent the origin-specific coefficients β_o obtained by estimating equation (3.4) at the origin level. Panel A shows the estimates for the immigrant concentration, and Panel B shows the estimates for the adjusted wage gap.

Figure 3: City size elasticity of immigrant concentration and wage gap of Mexicans



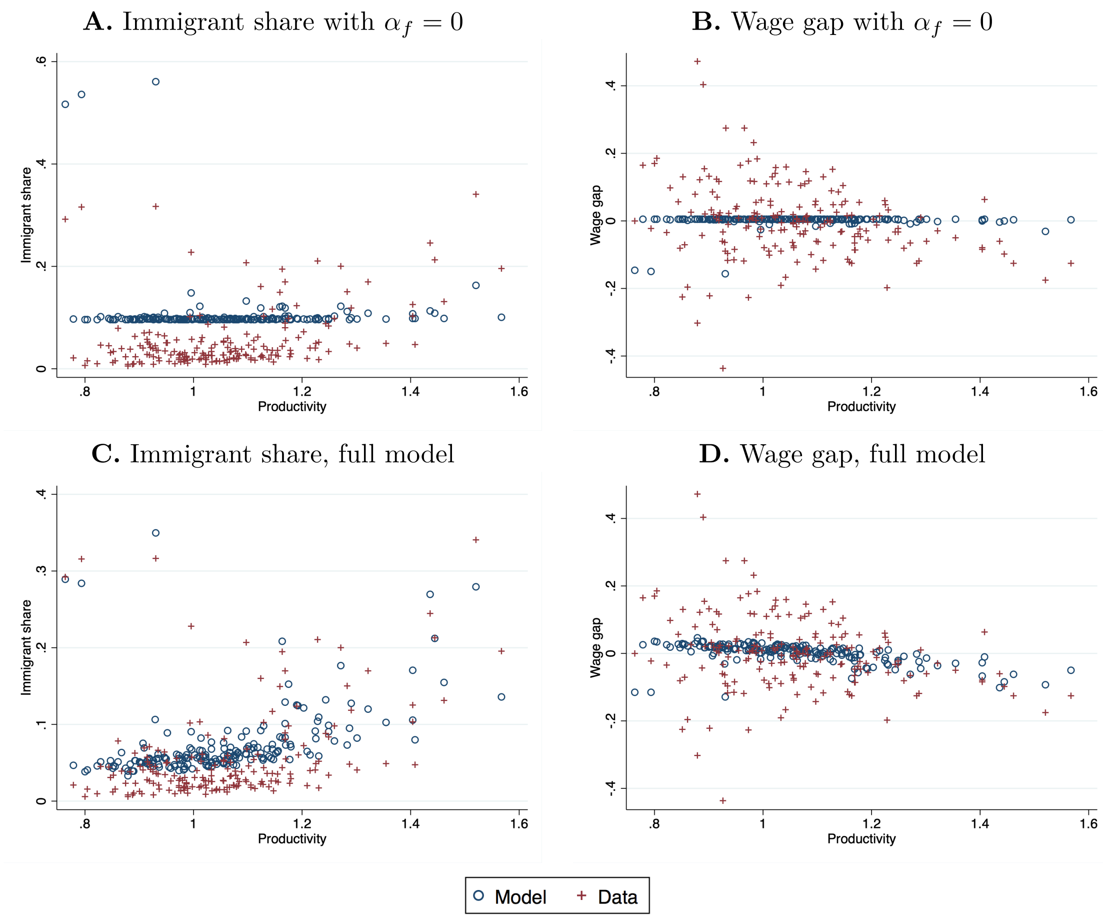
Notes: This figure uses data from the CPS 1994-2011 to show the heterogeneity in the relationships between the immigrant concentration and the (composition-adjusted) wage gap of Mexican immigrants who changed location during a year, and the city size elasticity. More specifically, we estimate the elasticity for each year and plot it against the average real exchange rate of the Mexican peso to the US dollar during that year. Hence, each dot represents an estimate of the coefficient β for the particular year based on equation (3.4). We can only compute city size elasticities, because price indices can only be computed using ACS and Census data, which is available only in selected years.

Figure 4: City price elasticity of the immigrant concentration and wage gap, by immigrants' years in the country



Notes: This figure uses 2000 Census data to show the relationships between the city price elasticity of the immigrant concentration and the native-immigrant wage gap as a function of the time spent in the United States. Each dot represents an estimate of the coefficient β for the indicated immigrant group based on a model similar to regression (3.4) but in which o denotes the groups defined by years spent in the US.

Figure 5: Model goodness of fit

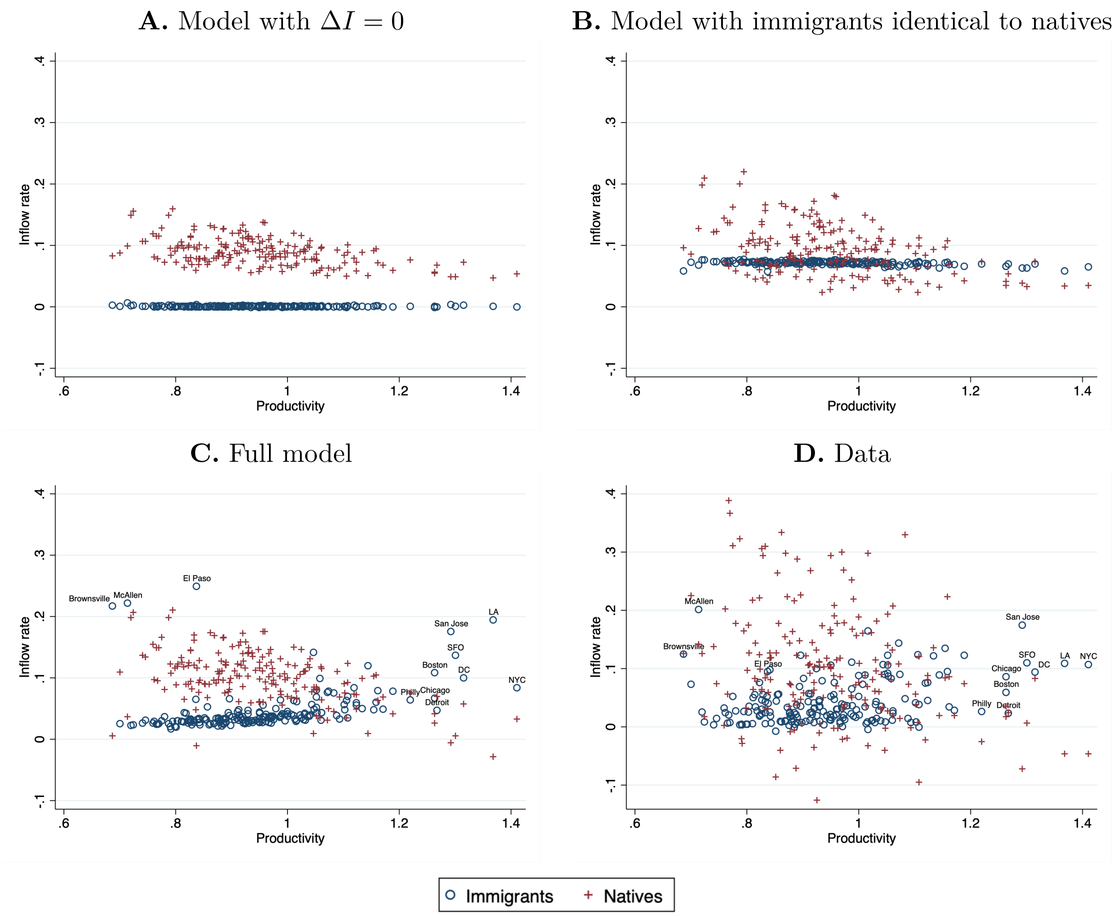


Notes: This figure compares immigrant shares and wage gaps predicted by the model and as measured using data for the year 1990. Each dot represents one MSA. Panels A and B assume that immigrants behave like natives (i.e. the weight of their home country goods is equal to 0), but value local immigrant networks. Panels C and D show the full model with the estimated α_f.

Figure 6: Non-targeted moments: Share of expenditures on local non-tradable goods

Notes: This figure plots the income share of rent expenses in the Census data by country of origin (adjusted for income, family size, age, marital status, and travel time to work) against the share of expenditures in non-tradable goods predicted by the model (both as deviations from the mean).

Figure 7: Model predictions and data of city population changes, 1990-2000



Notes: This figure shows the inflow rates of native and immigrants in each MSA in the data and as predicted by the model for different counterfactuals between 1990 and 2000. Inflow rates are defined as change in the MSA population of the respective group over total MSA population in 2000. Panel A shows the results without immigration. Panel B shows the results with immigrants behaving like natives. Panel C shows the results of the full model, where immigrants value home country goods and local networks. Panel D shows the data equivalent.

8 Tables

Table 1: List of top US cities by immigrant share in 2000

MSA	Immig. (%)	Size rank	Population	Weekly wage	Price index	Wage gap (%)
Miami-Hialeah, FL	63	22	1,141,666	300	1.13	-22
Los Angeles-Long Beach, CA	47	2	6,438,057	357	1.20	-24
McAllen-Edinburg-Pharr-Mission, TX	43	81	268,382	238	0.88	-9
San Jose, CA	43	25	934,163	491	1.52	-12
El Paso, TX	40	74	317,730	262	0.92	-15
Brownsville-Harlingen-San Benito, TX	37	124	158,981	248	0.90	-18
Stamford, CT	36	193	67,497	440	1.43	-18
New York, NY-Northeastern NJ	34	1	9,164,913	402	1.22	-21
Visalia-Tulare-Porterville, CA	33	115	180,277	276	0.95	-15
San Francisco-Oakland-Vallejo, CA	32	7	2,540,161	438	1.38	-16
Fort Lauderdale-Hollywood-Pompano Beach, FL	31	29	853,372	344	1.17	-13
Fresno, CA	30	61	388,063	294	0.98	-13
Santa Barbara-Santa Maria-Lompoc, CA	30	110	191,979	341	1.25	-15
San Diego, CA	29	15	1,398,599	360	1.19	-17
Ventura-Oxnard-Simi Valley, CA	28	60	393,320	396	1.23	-19
Riverside-San Bernardino, CA	28	13	1,593,796	336	1.07	-12
Stockton, CA	28	80	278,366	333	1.04	-14
Houston-Brazoria, TX	26	8	2,446,237	367	1.04	-15
Modesto, CA	25	93	229,426	317	1.03	-10
Santa Cruz, CA	25	135	139,133	395	1.32	-15

Notes: The statistics are based on the sample of prime-age workers (25-59) from the 2000 Census. Weekly wages are computed from yearly wage income and weeks worked. Local price indices are computed following Moretti (2013). The wage gap is the gap in wage earnings between natives and immigrants (a negative number means that natives' wages are higher), controlling for observable characteristics.

Table 2: Immigration, wage gaps and city price levels

	Immigrant concentration			Wage gap		
	(1) OLS	(2) IV	(3) OLS	(4) OLS	(5) IV	(6) OLS
(ln) Price	8.289***	10.167***	1.000**	-0.495***	-0.522***	-0.190**
	(0.746)	(0.947)	(0.459)	(0.075)	(0.088)	(0.073)
MSA FE	No	No	Yes	No	No	Yes
Observations	740	740	740	658	658	658
R-squared	0.589	0.559	0.964	0.419	0.418	0.782
F-stat 1st stage		62.49			61.11	

Notes: This table shows the results of estimating regression (3.1) in columns 1 to 3 and regression (3.3) in columns 4 to 6. Columns 2 and 5 instrument the local price index by the housing supply elasticity taken from Saiz (2010). Data come from Census data for 1980, 1990, and 2000 and the combined 2009-2011 ACS data. All specifications include year fixed effects. Standard errors are clustered at the MSA level. Observations are weighted by the population in a year-MSA cell. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 3: Immigration and heterogeneity: Exchange rate variation by country of origin

	Immigrant concentration			Wage gap		
	(1) PPML	(2) PPML	(3) PPML	(4) OLS	(5) OLS	(6) OLS
(ln) RER	0.175***	0.102*	0.079	0.035***	0.042***	0.042***
	(0.047)	(0.059)	(0.055)	(0.012)	(0.010)	(0.011)
(ln) Price	5.852***	5.192***	1.730***	-0.432***	-0.390***	-0.098
	(0.510)	(0.305)	(0.388)	(0.073)	(0.067)	(0.119)
(ln) Price × log RER	-1.423***	-1.641***	-1.440***	0.256***	0.263***	0.320***
	(0.489)	(0.383)	(0.352)	(0.069)	(0.072)	(0.067)
Immigrant network		10.497***	10.454***		-0.360***	-0.339**
		(0.559)	(0.837)		(0.071)	(0.170)
MSA FE	No	No	Yes	No	No	Yes
Observations	37740	37740	37740	19685	19685	19685
R-squared	0.076	0.1	0.232	0.587	0.596	0.638

Notes: This table shows regressions of the immigrant concentration and wage gaps on city prices, real exchange rates (RER), and their interaction. The regressions are based on 185 MSAs and 68 sending countries for the years 1990, 2000, and 2010. All specifications include origin and year fixed effects. Standard errors are clustered at the MSA-origin level. Observations are weighted by immigrant population in a year-MSA-origin cell. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 4: Immigration and heterogeneity: Mexican immigrant flows by state of origin and destination

	(1)	(2)	(3)	(4)
(ln) GDP pc US x (ln) GDP pc Mex	-0.300***	-0.293***	-0.224***	-0.230***
	(0.079)	(0.079)	(0.067)	(0.077)
(ln) GDP pc US	4.756***	2.598**		
	(0.905)	(1.009)		
(ln) GDP pc Mex	2.869***	2.798***		
	(0.991)	(0.994)		
Population controls	No	Yes	No	No
Orig./Dest. FE	No	No	Yes	Yes
Observations	1457	1457	1457	1426
R-squared	0.391	0.452	0.826	0.813
Sample	All	All	All	Without CA

Notes: This table shows the results of regressing the flows of immigrants from origin i to destination j relative to all migrants from Mexican state i on the GDP per capita at destination j, at origin i, and the interaction between the two. Population controls are the (log) total population at origin and destination. Data are from the Matricula Consular 2016 and cover 31 sending Mexican and 51 receiving US states (124 pairs omitted due to 0 bilateral flows). Column 4 excludes California from the regression. Robust standard errors clustered at the destination level are reported. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 5: Immigrants' expenditure

A. (ln) Total Expenditure (CEX)				
	(1)	(2)	(3)	(4)
Mexican	-0.330***	-0.382***	-0.149***	-0.224***
	(0.028)	(0.021)	(0.012)	(0.014)
State FE	No	Yes	Yes	Yes
Income	No	No	Yes	Yes
Characteristics	No	No	No	Yes
Observations	105975	105975	105975	105975
R-squared	0.031	0.063	0.301	0.314

B. (ln) Rent (CEX)				
	(1)	(2)	(3)	(4)
Mexican	-0.228***	-0.299***	-0.097***	-0.198***
	(0.060)	(0.025)	(0.023)	(0.019)
State FE	No	Yes	Yes	Yes
Income	No	No	Yes	Yes
Characteristics	No	No	No	Yes
Observations	105975	105975	105975	105975
R-squared	0.018	0.079	0.226	0.249

C. (ln) Rent (Census, ACS)				
	(1)	(2)	(3)	(4)
Immigrant	-0.018	-0.143***	-0.073***	
	(0.024)	(0.023)	(0.018)	
(ln) RER				0.068***
				(0.017)
(ln) HH income			0.277***	0.237***
			(0.004)	(0.005)
MSA FE	No	Yes	Yes	Yes
Observations	698277	698277	698277	149673
R-squared	0.295	0.443	0.539	0.513
Sample	All	All	All	Immigrants

Notes: Panels A and B use data from the Consumer Expenditure Survey 2003-2015 with the log total expenditure and monthly rent expenses as dependent variables. The income controls added in column 3 are dummies for household income bins. The characteristics added in column 4 include age, race, family size, marital status and a dummy indicating residence in an urban area. Panel C uses Census 1990 and 2000 and ACS 2009-2011 data for male household heads. Additional controls not shown include age, family size, and marital status. All specifications include year fixed effects. Standard errors are clustered at the state level in Panels A and B and at the MSA level in Panel C. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 6: Assumed and structurally estimated model parameters

Assumed parameters		Value	
Worker bargaining power	β	0.5	
Tradable goods spending share	α_t	0.4	
Estimated parameters		Value	s.e.
Elasticity of substitution local-origin goods	σ	1.86	0.50
Share of home goods consumption	α_f	0.17	0.015
Sensitivity to local conditions	λ	0.043	0.017
Sensitivity to immigrant networks	ϕ	1.49	0.46

Notes: This table shows the structural parameters of the model estimated by feasible generalized least squares using equations (4.7) and (4.8), given the assumed parameter values $\beta = 0.5$ and $\alpha_t = 0.4$. For the estimation, we use data on wages, the distribution of workers across locations at the country-of-origin level, local price levels, real exchange rates, immigrant networks and the aggregate remittance share. Bootstrapped standard errors with 500 repetitions are reported.

Table 7: Instrumental variables for immigrant shocks

	Dep. var.: Immigrant inflow rate ($\Delta Imm/Nat$)									
				By HS elasticity			By productivity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Bartik shock	0.991***			0.754	0.031	0.430***	0.636*	0.345	0.383***	0.627***
	(0.175)			(0.522)	(0.114)	(0.113)	(0.333)	(0.228)	(0.119)	(0.146)
Network IV		0.516***		-0.065	0.095	0.279***	0.019	0.371*	0.313***	0.117
		(0.156)		(0.161)	(0.165)	(0.095)	(0.140)	(0.192)	(0.087)	(0.102)
AM IV			0.317***	0.268***	0.572***	0.202***	0.259***	0.535***	0.211***	0.242***
			(0.031)	(0.077)	(0.150)	(0.069)	(0.067)	(0.104)	(0.061)	(0.022)
Observations	555	555	555	57	189	309	72	105	378	555
R-squared	0.399	0.531	0.600	0.616	0.686	0.783	0.617	0.806	0.660	0.645
Sample	All	All	All	HS<1	1≥HS≤2	HS>2	B>1.1	1≥B ≤1.1	B<1	All

Notes: The dependent variable is the immigrant inflow rate, defined as $\Delta Imm/Nat$. Data come from the 1990 and 2000 Census and the ACS in 2009-2011. The number of metropolitan areas is 185. "HS" indicates the housing supply elasticity. "B" indicates the estimated metropolitan area productivity. All specifications include year and Census region fixed effects. Robust standard errors clustered at the state level are reported. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 8: Immigration, native mobility, and housing prices

	A. Native relocation ($\Delta Nat/Nat$), main estimates					
	(1)	(2)	(3)	(4)	(5)	(6)
$\frac{\Delta \text{Imm}}{\text{Nat}}$	0.151	-0.793***	-0.866***	-0.782***	-0.746**	-0.819**
	(0.253)	(0.270)	(0.325)	(0.291)	(0.328)	(0.383)
Controls	No	No	No	No	Yes	Yes
Observations	555	555	555	555	555	555
R-squared	0.282	0.147	0.125	0.150	0.171	0.152
IV	-	AM	AM+Net	AM+Net+Bar	AM	AM+Net
F-stat 1st stage		108.0	298.5	118.1	124.1	259.6
p-value Sargan-Hansen			0.214	0.386		0.237
	B. Native relocation ($\Delta Nat/Nat$), heterogeneity					
	By HS elasticity			By productivity		
	(1)	(2)	(3)	(4)	(5)	(6)
$\frac{\Delta \text{Imm}}{\text{Nat}}$	-1.079***	0.162	-0.258	-0.923***	-0.091	0.262
	(0.332)	(0.447)	(0.347)	(0.323)	(0.637)	(0.165)
Observations	57	189	309	72	105	378
R-squared	0.057	0.397	0.277	0.156	0.392	0.285
Sample	HS<1	1≥HS≤2	HS>2	B>1.1	1≥B≤1.1	B<1
IV	AM	AM	Net	AM	AM	AM
F-stat 1st stage	15.85	48.72	14.87	17.32	36.67	72.24

Notes: The dependent variable is the native inflow rate defined as $\Delta Nat/Nat$. Data come from the 1990 and 2000 Census and the ACS in 2009-2011. The number of metropolitan areas is 185. "HS" indicates the housing supply elasticity. "B" indicates the estimated metropolitan area productivity. "AM" indicates the new instrument introduced in this paper. "Net" indicates the traditional Networks instrument. "Bar" indicates a Bartik labor demand shock. All specifications include year and Census region fixed effects. Robust standard errors clustered at the state level are reported. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table 9: Predicted aggregate TFP changes (%)

Counterfactual	HSE const	Native pop fixed	Baseline
1) No immigrant-native heterogeneity	-0.10	-0.24	-0.46
2) Only network amenity heterogeneity	-0.15	-0.27	-0.49
3) Network & consumption heterogeneity	0.42	0.15	-0.08
Immigrant contribution to TFP growth over the 1990s	0.52	0.39	0.38

Notes: This table shows the aggregate consequences of new immigration under different counterfactuals. In counterfactual 1), we assume that new immigrants consume like natives. In counterfactual 2) we assume that new immigrants only value immigrant networks, as estimated. Counterfactual 3) shows our baseline model. This table reports the aggregate consequences of these different counterfactuals on average labor productivity. We report results assuming three different scenarios, each in a separate column. Column 1 ("HSE const") assumes an (almost) perfectly elastic housing supply that is constant across MSAs. Column 2 assumes no native population growth over the 1990s. Column 3 takes into account both immigrant inflows and native population growth. We define the immigrant contribution to TFP as the difference between counterfactual 1) and counterfactual 3).

FOR ONLINE PUBLICATION

A Empirical appendix

A.1 Commuting zones

This section shows that the main relationship between location choices and relative wages between natives and immigrants documented throughout the paper is independent of using MSA-level data or commuting zone data. Commuting zones are equivalent to MSAs when we consider urban population. There are, however, many commuting zones that are rural areas, which are not covered in the MSA data. While in our context it seems quite natural to think about MSAs, for which we have estimates of housing supply elasticities, some papers have emphasized the use of local labor markets – typically measured by commuting zones – so that rural areas are also included.

Since we have price levels only for MSAs, we proxy commuting zone price levels by population, which is closely related to prices. The two graphs in Figure D.3 show that immigrants concentrate in large commuting zones and that their wages relative to those of natives are lower there.

A.2 Discussion of alternative mechanisms

For some of the evidence presented in the main text, one could think of alternative mechanisms as explanations. In this section, we discuss several such alternatives and check the robustness of the patterns we have uncovered by presenting additional empirical evidence.

First, Table E3 shows that results are unlikely to be driven by differential levels of human capital between natives and immigrants, nor by the fact that some immigrants are undocumented. We show that this is the case by restricting our main regressions, either for the immigrant concentration or the wage gaps, using OLS and IV, to particular groups of immigrants. We also show in Table E3 that all our results hold within education groups.

Another potential concern is that immigrants might settle in larger cities because it is easier to find a job there. However, this does not provide an explanation for why immigrants earn less relative to natives in these cities. To further discard this explanation, we show in Figure D.1 that there is no systematic relationship between city price and job finding rates or unemployment rates of immigrants. Thus, labor market opportunities of immigrants in large cities do not seem to be systematically better than those in smaller ones.

A third concern is that there might be differences in the distribution of jobs available in large cities. For instance, there could be more jobs in the lower part of the wage distribution, which might be disproportionately held by immigrants. This would be in line with the extreme skill complementarity hypothesis discussed in Eeckhout et al. (2014) and could generate a gap in wages between natives and immigrants, especially in large cities. To address this concern, we show in Figure D.2 that skewness and kurtosis of the *native* wage distribution do not seem to be systematically related to city price levels.[46]

[46] In fact, part of the wages in the tails of the overall wage distribution, particularly in the lower tail, within cities are driven by

A final concern is that the patterns might reflect the higher availability of products that immigrants value in larger cities (Handbury and Weinstein 2015). For instance, tradable goods from origin countries may only make it to the largest locations in the United States, if there is a fixed cost of entering local product markets. It is difficult, however, to explain the observed immigrant heterogeneity with tradable goods. For example, there is no reason why there should be more products from the poorest Mexican states of origin available in more expensive US states. Similarly, there is no intuition for why there should be relatively more Mexican products in places like New York City, than, for example, German or Canadian products, being that New York City is a relatively high-income area and Germany and Canada are countries specializing in relatively high-quality products.

A.3 Amnesty of 1986

In this section we explore in more detail wages and settlement patterns of undocumented immigrants. We argue that the patterns in the data can be interpreted by the fact that undocumented immigrants may be less attached to the host country and, hence, place a higher share of their consumption in the home country.

In particular, we exploit a sudden policy change introduced by the Reagan Administration in 1986. Among other things, the Immigration Reform and Control Act (IRCA) granted amnesty to undocumented immigrants that had arrived in the US before January 1st, 1982. Hence, immigrants who entered prior to this date are more likely to have have legal status and a work permit than those who arrived just after.

If gaining work permits means that immigrants are more likely to stay in the US and thus reduce the weight of the origin country in their consumption, we should see that the wage gap in expensive cities is larger for immigrants who arrived just after 1982, relative to those who arrived just before 1982; the year 1982 should generate a discontinuity in the relationship between the wages gaps and city price levels.

To test this idea we use the sample of immigrants who entered from 1976 to 1987 (i.e. the years around the reform), and we explore whether there is a discontinuity in the wages received by immigrants who arrived before and after 1982. To keep the two immigrant groups largely homogeneous, we control for individual heterogeneity by including, as previously, skill cell, race, marital status, and 82 occupation fixed effects. Additionally, we also include a cubic term in years since arrival to control for wage convergence over time in the US.

More specifically, we use the following regression:

$$\ln w_{i,c,t} = \alpha_1 before82_{i,c,t} + \alpha_2 \ln P_{c,t} + \alpha_3 \ln P_{c,t} \times before82_{i,c,t} + \phi X_{i,c,t} + \varepsilon_{i,c,t}, \tag{A.1}$$

where $\ln w_{i,c,t}$ is the (log) weekly wage in location c at time t of individual i, $before82_{i,c,t}$ is a dummy variable indicating that an immigrant entered the US before 1982, $P_{c,t}$ is the MSA price level, and $X_{i,c,t}$ is the vector of controls.

If immigrants who arrived after 1982 are (discontinuously) less likely to have gained legal status, we

immigrants, which provides alternative explanations for some of the findings in prior literature.

expect $\alpha_3 > 0$; hence, the gap in wages to immigrants that arrived earlier increases with city price; and, hence, the gap in wages between immigrants and natives as a function of city prices is larger for immigrants who are more likely to be undocumented.

A problem of using the full sample of immigrants is that only a fraction of them were (potentially) affected by the amnesty, namely those who entered the US illegally. Thus, the pooling of both legal and undocumented immigrants at entry attenuates the coefficients α_3 towards 0. We try to take this into account by identifying likely undocumented immigrants as suggested by Borjas (2017). With this identifier, we then extend the previous regression equation by estimating:

$$\begin{aligned}\ln w_{i,c,t} =& \beta_1 before82_{i,c,t} + \beta_2 \ln P_{c,t} + \beta_3 \ln P_{c,t} \times before82_{i,c,t} + \beta_4 undoc_{i,c,t} + \beta_5 undoc_{i,c,t} \times before82_{i,c,t} \\ &+ \beta_6 \ln P_{c,t} \times undoc_{i,c,t} + \beta_7 \ln P_{c,t} \times undoc_{i,c,t} \times before82_{i,c,t} + \phi X_{i,c,t} + \varepsilon_{i,c,t},\end{aligned} \tag{A.2}$$

where $undoc_{i,c,t}$ is the indicator for potentially undocumented status. Note that this is an imperfect proxy for the entry status because it does not include immigrants that entered illegally but rather that fulfill the criteria to be classified as legal at the time of observation.[47] Hence, it is probable the coefficient β_6 suffers from an endogeneity bias, as immigrants more attached to the US might, for instance, be more likely to apply for citizenship or marry a US citizen and thus be captured by the *undoc* indicator. However, if this bias is the same for those that entered before and those that entered after 1982, we can still obtain an unbiased estimate of our coefficient of interest β_7. We expect this coefficient to be positive. This indicates that gaining legal status due to the 1982 amnesty increases the attachment to the US and thus makes the relationship between wages and local price levels stronger.

We report the results of estimating equation (A.1) in column 1 of Table E4 and of equation (A.2) in column 2. In column 3, we drop individuals with a college degree, as Albert (Forthcoming) shows that the identification of undocumented immigrants is much less accurate for the high skilled. In column 4, we restrict the sample to 1990, the first Census after the amnesty, when the undocumented indicator should be a more accurate proxy for status at entry. We find that there is a positive but insignificant effect of having entered before 1982 on the price elasticity of immigrants' wages in column 1. However, when we estimate equation (A.2) allowing the elasticity to be different for potentially undocumented immigrants, we obtain a negative coefficient β_6; hence, the wage gap for those classified as immigrants who have potentially entered illegally to those are not classified as undocumented increases with city size. As noted above, this coefficient does not have a causal interpretation but is consistent with the notion that, due to unobserved factors, this group is less attached to the US and origin consumption plays a larger role. As expected, we find β_7 to be positive, which indicates that arriving prior to 1982 increases the price elasticity of wages more for potentially undocumented immigrants, while it has no effect on the elasticity for immigrants classified as legal.

[47] An immigrant is classified as legal if at least one of these conditions is fulfilled: is a US citizen, receives social security benefits or public health insurance, is a veteran or in the armed forces, works in the government sector or in occupations that require licensing, is Cuban, or is married to a legal immigrant or US citizen.

Altogether, the evidence presented in this section is consistent with the hypothesis that attachment to the US affects how much immigrants' wages are related to local price levels. We argue that the channel behind this finding is that the importance of the goods purchased in the country of origin in immigrants' consumption bundle decreases with higher attachment to the US.

A.4 Homeownership

If immigrants plan on returning to their countries of origin it is likely that ownership rates are lower among them. Ownership rates vary considerably by income and other characteristics. Thus, it may be useful to see if it is indeed the case that homeownership rates are lower among immigrants than *similar-looking* natives. This can be shown with the following regression:

$$\text{Owner}_i = \alpha + \beta \text{Immigrant}_i + \gamma \ln \text{Household Income}_i + \eta_c X_i + \varepsilon_i \tag{A.3}$$

where "Owner" indicates whether the head of household i is a homeowner or not, Immigrant_i is a dummy indicating that household i has at least one immigrant, and X_i denotes various household characteristics, like the education level of the head of the household, marital status, the race of the head of the household, the size of the household, MSA fixed effects, occupation fixed effects, and time fixed effects. Thus, β identifies whether immigrants tend to rent rather than own the house in which they live, relative to similar-looking natives.

The results are shown in Table E5, using Census and ACS data. It is apparent that immigrants are around 10 percentage points less likely to own the house in which they reside.

B Theory appendix

B.1 Derivation of indirect utility

The utility in location c for an individual i from country of origin j can be written as:

$$\ln U_{ijc} = \rho + \ln A_{jc} + \alpha_t \ln C_T + (1-\alpha_t)\frac{\sigma}{\sigma-1} \ln\left(\bar{\alpha}_l C_H)^{\frac{\sigma-1}{\sigma}} + \bar{\alpha}_f (C_F)^{\frac{\sigma-1}{\sigma}}\right) + \varepsilon_{ijc}$$
$$\text{s.t. } C_T + p_c C_H + p_j C_F \leq w_{jc}$$

We can solve the maximization problem in two stages.

- **Stage 1**: Define an auxiliary variable E and find the optimal decisions $C_{NT*}(p_c, p_j, E)$ and $C_{F*}(p_c, p_j, E)$ to the following maximization problem:

$$\max \ (1-\alpha_t)\frac{\sigma}{\sigma-1} \ln\left(\bar{\alpha}_l (C_H)^{\frac{\sigma-1}{\sigma}} + \bar{\alpha}_f (C_F)^{\frac{\sigma-1}{\sigma}}\right)$$
$$\text{s.t. } p_c C_H + p_j C_F = E.$$

Let

$$\tilde{V}(p_c, p_j, E) = (1-\alpha_t)\frac{\sigma}{\sigma-1}\ln\left(\bar{\alpha}_l (C_{NT*})^{\frac{\sigma-1}{\sigma}} + \bar{\alpha}_f (C_{F*})^{\frac{\sigma-1}{\sigma}}\right).$$

- **Stage 2**: Solve for $C_j^{T*}(p_c, p_j, w_{jc})$ and $E^*(p_c, p_j, w_{jc})$ of the maximization problem

$$\begin{aligned}
&\max\ \rho + \ln A_{jc} + \alpha_t \ln C_T + \tilde{V}(p_c, p_j, E)\\
&\text{s.t. } C_T + E \leq w_{jc}.
\end{aligned}$$

Stage 1

$$\begin{aligned}
&\max\ \ (1-\alpha_t)\frac{\sigma}{\sigma-1}\ln\left(\bar{\alpha}_l (C_H)^{\frac{\sigma-1}{\sigma}} + \bar{\alpha}_f (C_F)^{\frac{\sigma-1}{\sigma}}\right)\\
&\text{s.t. } p_c C_H + p_j C_F = E
\end{aligned}$$

The associated Lagrangian is

$$\mathcal{L} = (1-\alpha_t)\frac{\sigma}{\sigma-1}\ln\left(\bar{\alpha}_l (C_H)^{\frac{\sigma-1}{\sigma}} + \bar{\alpha}_f (C_F)^{\frac{\sigma-1}{\sigma}}\right) + \lambda(E - p_c C_H - p_j C_F).$$

First-order conditions are given by

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial C_H} &: \ \frac{(1-\alpha_t)\bar{\alpha}_l (C_H)^{\frac{-1}{\sigma}}}{\bar{\alpha}_l (C_H)^{\frac{\sigma-1}{\sigma}} + \bar{\alpha}_f (C_F)^{\frac{\sigma-1}{\sigma}}} - p_c\lambda = 0\\
\frac{\partial \mathcal{L}}{\partial C_F} &: \ \frac{(1-\alpha_t)\bar{\alpha}_f (C_F)^{\frac{-1}{\sigma}}}{\bar{\alpha}_l (C_H)^{\frac{\sigma-1}{\sigma}} + \bar{\alpha}_f (C_F)^{\frac{\sigma-1}{\sigma}}} - p_j\lambda = 0.
\end{aligned}$$

Dividing the two first-order conditions, we obtain the following relationship:

$$\frac{\bar{\alpha}_l}{\bar{\alpha}_f}\left(\frac{C_H}{C_F}\right)^{\frac{-1}{\sigma}} = \frac{p_c}{p_j} \Rightarrow C_H = \left(\frac{\bar{\alpha}_f p_c}{\bar{\alpha}_l p_j}\right)^{-\sigma} C_F.$$

Using this relationship and the budget constraint, we find

$$\begin{aligned}
C_H &= \frac{\left(\frac{p_c}{\bar{\alpha}_l}\right)^{-\sigma}}{p_c\left(\frac{p_c}{\bar{\alpha}_l}\right)^{-\sigma} + p_j\left(\frac{p_j}{\bar{\alpha}_f}\right)^{-\sigma}} E\\
C_F &= \frac{\left(\frac{p_j}{\bar{\alpha}_f}\right)^{-\sigma}}{p_c\left(\frac{p_c}{\bar{\alpha}_l}\right)^{-\sigma} + p_j\left(\frac{p_j}{\bar{\alpha}_f}\right)^{-\sigma}} E.
\end{aligned}$$

Thus, the maximized objective function is

$$
\begin{aligned}
\tilde{V} &= (1-\alpha_t)\frac{\sigma}{\sigma-1}\ln\left(\bar{\alpha}_l\left(\frac{\left(\frac{p_c}{\bar{\alpha}_l}\right)^{-\sigma}}{p_c\left(\frac{p_c}{\bar{\alpha}_l}\right)^{-\sigma}+p_j\left(\frac{p_j}{\bar{\alpha}_f}\right)^{-\sigma}}E\right)^{\frac{\sigma-1}{\sigma}}+\bar{\alpha}_f\left(\frac{\left(\frac{p_j}{\bar{\alpha}_f}\right)^{-\sigma}}{p_c\left(\frac{p_c}{\bar{\alpha}_l}\right)^{-\sigma}+p_j\left(\frac{p_j}{\bar{\alpha}_f}\right)^{-\sigma}}E\right)^{\frac{\sigma-1}{\sigma}}\right) \\
&= (1-\alpha_t)\ln E + (1-\alpha_t)\frac{1}{\sigma-1}\ln\left(p_c\left(\frac{p_c}{\bar{\alpha}_l}\right)^{-\sigma}+p_j\left(\frac{p_j}{\bar{\alpha}_f}\right)^{-\sigma}\right) \\
&= (1-\alpha_t)\ln E - (1-\alpha_t)\ln \bar{p}(\bar{\alpha}_l,\bar{\alpha}_f).
\end{aligned}
$$

where $\bar{p}_{jc}(\bar{\alpha}_l,\bar{\alpha}_f) = (\bar{\alpha}_l^{\sigma} p_c^{1-\sigma} + \bar{\alpha}_f^{\sigma} p_j^{1-\sigma})^{\frac{1}{1-\sigma}}$.

Stage 2

$$
\begin{aligned}
&\max\ \rho + \ln A_{jc} + \alpha_t \ln C_T + (1-\alpha_t)\ln E + (1-\alpha_t)\frac{1}{\sigma-1}\ln \bar{p}(\bar{\alpha}_l,\bar{\alpha}_f) \\
&\text{s.t. } C_T + E \le w_{jc}
\end{aligned}
$$

The associated Lagrangian is

$$
\mathcal{L} = \rho + \ln A_{jc} + \alpha_t \ln C_T + (1-\alpha_t)\ln E - (1-\alpha_t)\ln \bar{p}(\bar{\alpha}_l,\bar{\alpha}_f) + \lambda(w_{jc} - C_T - E).
$$

The first-order conditions are

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial C_T} &: \frac{\alpha_t}{C_T} - \lambda = 0 \\
\frac{\partial \mathcal{L}}{\partial E} &: \frac{(1-\alpha_t)}{E} - \lambda = 0.
\end{aligned}
$$

Using these first-order conditions and budget constraints,

$$
\begin{aligned}
C_T &= \alpha_t w_{jc} \\
E &= (1-\alpha_t) w_{jc}.
\end{aligned}
$$

Thus, the optimal choices for consumption can be written as

$$
\begin{aligned}
C_T &= \alpha_t w_{jc} \\
C_H &= \frac{\bar{\alpha}_l^\sigma p_c^{-\sigma}}{\bar{\alpha}_l^\sigma p_c^{1-\sigma} + \bar{\alpha}_f^\sigma p_j^{1-\sigma}}(1-\alpha_t) w_{jc} \\
C_F &= \frac{\bar{\alpha}_f^\sigma p_j^{-\sigma}}{\bar{\alpha}_l^\sigma p_c^{1-\sigma} + \bar{\alpha}_f^\sigma p_j^{1-\sigma}}(1-\alpha_t) w_{jc}.
\end{aligned}
$$

This solution can be shown to satisfy the first-order conditions of the original problem. If we let ρ be a constant such that the indirect utility function has no constant, the indirect utility function can be written as

$$
\ln V_{ijc} = \ln V_{jc} + \ln \varepsilon_{ijc} = \ln A_{jc} + \ln w_{jc} - (1-\alpha_t) \ln \bar{p}_{jc}(\bar{\alpha}_l, \bar{\alpha}_f) + \ln \varepsilon_{ijc}
$$

with $\bar{p}_{jc}(\bar{\alpha}_l, \bar{\alpha}_f) = (\bar{\alpha}_l^\sigma p_c^{1-\sigma} + \bar{\alpha}_f^\sigma p_j^{1-\sigma})^{\frac{1}{1-\sigma}}$.

B.2 Proofs of propositions of the quantitative model

Assumptions Natives only care about local price indices so that $\alpha_f = 0$ and $\alpha_l = \alpha$. Immigrants care about local and foreign price indices so that $\alpha_f \neq 0$ and $\alpha_l + \alpha_f = \alpha$. Moreover, $\ln A_{jc} = \ln A_c + \ln \tilde{A}_{jc}$ and $\ln \tilde{A}_{Nc} = 0$.

Proof. **Proposition 1**

- $\ln w_{jc} = -(1-\beta)(\ln A_c + \ln \tilde{A}_{jc}) + \beta \ln \tilde{B}_c + (1-\beta)(1-\alpha_t) \ln \bar{p}_{jc}$
- $\ln w_{Nc} = -(1-\beta) \ln A_c + \beta \ln \tilde{B}_c + (1-\beta)(1-\alpha_t) \ln p_c$

Taking the difference, we get

$$
\ln w_{Nc} - \ln w_{jc} = (1-\beta)(1-\alpha_t) \ln \frac{p_c}{\bar{p}_{jc}} + (1-\beta) \ln \tilde{A}_{jc}
$$

Denote $W = \ln w_{Nc} - \ln w_{jc}$. We are interested in the sign of $\frac{\partial W}{\partial \ln p_c}$.

$$
\begin{aligned}
W =& (1-\beta)(1-\alpha_t) \ln p_c - (1-\beta)(1-\alpha_t) \frac{1}{1-\sigma} \ln(\bar{\alpha}_l^\sigma p_c^{1-\sigma} + \bar{\alpha}_f^\sigma p_j^{1-\sigma}) \\
\frac{\partial W}{\partial p_c} =& \frac{(1-\beta)(1-\alpha_t)}{p_c} - \frac{(1-\beta)(1-\alpha_t)\bar{\alpha}_l^\sigma p_c^{-\sigma}}{\bar{\alpha}_l^\sigma p_c^{1-\sigma} + \bar{\alpha}_f^\sigma p_j^{1-\sigma}} \\
\frac{\partial W}{\partial p_c} p_c =& (1-\beta)(1-\alpha_t) \frac{\bar{\alpha}_f^\sigma p_j^{1-\sigma}}{\bar{\alpha}_l^\sigma p_c^{1-\sigma} + \bar{\alpha}_f^\sigma p_j^{1-\sigma}}.
\end{aligned}
$$

As $\partial \ln p_c / \partial p_c = 1/p_c$, we obtain the wanted derivative, which is always positive:

$$\frac{\partial W}{\partial \ln p_c} = (1-\beta)(1-\alpha_t)\frac{\bar{\alpha}_f^{\sigma} p_j^{1-\sigma}}{\bar{\alpha}_l^{\sigma} p_c^{1-\sigma} + \bar{\alpha}_f^{\sigma} p_j^{1-\sigma}} > 0.$$

Also,

$$\frac{\partial^2 W}{\partial \ln p_c \partial p_j} = (1-\beta)(1-\alpha_t)\frac{(\bar{\alpha}_l^{\sigma} p_c^{1-\sigma} + \bar{\alpha}_f^{\sigma} p_j^{1-\sigma})(1-\sigma)\bar{\alpha}_f^{\sigma} p_j^{-\sigma} - \bar{\alpha}_f^{\sigma} p_j^{1-\sigma}(1-\sigma)\bar{\alpha}_f^{\sigma} p_j^{-\sigma}}{(\bar{\alpha}_l^{\sigma} p_c^{1-\sigma} + \bar{\alpha}_f^{\sigma} p_j^{1-\sigma})^2}.$$

Multiplying both sides with p_j and again using $\partial \ln p_j / \partial p_j = 1/p_j$ we get

$$\frac{\partial^2 W}{\partial \ln p_c \partial \ln p_j} = (1-\beta)(1-\alpha_t)(1-\sigma)\frac{\bar{\alpha}_f^{\sigma} p_j^{1-\sigma} \bar{\alpha}_l^{\sigma} p_c^{1-\sigma}}{(\bar{\alpha}_l^{\sigma} p_c^{1-\sigma} + \bar{\alpha}_f^{\sigma} p_j^{1-\sigma})^2} < 0$$

Thus, the gap in wages between natives and immigrants is increasing in the local price index. Furthermore, the effect of the local price index on the wage gap is larger for lower p_j when $\sigma > 1$.

□

Proof. **Proposition 2**

We have that

$$\pi_{jc} = \frac{V_{jc}^{1/\lambda}}{\sum_k V_{jk}^{1/\lambda}} = \left(\frac{V_{jc}}{V_j}\right)^{1/\lambda}.$$

Thus,

$$\ln \pi_{jc} - \ln \pi_{Nc} = \frac{1}{\lambda}(\ln V_{jc} - \ln V_{Nc}) - \frac{1}{\lambda}(\ln V_j - \ln V_N).$$

Using the definition of $\ln V_{jc}$ and the expression for the wage gap obtained above, we have

$$\begin{aligned} \ln V_{jc} - \ln V_{Nc} &= \ln w_{jc} - \ln w_{Nc} - (1-\alpha_t)\left(\ln \bar{p}_{jc} - \ln p_c\right) + \ln \tilde{A}_{jc} \\ &= \beta(1-\alpha_t)\ln p_c - \beta(1-\alpha_t)\ln \bar{p}_{jc} + \beta \ln \tilde{A}_{jc}. \end{aligned}$$

Note that

$$\begin{aligned} \ln V_{jc} &= \ln A_{jc} + \ln w_{jc} - (1-\alpha_t)\ln \bar{p}_{jc} \\ &= \ln A_{jc} + \big(-(1-\beta)\ln A_{jc} + \beta \ln \tilde{B}_c + (1-\beta)(1-\alpha_t)\ln \bar{p}_{jc}\big) - (1-\alpha_t)\ln \bar{p}_{jc} \\ &= \beta \ln A_{jc} + \beta \ln \tilde{B}_c - \beta(1-\alpha_t)\ln \bar{p}_{jc} \end{aligned}$$

Thus,

$$V_{jc} = A_{jc}^{\beta} \tilde{B}_c^{\beta} \bar{p}_{jc}^{\beta(\alpha_t - 1)}.$$

Then,

$$V_j = \left(\sum_k \left(A_{jk} \tilde{B}_k \bar{p}_{jk}^{(\alpha_t - 1)} \right)^{\frac{\beta}{\lambda}} \right)^{\lambda}$$

$$V_N = \left(\sum_k \left(A_k \tilde{B}_k p_k^{(\alpha_t - 1)} \right)^{\frac{\beta}{\lambda}} \right)^{\lambda} = \left(\sum_k \left(\frac{A_k \tilde{B}_k}{p_c^{1-\alpha_t}} \right)^{\frac{\beta}{\lambda}} \right)^{\lambda}.$$

And from this we have the expression we needed.

□

Proof. **Proposition 3**

Note that

$$\pi_{jc} = \frac{L_{jc}}{L_j} = \left(\frac{V_{jc}}{V_j} \right)^{\frac{1}{\lambda}} = \left(\frac{A_{jc}^{\beta} \tilde{B}_c^{\beta} / \bar{p}_{jc}^{\beta(1-\alpha_t)}}{V_j} \right)^{\frac{1}{\lambda}}.$$

Then, the total immigrant population in city c is

$$L_{Ic} = \sum_j L_{jc} = \sum_j L_j \frac{L_{jc}}{L_j} = \sum_j L_j \left(\frac{A_{jc}^{\beta} \tilde{B}_c^{\beta} / \bar{p}_{jc}^{\beta(1-\alpha_t)}}{V_j} \right)^{\frac{1}{\lambda}}.$$

Substituting the expression for V_j, we get

$$L_{Ic} = (A_{jc} \tilde{B}_c)^{\frac{\beta}{\lambda}} \sum_j \frac{L_j / \bar{p}_{jc}^{(1-\alpha_t)\frac{\beta}{\lambda}}}{\sum_k (A_{jk} \tilde{B}_k / \bar{p}_{jk}^{(1-\alpha_t)})^{\frac{\beta}{\lambda}}}.$$

For natives,

$$L_{Nc} = \frac{(A_c \tilde{B}_c)^{\frac{\beta}{\lambda}}}{\sum_k (A_k \tilde{B}_k / p_k^{(1-\alpha_t)})^{\frac{\beta}{\lambda}}} L_N / p_c^{(1-\alpha_t)\frac{\beta}{\lambda}} = \frac{(A_c \tilde{B}_c / p_c^{(1-\alpha_t)})^{\frac{\beta}{\lambda}}}{\sum_k (A_k \tilde{B}_k / p_c^{(1-\alpha_t)})^{\frac{\beta}{\lambda}}} L_N,$$

and $L_c = L_{Ic} + L_{Nc}$. □

C Estimation details and robustness to alternative calibrations

C.1 Tradable and origin goods expenditure shares

Tradable goods expenditure share

Our chosen tradable goods expenditure share of 40% is based on the sum of the weights of those goods in the Bureau of Labor Statistics (BLS) consumer price index (CPI-U) that are either tradable or difficult to substitute with goods from the origin country because they are mainly consumed locally. Specifically, these are *Food and beverages*, *Apparel*, *Transportation*, and *Education and communication*. The remaining components of the CPI-U are *Housing*, *Medical care*, *Recreation*, and *Other goods and services*. While housing is also consumed locally, unlike transportation or education, it can be partially substituted with home consumption; for example, by renting a low-quality apartment in the host country while sending remittances or saving to rent or buy an apartment of higher quality in the home country. This notion is consistent with our findings on differences in housing expenditure between natives and immigrants in Section 3.3. In Section C.2, we test the robustness of the model parameters to setting the tradable share to a lower bound of 20% or an upper bound of 60%. The latter value implies that non-tradable goods roughly have the same weight as the *Housing* component in the CPI-U. We view this as an upper bound for the actual share of tradable consumption, because non-tradable goods usually include more than just housing.

Origin goods expenditure share

We proxy the share of expenses in the origin country by the annual amount of remittances sent by immigrants divided by their annual income. We believe this share is rather a lower bound for the actual share of expenditures in origin goods because it does not include savings or income spent during visits in the origin country.

We compute the total amount of outgoing remittances from the US using the Bilateral Remittance Matrix provided by the World Bank.[48] The earliest available years around 2010, so we base our calculation of the remittance share of income on the total remittances sent from the US during that year, which amounts to around $110 billion. We calculate the total income of immigrants by summing up their total wage income in the three-year 2010 ACS, restricting the sample to labor force participants aged 18 to 65. This amounts to $540 billion. Thus, the resulting remittance share of income is 20%.

C.2 Estimates of on α_f and σ under alternative assumptions for β and α_t

In our model estimation based on the structural equations (4.7) and (4.8), we are not able to separately identify β and α_t from the remaining parameters. Thus, the vector of structurally estimated parameters depends on the assumption on these two parameters, while the fit and the predictions of the model remain (virtually) the same regardless of their values.[49]

To document the sensitivity of the key parameters α_f and σ, we estimate the model for a range of combinations of β and α_t within a reasonable interval. In particular, we create a grid for β bounded between 0.2 and 0.8 and a grid for α_t bounded between 0.2 and 0.6. Grid values are evenly spaced with an increment of 0.05, yielding 117 parameter combinations. Figure D.4 shows the densities of the estimates of α_f and σ that we obtain by estimating the model for each of these combinations. The plots suggest that the estimates lie within a relatively close range around the baselines estimates (indicated by the red vertical line). Importantly, the elasticity of substitution between local and origin goods lies well above 1, the limiting case in which there would not be any heterogeneity across countries of origin. The estimate of α_f varies between 0.129 and 0.195, while that of σ varies between 1.53 and 2.86.

The highest estimate for σ, which coincides with the lowest one for α_f, is obtained when setting both

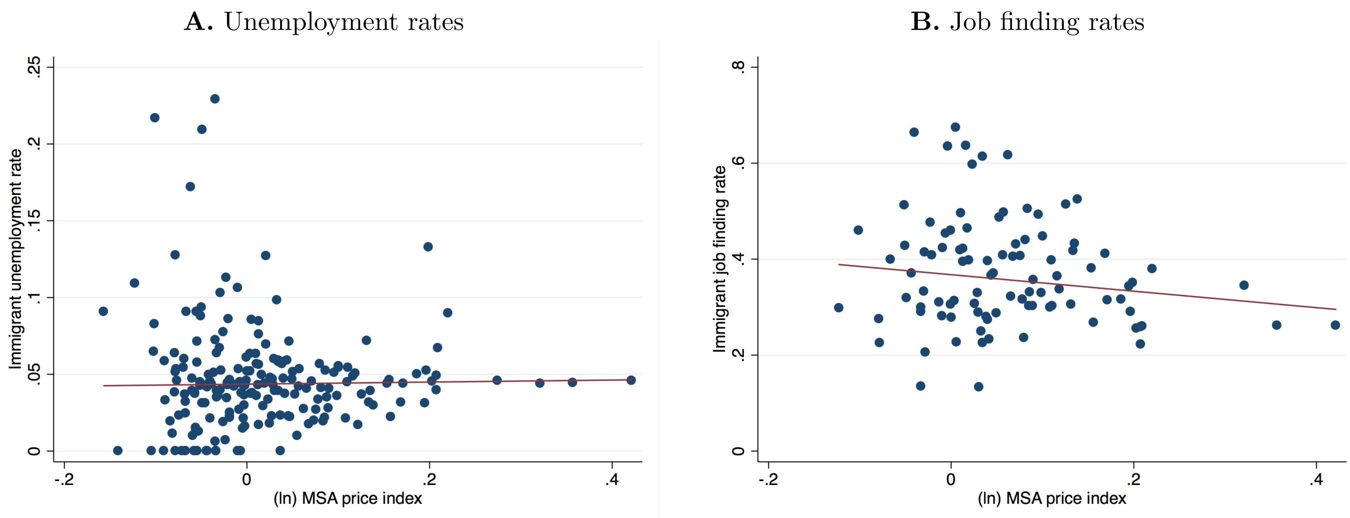
[48] `https://www.worldbank.org/en/topic/migrationremittancesdiasporaissues/brief/migration-remittances-data`

[49] This is true as long as β is neither equal to its lower bound of zero nor its upper bound of one. If $\beta = 0$, the model cannot generate any wage gap between natives and immigrants, while if $\beta = 1$, it cannot generate any difference in their distribution across cities. Further, if $\alpha_t = 0$ and thus there is no non-tradable consumption, the model cannot generate neither wage gaps nor different spatial distributions.

α_t and β to their upper bounds. Intuitively, this is because the sensitivity of the wage gap to origin prices falls in both parameters as seen in equation (4.7). Hence, this is needs to be compensated by a higher elasticity of substitution between local and origin goods in order to fit the data. A higher elasticity implies a larger share of spending on origin goods; this in turn needs to be compensated by a lower α_f, which ensures that the model yields an average origin spending share of 0.2. Accordingly, the lowest and highest estimates of σ and α_f, respectively, are obtained when α_t and β are at their lower bounds.

D Figures appendix

Figure D.1: City size, price index, and immigrants' unemployment and job finding

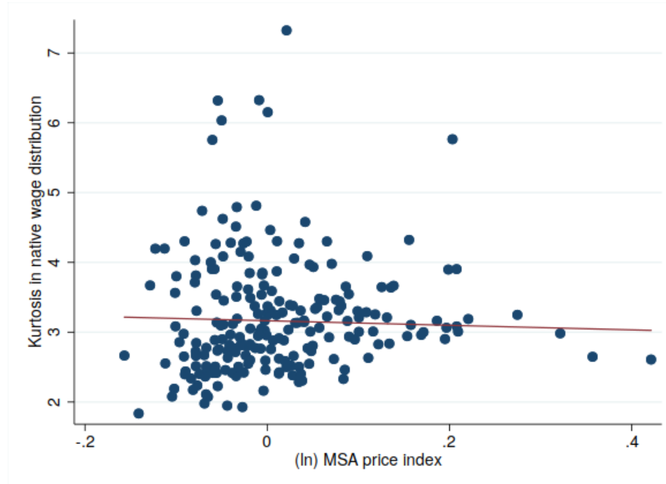


Notes: This figure uses city price data from the 2000 Census and data for immigrant workers aged 25 to 59 from the CPS basic monthly files. The unemployment and job finding rates are calculated for each city that can be matched to the Census data and are computed as the average of the variable over the period 1995-2005. The job finding rate is the monthly share of unemployed job searchers transitioning to employment.

Figure D.2: City size, price index, and skewness and kurtosis

A. Skewness

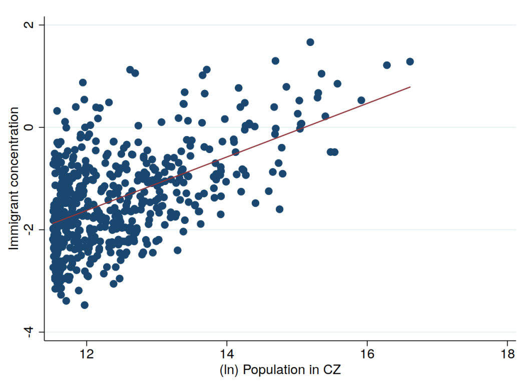
B. Kurtosis



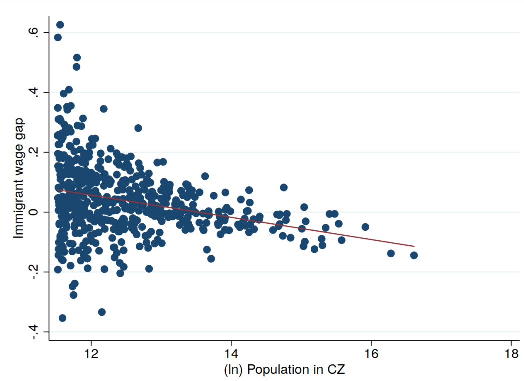
Notes: This figure uses data from the 2000 Census and computes the within-MSA wage skewness and kurtosis, which are plotted against city level prices.

Figure D.3: Commuting zone size and immigrant distribution

A. Immigrant concentration

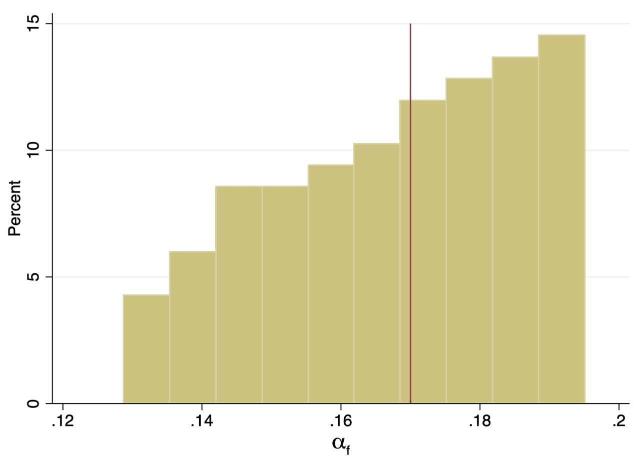


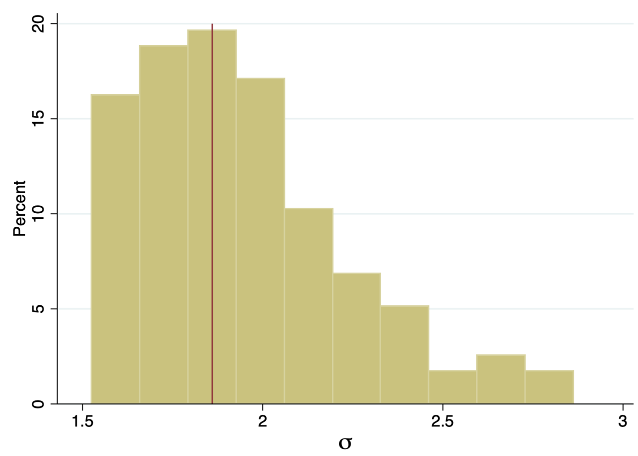
B. Wage gap



Notes: The figure is based on the sample of prime-age workers (25-59) from the 2000 Census. Each dot represents a different commuting zone (CZ). There are 541 different CZs in our sample, which are the ones with positive numbers of immigrants out of 741 CZs. The red line is the fitted line of a linear regression. The left panel shows the relationship between the immigrant concentration and CZ size, while the right panel shows the relationship between native-immigrant wage gaps and CZ size.

Figure D.4: Robustness of estimates of α_f and σ to different β and α_t





Notes: These histograms show the percentages of parameter estimates of α_f and σ lying in the respective bins. We obtain these values by estimating the model for a 117-point uniform grid $\beta \times \alpha_t \in [0.2, 0.8] \times [0.2, 0.6]$. The baseline estimates obtained when assuming $\beta = 0.5$ and $\alpha_t = 0.4$ are indicated by the vertical red lines.

E Tables appendix

Table E1: Countries with the highest and the lowest real exchange rates

Highest		Lowest	
Norway	1.36	Vietnam	0.20
Japan	1.34	Pakistan	0.21
Bermuda	1.34	Yemen	0.23
Denmark	1.30	Egypt	0.24
Switzerland	1.29	Sierra Leone	0.24
Sweden	1.25	Sri Lanka	0.24
Finland	1.22	Nepal	0.25
United Kingdom	1.10	Indonesia	0.26
France	1.07	Tanzania	0.26
Australia	1.07	Nigeria	0.26

Notes: This table lists the top and bottom 10 countries with the highest and the lowest average real exchange rate over 1990, 2000 and 2010 with respect to the United States according to real exchange rate data from the World Bank.

Table E2: Relative labor supply and the wage gap

VARIABLES	(1) OP	(2) BGH	(3) Card	(4) Eq. (3.2)	(5) Eq. (3.2)	(6) Eq. (3.2)
Rel. labor supply	-0.0355**	-0.0199	-0.0308***	-0.0529***	-0.0135**	-0.00571
	(0.0150)	(0.0150)	(0.00622)	(0.00398)	(0.00555)	(0.00712)
Observations	120	120	124	5,571	5,571	5,570
R-squared	0.791	0.763	0.809	0.287	0.391	0.473
Skill FE	yes	yes	no	yes	yes	yes
MSA FE	no	no	no	no	yes	yes
MSA-year FE	no	no	no	no	no	yes
F-stat 1st stage			696.9	1103	396.3	283.3

Notes: This table reports results of running regressions of relative immigrant wages on relative immigrant supplies based on data from the Census 1980 to 2000 and the combined ACS 2009-2011. Column 1 uses variation across experience, education groups, and decades (8 x 4 x 4, with 8 cells that are missing because of the experience definition base on age and years of education). In column 1 we use the log of the mean of raw wages, as in Ottaviano and Peri (2012) (indicated by “OP”), to compute relative wages. Column 2 replicates the specification of column 1, but we use as dependent variable the difference in the mean of the log of (raw) wages between natives and immigrants, as recommended in Borjas et al. (2012) (indicated by “BGH”). Column 3 uses variation across 124 metropolitan areas following Card (2009), Table 6, adjusting wages for composition. Columns 4 to 6 use variation across experience, education, metropolitan areas, and decade (4 x 2 x 212 x 4, with 1213 missing observations due to 0 immigrants and (especially) lower coverage of metropolitan areas in 1970, which is used to build the IV for 1980). IV estimates are reported in columns 3 to 6 using the networks IV with the preceding decade immigrant distribution to assign flows. Robust standard errors clustered at the experience-education level (columns 1 and 2) and clustered at the metropolitan area (columns 3 to 6) are reported. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table E3: Immigration, wage gaps and city price levels, robustness

	A. Immigrant concentration, OLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(ln) Price	8.095***	9.321***	10.349***	9.984***	8.243***	6.255***	7.673***
	(0.670)	(0.884)	(1.209)	(0.716)	(0.607)	(0.516)	(0.431)
Observations	740	739	739	740	740	740	740
R-squared	0.606	0.536	0.482	0.634	0.670	0.651	0.738
	B. Immigrant concentration, IV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(ln) Price	10.272***	10.685***	12.898***	12.215***	9.981***	7.683***	9.407***
	(0.876)	(1.186)	(1.536)	(0.933)	(0.777)	(0.686)	(0.549)
Observations	740	739	739	740	740	740	740
R-squared	0.562	0.525	0.454	0.603	0.640	0.617	0.700
	C. Wage gap, OLS						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(ln) Price	-0.591***	-1.141***	-0.737***	-0.395***	-0.539***	-0.753***	-1.022***
	(0.065)	(0.153)	(0.143)	(0.050)	(0.061)	(0.112)	(0.087)
Observations	657	638	630	641	646	655	657
R-squared	0.472	0.474	0.406	0.301	0.357	0.394	0.627
	D. Wage gap, IV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(ln) Price	-0.663***	-1.240***	-0.937***	-0.284***	-0.481***	-0.908***	-1.156***
	(0.080)	(0.195)	(0.152)	(0.090)	(0.099)	(0.134)	(0.109)
Observations	657	638	630	641	646	655	657
R-squared	0.466	0.472	0.386	0.292	0.354	0.377	0.616
Sample	Docum.	Undocum.	<HS	HS	SC	C	No Lat. Am.

Notes: This table shows regressions of the immigrant concentration, in Panels A and B, and the gap in wages between natives and immigrants adjusted for imperfect substitutability between natives and immigrants ($\hat{\phi}_{c,t}$), in Panels C and D, on the city price index. City price indices are computed following Moretti (2013). Panels A and C show OLS regressions, while panels B and D instrument the local price index by the housing supply elasticity estimated in Saiz (2010). This table uses data from the Census 1980 to 2000 and the combined 2009-2011 ACS data. Column 1 only includes documented immigrants. Column 2 only includes undocumented immigrants. Undocumented immigrants are identified following Borjas (2017). Column 3 restricts the sample to high-school dropouts. Column 4 restricts the sample to high-school graduates. Column 5 restricts the sample to workers with some college. Column 6 restricts the sample to workers with a college degree or more. Column 7 excludes immigrants from Latin American countries. Robust standard errors clustered at the metropolitan area are reported. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table E4: Wages of immigrants arriving before and after 1982

	(1)	(2)	(3)	(4)
(ln) Price	0.474***	0.565***	0.500***	0.677***
	(0.109)	(0.086)	(0.106)	(0.178)
(ln) Price × before 82	0.026	-0.015	0.037	-0.013
	(0.021)	(0.025)	(0.029)	(0.084)
(ln) Price × undocumented		-0.281***	-0.238***	-0.492***
		(0.081)	(0.073)	(0.111)
(ln) Price × undocumented × before 82		0.104***	0.074**	0.282***
		(0.029)	(0.031)	(0.109)
Observations	227875	227875	170940	50665
R-squared	0.369	0.369	0.207	0.218
Sample	All	All	$<$ College	$<$ College 1990

Notes: This table shows how wages of immigrants increase with city prices for immigrants arriving before 1982 and after 1982. IRCA 1986 allowed immigrants arriving prior to 1982 to gain legal status. Likely undocumented immigrants are identified following Borjas (2017). Robust standard errors clustered at the metropolitan area level are reported. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table E5: Immigrants' homeownership rates

	(1) Ownership	(2) Ownership	(3) Ownership	(4) Ownership
Immigrant	-0.193***	-0.116***	-0.137***	
	(0.00937)	(0.00993)	(0.0178)	
(ln) HH income		0.213***	0.215***	0.231***
		(0.00411)	(0.00431)	(0.00686)
(ln) RER				-0.0135**
				(0.00573)
Observations	3,464,960	3,464,960	3,464,960	444,254
Year FE	yes	yes	yes	yes
MSA FE	yes	yes	no	yes
Sample	all	all	all	Imm. only

Notes: This table shows regressions with a dummy for home ownership as dependent variable using Census 1990 and 2000 and ACS data for 2009-2011 data. Additional controls include dummies for the number of family members living in the household, marital status and age. Standard errors clustered at the MSA level. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table E6: Immigrants' expenditures on housing, alternative specifications

A.

	(1) ln (rent / HH Income)	(2) ln (rent / HH Income)	(3) ln (rent / HH Income)	(4) ln (rent / HH Income)
Immigrant	0.118***	-0.0741***	-0.110***	
	(0.0130)	(0.0146)	(0.0236)	
Immigrant x (ln) City price			0.0383	
			(0.0318)	
(ln) RER				0.0711***
				(0.0155)
(ln) HH income		-0.729***	-0.730***	-0.764***
		(0.00486)	(0.00549)	(0.00551)
Observations	903,825	903,825	903,825	204,276
Year FE	yes	yes	yes	yes
MSA FE	yes	yes	no	yes
Sample	all	all	all	Imm. only

B.

	(1) (rent / HH Income)	(2) (rent / HH Income)	(3) (rent / HH Income)	(4) (rent / HH Income)
Immigrant	0.0407***	-0.0171***	-0.0515***	
	(0.00338)	(0.00343)	(0.00673)	
Immigrant x (ln) City price			0.0374***	
			(0.00790)	
(ln) RER				0.0146***
				(0.00548)
(ln) HH income		-0.220***	-0.219***	-0.251***
		(0.00411)	(0.00384)	(0.00567)
Observations	903,825	903,825	903,825	204,276
Year FE	yes	yes	yes	yes
MSA FE	yes	yes	no	yes
Sample	all	all	all	Imm. only

Notes: These regressions follow Table 5 with the dependent variable in levels. Standard errors are clustered at the MSA level. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table E7: Immigrant IV, first step

VARIABLES	(1) Immigration Concetration	(2) Immigration Concetration	(3) Immigration Concetration
Housing supply elasticity (Saiz, 2010)	-0.640***	-0.656***	
	(0.103)	(0.0251)	
Real exchange rate, WB		-0.474	-0.337***
		(0.231)	(0.0872)
Real exchange rate x HS elasticity		0.325*	0.263***
		(0.0797)	(0.0565)
Observations	740	30,750	30,750
R-squared	0.394	0.217	0.471
Year FE	yes	yes	yes
CO variation and FE	no	yes	yes
Metropolitan area FE	no	no	yes

Notes: Data from the 1980, 1990, and 2000 Censuses and the ACS in 2009-2011. The number of metropolitan areas is 185 (4 x 185 = 740). Columns 2 and 3 use data from 68 countries of origin. Note that there are a few country of origin-metropolitan area-year cells with 0 observations. Robust standard errors clustered at the metropolitan are reported. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.

Table E8: Immigration and native mobility, checking for pre-trends

	(1)	(2)	(3)	(4)
	Dep. Var.: Native relocation ($\frac{\Delta \text{Nat}}{\text{Nat}}$)			
VARIABLES	IV	IV	IV	IV
$\frac{\Delta \text{Imm}}{\text{Nat}}$, Lagged 1 Period	-0.192	-0.177	-0.0573	0.0871
	(0.177)	(0.162)	(0.247)	(0.287)
$\frac{\Delta \text{Imm}}{\text{Nat}}$, Lagged 2 Periods			0.198	0.0366
			(0.127)	(0.185)
Observations	370	370	185	185
Year FE	yes	yes	yes	yes
Region FE	yes	yes	yes	yes
Controls	yes	yes	yes	yes
IV	AM + Net	AM	AM + Net	AM
F-stat First Stage	9.103	210.5	7.249	14.80

Notes: Data come from the 1990 and 2000 Censuses and the ACS in 2009-2011. The number of metropolitan areas is 185. Robust standard errors clustered at the state level are reported. * significant at the 0.10 level; ** significant at the 0.05 level; *** significant at the 0.01 level.