Homeownership, Polarization, and Inequality*

Andrii Parkhomenko†

May 19, 2025

Abstract

Why are job polarization and income inequality higher in large U.S. cities? I offer a new explanation: when house prices grow faster in large cities, middle-income households increasingly cannot afford to own a house there. They move to smaller cities and the middle of the income distribution in large cities hollows out, making them more polarized and unequal. I document that (1) cities with higher price growth experienced larger polarization and increase in inequality since 1980 and (2) middle-income households migrate more often to cheaper locations for housing-related reasons than low- or high-income households. Using a spatial equilibrium model with tenure choice and skill heterogeneity, I find that excess growth of prices relative to incomes and rents in large cities accounts for nearly all of the gap in polarization and almost one-half of the gap in inequality growth between large and small cities from 1980 to 2019.

Key Words: homeownership, job polarization, income inequality, house prices, spatial equilibrium.

JEL Classification: J24, J31, R21, R23, R31.

*I thank Daniel Angel for outstanding research assistance. I also thank Kurt Mitman (editor), four anonymous referees, David Albouy, Nate Baum-Snow, Eirik Brandsaas, Victor Couture, Don Davis, Jorge De la Roca, Matt Delventhal, Jan Eeckhout, Heidi Falkenbach, Carlos Garriga, Andra Ghent, Laurent Gobillon, James Graham, Aaron Hedlund, Christian Hilber, Ayse Imrohoroglu, Erica Jiang, Matt Kahn, Wenhao Li, Giacomo Ponzetto, Frederic Robert-Nicoud, Stijn Van Nieuwerburgh, Jaume Ventura, Weifeng Wu, and Yuxi Yao, as well as seminar and conference participants at USC Marshall, ASSA, UEA European Meeting, AREUEA National Meeting, SED, EEA-ESEM, Stanford Institute for Theoretical Economics: Housing & Urban, UEA North American Meeting, NYU Furman Center, Southern Methodist University, Workshop on Structural Transformation and Macroeconomic Dynamics in Cagliari, WEAI Meeting in Portland, SAET, IHA Summit at the University of Utah, AREUEA virtual seminar, AEA Meeting in New Orleans, Universitat Autonoma de Barcelona, Conference on Economics of Housing and Housing Policies (ECHOPPE) in Toulouse, Universitat Pompeu Fabra, AREUEA International Meeting in Cambridge, University of Illinois Urbana-Champaign, SEA Meeting in New Orleans, Conference on Low-Income Housing Supply and Housing Affordability at UCLA, and Workshop on Urban Economics in Barcelona for comments and suggestions. First draft: May 26, 2021.

†University of Southern California, Marshall School of Business, Department of Finance and Business Economics; Los Angeles, CA 90089. Email: parkhome@marshall.usc.edu

1 Introduction

Since the 1980s, the labor market in the United States has become more unequal and polarized. Individual differences in labor earnings have widened, while the shares of low- and high-income jobs have increased at the expense of middle-income jobs. Moreover, this labor market polarization and the increase in income inequality have been more pronounced in large cities (Baum-Snow and Pavan, 2013; Autor, 2019).

Why is the middle-class leaving large cities and why have those cities become more unequal? These are important questions for economic growth, human capital investment, and public policy. The leading explanations in the literature emphasized the role of production technology and offered explanations that rely on skill-biased technical change (Baum-Snow, Freedman, and Pavan, 2018; Cerina, Dienesch, Moro, and Rendall, 2023), external labor demand shocks (Davis, Mengus, and Michalski, 2020), or the displacement of jobs with information technology (Eeckhout, Hedtrich, and Pinheiro, 2024).

In this paper, I propose a novel explanation for the disproportionate rise in income inequality and employment polarization in large cities. It emphasizes the role of the housing market and, unlike the previous literature, does not rely on features of the production technology. Since 1980 large cities have experienced faster growth in house prices, both in absolute terms and relative to wages and rents. Rapid price growth has made homeownership out of reach for more and more middle-income households. These households have increasingly chosen smaller and more affordable cities where they could buy a house. At the same time, low-income households who struggle to buy even in affordable cities and high-income households who can buy even in expensive cities were not affected as much as the middle class. This hollowed out the middle of the income distribution of big cities, making them more polarized and unequal.

Empirical evidence supports this mechanism. I document that the decline in middle-income employment shares was more pronounced in commuting zones (CZ) where prices, price-wage, and price-rent ratios rose faster between 1980 and 2019. I also find that such cities saw higher increases in wage inequality. These results do not merely pick up the fact that prices grew more in large cities and hold even after controlling for the CZ population in 1980, as well as other factors. In addition, I find that greater polarization and higher inequality growth in big cities can be attributed to out-migration of middle-income households. Using interstate migration data, I show that a doubling of house prices, price-rent, or price-wage ratios in the state of origin relative to destination raises the probability that a middle-income household migrates for housing-related reasons by 50–80% relative to the probability that a low- or high-income household makes a housing-related move.

Next, I show that a standard spatial equilibrium model with two extensions–skill heterogeneity and tenure choice–can account for the empirical evidence outlined above. In the model, households are heterogeneous in their skill level which determines their income. They choose consumption of a traded good and housing, and also decide in which city to live (*location choice*) and whether to own or rent housing (*tenure choice*). Housing is supplied by developers who either sell it to homeowners or to real estate managers, who in turn lease it to renters. Local prices and rents depend on the productivity of developers, housing supply elasticity, and the purchasing power of households. Prices are equal to the discounted sum of rents and depend on the user cost of housing. Buying a house has financial advantages but is subject to minimum-size and payment-to-income constraints. As a result, only households with sufficiently high income can own a house.

The relationship between location and tenure choices leads to a peculiar sorting pattern of households across cities. The income of low-skilled households is insufficient to buy a house even in cities with low prices. The high-skilled can afford to buy a house in all cities, even the most expensive ones. Thus, location and tenure choices of these two skill groups are independent. It is the middle-skilled households who can buy a house in a city with low prices but cannot do so in a more expensive one. As a result, location and tenure choices of the middle-skilled depend on each other. Subsequently, many middle-skilled households choose to settle in cities where house prices are lower relative to wages or rents. This empties out the middle of the income distribution in expensive cities. Thus, larger polarization and inequality in locations with high price-wage and price-rent ratios is an equilibrium outcome of the model.

To understand the role of housing markets and tenure choice in shaping the differences in polarization and inequality between large and small cities, I build a quantitative version of the model for years 1980 and 2019. The model has two locations that represent large and small CZs. In the model, large CZs experience faster growth in house prices as a result of a lower housing supply elasticity and higher labor productivity growth. The growth of labor productivity in large CZs occured primarily at the top of the skill distribution, which represents skill-biased technical change (SBTC).

Then I run two sets of counterfactual experiments. In the first one, I fix parameters that govern local returns to skills at the level of 1980, thereby shutting down SBTC, and compute a counterfactual equilibrium in 2019. In this experiment, the difference in the decline in the middle-skilled share between small and large CZs from 1980 to 2019 is 54% smaller and the gap in the increase in the variance of log wages is 73% smaller. This aligns with the results in the earlier literature that finds that SBTC can account for most of the disproportionate polarization and the rise of inequality in big cities. In the second set of counterfactuals, I allow for SBTC but keep either price-wage or price-rent ratios at their 1980 levels by allowing faster housing construction productivity in large cities and adjusting the user cost of housing. In these experiments, the excess polarization in big cities from 1980 to 2019 falls by 89–96% and the excess rise of inequality falls by 27–40%. In other words, even if SBTC is a major driver of disproportionate polarization and inequality in big cities, its effect is significantly *amplified* by declining housing affordability in these cities and would be significantly smaller if price-wage and price-rent ratios remained at their 1980 levels. The largest losers from rising price-rent and price-wage ratios in big CZs are middle-skilled workers who either lost the opportunity to own a home or moved to less productive CZs.

Why should we be concerned about polarization and inequality *within* cities? First, understanding these phenomena at the local level may help our understanding of the mechanisms that are responsible for polarization and inequality at the aggregate level. Second, the skill mix at the city level matters for its productivity. Rossi-Hansberg, Sarte, and Schwartzman (2019) and Fajgelbaum and Gaubert (2020) show that local productivity spillovers depend on the skill mix and may produce inefficient skill sorting across cities. In turn, the distribution of productivity levels across cities determines aggregate productivity (Hsieh and Moretti, 2019). Third, greater job polarization in large cities implies that some essential middle-income workers, such as teachers, may be undersupplied (Florida, 2017). Finally, local polarization and inequality may determine other important outcomes. For instance, Glaeser, Resseger, and Tobio (2009) show that more unequal cities have higher crime rates and lower self-reported happiness.

Four recent studies are most related to this paper. Each of them proposes a theoretical explanation why inequality or polarization are greater in large cities. While these explanations are based on features of the production technology, my paper provides a complementary explanation that is based on features of the housing market. First, Baum-Snow, Freedman, and Pavan (2018) build a model where local productivity depends on skill-specific agglomeration externalities. They find that the increase in the bias of agglomeration economies toward high-skilled workers (i.e., SBTC) accounts for about 80% of the disproportionate increase in wage inequality in large cities between 1980 and 2007. Second, Cerina, Dienesch, Moro, and Rendall (2023) argue that the simultaneous increase in low and high-skilled employment in a given location is driven by the complementarity of low- and high-skilled workers in job tasks, based on the evidence from Eeckhout, Pinheiro, and Schmidheiny (2014), as well as consumption complementarities that link the income of the high-skilled with the demand for services performed by the low-skilled. They find that SBTC accounts for 67% of observed disproportionate polarization in big cities. I also find that SBTC generates large differences in polarization and the rise in inequality between big and small cities, but argue that its effect would be much smaller if house prices did not grow faster in big CZs. Third, Davis, Mengus, and Michalski (2020) develop a framework that simultaneously produces higher labor market polarization and greater skill concentration in large cities in response to an external labor demand shock, without relying on SBTC. Fourth, Eeckhout, Hedtrich, and Pinheiro (2024) argue that polarization is more pronounced in large cities because local firms, facing high labor costs, have greater incentives to invest in information technology that replaces middle-skilled routine workers.

The paper is also highly related to the work on the increasing spatial dispersion of prices, rents, and price-rent ratios (Van Nieuwerburgh and Weill, 2010; Gyourko, Mayer, and Sinai, 2013; Hilber and Mense, 2021; Howard and Liebersohn, 2023). This strand of literature argues that some locations experienced faster growth in housing costs as a result of a combination of inelastic supply and extra demand, where the latter comes primarily from high-income workers. While this mechanism is present in my model, I argue that, when combined with tenure choice, it can also explain high inequality and polarization *within* large cities. At the same time, Karlman (2022) and Amaral, Dohmen, Kohl, and Schularick (2024) attribute the increase in spatial dispersion of prices to falling interest rates. This paper is also related to several other strands of literature.[1]

The paper is organized as follows: Section 2 presents empirical evidence on housing price growth, job polarization, and income inequality. Section 3 describes the theoretical framework and explains the mechanism that relates house price growth to local job polarization and income inequality. Section 4 builds a quantitative version of the model. Section 5 presents counterfactual experiments that evaluate the impact of rising price-wage and price-rent ratios on polarization and inequality in large cities. Section 6 concludes.

2 Empirical Evidence

This section documents empirical relationships between the growth in housing prices, job polarization, and income inequality. First, however, I revisit the evidence documented by the literature and confirm that larger commuting zones (CZs) experienced greater job polarization, higher increase in income inequality, and also faster housing price growth. Then I show that CZs where prices advanced faster since 1980 also saw a greater increase in polarization and inequality. Finally, I provide evidence for a migration mechanism that links housing price growth to polarization and the increase in inequality.

2.1 Data

I perform empirical analysis at the level of commuting zones (CZs) or states. To study polarization and inequality, I use the Census data from 1980, 1990, and 2000, and the 5-year American Community Survey (ACS) samples from 2006–2010 and 2015–2019. I focus

[1]It is related to the literature on: job polarization (Autor and Dorn, 2013; Goos, Manning, and Salomons, 2014) and rising inequality (Katz and Murphy, 1992; Piketty and Saez, 2003); spatial divergence in economic outcomes–termed by Moretti (2012) as the "Great Divergence"–such as productivity and wages (Moretti, 2011; Gaubert, Kline, Vergara, and Yagan, 2021; Giannone, 2022), and college shares (Costa and Kahn, 2000; Berry and Glaeser, 2005; Diamond, 2016); sorting of workers into large cities (Behrens, Duranton, and Robert-Nicoud, 2014; De la Roca and Puga, 2017); causes and consequences of rising income inequality within cities (Glaeser, Resseger, and Tobio, 2009; Baum-Snow and Pavan, 2013; Truffa, 2017; Couture, Gaubert, Handbury, and Hurst, 2020). relationship between local housing markets and the spatial distribution of labor (Ganong and Shoag, 2017; Herkenhoff, Ohanian, and Prescott, 2018; Hsieh and Moretti, 2019; Parkhomenko, 2023; Duranton and Puga, 2023); as well as the small but growing literature that incorporate housing tenure choice into location choice models (Oswald, 2019; Giannone, Li, Paixao, and Pang, 2023; Mabille, 2023; Favilukis, Mabille, and Van Nieuwerburgh, 2023; Greaney, 2023; Greaney, Parkhomenko, and Van Nieuwerburgh, 2024).

on changes between years 1980 and 2019.[2] I restrict analysis to 465 larger CZs that have sufficient number of individual observations to build measures of polarization, inequality, wages, house prices, and rents. Online Appendix Section A.1 provides more details.

To study changes in housing costs, I compute hedonic price and rent indices using the Census and the ACS data, as well as median annual wages for each CZ and year, and then construct three measures of housing costs: prices, price-wage ratios, and price-rent ratios. Online Appendix Sections A.2 and A.3 discuss the details.

To measure polarization, I follow the methodology of Autor and Dorn (2013) and assign 3-digit occupations into income percentiles in 1980. I label occupations in the 1st–20th income percentile as "low-skilled," those in the 21st–80th percentile as "middle-skilled," and those in the 81st–100th percentile as "high-skilled." Then, using a consistent definition of occupations and keeping the assignment of each occupation into skill group constant over time, I compute the shares of each group for each CZ in years 1980 and 2019.[3] Finally, I calculate the difference in shares between 1980 and 2019. Online Appendix Section A.4 provides additional details.

To measure inequality, I compute the Gini coefficient of annual wages and the variance of log annual wages for each CZ in years 1980 and 2019. More details can be found in Online Appendix Section A.5.

To study the relationship between migration and house prices, I use the 2001–2019 data from the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS). The CPS can only identify migration between states, not CZs. However, a key benefit of the CPS is that, unlike the ACS, it reports reasons for moving. This allows me to focus on households who move primarily for housing-related reasons. They constitute over 12% of interstate migrants. Online Appendix Sections A.6 and A.7 provide details.

2.2 Polarization, Inequality, and Housing Prices by CZ Size

Previous research showed that large U.S. cities experienced greater job polarization (Autor, 2019; Cerina, Dienesch, Moro, and Rendall, 2023) and higher growth in income inequality (Baum-Snow and Pavan, 2013) in recent decades. To verify these findings, I compute differences in skill shares, the Gini coefficients, and the variance of log wages between 1980 and 2019, and regress the change in skill shares and income inequality on log CZ population in 1980. The results are shown in panel A of Table 1. In column (1), we can see that larger CZs had a more pronounced decline in the middle-skilled share, i.e., they experienced greater job polarization. Doubling CZ size is associated with a 1.9 percentage

[2] As documented by previous literature, polarization and income inequality started around 1980 (Cutler and Katz, 1992; Piketty and Saez, 2003; Autor and Dorn, 2013). My analysis stops in 2019 to avoid possible effects of the Covid-19 pandemic.

[3] Hereafter, 1980 variables were constructed using the 5% sample of the 1980 Census, while 2019 variables were constructed using the 5-year combination of 1% ACS samples from 2015 to 2019.

Table 1: Relationships between city size, polarization, inequality, and house prices

Panel A: Polarization and inequality

	(1) Mid-skl. chg.	(2) Gini coeff. chg.	(3) Var. log w chg.
Log initial population	-1.899***	1.254***	2.745***
	(0.133)	(0.0765)	(0.182)
R-squared	0.298	0.363	0.405

Panel B: Housing costs

	(1) Log price chg.	(2) Log p/w chg.	(3) Log p/r chg.
Log initial population	0.0339***	0.0311***	0.101***
	(0.0114)	(0.00916)	(0.00955)
R-squared	0.0310	0.0380	0.187

The table shows the results from first-difference OLS regressions for the 1980–2019 period. In panel A, column (1) reports the coefficient from the regression of the change in 100× the middle-skilled share on log CZ population in 1980. Columns (2) and (3) report the coefficients for the change in 100× the Gini coefficient of annual wages and 100× the variance of log annual wages, respectively. In panel B, columns (1) to (3) report the coefficients for the change in log house prices, price-wage ratios, and price-rent ratios. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

point greater fall in the middle-skilled share. In columns (2) and (3), we can see that larger CZs also experienced greater increase in income inequality. Doubling CZ size is associated with a 1.3 point higher growth in 100× the Gini coefficient and a 2.7 point higher growth in 100× the variance of log wages.

To complement these results, I also study the evolution of house prices as a function of CZ size. I compute differences in log prices, price-wage ratios, as well as price-rent ratios from 1980 to 2019, and regress these differences on the CZ population in 1980. Panel B in Table 1 demonstrates that larger CZs experienced faster price growth: doubling CZ size is associated with 3.4 percent higher price growth, 3.1 percent higher price-wage ratio growth, and 10 percent greater increase in price-rent ratios. In Online Appendix Section B.1, I also look at the relationship between rent growth and city size.

In addition to examining the relationship between the growth of average CZ prices and city size, I also look at the variation across neighborhoods within CZs in Online Appendix Section B.2. Online Appendix Table B.3 shows that price-rent ratios everywhere from the 10th to the 90th percentile of the within-CZ distribution grew more in big cities. In addition, prices and price-wage ratios from the 25th to the 90th percentile grew more in large CZs. That is, large CZs not only experienced faster price growth on average, but the growth of prices was faster in nearly all neighborhoods of large CZs.

2.3 House Prices, Polarization, and Inequality

2.3.1 House Prices and Polarization

Has disproportionately faster price growth in large cities displaced middle-skilled households? To answer this question, I estimate the following relationship between the change in prices from 1980 to 2019 and the change in the shares of middle-skilled workers:

$$\Delta n_i^M = \beta_0 + \beta_1 \Delta Q_i + \beta_2 \ln N_{i,1980} + \beta_3 \mathcal{X}_{i,1980} + \varepsilon_i. \tag{1}$$

In this expression, $\Delta n_i^M \equiv n_{i,2019}^M - n_{i,1980}^M$, and $n_{i,t}^M$ is the employment share of middle-skilled workers in CZ i and year t. The change in prices is denoted by $\Delta Q_i \equiv Q_{i,2019}/Q_{i,1980} - 1$, and $Q_{i,t}$ is either the price index, the price-wage ratio, or the price-rent ratio in CZ i and year t. The population of CZ i in year t is denoted by $N_{i,t}$. Finally, $\mathcal{X}_{i,t}$ is a vector of controls that includes the 1980 share of manufacturing employment, share of female employment, share of college workers, share of foreign-born workers, and state fixed effects.[4]

OLS results. Table 2 shows the estimates. In the first three columns, I report OLS results for prices (panel A), price-rent ratios (panel B), and price-wage ratios (panel C). First, I omit initial population levels and additional controls from equation (1), and regress the change in the middle-skilled share on the price change. Column (1) shows that there is a statistically significant negative relationship between the growth in all three measures of housing prices and the middle-skilled share.[5] The coefficient values are sizable. Doubling of prices is associated with a 1 percentage point decline in the middle-skilled share, while doubling of the price-rent and price-wage ratio are associated with 11.3 and 5.3 reduction in the middle-skilled share.[6]

These findings could mask the effect of city size on polarization that has been documented in previous literature and confirmed in Table 1. Nonetheless, the results in column (2) show that, even when I control for initial CZ size, the relationship between price growth

[4]This set of controls is similar to the one used in Autor, Dorn, and Hanson (2013)'s study of employment polarization in the U.S. See Online Appendix Section A.8 for more details.

[5]Lindley and Machin (2014) document a similar relationship between polarization and housing costs, showing that U.S. states that experienced a greater increase in employment polarization between 1980 and 2010 also saw a larger change in house prices. Schubert (2021) shows that cities with high prices experienced greater displacement of non-college workers by college workers. On the other hand, Feng, Jaimovich, Rao, Terry, and Vincent (2023) find that house price growth has been slower in manufacturing-heavy U.S. regions, though the decline in manufacturing is one of the major drivers of job polarization in these and other areas.

[6]To put these results in perspective, note that between 1980 and 2019 prices changed by 348% on average (with 282% and 399% at the 25th and the 75th percentiles, respectively), price-rent ratios by −35.2% (−46.8% and −24.3%), and price-wage ratios by 31.4% (14.5% and 43.1%). The OLS coefficient values suggest that a CZ at the 75th percentile of price changes would experience a $1.17 = 1.002 \times (3.99 - 2.82)$ larger decline in the middle-skilled share than a CZ at the 25th percentile. A CZ at the 75th percentile of price-rent ratio changes would have a $2.53 = 11.26 \times (0.468 - 0.243)$ larger decline, while a CZ at the 75th percentile of price-wage ratio changes would experience a $1.53 = 5.335 \times (0.431 - 0.145)$ larger decline. For comparison, the mean change in the middle-skilled share across the 465 CZs is −2.77 percentage points, while the 25th and the 75th percentiles are −5.65 and −0.001 percentage points, respectively.

Table 2: Change in the middle-skilled share and house price growth

Panel A: Prices

	(1)	(2)	(3)	(4)	(5)	(6)
Price change	-1.002***	-0.585***	-1.006***	-2.899***	-2.387***	-4.052***
	(0.183)	(0.165)	(0.197)	(0.558)	(0.524)	(1.201)
Log initial population		-1.784***	-1.275***		-1.430***	-1.596***
		(0.135)	(0.161)		(0.208)	(0.222)
Mean of dependent variable	-2.860	-2.860	-2.860	-2.860	-2.860	-2.860
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.0700	0.320	0.595			
1st-stage F-statistic				53.66	52.61	17.98

Panel B: Price-rent ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-rent ratio change	-11.26***	-7.127***	-3.135**	-27.25***	-26.89***	-47.84**
	(1.110)	(1.181)	(1.426)	(5.375)	(6.862)	(22.45)
Log initial population		-1.424***	-1.077***		-0.107	0.232
		(0.141)	(0.177)		(0.520)	(0.745)
Mean of dependent variable	-2.860	-2.860	-2.860	-2.860	-2.860	-2.860
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.240	0.375	0.574			
1st-stage F-statistic				22.74	15.53	4.319

Panel C: Price-wage ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-wage ratio change	-5.335***	-3.832***	-4.458***	-9.688***	-7.830***	-14.06***
	(0.656)	(0.652)	(0.845)	(1.656)	(1.486)	(3.974)
Log initial population		-1.729***	-1.201***		-1.552***	-1.271***
		(0.134)	(0.160)		(0.173)	(0.176)
Mean of dependent variable	-2.860	-2.860	-2.860	-2.860	-2.860	-2.860
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.120	0.357	0.600			
1st-stage F-statistic				73.42	68.89	21.85

Notes: The table shows the results from first-difference regressions for the 1980–2019 period. Panel A shows results for the house price index, panel B shows results for price-rent ratios, and panel C shows results for price-wage ratios. Column (1) reports the coefficients from the OLS regression of the change in 100× the middle-skilled share on the change in prices. Column (2) includes initial CZ population as a control. Column (3) adds manufacturing share, female share, college share, foreign-born share, and state dummy as additional controls. Columns (4)–(6) report the results from 2SLS estimation. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

and the change in the middle-skilled share remains negative, albeit becomes smaller in magnitude. That is, greater polarization is not just a feature of big cities but also a feature of expensive cities. Surely, large cities experienced faster price growth. But regardless of city size, the places where prices increased more have seen stronger polarization. As shown in column (3), additionally controlling for manufacturing share, female share, college share, foreign-born share, and state effects does not change the relationship between price growth and the decline in the middle-skilled share.

Instrumental variables. OLS estimates of β_1 are almost certainly biased due to the omitted variable bias and reverse causality, as price changes may depend on changes in skill shares.[7] Thus, I instrument for changes in prices, price-rent, and price-wage ratios using a variable that represents local geographic features and long-run changes in the interest rate. Previous work showed that locations with less land available for construction experience faster price growth in recent decades (Lutz and Sand, 2023) and that the decline of interest rates has amplified price growth differences across locations (Karlman, 2022; Amaral, Dohmen, Kohl, and Schularick, 2024).[8] Therefore, I construct a variable $Z_t = U^{-r_t}$ where U is the share of land unavailable for construction for each CZ from Lutz and Sand (2023) and r_t is the real interest rate in year t from the FRED database, equal to 6.93% in 1982 and 0.53% in 2019.[9,10] The functional form of Z_t implies that the impact of the differences in land use availability is amplified when interest rates are lower. Then, as the instrumental variable I use the ratio of Z_t in 2019 to its value in 1980, i.e., $U^{-(r_{2019}-r_{1980})}$.

The value of the instrument is solely determined by plausibly exogenous first-nature geography and national-level interest rates which are unlikely to be substantially affected by economic conditions in any given CZ. Housing supply is more constrained in CZs where land available for construction is scarce and housing demand shocks in such CZs will result in larger price growth. When interest rates are low, the cost of borrowing is lower and price growth differences across CZs are larger. While geographic constraints may affect local job polarization and income distribution, they are most likely to do so via housing and land

[7]For instance, a large share of high earners in a given city may drive up house prices via housing investment demand or by supporting restrictive land use regulations.

[8]The explanation that rising differences in prices are due to falling interest rates is consistent with the role of declining interest rates in the rise in wealth inequality (Greenwald, Leombroni, Lustig, and Van Nieuwerburgh, 2021).

[9]While a more widely used measure of land unavailability is from Saiz (2010), there are two benefits of using the Lutz and Sand (2023)'s measure. First, it provides data at the county level for nearly all U.S. counties, which I aggregate to the CZ level, whereas Saiz (2010) provides data at the metropolitan area level. Second, Lutz and Sand (2023) show that their measure of land unavailability has a stronger first-stage predictive power for house price growth than Saiz (2010)'s measure and that it is also not positively correlated with several proxies of local housing demand. Meanwhile, Saiz (2010)'s measure was criticized by Davidoff (2016) due to its potential correlation with housing demand proxies, perhaps invalidating it as an instrument for house prices.

[10]I use the variable https://fred.stlouisfed.org/series/REAINTRATREARAT10Y#0. Since the series start in 1982, I use the 1982 value for 1980.

Table 3: First-stage regressions

	(1) Price chg.	(2) P/r chg.	(3) P/w chg.
Land and interest rate IV	4.087***	0.435***	1.223***
	(0.558)	(0.0912)	(0.143)
R-squared	0.132	0.0550	0.195

The table shows the results from first-difference regressions for the 1980–2019 period. Columns (1), (2), and (3) report the coefficients from regressions of the change in housing prices, price-rent indices, and price-wage indices, respectively, on the instrumental variable that interacts the Lutz and Sand (2023) land unavailability index with the change in real interest rate. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

prices and not other channels.[11] Table 3 presents the results of first-stage regressions of the instrumental variable on changes in the price index, price-rent ratio, and price-wage ratio. They demonstrate that the interaction between less available land and lower interest rates is associated with larger growth in housing prices.

IV results. In columns (4)–(6) of Table 2, I present the results of the two-stage least squares (2SLS) estimation. The coefficients on changes in prices, price-rent ratios, and price-wage ratios remain negative and significant, although the F-statistic falls below the conventional threshold of 10 in the regression with price-rent ratios. This suggests that polarization and price growth may not be simply negatively related, but that stronger price growth may lead to more polarization, regardless of city size and other city characteristics.

Note that in all specifications, the 2SLS coefficients are larger than the OLS coefficients. There are at least two possible reasons. First, there may be important variables that are correlated with the change in the middle-skilled share and that were omitted from OLS regressions. Second, there may be strong reverse causality. In particular, if larger reductions in the fraction of middle-skilled workers lead to lower price growth due to lower demand for housing, OLS coefficients could be attenuated.

In Online Appendix Section B.3, I show that the relationship between price growth and the decline in the middle-skilled share is robust to using different thresholds to split employment into low-, middle-, and high-skilled groups, and using two separate time intervals: 1980–2000 and 2000–2019.

2.3.2 House Prices and Income Inequality

Job polarization is mechanically related to changes in income inequality, as a larger number of individuals in the tails of the income distribution may lead to a greater dispersion of

[11] The instrumental variable is not highly correlated with the change in the middle-skilled share: the correlation is 0.135. It is also not highly correlated with the change in the variance of log wages studied below: the correlation is 0.19.

income. Hence, if cities with faster growth in housing prices experienced more pronounced polarization, we should expect that they also had a larger increase in income inequality. To study this hypothesis, I estimate the following relationship between changes in prices and income inequality:

$$\Delta I_i = \beta_0 + \beta_1 \Delta Q_i + \beta_2 \ln N_{i,1980} + \beta_3 \mathcal{X}_{i,1980} + \varepsilon_i, \tag{2}$$

where ΔI_i is the difference in the variance of log annual wages between 1980 and 2019, and other variables are the same as in equation (1).

OLS results. Column (1) in Table 4 shows that there is a statistically significant positive relationship between the growth in housing costs and the increase in the variance of log wages.[12] Doubling of prices is associated with a 2.3 point increase in 100× the variance of log wages, whereas doubling of the price-rent and price-wage ratio are associated with 11.7 and 7.7 point increase.[13] The magnitudes of these relationships are somewhat lower when I control for initial CZ size (column 2) and include additional regressors (column 3), but they remain sizable and statistically significant. As in the case of polarization, these results mean that rising prices is a separate channel that is related to larger increases in inequality, regardless of city size.

IV results. In columns (4)–(6) of Table 4, I estimate the relationship between prices and inequality using the 2SLS estimation and the same instrument for price growth as in the case of polarization. The coefficients on price growth remain positive, although statistical significance weakens as I include additional controls in column (6). Nonetheless, these findings suggest that there may be a causal relationship between price growth, at least for prices and price-wage ratios, and changes in income inequality at the CZ level, regardless of the city size and other city characteristics. As in the case of polarization, IV coefficients are somewhat larger than OLS coefficients, likely for the same reasons as outlined in Section 2.3.1.

Online Appendix Section B.4 shows that these findings are robust to measuring inequality using the Gini coefficient, using hourly income, and using two separate time intervals: 1980–2000 and 2000–2019.

[12] These findings are related to Moretti (2013) who shows that metropolitan statistical areas (MSAs) where the share of college graduates grew the most in 1980–2000 also had faster increases in the college wage gap.

[13] To put these results in perspective, note that between 1980 and 2019 prices changed by 348% on average (with 282% and 399% at the 25th and the 75th percentiles, respectively), price-rent ratios by −35.2% (−46.8% and −24.3%), and price-wage ratios by 31.4% (14.5% and 43.1%). The OLS coefficient values suggest that a CZ at the 75th percentile of price changes would experience a $2.68 = 2.29 \times (3.99 - 2.82)$ larger increase in 100× the variance of log wages than a CZ at the 25th percentile. A CZ at the 75th percentile of price-rent ratio changes would have a $2.64 = 11.72 \times (0.468 - 0.243)$ larger increase, while a CZ at the 75th percentile of price-wage ratio changes would experience a $2.21 = 7.735 \times (0.431 - 0.145)$ larger increase. For comparison, the mean growth of 100× the variance of log wages across the 465 CZs is 9.93 points, while the 25th and the 75th percentiles are 6.41 and 12.59 points, respectively.

Table 4: Change in income inequality and house price growth

Panel A: Prices

	(1)	(2)	(3)	(4)	(5)	(6)
Price change	2.290***	1.728***	1.709***	2.350***	1.470***	2.677**
	(0.269)	(0.204)	(0.230)	(0.524)	(0.450)	(1.090)
Log initial population		2.405***	1.711***		2.456***	1.813***
		(0.146)	(0.203)		(0.166)	(0.229)
Mean of dependent variable	10.02	10.02	10.02	10.02	10.02	10.02
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.237	0.534	0.666			
1st-stage F-statistic				53.66	52.61	17.98

Panel B: Price-rent ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-rent ratio change	11.72***	4.641***	5.816***	22.09***	16.56***	31.61*
	(1.641)	(1.360)	(1.637)	(5.427)	(5.963)	(17.30)
Log initial population		2.436***	1.361***		1.642***	0.605
		(0.169)	(0.225)		(0.425)	(0.576)
Mean of dependent variable	10.02	10.02	10.02	10.02	10.02	10.02
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.169	0.426	0.628			
1st-stage F-statistic				22.74	15.53	4.319

Panel C: Price-wage ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-wage ratio change	7.735***	5.563***	4.314***	7.853***	4.822***	9.292**
	(1.241)	(0.864)	(1.113)	(1.867)	(1.535)	(4.027)
Log initial population		2.499***	1.562***		2.531***	1.599***
		(0.154)	(0.219)		(0.164)	(0.220)
Mean of dependent variable	10.02	10.02	10.02	10.02	10.02	10.02
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.164	0.487	0.634			
1st-stage F-statistic				73.42	68.89	21.85

Notes: The table shows the results from first-difference regressions for the 1980–2019 period. Panel A shows results for the house price index, panel B shows results for price-rent ratios, and panel C shows results for price-wage ratios. Column (1) reports the results from the OLS regression of the change in 100× the variance of log annual wages on the change in prices. Column (2) includes initial CZ population as a control. Column (3) adds manufacturing share, female share, college share, foreign-born share, and state dummy as additional controls. Columns (4)–(6) report the results from 2SLS estimation. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

2.3.3 Discussion: OLS vs. IV Results

The evidence from 2SLS regressions suggests that there may be a causal link from price growth to polarization and the growth in inequality. However, the evidence from OLS regressions is no less important. It underscores the relationship between price growth, polarization, and inequality that is independent from city size. In the model presented below in Section 3, the relationship runs both ways. On the one hand, higher prices push out the middle-skilled and result in more polarization and inequality. On the other hand, changes in local skill distribution and especially larger share of high-skilled workers in more polarized locations affects prices via housing demand. Finally, note that the purpose of this section is to examine the relationship between price growth, polarization, and inequality; however, none of the estimates are used in the quantitative model or counterfactual analysis in Sections 4 and 5.

2.4 Interstate Migration

A natural mechanism that may link polarization and changes in inequality to changes in house prices is migration. If households in the middle of the income distribution are more sensitive to differences in house prices, they will be more likely to migrate out of expensive locations and these locations should experience greater polarization and inequality.

To study the effect of being in a certain position in the income distribution on the probability of moving for housing-related reasons, I turn to the CPS migration data. I split households into five quintiles by income in the state of origin and estimate a logit regression $\mathbb{M}_{h,ij,t+1} = \Lambda(\mathbf{X}'\boldsymbol{\delta})$.[14] Indicator variable $\mathbb{M}_{h,ij,t+1}$ is equal to 1 if household h moved for housing-related reasons from state i to state $j \neq i$ between years t and $t+1$, and

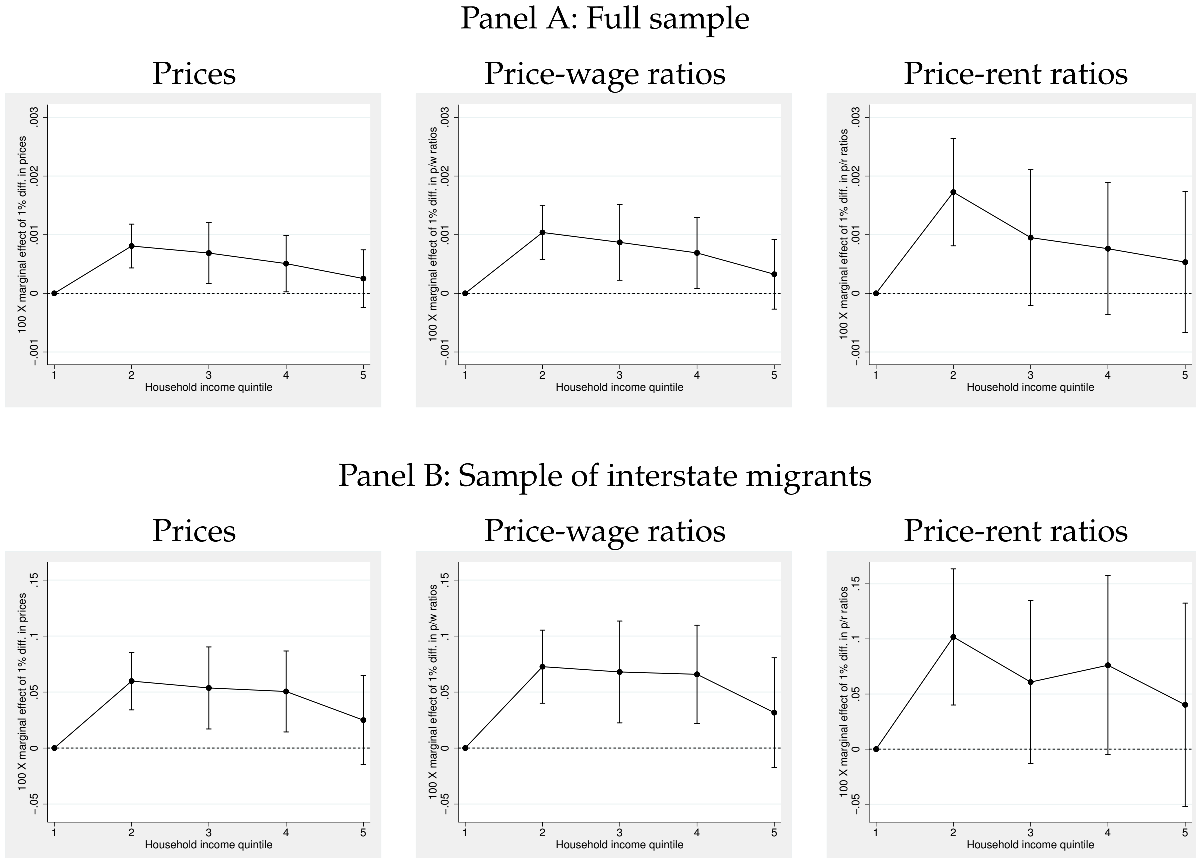
$$\mathbf{X}'\boldsymbol{\delta} = \sum_{q=1}^{5} \delta_1^q \mathbb{I}_{h,it}^q + \sum_{q=1}^{5} \delta_2^q \ln\left(\frac{Q_{it}}{Q_{jt}}\right) \times \mathbb{I}_{h,it}^q + \delta_3 \ln\left(1 + D_{ij}\right) + \boldsymbol{\delta}_4 \mathcal{X}_{h,ij,t+1} + \varphi_i + \varphi_j + \varphi_{t+1} + \varepsilon_{h,ij,t+1}. \quad (3)$$

In this specification, $\mathbb{I}_{h,it}^q$ indicates whether household h belonged to the q-th quintile by household income in state i and year t; $Q_{i,t}$ is a measure of housing prices, either the house price index, the price-wage ratio, or the price-rent ratio; D_{ij} is the great-circle distance between population-weighted centroids of states i and j; $\mathcal{X}$ is a set of controls which include information about gender, race, household composition, education, and age; and parameters φ_i, φ_j, and φ_t are origin, destination, and year fixed effects, respectively. Since specification (3) includes income and housing cost variables in the year prior to the observed new state of residence and because none of the control variables are likely to be affected by the decision to migrate, all right-hand side variables are plausibly exogenous.[15]

[14] Λ denotes the logistic cdf, i.e., $\Lambda(\mathbf{X}'\boldsymbol{\delta}) = \exp(\mathbf{X}'\boldsymbol{\delta})/(1 + \exp(\mathbf{X}'\boldsymbol{\delta}))$.

[15] The CPS respondents report their income and state of residence in the previous year and also the state

Figure 1: Marginal effects on migration for housing-related reasons



Note: The figure reports marginal effects from coefficients δ_2^q from regression (3) on the full sample of 1,265,832 observations (panel A) or the sample of 15,815 interstate migrants (panel B). The left plot shows 100× marginal effects on the probability of moving for housing-related reasons of the log ratio of house prices in the location of origin to the prices in the location of destination for each quintile in the household income distribution at the location of origin. The center plot shows the marginal effects of price-wage ratios, and the right plot shows the marginal effects of price-rent ratios. Vertical bars represent the 95% confidence interval. Standard errors of marginal effects are computed using the Delta method. Standard errors of the underlying logit regression are clustered by the state of origin.

The main coefficients of interest are those on the interaction variable between the income quintile and the ratio of housing costs, δ_2^q. They measure how much more likely a household is to move for housing-related reasons from a less affordable to a more affordable U.S. state depending on their position in the income distribution.

Results. Panel A of Figure 1 plots 100× marginal effects associated with δ_2^q for each quintile q relative to the first quintile. The marginal effects represent the percentage point change in the probability of moving for housing-related reasons in response to a 1% increase in prices, price-wage, or price-rent ratios in state i relative to state j. The figure shows that, compared to households in the 1st or the 5th income quintile, those in the 2nd to the 4th

of residence in the current year.

quintiles are more likely to move for housing-related reasons from state i to j when state j has lower prices, price-wage, or price-rent ratios. In case of price-rent ratios, the marginal effects have larger standard errors but are quantitatively similar to those for prices and price-wage ratios.[16]

Panel B of Figure 1 focuses on the subsample of interstate migrants. It shows that, even conditional on moving across states, middle-income households are more likely to move for housing-related reasons than low- or high-income households. A doubling of the price index in state i relative to j increases the probability that a household in the 2nd to the 4th quintile moves to state j for housing-related reasons by 5–6 percentage points. Similarly, a doubling of the price-wage ratio raises the moving probability by 6–7 percentage points for households in the middle of the income distribution, whereas a doubling of the price-rent ratio increases it by 7–10 percentage points. Given that the probability of moving for housing-related reasons is 0.12 (the fraction of migrants who moved for housing reasons), these marginal effects imply that a twofold increase in prices and price-wage ratios may increase the probability of a housing-related relocation for middle-income households by about 50%, while a similar increase in the price-rent ratio may increase it by 60% to 80%.[17]

In Online Appendix Section B.5, I show that this relationship is unique to housing-related migration and does not hold for non-housing reasons.

2.5 Homeownership Rates

One may conjecture that local changes in prices should be correlated with changes in the homeownership rate. It is not clear, however, whether any such relationship should exist. On the one hand, rising prices may deter some households from buying a house. On the other hand, prices may go up precisely because many households are buying. Moreover, local homeownership rates are affected by compositional changes if those who cannot buy a house move elsewhere, as shown in Section 2.4. Indeed, Online Appendix Table B.2 shows that there is no statistically significant relationship between the change in homeownership and the growth in prices, price-rent, and price-wage ratios from 1980 to 2019.

3 Theoretical Framework

In this section, I construct a parsimonious spatial equilibrium model consistent with the evidence presented in the previous section. The model builds on the standard system-of-

[16] Noise is expected since migration is measured at the state level. In some states, such as New York and California, the differences in housing costs between CZs are large and comparable to cross-state differences.

[17] Since regression (3) includes state and year fixed effects, the results should be interpreted as those that arise from doubling of prices *relative* to the predicted value for the same state and year. While such doubling in just one year is infeasible, repeated migration may compound to create sizable changes in the composition of population from 1980 to 2019, the period studied in the counterfactual experiments in Section 5.

cities model (Henderson, 1974; Rosen, 1979; Roback, 1982) and only makes two extensions to it: skill heterogeneity and housing tenure choice.

The economy consists of $\mathcal{I}$ cities, indexed by i, and is populated by a measure 1 of households that live for T years, as well as infinitely-lived real estate managers.[18] Each year a measure $1/T$ of workers exit the economy and the same number of workers enter the economy. The economy evolves over time as a sequence of spatial equilibria. I restrict my analysis to stationary equilibria in which all equilibrium variables are time-invariant.

3.1 Preferences and Choices

Newborn households draw skill s from the distribution $\Phi(s)$ that has full support on $(0, 1]$ and density function $\phi(s)$. The skill level is fixed until the end of their life and each period households supply labor inelastically to the firms that produce the final consumption good. Workers live hand-to-mouth, i.e., they cannot save or borrow.

Households consume positive quantities of the final good, c, and housing, h. The demand for housing can be satisfied by either renting or owning a dwelling. Households also derive utility from the amenities of the city they live in, X_i. In addition, each household n has an idiosyncratic preference for city i, denoted by ξ_{ni}. Location preferences are drawn at the beginning of life from the Fréchet distribution $F(\xi) = \exp\left(-\xi^{-\epsilon}\right)$ with $\epsilon > 1$, and remain fixed throughout the life cycle. The utility of worker n who lives in city i is given by $u_{ni}(c, h) = \xi_{ni} v_i(c, h)$, where

$$v_i(c, h) \equiv \left(\frac{c}{1-\gamma}\right)^{1-\gamma} \left(\frac{h}{\gamma}\right)^{\gamma} X_i \tag{4}$$

is the common component of the utility function and $0 < \gamma < 1$.

Households make three choices: location, housing tenure, and consumption of goods and housing. Location and tenure choices are made at the beginning of life and households cannot change their decision afterwards. Because the economy is assumed to be in a stationary equilibrium, because workers receive all shocks (skill and location preference) once, because they make location and tenure choices at the beginning of life, and because they live hand-to-mouth, their problem is essentially static. Therefore, unless noted otherwise, all flow variables and parameters are meant to represent present discounted values over T years and are written without a time subscript, whereas variables that can change at annual frequency have a t-subscript.

3.1.1 Consumption of Goods and Housing

I now solve the household utility maximization problem, conditional on the tenure choice.

[18] I use the terms "household" and "worker" interchangeably.

Renters. A household that chooses to rent solves

$$\max_{c,h} v_i(c,h) \quad \text{subject to: } w_i(s) = c + r_i h, \tag{5}$$

where $w_i(s)$ is the wage of a worker with skill s in city i and r_i is the rent per square foot in city i. The housing demand function is $h = \gamma w_i(s)/r_i$ and the indirect utility function of a renter is

$$v_i^R(w_i(s), r_i) = \frac{w_i(s)X_i}{r_i^{\gamma}}. \tag{6}$$

Homeowners. A household that chooses to own buys a house at the beginning of their life at price p_i and sells it at the end of their life. Since the equilibrium is stationary, they sell at the same price p_i. Owning a house has a city-specific user cost δ_i that represents depreciation, property taxes, mortgage interest payments, etc. Thus, the budget constraint of a homeowner is $w_i(s) + p_i h = c + (1+\delta_i)p_i h$ which can be simplified as $w_i(s) = c + \delta_i p_i h$. Besides the budget constraint, homeowners are subject to two additional constraints. First, owner-occupied houses cannot be smaller than $\bar{h} > 0$. This is the *minimum-size constraint*. Second, a household cannot spend more than a fraction $\lambda < 1$ of their labor earnings on purchasing a house. This is the *payment-to-income (PTI) constraint*. Thus, owners solve

$$\max_{c,h} v(c,h) \quad \text{subject to: } w_i(s) = c + \delta_i p_i h, \tag{7}$$

$$h \geq \bar{h} \ \ \text{(minimum-size constraint)}, \tag{8}$$

$$p_i h \leq \lambda w_i(s) \ \ \text{(PTI constraint)}. \tag{9}$$

The solution to the homeowner's problem yields the following housing consumption function:

$$h = \begin{cases} \frac{\gamma}{\delta_i}\frac{w_i(s)}{p_i} & \text{if } \lambda > \frac{\gamma}{\delta_i} \text{ and } w_i(s) \geq \frac{\delta_i p_i \bar{h}}{\gamma}, \\ \bar{h} & \text{if } \lambda > \frac{\gamma}{\delta_i} \text{ and } \delta_i p_i \bar{h} < w_i(s) < \frac{\delta_i p_i \bar{h}}{\gamma}, \\ \lambda \frac{w_i(s)}{p_i} & \text{if } \lambda \leq \frac{\gamma}{\delta_i} \text{ and } w_i(s) \geq \frac{p_i \bar{h}}{\lambda}, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

This function describes four cases. First, a homeowner's wage may be high enough to spend just a fraction γ/δ_i of income on housing and be able to afford a house at least as large as $\bar{h}$. Second, the wage may be relatively low and buying even the minimum-size house requires spending a fraction of income greater than γ/δ_i. These two cases occur when the PTI constraint is sufficiently lax, $\lambda > \gamma/\delta_i$. Third, if $\lambda \leq \gamma/\delta_i$, households spend a fraction λ of income on housing.[19] In the fourth case, the wage is so low that even if

[19]Note that households cannot spend more than a fraction λ of their income on housing. The only household who buys a house of size $\bar{h}$ is the one whose income is equal to $p_i\bar{h}/\lambda$, and this situation is already captured by the third case.

the household spends all of their income on housing, they would not be able to afford the minimum-size property. Their housing consumption is zero, which violates the positive consumption requirement. These four cases result in the following indirect utility function:

$$v_i^O\left(w_i(s), p_i\right) = \begin{cases} \frac{w_i(s)}{\delta_i p_i^{\gamma}} \left(\frac{\delta_i - \gamma}{1-\gamma}\right)^{1-\gamma} X_i & \text{if } \lambda > \frac{\gamma}{\delta_i} \text{ and } w_i(s) \geq \frac{\delta_i p_i \bar{h}}{\gamma}, \\ \left(\frac{w_i(s) - \delta_i p_i \bar{h}}{1-\gamma}\right)^{1-\gamma} \left(\frac{\bar{h}}{\gamma}\right)^{\gamma} X_i & \text{if } \lambda > \frac{\gamma}{\delta_i} \text{ and } \delta_i p_i \bar{h} < w_i(s) < \frac{\delta_i p_i \bar{h}}{\gamma}, \\ \frac{w_i(s)}{p_i^{\gamma}} \left(\frac{1-\delta_i \lambda}{1-\gamma}\right)^{1-\gamma} \left(\frac{\lambda}{\gamma}\right)^{\gamma} X_i & \text{if } \lambda \leq \frac{\gamma}{\delta_i} \text{ and } w_i(s) \geq \frac{p_i \bar{h}}{\lambda}, \\ -\infty & \text{otherwise.} \end{cases} \tag{11}$$

3.1.2 Housing Tenure Choice

Indirect utility functions (6) and (11) result in different utility levels for the same wage, depending on the housing tenure choice. A worker with skill s in city i chooses to rent if it yields higher utility and to own otherwise. Hence, the indirect utility function of such a worker is

$$v_i\left(w_i(s), p_i, r_i\right) \equiv \max\left\{v_i^O\left(w_i(s), p_i\right), v_i^R\left(w_i(s), r_i\right)\right\}. \tag{12}$$

The next two results characterize tenure choice. Lemma 1 provides necessary and sufficient conditions for homeownership to be an optimal choice for some households. If any of these conditions does not hold, all city residents will rent. Lemma 2 describes who owns and who rents.

Lemma 1 (necessary and sufficient conditions for homeownership). Homeownership with housing consumption $h > \bar{h}$ is the optimal tenure choice for some households in city i if and only if the minimal house size is sufficiently small,

$$\bar{h} < \min\left\{\frac{\gamma}{\delta_i}, \lambda\right\} \frac{w_i(s=1)}{p_i}, \tag{13}$$

and the price-rent ratio is sufficiently low,

$$\frac{p_i}{r_i} \leq \begin{cases} \left(\frac{\delta_i - \gamma}{1-\gamma}\right)^{\frac{1-\gamma}{\gamma}} \frac{1}{\delta_i^{1/\gamma}} & \text{if } \lambda > \frac{\gamma}{\delta_i}, \\ \left(\frac{1-\delta_i \lambda}{1-\gamma}\right)^{\frac{1-\gamma}{\gamma}} \frac{\lambda}{\gamma} & \text{if } \lambda \leq \frac{\gamma}{\delta_i}. \end{cases} \tag{14}$$

Proof. See Online Appendix Section C.1 □

Lemma 2 (skill thresholds for homeownership). Let the conditions (13) and (14) hold. Then:

(a) there exists a unique skill threshold $s_i^* \in (0, 1)$ such that all workers in city i with skill $s < s_i^*$ choose to rent, and all those with skill $s \geq s_i^*$ choose to buy;

(b) there exists a unique skill threshold $s_i^{**} \in [s_i^*, 1)$ such that all workers with skill $s > s_i^{**}$ choose to buy a house larger than the minimum size $\bar{h}$;

(c) conditional on prices p_i and rents r_i, skill thresholds s_i^* and s_i^{**} are decreasing in local labor productivity, A_i.[20]

Proof. See Online Appendix Section C.2 □

The previous two results stress the fact that the decision to own a home depends on the price-rent ratio (Lemma 1) and the price-wage ratio (Lemma 2). The lower these ratios are, the more households will choose to buy instead of renting.

Lemma 2 also indicates that there are two types of homeowners. First, there are owners whose income is large enough that they can afford to purchase a house larger than $\bar{h}$ without spending more than a fraction γ/δ_i of their income. I call them "full owners." Second, there are owners whose income is relatively low and they must spend a fraction greater than γ/δ_i on buying a house. Nonetheless, their utility of owning exceeds the utility of renting and they find it optimal to buy a house of size $\bar{h}$. I dub these homeowners as "marginal owners." Note that marginal owners only exist if the PTI constraint is not binding. If it is binding, full owners are already spending the maximum allowed share of income on housing and no one can spend a share higher than λ in order to buy a house of minimum size. In this case, the two ownership thresholds coincide, i.e., $s_i^* = s_i^{**}$.

Figure 2 demonstrates an example of tenure choice and housing consumption when the PTI constraint does not bind. In this example, the value of renting exceeds the value of owning for all skill levels below s^*. At this skill level and above, the value of a owning is higher but households with skills below s^{**} can only afford to buy a minimum-size house. Housing consumption jumps at s^* and remains constant up to s^{**}. Starting from s^{**}, it is optimal to buy a house larger than $\bar{h}$. Thus, workers with skill $s \in (0, s^*)$ are renters, those with $s \in [s^*, s^{**})$ are marginal owners, and workers with $s \in [s^{**}, 1]$ are full owners.

The reason why households prefer to own is purely financial: ownership reduces the cost of housing consumption. Given that workers make location and tenure decisions in the beginning of the life cycle, the model is essentially static and many other reasons why households prefer to own (e.g., wealth accumulation or risk insurance) are absent.

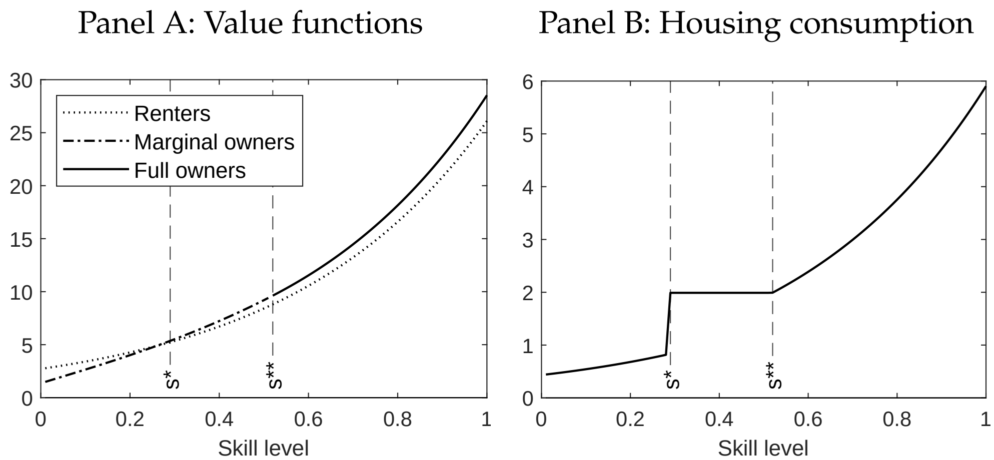
3.1.3 Location Choice

Each worker n chooses city i that maximizes her utility. The location choice problem is characterized by

$$\max_i \{\xi_{ni} v_i (w_i(s), p_i, r_i)\}. \tag{15}$$

[20]Note that s_i^* and s_i^{**} are endogenous variables: s_i^* is a function of p_i and r_i, and s_i^{**} is a function of p_i. They are implicitly defined by equations (37) and (38) in Online Section C.2.

Figure 2: Tenure choice and housing consumption



Notes: Panel A shows utility levels for each skill level implied by value functions of renters, marginal owners, and full owners. Note that full ownership is not a feasible choice for $s < s^{**}$; hence, the value of a full owner is only shown for $s \geq s^{**}$. Panel B shows housing consumption for each skill level. Both panels display skill thresholds for marginal ownership (s^{*}) and full ownership (s^{**}).

Since the location preference shocks follow the Fréchet distribution with shape parameter ϵ, the probability that a worker with skill s chooses to live in city i is given by

$$\pi_i(s) = \frac{v_i\left(w_i(s), p_i, r_i\right)^{\epsilon}}{\sum_{j\in\mathcal{I}} v_j\left(w_j(s), p_j, r_j\right)^{\epsilon}}. \tag{16}$$

Thus, a worker is more likely to choose a city that offers high wages, has low prices and rents, and has high amenities. The equilibrium supply of workers with skill s in city i is equal to

$$N_i(s) = \pi_i(s)\phi(s), \tag{17}$$

and the total employment in the city is $N_i \equiv \int_0^1 N_i(s)\mathrm{d}s$.

3.1.4 Relationship Between Location and Tenure Choices

Consider two cities, i and j, and suppose that the skill threshold for homeownership is higher in city i, i.e., $s_i^* > s_j^*$. Then, when choosing city, households with $s \in [s_j^*, s_i^*)$ compare not only wages, housing costs, and amenities, but they also compare between owning a house in city j and renting in city i. Their relative probability of choosing city i incorporates the comparison of the two tenure choices and is given by

$$\frac{\pi_i(s)}{\pi_j(s)} = \left[\frac{v_i^R\left(w_i(s), r_i\right)}{v_j^O\left(w_j(s), p_j\right)}\right]^{\epsilon} \quad \text{for } s \in [s_j^*, s_i^*). \tag{18}$$

Thus, for households whose skills are high enough to buy in city j but insufficient to buy in city i, location choice and tenure choice depend on each other. At the same time, for households with skills below s_j^* or above s_i^*, location and tenure choices are independent.

3.1.5 Welfare

The location choice probabilities lead to an expression for welfare. The expected utility of an s-skilled worker prior to making any choices and knowing the value of the location preference shock is[21]

$$V(s) = \left[\sum_{i \in \mathcal{I}} \left(v_i \left(w_i(s), p_i, r_i \right) \right)^{\epsilon} \right]^{\frac{1}{\epsilon}}. \tag{19}$$

Aggregate welfare is defined as the weighted average of expected utilities for workers of all skill levels,

$$V = \int_0^1 V(s) \mathrm{d}\Phi(s). \tag{20}$$

3.2 Final Goods Production

In each city, there is a representative firm that produces the final numeraire good that is traded across cities at no cost. The production function combines workers of different skills as perfect substitutes:

$$Y_i = A_i \int_0^1 a(s) N_i(s) \mathrm{d}s. \tag{21}$$

Here, A_i represents labor productivity of city i, $a_i(s)$ is a city-specific strictly increasing continuous function that determines the productivity of a worker with skill s in city i, while $N_i(s)$ is the employment of s-skilled workers in the city. Productivity A_i depends on total local employment N_i via agglomeration externalities:

$$A_i = \bar{A}_i N_i^{\theta}, \tag{22}$$

where $\bar{A}_i$ is the exogenous part of productivity. Perfect competition implies that the wage of an s-skilled worker is

$$w_i(s) = A_i a_i(s). \tag{23}$$

[21]The expressions for the location choice probability (16) and expected welfare (19) arise from the properties of the Fréchet distribution and are standard in the literature. Detailed derivations can be found, for example, in the appendix to Monte, Redding, and Rossi-Hansberg (2018).

3.3 Housing Markets

Housing is produced by perfectly competitive developers with technology that uses land L_i and the numeraire good K_i,

$$H_i = \left(\phi_i L_i\right)^{\eta_i} K_i^{1-\eta_i}. \tag{24}$$

Land input is augmented by productivity ϕ_i and η_i is the share of land in construction. Each city is endowed with an exogenous quantity of land Λ_i owned by absentee landowners. There is no alternative use of land; hence, landowners are willing to sell land at any positive price and developers optimally buy all available land, $L_i = \Lambda_i$.

A part of the housing stock is owned by homeowners, while the remainder is owned by a large number of infinitely lived real estate managers and leased to renters. Every year, the managers earn $r_{i,t}$ from leasing housing to renters and incur user costs $\tilde{\delta}_i/T$. I assume that housing stock does not depreciate thanks to maintenance expenditures that are included in δ_i and $\tilde{\delta}_i$.[22] At the end of each year, homeowners at the end of their life cycle sell their houses to the real estate managers at price $p_{i,t}$. At the turn of two years, the managers own this housing stock. Then, at the start of year $t+1$, newborn homeowners buy a part of the housing stock from the managers at price $p_{i,t+1}$, which is equal to $p_{i,t}$ given that the economy is in a stationary equilibrium. In each period, real estate managers earn rental income and spend it on the consumption of the final good.[23]

Real estate managers discount future with factor $\beta < 1$. In equilibrium, rents and prices adjust so that each manager is indifferent between selling a property or keeping it and earning rents perpetually. This no-arbitrage condition implies that house prices are equal to the expected discounted sum of rents and that the price-rent ratio is[24]

$$\frac{p_{i,t}}{r_{i,t}} = \sum_{\tau=t}^{\infty} \left[\beta(1-\tilde{\delta}_i/T)\right]^{\tau-t} = \frac{1}{1-\beta(1-\tilde{\delta}_i/T)}. \tag{25}$$

The previous expression implies that the price-rent ratio is higher when the user cost $\tilde{\delta}_i$ is low. The model is silent regarding why $\tilde{\delta}_i$ can vary across cities or over time, and important sources of variation may include interest or depreciation rates.

Equilibrium house price must clear the market by satisfying

$$\phi_i \Lambda_i \left(1-\eta_i\right)^{\frac{1-\eta_i}{\eta_i}} p_i^{\frac{1-\eta_i}{\eta_i}} = \frac{\gamma}{r_i} \int_0^{s_i^*} w_i(s) N_i(s) \mathrm{d}s + \bar{h} \int_{s_i^*}^{s_i^{**}} N_i(s) \mathrm{d}s + \min\left\{\frac{\gamma}{\delta_i}, \lambda\right\} \frac{1}{p_i} \int_{s_i^{**}}^{1} w_i(s) N_i(s) \mathrm{d}s, \tag{26}$$

[22]The user cost of real estate managers $\tilde{\delta}_i$ may differ from the user cost incurred by homeowners δ_i. For example, institutional real estate investors do not benefit from mortgage tax deductions but may save on other costs thanks to scale economies.

[23]In Section 5.3, I consider an alternative model in which rental income is redistributed to homeowners.

[24]Because the economy is assumed to be in a stationary equilibrium, the price-rent ratio does not have a term that describes the growth of rents.

where the left-hand side is housing supply and the right-hand side is the demand, which consists of three components: the demand from renters, the demand from marginal owners, and the demand from full owners. Note that higher values of the land share parameter η_i imply a greater price elasticity of housing supply. Also, note that s_i^* is a function of p_i and r_i, and s_i^{**} is a function of p_i, as discussed in detail in footnote 20.

Given that the model is essentially static, it is not well-suited to account for short-run fluctuations in house prices. Instead, I focus on long-run changes in prices, and what these changes imply for polarization and inequality within cities.

3.4 Equilibrium

The following definition describes a stationary spatial equilibrium for this model:

Definition 1 (stationary spatial equilibrium). Conditional on local and economy-wide parameters, a *stationary spatial equilibrium* is given by skill-specific local labor supply, $N_i(s)$, rents, r_i, and prices, p_i, such that equations (17), (25), and (26) are satisfied for all i.

3.5 Main Mechanism: Homeownership, Polarization, and Inequality

Even though the nationwide distribution of skills is exogenously given, endogenous location choices determine local distribution of skills and wages in each city.[25] In this section, I discuss how the choice to rent or own shapes the skill distribution within a city.

In order to study labor market polarization analytically, I define low-, middle-, and high-skilled employment shares as follows.

Definition 2 (skill shares). Consider arbitrary skill levels s' and s'' that satisfy $s'' > s'$. Define the *s'-low-skilled share* as the fraction of workers with skills below s' in city i,

$$n_i^L(s') \equiv \frac{1}{N_i} \int_0^{s'} N_i(s)\mathrm{d}s, \tag{27}$$

and the *s''-high-skilled share* as the fraction of workers with skills above s'',

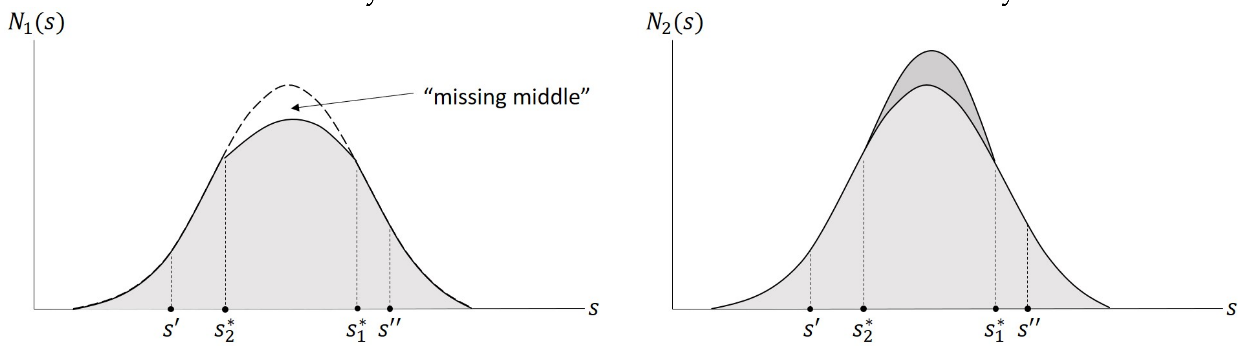
$$n_i^H(s'') \equiv \frac{1}{N_i} \int_{s''}^{1} N_i(s)\mathrm{d}s. \tag{28}$$

The (s', s'')-*middle-skilled share* is given by $n_i^M(s', s'') \equiv 1 - n_i^L(s') - n_i^H(s'')$.

To facilitate analysis, consider a simplified version of the model with two cities, 1 and 2, exogenous housing supply $\bar{H}_i$, no individual ownership costs (i.e., $\delta_i = 0$), and no

[25] Wages only depend on a function that is strictly increasing in skills and an exogenous productivity term. Hence, local skill and wage distributions are isomorphic and I use these two terms interchangeably.

Figure 3: Illustration of Proposition 1



Panel A shows the equilibrium skill distribution in city 1, and panel B shows the distribution in city 2. Workers with skill $s \in [s_2^*, s_1^*)$ can buy a house in city 2 but not in city 1. Some of them–the "missing middle"–choose to locate in city 2 only because they can buy a house there. This increases the s'-low-skilled and the s'-high-skilled employment shares in city 1.

agglomeration externalities (i.e., $\theta = 0$).[26] Furthermore, in order to verify that theoretical results do not depend on local differences in skill returns, assume that the returns to skill are the same in both cities, i.e., $a_1(s) = a_2(s)$.

The following proposition constitutes the central theoretical result of the paper. It demonstrates that cities with higher price-wage and price-rent ratios have a higher low-skilled share and, under an additional condition, also have a larger high-skilled share than cities with lower ratios. This implies that such cities have smaller middle-skilled share, i.e., they exhibit greater employment polarization.

Proposition 1 (larger polarization in cities with higher price-wage and price-rent ratios). Let city 1 have higher price-wage and price-rent ratios, i.e., $p_1/A_1 > p_2/A_2$ and $p_1/r_1 > p_2/r_2$.[27] Consider arbitrary skill levels s' and s'' such that s' is below the level required to own a house in city 2 and s'' is above the level required to own in city 1, i.e., $0 < s' < s_2^*$ and $s_1^* < s'' < 1$. Also, let the conditions of Lemma 1 hold. Then city 1 has a larger s'-low-skilled share, $n_1^L(s') > n_2^L(s')$. Furthermore, if the difference in price-rent ratios is bounded by a constant $\mathcal{B} > 1$, i.e., $\frac{p_1/r_1}{p_2/r_2} < \mathcal{B}$, then city 1 has a larger s'-high-skilled share, $n_1^H(s'') > n_2^H(s'')$, and therefore lower (s', s'')-middle-skilled share, $n_1^M(s', s'') < n_2^M(s', s'')$.[28]

Proof. See Online Appendix Section C.3 □

Why do higher price-wage and price-rent ratios in city 1 lead to larger employment polarization there? Figure 3 depicts the intuition behind this result. All households with

[26]Propositions 1 and 2 derived below will most likely hold with $\delta_i > 0$ and $\theta = 0$ but assuming that those terms are zero significantly simplifies the proofs.

[27]I refer to p_i/A_i as the price-wage ratio, since A_i is the only local component of wages (see equation 23).

[28]The constant $\mathcal{B}$ is defined in the proof of the proposition.

skill levels below s' cannot afford a house in any city, whereas those with skills above s'' have sufficient income to buy a house in any city. As discussed in Section 3.1.4, location choice of these two groups is independent of their tenure choice. At the same time, some households with skills between s' and s'' can afford to buy a house in city 2, but not in city 1. Since ownership has financial advantages, workers in this skill interval have an extra reason to live in city 2. This empties out the middle of the income distribution in city 1, thereby resulting in *higher polarization*. The reduction in the number of workers in the middle also leads to greater dispersion of income in city 1, i.e., *higher income inequality*.

The next proposition shows that the presence of both renters and owners is crucial to produce differences in polarization across cities. In an economy with renters only or owners only, skill shares would be the same in both cities, regardless of the differences in price-wage or price-rent ratios.

Proposition 2 (no differences in polarization without heterogeneity in housing tenure). Consider the following two scenarios:

1. One of the conditions of Lemma 1 is not satisfied in each city. In this case, every household in each city chooses to rent.
2. Both conditions of Lemma 1 are satisfied in each city, and there is no minimum size constraint (i.e., $\bar{h} = 0$). In this case, every household in each city chooses to own.

Then, in each scenario and for any finite and positive price-wage and price-rent ratios in each city, as well as for any values s' and s'' in $(0, 1)$, cities 1 and 2 have the same s'-low-skilled, s''-high-skilled, and (s', s'')-middle-skilled shares.

Proof. See Online Appendix Section C.4 □

3.5.1 The Role of Minimum Size and PTI Constraints

The main mechanism hinges on the result that households with sufficiently low income are excluded from homeownership. In turn, the inability of low-income households to buy a home depends on two parameters: the minimum size constraint, $\bar{h}$, and the PTI constraint, λ. Note that in the extreme case of $\bar{h} = 0$, the necessary and sufficient conditions for ownership always hold and, if in addition the conditions for Lemma 1 hold, everyone chooses to buy a house.

How empirically plausible are these two constraints?[29] While it is certainly possible to rent a fraction of a property (e.g., rent a room in a house), fractional ownership of residential units is rare. Moreover, even though there are small-sized properties that one could buy,

[29]The minimum-size constraint is commonly used in quantitative models of homeownership (Davis and Van Nieuwerburgh, 2015; Imrohoroğlu, Matoba, and Tüzel, 2018; Garriga and Hedlund, 2020). The PTI constraint is less common but also used (Greenwald, 2018; Kaplan, Mitman, and Violante, 2020). In this paper, it also indirectly plays the role of the loan-to-value constraint, which is more common. However, since tenure choice is essentially static, there is no explicit downpayment or loan-to-value requirement.

especially in dense cities such as New York, the parameter $\bar{h}$ would likely depend on the household type (e.g., a large family with children requires a larger minimal size than a single adult).

In addition, the minimum-size constraint implies that rental and owner-occupied markets are segmented by size. Size segmentation is a common feature of many models with tenure choice (Davis and Van Nieuwerburgh, 2015) and, as Online Appendix Figure F.1 demonstrates, the U.S. housing market exhibits strong segmentation by size. While 23% of owner-occupied units and 28% of rental units are between 1000 and 1500 sqft, there are very few owner-occupied units are smaller than 1000 sqft and very few rental properties are larger than 1500 sqft.

Greenwald and Guren (2021) provide another piece of evidence for segmentation by showing that homeownership rates are much less sensitive to credit supply shocks than price-rent ratios. Using the quantitative version of the model described later in Section 4, I show in Online Figure F.2 that in my model too, interest rate shocks have a much larger effect on the price-rent ratio than on homeownership.[30]

As for the PTI constraint, the vast majority of home purchases in the data are financed through mortgages which are typically conditioned on the borrower's income at the time of origination, although as shown in Greenwald (2018) the constraint may often not bind.

3.5.2 Comparison to Existing Mechanisms

The vast majority of spatial equilibrium models, including those that have previously studied local differences in labor market polarization and inequality (Baum-Snow, Freedman, and Pavan, 2018; Davis, Mengus, and Michalski, 2020; Cerina, Dienesch, Moro, and Rendall, 2023; Eeckhout, Hedtrich, and Pinheiro, 2024), do not distinguish between renting and owning and cannot produce local differences in low- or high-skilled shares without relying on various features of the production function, such as skill-biased productivity, skill complementarities, or task automation. In contrast, the model in this paper can generate differences in low and high-skilled employment shares across locations using a production function with labor only, perfectly substitutable skills, and without local differences in returns to skills.

4 Quantitative Model

This section describes the quantitative version of the model that is later used in counterfactual experiments. All model parameters are listed in Table 5.

[30]The reduction in the interest rate is engineered via changes in user costs δ_i and $\tilde{\delta}_i$. A lower interest rate increases homeownership by lowering δ_i, but also reduces it by raising price-rent ratios. In my quantitative model, the first effect dominates when interest rates reductions are large, and the second effect dominates when they are small. The notes to Online Appendix Figure F.2 provide more details.

Time. The model is calibrated separately for 1980 and 2019. Each calibrated model is assumed to be a separate stationary spatial equilibrium. All individual-level data used to calibrate the model is taken from individuals aged 25–64. This implies that the length of the life cycle is $T = 40$ years. The annual discount factor β is set to 0.96.

Locations and amenities. The model has two locations that represent two groups of cities: large and small. In order to aggregate 465 CZs into two location groups, I first sort them by the size of their labor force in 2019. Then I assign the 30 largest CZs into the "large" group and the remaining CZs into the "small" group.[31] Splitting the CZs into large and small at the 30th rank, I obtain two roughly equally-sized groups: the employment share of the 30 largest CZs is 49.3%. All CZ group-level empirical moments that are used in the quantitative model (e.g., wages, house prices, rents, etc.) are employment-weighted averages of CZ-level moments.

Local amenities, X_{it}, are calibrated to match employment in each location group and year. As can be seen in Table 5, model-implied amenities are slightly smaller in big cities in 1980 but larger in 2019.

Skill distribution and labor productivity. In the quantitative model, the skill distribution is discretized into a 100-point grid so that $s \in \{0.01, 0.02, ..., 1\}$. Skill groups are labeled "low," "middle," or "high" and are separated at the 20th and the 80th percentiles of the skill distribution, as described in Section 2.1.[32] In each calibration year, aggregate shares of low-, middle-, and high-skilled workers are taken from the data to reflect economy-wide changes in each skill group's employment shares. Furthermore, to match the share of each skill group in each CZ group, I introduce adjustment factors that multiply the amenity value of a city for middle- and high-skilled workers in each year and CZ group, x_{it}^M and x_{it}^H. As shown in Table 5, these adjustment factors are very close to 1. Returns to skill in each year and CZ group are described by the exponential function

$$a_{it}(s) = e^{\alpha_{it} s}. \tag{29}$$

Parameter α_{it} governs the dispersion of returns to skill and is calibrated separately for each CZ group and year so as to match the observed variance of log hourly wages in each location-year combination.

The exogenous component of city-level labor productivity, $\bar{A}_{it}$, is calibrated to match mean hourly wages in each location group and year. As can be seen in Table 5, $\bar{A}_{it}$ is 8.8% higher in large cities in 1980 but only 2.3% higher in 2019. At the same time, α_{it} increased

[31] A CZ is named after the most populous municipality it contains. The 30 largest CZs are: Los Angeles, CA; New York, NY; Chicago, IL; Washington, DC; Houston, TX; Newark, NJ; Philadelphia, PA; Boston, MA; San Francisco, CA; Atlanta, GA; Dallas, TX; Seattle, WA; Detroit, MI; Miami, FL; Phoenix, AZ; Minneapolis, MN; Denver, CO; Bridgeport, CT; San Diego, CA; Tampa, FL; Baltimore, MD; Sacramento, CA; San Jose, CA; Orlando, FL; Fort Worth, TX; St. Louis, MO; Cleveland, OH; Pittsburgh, PA; Austin, TX; and Portland, OR.

[32] That is, low-skilled workers have $s \in [0.01, 0.2]$, the medium-skilled have $s \in (0.2, 0.8]$, and the high-skilled have $s \in (0.8, 1]$.

Table 5: Model Parameters

Panel A: Economy-wide time-invariant parameters

Parameter	
Fréchet elasticity	$\epsilon = 6.1$
Agglomeration externality	$\theta = 0.04$
PTI constraint	$\lambda = 0.308$
Annual discount factor	$\beta = 0.96$

Panel B: Local time-invariant parameters

Parameter	CZ group	
Land share, η_i	Small	0.2239
	Large	0.3965

Panel C: Economy-wide time-varying parameters

Parameter	1980	2019
Elasticity of utility w.r.t. housing, γ	0.204	0.255
Minimum owner-occupied size, $\bar{h}$	2.68	1.99

Panel D: Local time-varying parameters

Parameter	CZ group	1980	2019
Productivity, A_{it}	Small	1.000	1.000
	Large	1.088	1.023
Amenities, X_{it}	Small	1.00	1.00
	Large	0.98	1.12
Construction productivity X land, $(\phi_{it}\Lambda_{it})^{\eta_i}$	Small	1.00	1.00
	Large	0.67	0.65
Owners' expenditures (annualized), δ_{it}	Small	0.032	0.046
	Large	0.039	0.035
Real estate managers' expenditures (annualized), $\tilde{\delta}_{it}$	Small	0.000	0.021
	Large	0.007	0.010
Middle-skilled amenity shifter, x_{it}^{M}	Small	0.99	1.04
	Large	1.02	0.97
High-skilled amenity shifter, x_{it}^{H}	Small	0.97	1.05
	Large	1.02	0.94
Skill dispersion, α_{it}	Small	1.81	2.09
	Large	1.84	2.26

Notes: Panel A lists model parameters common to all years and locations. Panel B shows parameters that differ across locations but are constant over time. Panel C shows parameters that vary over time but are common to all locations. Panel D lists parameters that vary by year and location.

much more rapidly in large CZs which is consistent with more pronounced SBTC in big cities. I set the value of the agglomeration externality to $\theta = 0.04$, the average of the estimates in the literature in the meta-study of Ahlfeldt and Pietrostefani (2019).

Fréchet elasticity. It is important that the model produces realistic migration responses to shocks. The estimates of labor supply elasticity to productivity shocks in the literature range from 1.52 in Monte, Redding, and Rossi-Hansberg (2018) to 4.03 in Hornbeck and Moretti (2024). I pick the midpoint of 2.8, and then simulate 1% shocks to A_i in each location and calibrate the scale parameter of the Fréchet distribution ϵ such that the average elasticity of employment with respect to the shock in the model is 2.8. I obtain $\epsilon = 6.1$. In Section 5.3, I examine the sensitivity of results to lower and higher values of ϵ.

Housing demand. In the model, the elasticity of utility with respect to housing consumption, γ, is also the housing expenditure share for renters. The expenditure share of full owners is $\min\{\gamma/\delta_i, \lambda\}$ (see equation 10). The expenditures of real estate managers are $\tilde{\delta}_i$. I assume that, while δ_i and $\tilde{\delta}_i$ may differ, the difference between the two is not city-specific, i.e., $\delta_i - \tilde{\delta}_i = \hat{\delta}$ for all i.[33] Then, I calibrate γ to be equal to the share of renters' expenditure on shelter from the Consumption and Expenditure Survey and $\tilde{\delta}_i$ to the local price-rent ratio. Finally, δ_i is calibrated such that their weighted average across cities is equal to the homeowners' expenditure share on shelter. Online Appendix Section D.1 provides more details. The annualized (i.e., divided by T) values of these parameters are shown in Table 5: γ increases between 1980 and 2019 to reflect rising expenditure share of renters on shelter; δ_i increases in small CZs and fall in large CZs due to the differential evolution of price-rent ratios in these two groups of locations (see Figure 4).

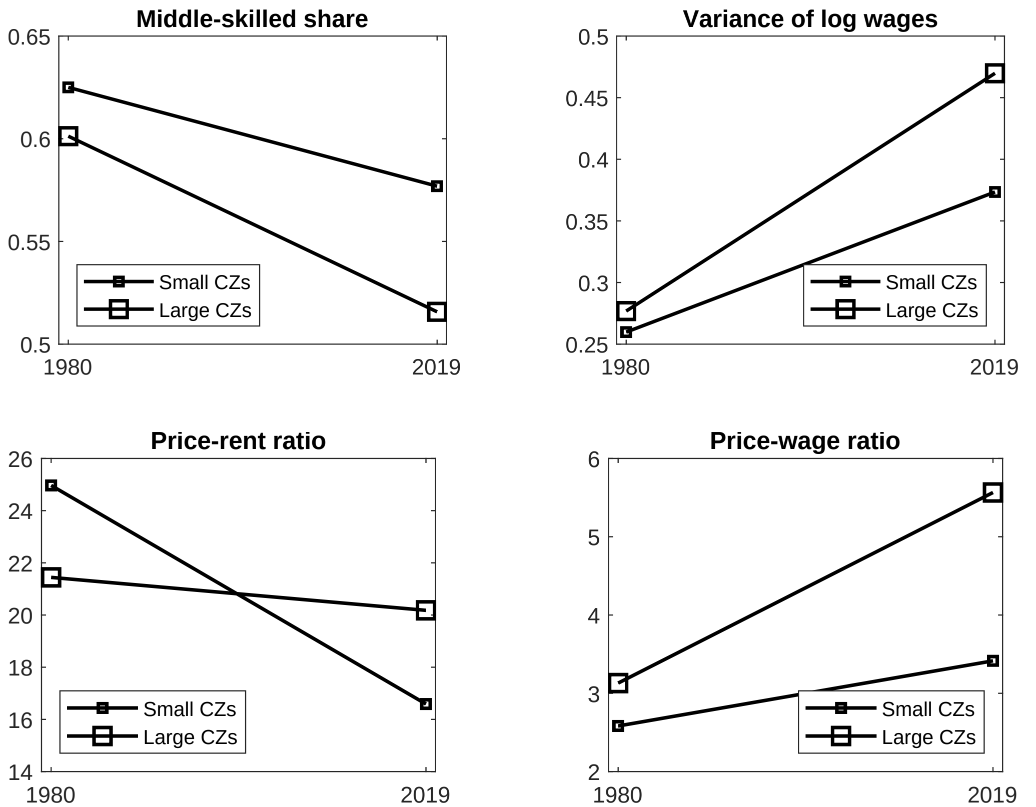
The minimum size of an owner-occupied house, $\bar{h}$, is allowed to vary over time and is calibrated to the observed nationwide homeownership rate in each year. The PTI constraint, λ, is set as follows. I follow Greenwald (2018) who provides evidence for a PTI constraint of 0.5.[34] However, I need to adjust this number to make it consistent with the model. The life cycle lasts 40 years and I assume that houses are bought in the beginning of the period. At the same time, mortgage contracts in the United States are typically underwritten for a 30-year period. Therefore, the PTI constraint has to be multiplied by 30 and then by the fraction of lifetime income an individual earns in the first year of the life cycle (estimated to be 0.0205 from the ACS data), which implies $\lambda = 0.5 \times 30 \times 0.0205 = 0.308$. See Online Appendix Section D.2 for more details on this calculation. Since $\gamma/\delta_i < \lambda$ in all CZs and years, the PTI constraint does not bind for full owners, although it may still bind for marginal owners.

Housing supply. The land share in the developers' production function, η_i, is constructed from county-level estimates of the average land share in single-family house values from Davis, Larson, Oliner, and Shui (2021), separately for large and small CZs. Land value shares are 0.2239 in small CZs and 0.3965 in large CZs. Higher land share in large CZs also

[33]It is likely that the variation in δ_i and $\tilde{\delta}_i$ across cities is explained by factors that are important for both homeowners and real estate managers, such as differences in property taxes or depreciation rates. The difference between δ_i and $\tilde{\delta}_i$ is likely explained by differences in preferences and risk, which are unlikely to differ by city.

[34]Another study by Kaplan, Mitman, and Violante (2020) use a PTI constraint of 0.25. Greenwald (2018) shows that a large fraction of new mortgages in 2014 exceed 0.25 but almost all are below 0.5.

Figure 4: Polarization, inequality, and house prices in large and small CZs



Notes: The top two panels plot the middle-skilled share and the variance of log wages in the groups of large and small CZs in the model economy. The bottom two panels plot price-rent and price-wage ratios.

implies that they have a lower housing supply elasticity.

For the quantitative model, it is not necessary to separately identify the productivity of developers and land area; hence, I calibrate the product $\phi_{it}\Lambda_{it}$ to the observed price-wage ratio in each CZ group and year. As Table 5 demonstrates, $(\phi_{it}\Lambda_{it})^{\eta_i}$ is smaller in large CZs in both 1980 and 2019. This is consistent with the evidence that housing supply in many large cities is more regulated (Glaeser and Gyourko, 2018) and that developers in more heavily regulated cities tend to be less productive (D'Amico, Glaeser, Gyourko, Kerr, and Ponzetto, 2023).

Evolution of polarization, inequality, and house prices. Figure 4 shows the evolution of the middle-skilled share and the variance of log wages in large and small CZs in the data and, since both are calibration targets, in the model. The figure corroborates the evidence shown in Section 2. In 1980, large CZs already had a lower middle-skilled share but the difference between large and small CZs was only 2.5 percentage points. Between 1980 and 2019, the middle-skilled share declined in all cities, a result of aggregate labor market polarization, however it fell faster in large cities and the gap between the two groups of CZs widened to over 6 percentage points. Similarly, in 1980 the variance of log wages was nearly the same in both groups of CZs: 0.26 in small and 0.28 in large ones.[35] By 2019, it increased much faster in large CZs and the gap widened from 0.02 to nearly 0.1.

Figure 4 also shows the evolution of price-rent and price-wage ratios in the two groups of cities. The model reproduces these ratios exactly because they are calibration targets for $\hat{\delta}_{it}$ and $\phi_{it}\Lambda_{it}$, respectively. The price-rent ratio was relatively stable in large cities between 1980 and 2019 but it fell in small cities.[36] At the same time, the price-wage ratio increased much more in large cities. Both indicators imply that purchasing a house became relatively more expensive in large CZs.

Why did prices grow more in big cities? Table 5 suggests that it was a combination of faster demand growth and lower supply elasticity.[37] On the one hand, returns to skill, as represented by α_{it}, increased much more in large CZs. This means that large CZs saw a disproportionate increase in high-income households that fueled the demand for housing. On the one hand, housing supply in large CZs is less elastic due to higher land share η_i, which means that a demand shock will have a large impact on prices.

5 Counterfactual Experiments

As discussed earlier, a common explanation for greater polarization and rise in inequality in large U.S. cities in the previous literature is skill-biased technical change (SBTC). This paper offers a novel explanation that relies on the interaction between faster house price growth in large cities and the desire to own a house. In this section, I use the quantitative model to compare these two explanations by first shutting down SBTC and then the factors that changed price-wage and price-rent ratios. I find that, while SBTC is a powerful force that accounts for the bulk of the disproportionate polarization and inequality in big cities, the interaction of price growth and homeownership significantly amplifies the effect of SBTC on polarization and inequality in large cities.

5.1 Setup

Each counterfactual experiment is unanticipated by real estate managers and therefore does not affect pre-counterfactual price-rent ratios. I study each experiment by comparing pre- and post-counterfactual steady states.

No SBTC. In the model, SBTC is generated by changes in skill dispersion parameters α_{it} which represent the elasticity of wages with respect to the skill level. Faster growth in α_{it} in large cities between 1980 and 2019 should produce greater polarization and inequality in large CZs. To quantify the role of SBTC in driving greater polarization and inequality in

[35] The variance of log wages in each group of cities is the weighted-average of CZ-specific variances.

[36] Other studies showed that the aggregate price-rent ratio in the U.S. is stable over the long run, though it varies over the business cycle (Davis and Van Nieuwerburgh, 2015; Piazzesi and Schneider, 2016).

[37] This explanation is similar to the explanation proposed by Gyourko, Mayer, and Sinai (2013).

large cities, I keep α_{it} in each CZ group at their 1980 levels and compute a counterfactual equilibrium for 2019.

Same p/w **and** p/r**.** To study the importance of higher price growth in large cities, I run counterfactual experiments in which price-wage and price-rent ratios in each group of cities remained at their 1980s levels. Note that prices, rents, and wages are endogenous.

In order to keep price-wage ratios at their 1980 levels, I obtain a reduction prices by increasing the combined developers' productivity and land supply, $\phi_{it}\Lambda_{it}$.[38] In other words, this counterfactual assumes that the productivity of developers in big cities grew sufficiently to ensure that price-wage ratios did not change from 1980.

To keep price-rent ratios constant, I keep the real estate managers' expenditure parameters $\tilde{\delta}_{it}$ at their 1980 level. Since the difference between $\tilde{\delta}_{it}$ and δ_{it} is fixed, this means that the difference in owners' expenses δ_{it} between the two cities is the same as in 1980, although their levels still grow to reflect rising homeowners' expenditures on shelter between 1980 and 2019 (see Section 4 for more details on how $\tilde{\delta}_{it}$ and δ_{it} are calibrated). Nonetheless, unlike in the benchmark model, in the experiments with constant price-rent ratios, δ_{it} is greater in big than in small CZs in 2019. This means that owning a house is less attractive in big than in small cities and therefore the demand for housing from potential homeowners in large CZs is lower. In other words, this counterfactual assumes that owning a house in a big city did not become relatively more attractive in 2019, be it because of interest rates, growth expectations, or any other factor that affects local price-rent ratios.

First, I fix price-wage ratios only, then price-rent ratios only, and finally both price-wage and price-rent ratios. Note that in each of these counterfactuals I allow for SBTC, i.e., parameters α_{it} change over time as they do in the benchmark economy.

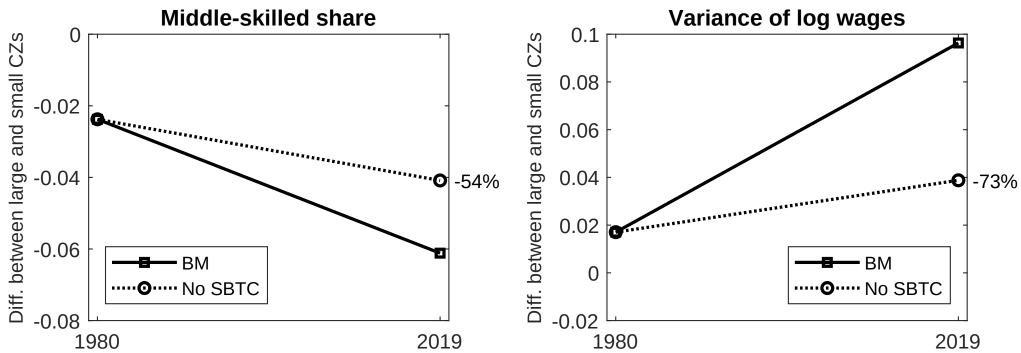
5.2 Results

Panel A of Figure 5 reports the results of the experiment where the returns to skill did not change since 1980. It shows that if skill returns remained constant, we would see a 64% smaller difference in polarization and a 76% lower increase in the difference in inequality between large and small CZs in 2019. That is, greater SBTC in large cities accounts for most of the excess polarization and rise in income inequality in large cities. These numbers are close to the findings of the previous literature discussed earlier. Cerina, Dienesch, Moro, and Rendall (2023) find that 67% of the excess job polarization in big cities is due to SBTC, while Baum-Snow, Freedman, and Pavan (2018) find that SBTC accounts for about 80% of the excess rise of inequality in large cities.

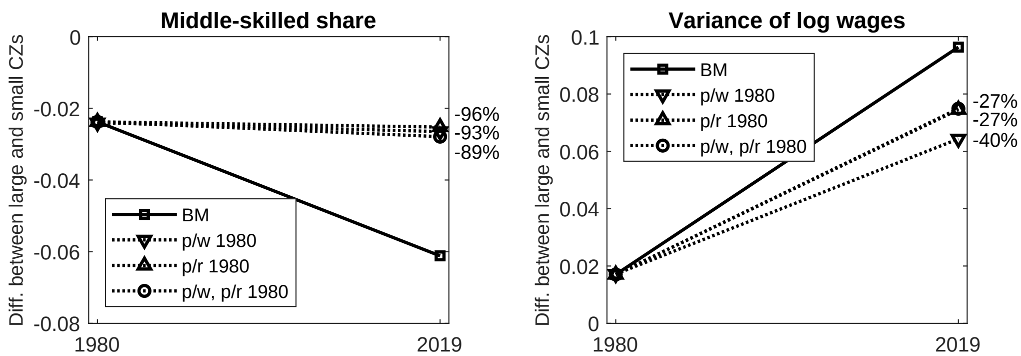
[38]Another possibility is to increase wages, e.g., by raising local productivity A_i. But wage increases capitalize in house prices. In my numerical experiments, wage changes did not have a noticeable effect on the price-wage ratio and I was not able to obtain the same price-wage ratio growth in large CZs as in small CZs by only recalibrating local productivity.

Figure 5: Counterfactual Results

Panel A: Role of SBTC



Panel B: Role of Price-Wage and Price-Rent Ratios



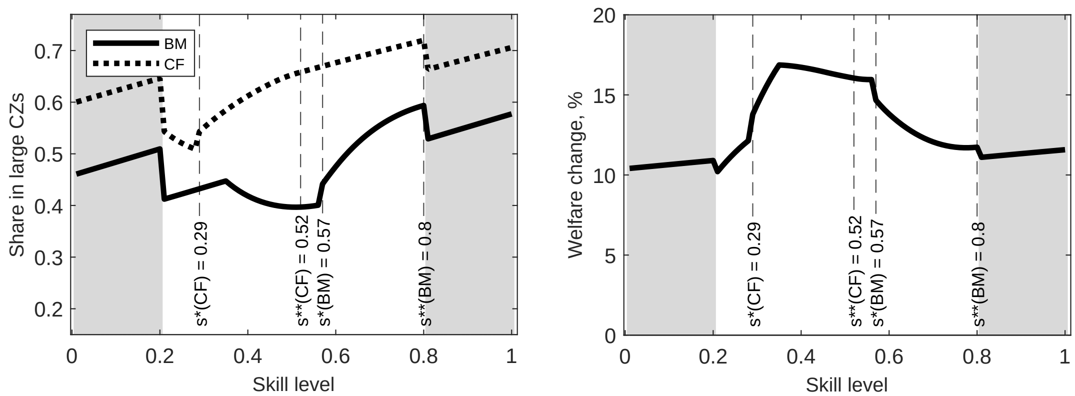
Notes: The left figure in panel A shows the difference between the share of middle-skilled employment in large CZs and the share in small CZs in the benchmark economy (solid line) and in the counterfactual (dotted line) where SBTC is shut down. The right figure in panel A shows the difference between the variance of log wages in large CZs and the variance in small CZs for the benchmark and the counterfactual economies. Panel B shows results for the counterfactual where the levels of price-wage and price-rent ratios are fixed at their 1980 levels. The numbers to the right of each panel report the percentage change from the benchmark to the counterfactual economy in the gap between large and small CZs.

Panel B of Figure 5 shows the results of the experiments where price-wage and price-rent ratios remain at their 1980 levels. When only price-wage ratios are constant, there is a 93% smaller difference in the decline in the middle-skilled share between the CZ groups and also a 40% slower increase in the difference in inequality, even though SBTC is in full force. When only price-rent ratios are constant, the polarization gap is 96% smaller and the gap in inequality between small and large cities grows by 27% less. When both price-wage and price-rent ratios are constant, the polarization gap is 89% smaller and the difference in inequality grows by 27% less.

These results mean that, while SBTC is an important determinant of the rising gap in polarization and inequality between large and small CZs, its effect depends on and is significantly amplified by faster growth of prices relative to rents and wages in big

Figure 6: Results by skill

Panel A: Skill distribution in large CZs Panel B: Welfare gains



Notes: Panel A shows the fraction of workers of each skill level that resides in large CZs in the benchmark economy (solid line) and the counterfactual economy (dashed-dotted line) where price-wage ratios remain constant at the 1980 level. Panel B shows welfare gains in the counterfactual scenario relative to the benchmark for workers at each skill level. In both panels, vertical lines show the skill thresholds for homeownership in benchmark and counterfactual economies.

cities. Without this amplification mechanism, the effect of SBTC on polarization would be 89–96% smaller and its effect on inequality about 27–40% smaller. Notably, changes in price-wage and price-rent ratios are independently important in producing the polarization and inequality gaps between large and small cities.

One could argue that faster price growth in big CZs is merely a consequence of SBTC and cannot be studied as a separate channel. However, note that I examine the growth of prices *relative* to wages and rents. In a simpler spatial equilibrium model without tenure choice and without distinction between prices and rents, price-rent ratios would not change simply because they are not defined. Price-wage ratios could change, but this would not differentially affect workers across the skill distribution because they would not be making the choice to rent or own. A model without tenure choice would still be able to account for much of the increase in inequality and polarization in large cities via the SBTC channel; however, in such model the skill-returns parameters would have to change by more because changes in housing parameters would not play any role.

Why was there more polarization and increase in income inequality in big cities in the benchmark economy compared to counterfactual experiments? It is because large cities were becoming increasingly unattractive for middle-income workers between 1980 and 2019. To understand better how the middle class was squeezed out of big cities due to higher house prices, it is useful to look at the distribution of skills in large CZs in the benchmark and counterfactual economies. For this illustration, I will turn to the counterfactual in which price-wage ratios are fixed at their 1980 levels. Panel A of Figure 6 shows that, while workers of all skill levels would move to large CZs if price-wage ratios

Table 6: Counterfactual changes

	Bench-mark	Const. p/w	Const. p/r	Const. p/w, p/r
Homeownership rate, %	61.0	78.2	52.3	68.1
Employment in large CZs, %	49.2	64.7	40.3	50.2
Output	100	102.1	98.4	99.9
Welfare	100	112.4	98.6	112.4

Notes: The table shows the values of several variables of interest in the benchmark economy calibrated to 2019 and the counterfactual economy.

there remained at the 1980 level, the relocation among middle-skilled workers with skill levels between 0.29 and 0.8 would be the largest.[39] This is because the skill threshold for full owners in large CZs falls from $s^{**} = 0.8$ to $s^{**} = 0.52$ and the threshold for marginal owners falls from $s^* = 0.57$ to $s^* = 0.29$. This means that workers with skill levels from 0.29 to 0.57 can now own a house in a big city instead of renting and those with skill levels from 0.57 to 0.8 can spend a smaller share of income on housing. In line with these results, panel B of Figure 6 shows that counterfactual welfare gains are concentrated among middle-skilled workers with skill levels between 0.29 and 0.8.[40] This suggests that middle-skilled workers were the largest losers from disproportionately rising prices in large cities from 1980s.

Table 6 shows several other results. First, while fixing price-wage ratios has a massive positive effect on homeownership (because they grew between 1980 and 2019), fixing price-rent ratios has the opposite effect (because they fell between 1980 and 2019). Second, the effect of counterfactual experiments on the city size distribution differs across the three experiments. In counterfactuals where price-rent ratios are constant, large CZs lose population because the user cost goes up to lower the price-rent ratio. This means that large cities become less attractive to prospective homeowners. The effect on output depends on whether large cities gain or lose population. This is because large cities are more productive (thanks to a larger dispersion of α_{it}) and when more workers live there, the whole economy is more productive.[41] Welfare grows the most in the counterfactuals where price-wage ratios are fixed at their low 1980 levels, which allows higher housing consumption and homeownership.[42]

Note that the magnitudes of aggregate changes reported in Table 6 are large and,

[39]The skill distributions have jumps at $s = 0.2$ and $s = 0.8$, since these thresholds separate low-, middle- and high-skilled groups of workers, and shifters x_{it}^M and x_{it}^H are calibrated to match local fractions of these groups in the data.

[40]Welfare gains for workers are computed as consumption-equivalent percentage changes of the expression in equation (20).

[41]Herkenhoff, Ohanian, and Prescott (2018), Hsieh and Moretti (2019), Parkhomenko (2023), and Duranton and Puga (2023) also find sizable aggregate productivity gains from lowering housing costs in the most productive locations and the resulting relocation of workers there.

[42]Real estate managers are excluded from welfare calculations.

potentially, unrealistic. This is because the counterfactual experiments do not describe a particular set of policies but rather completely shut down changes in prices relative to rents and wages over the course of nearly forty years. Such a scenario is highly unlikely and thus the numbers in Table 6 should be taken as mere indications of which direction the effects could go if a specific policy could alter price-wage and price-rent ratios in big versus small cities.

5.3 Robustness and Sensitivity

Multiple locations within a city. The model abstracts from internal city structure and the price of housing is the same everywhere in a CZ. In practice, however, there is large price heterogeneity within CZs. Moreover, just as price and rent dispersion *across* local labor markets has gone up in recent decades (Van Nieuwerburgh and Weill, 2010), *within*-city differences in prices have increased too (Guerrieri, Hartley, and Hurst, 2013; Albouy and Zabek, 2016). One reason why modeling internal city structure may be second-order is that, as discussed in Section 2.2 and shown in Online Appendix Table B.3, prices grew faster in most neighborhoods of large CZs, not only on average. Nonetheless, it is possible that in response to rapid growth in price-wage and price-rent ratios in some areas of the city, the priced-out middle class workers move to more affordable neighborhoods in the same city instead of leaving it altogether.

To examine the role of differences in prices and migration within cities, in Online Appendix Section E.1, I extend the model to allow for two neighborhoods in each city, calibrated to represent the areas with prices and price-rent ratios above and below the CZ median. Then, I run the same counterfactuals and show that the results for inequality and polarization are, if anything, larger than in the main model. These findings, together with the evidence in Online Appendix Table B.3, suggest that neighborhoods with affordable housing in large expensive cities are rare and middle-income workers who want to own a house often simply have to move to a different city.

Labor supply elasticity. In the main model, I calibrate the Fréchet scale parameter ϵ to match the labor supply elasticity with respect to productivity shocks equal to 3. As discussed in Section 4, the estimates of this value in the literature vary. Thus, I recalibrate the model to match a lower value of elasticity of 2 and a higher value of 4. This yields $\epsilon = 4.7$ and $\epsilon = 8.4$, respectively. Online Appendix Figure F.3 shows that the counterfactual results are smaller when the elasticity is lower and larger when it is higher, as should be expected given that this elasticity governs the decision of households to relocate in response to changes in local fundamentals. In both cases, however, keeping price-wage and price-rent ratios constant at the 1980 level reduces the gap in polarization and inequality between large and small cities.

Redistribution of rental revenue. In the main model, I assume that rental revenues are earned by absentee real estate managers. What if instead they were earned by workers? To examine whether the assumption of absentee landlords has any effect on the results, I calibrate two alternative models and re-run the counterfactual experiments. First, I redistribute local rents proportionally to local full homeowners, i.e., those with skill level $s > s_i^{**}$.[43] Second, I redistribute the sum of nationwide rental earnings in equal proportion to all full homeowners. Panel A in Online Appendix Figure F.4 shows that allowing for redistribution of rents does not have a significant effect on the main counterfactual results. Online Appendix Table F.1 compares welfare gains with and without rent redistribution. In the counterfactual where only price-wage ratios are kept constant, welfare gains are similar to the main counterfactual. However, in the experiment where price-rent ratios are kept constant, the welfare gains are larger because rents in this counterfactual are higher than in the baseline.

PTI constraint. In the main model, I use the PTI constraint of 0.5, following the evidence from Greenwald (2018) for the year 2014. This implies $\lambda = 0.308$. While he shows that PTI ratios fell between 2006 and 2014, they might have been even lower in 1980. As a robustness check, I use a PTI constraint of 1/3 which implies $\lambda = 0.205$. Online Appendix Figure F.5 demonstrates that in this case, the counterfactual reduction in the polarization and inequality gaps between large and small cities is significantly larger. When $\lambda = 0.205$, the PTI constraint binds for many marginal homeowners, and especially so in large cities. This creates an extra incentive for middle-income households to move to smaller CZs and results in larger polarization and inequality in large CZs.

No agglomeration externalities. To examine the role of agglomeration, I let the labor productivity to be exogenous and set $\theta = 0$. Then I recalibrate the model and re-run the counterfactuals. Online Appendix Figure F.6 shows that the results barely change.

Housing expenditure shares. In the quantitative model, I let the housing expenditure share increase over time for both renters and owners in line with the evidence from the CEX data. If the shares remained unchanged, the demand for housing would be weaker and, perhaps, the effect of rising prices on polarization and inequality in large cities would be smaller. I run a counterfactual where γ is fixed at its 1980 level but, as Online Appendix Figure F.7 shows, the counterfactual results remain sizable and are even larger for polarization.

Minimum owner-occupied size. In order to match the homeownership rate in both periods, I calibrate the minimum owner-occupied size $\bar{h}$ and, as Table 5 shows, it goes down from 1980 to 2019. While there is no evidence that smallest owner-occupied houses became smaller, what is the role of this parameter in generating counterfactual results? To answer this question, I fix $\bar{h}$ at its 1980 level and rerun counterfactual experiments. Online Appendix Figure F.8 shows that the results remain similar.

[43]Presumably, since full homeowners can afford a house size greater than the minimum, they also have resources to invest in other assets, such as rental housing.

5.4 Discussion

5.4.1 Implications for Housing Policies

The findings of the counterfactual experiments provide important insights for understanding the implications of housing policies in the United States. In particular, the results suggest that policies that could increase housing supply in large but unaffordable cities, such as zoning reforms, could not only lead to a more efficient spatial allocation of labor and greater aggregate productivity (Herkenhoff, Ohanian, and Prescott, 2018; Hsieh and Moretti, 2019; Parkhomenko, 2023; Duranton and Puga, 2023) but also make these cities more attractive to middle-income workers, less economically polarized and unequal, as well as reduce the wealth gap between owners and renters (Hilber and Turner, 2024). The results also suggest that such housing policies have a potential to substitute for labor market policies designed to deal with negative consequences of polarization and inequality.

This contrasts with the potential effect of policies that promote homeownership by reducing the cost of owning a home (represented in the model by δ_i and $\tilde{\delta}_i$) but without raising housing supply (represented by $\phi_i \Lambda_i$). As the results of the counterfactual where only price-rent ratios are kept constant show, such policies may reduce local polarization and inequality, but can also lower aggregate output and, ironically, do not necessarily increase homeownership (Hilber and Turner, 2014; Sommer and Sullivan, 2018).

5.4.2 Possible Extensions

The model is rather stylized and omits several features that may affect the interaction between homeownership, polarization, and inequality. First, since workers make location and tenure decisions in the beginning of the life cycle, the model is essentially static. If these choices could take place any time, it would be possible to incorporate other reasons why households may prefer to own rather than rent, such as wealth accumulation or insurance against labor or housing market risk. It would also be possible to study how changes in location choices at different stages of the life cycle interact with tenure choices.[44]

Second, labor is the only factor of production and skills are perfect substitutes. A model that features interactions between labor and other production inputs as well as between different skills could provide additional nuance in understanding the relationship between house prices, job polarization, and income inequality at the local level.[45]

Third, I assume that households must live and work in the same CZ. However, since 2020 we saw an explosion of remote and hybrid working arrangements that decoupled

[44] Examples of dynamic models with housing include Henderson and Ioannides (1983), Chambers, Garriga, and Schlagenhauf (2009), Favilukis, Ludvigson, and Van Nieuwerburgh (2017), Kaplan, Mitman, and Violante (2020), and many others. Examples of dynamic migration models include Kennan and Walker (2011), Howard (2020), and Garriga, Hedlund, Tang, and Wang (2023), among others.

[45] At the same time, Baum-Snow, Freedman, and Pavan (2018) find that capital-skill complementarity explains little of the excess growth in income inequality in large cities.

the choice of residence and the choice of workplace for many workers (Barrero, Bloom, and Davis, 2021). Since high-skilled workers are more often able to work from home than middle- and low-skilled workers, they may become less attached to their workplaces, and disproportionate polarization and increase in inequality in big cities may fade.[46]

6 Conclusions

In this paper, I propose a novel mechanism that explains why jobs have become more polarized and income distribution has become more unequal in large and expensive cities. While previous studies emphasized the role of skill-biased technical change, external labor demand shocks, and displacement of routine jobs with information technology, I argue that housing markets play a key role. When local price-wage and price-rent ratios are high, some middle-income households relocate to more affordable cities where they can buy a house. This hollows out the middle of the income distribution in expensive cities, which are typically also large cities, and results in higher polarization there.

I provide empirical evidence that supports this hypothesis. First, I show that polarization and the rise in income inequality were stronger in CZs where prices, price-rent, and price-rent ratios increased more since 1980, even when controlling for CZ size and other CZ characteristics. Second, I show that middle-income households are more likely than low- and high-income households to move for housing-related reasons to more affordable states. I build a parsimonious spatial equilibrium model that is consistent with this evidence and in which greater polarization and inequality in locations with high price-wage and price-rent ratios is an equilibrium outcome.

Quantitative exercises corroborate the findings of the previous literature that SBTC is an important driver of greater polarization and inequality in big cities. However, the effect of SBTC is significantly amplified by higher growth or price-rent and price-wage ratios in large cities. Absent this amplification in the housing market, the effect of SBTC on disproportionate polarization in big cities would be 63–81% smaller and the effect on the inequality gap between large and small cities 18–36% smaller. These results suggest that policies that constrained housing supply and contributed to high housing costs in many large cities also led to greater polarization and inequality there.

This paper also highlights the benefits of studying location choice and housing tenure choice jointly, as these two choices are interconnected for many households. Most models with location choice do not have tenure choice, while most models of tenure choice do not have location choice. As this paper shows, a model that combines these two choices goes a long way in explaining differences in economic outcomes across locations.

[46] An example of such re-sorting can be found in Delventhal and Parkhomenko (2023). They build a spatial equilibrium model with work from home and show that residential locations of college graduates since 2020 partly reversed the trends observed from the 1980s until the beginning of the Covid-19 pandemic.

Bibliography

Ahlfeldt, G. M., and E. Pietrostefani (2019): "The Economic Effects of Density: A Synthesis," *Journal of Urban Economics*, 111, 93–107.

Albouy, D., and M. Zabek (2016): "Housing Inequality," Working Paper 21916, National Bureau of Economic Research.

Amaral, F., M. Dohmen, S. Kohl, and M. Schularick (2024): "Interest rates and the spatial polarization of housing markets," *American Economic Review: Insights*, 6(1), 89–104.

Autor, D. H. (2019): "Work of the Past, Work of the Future," *AEA Papers and Proceedings*, 109, 1–32.

Autor, D. H., and D. Dorn (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103(5), 1553–97.

Autor, D. H., D. Dorn, and G. H. Hanson (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103(6), 2121–68.

Barrero, J. M., N. Bloom, and S. J. Davis (2021): "Why Working from Home Will Stick," Working Paper 28731, National Bureau of Economic Research.

Baum-Snow, N., M. Freedman, and R. Pavan (2018): "Why Has Urban Inequality Increased?," *American Economic Journal: Applied Economics*, 10(4), 1–42.

Baum-Snow, N., and L. Han (2024): "The Microgeography of Housing Supply," *Journal of Political Economy*, 132(6), 1897–1946.

Baum-Snow, N., and R. Pavan (2013): "Inequality and City Size," *The Review of Economics and Statistics*, 95(5), 1535–1548.

Behrens, K., G. Duranton, and F. Robert-Nicoud (2014): "Productive Cities: Sorting, Selection, and Agglomeration," *Journal of Political Economy*, 122(3), 507 – 553.

Berry, C. R., and E. L. Glaeser (2005): "The divergence of human capital levels across cities," *Papers in Regional Science*, 84(3), 407–444.

Cerina, F., E. Dienesch, A. Moro, and M. Rendall (2023): "Spatial Polarisation," *The Economic Journal*, 133(649), 30–69.

Chambers, M. S., C. Garriga, and D. Schlagenhauf (2009): "The loan structure and housing tenure decisions in an equilibrium model of mortgage choice," *Review of Economic Dynamics*, 12(3), 444–468.

Costa, D. L., and M. E. Kahn (2000): "Power Couples: Changes in the Locational Choice of the College Educated, 1940-1990," *The Quarterly Journal of Economics*, 115(4), 1287–1315.

Couture, V., C. Gaubert, J. Handbury, and E. Hurst (2020): "Income Growth and the Distributional Effects of Urban Spatial Sorting," Working Paper.

Cutler, D. M., and L. F. Katz (1992): "Rising Inequality? Changes in the Distribution of Income and Consumption in the 1980's," *The American Economic Review*, 82(2), 546–551.

D'Amico, L., E. L. Glaeser, J. Gyourko, W. Kerr, and G. A. M. Ponzetto (2023): "Why Has Construction Productivity Stagnated? The Role of Land-Use Regulation," Working Paper.

Davidoff, T. (2016): "Supply Constraints Are Not Valid Instrumental Variables for Home Prices Because They Are Correlated With Many Demand Factors," *Critical Finance Review*, 5(2), 177–206.

Davis, D. R., E. Mengus, and T. K. Michalski (2020): "Labor Market Polarization and The Great Urban Divergence," Working Paper 26955, National Bureau of Economic Research.

Davis, M. A., W. D. Larson, S. D. Oliner, and J. Shui (2021): "The price of residential land for counties, ZIP codes, and census tracts in the United States," *Journal of Monetary Economics*, 118, 413–431.

Davis, M. A., and F. Ortalo-Magné (2011): "Household Expenditures, Wages, Rents," *Review of Economic Dynamics*, 14(2), 248 – 261.

Davis, M. A., and S. Van Nieuwerburgh (2015): "Housing, Finance, and the Macroeconomy," *Handbook of Regional and Urban Economics*, 5, 753–811.

De la Roca, J., and D. Puga (2017): "Learning by Working in Big Cities," *Review of Economic Studies*, 84(1), 106–142.

Delventhal, M. J., and A. Parkhomenko (2023): "Spatial Implications of Telecommuting," Working Paper.

Diamond, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106(3), 479–524.

Duranton, G., and D. Puga (2023): "Urban Growth and Its Aggregate Implications," *Econometrica*, 91(6), 2219–2259.

Eckert, F., A. Gvirtz, J. Liang, and M. Peters (2020): "A Method to Construct Geographical Crosswalks with an Application to US Counties since 1790," Working Paper 26770, National Bureau of Economic Research.

Eeckhout, J., C. Hedtrich, and R. Pinheiro (2024): "IT and Urban Polarization," Working Paper.

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

Fajgelbaum, P. D., and C. Gaubert (2020): "Optimal Spatial Policies, Geography, and Sorting," *The Quarterly Journal of Economics*, 135(2), 959–1036.

Favilukis, J., S. C. Ludvigson, and S. Van Nieuwerburgh (2017): "The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium," *Journal of Political Economy*, 125(1), 140–223.

Favilukis, J., P. Mabille, and S. Van Nieuwerburgh (2023): "Affordable Housing and City Welfare," *The Review of Economic Studies*, 90(1), 293–330.

Feng, X., N. Jaimovich, K. Rao, S. J. Terry, and N. Vincent (2023): "Location, Location, Location: Manufacturing and House Price Growth," *The Economic Journal*, 133(653), 2055–2067.

Flood, S., M. King, R. Rodgers, S. Ruggles, and J. R. Warren (2020): "Integrated Public Use Microdata Series, Current Population Survey: Version 8.0 [dataset]," .

Florida, R. (2017): *The New Urban Crisis*. Basic Books, New York.

Ganong, P., and D. Shoag (2017): "Why has regional income convergence in the U.S. declined?," *Journal of Urban Economics*, 102(C), 76–90.

Garriga, C., and A. Hedlund (2020): "Mortgage Debt, Consumption, and Illiquid Housing Markets in the Great Recession," *American Economic Review*, 110(6), 1603–34.

GARRIGA, C., A. HEDLUND, Y. TANG, AND P. WANG (2023): "Rural-urban migration, structural transformation, and housing markets in China," *American Economic Journal: Macroeconomics*, 15(2), 413–440.

GAUBERT, C., P. KLINE, D. VERGARA, AND D. YAGAN (2021): "Trends in U.S. Spatial Inequality: Concentrating Affluence and a Democratization of Poverty," in *AEA Papers and Proceedings*, vol. 111, pp. 520–525.

GIANNONE, E. (2022): "Skill-Biased Technical Change and Regional Convergence," Working Paper.

GIANNONE, E., Q. LI, N. PAIXAO, AND X. PANG (2023): "Unpacking Moving," Working paper.

GLAESER, E., AND J. GYOURKO (2018): "The Economic Implications of Housing Supply," *Journal of Economic Perspectives*, 32(1), 3–30.

GLAESER, E. L., M. RESSEGER, AND K. TOBIO (2009): "Inequality in Cities," *Journal of Regional Science*, 49(4), 617–646.

GOOS, M., A. MANNING, AND A. SALOMONS (2014): "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, 104(8), 2509–26.

GREANEY, B. (2023): "Homeownership and the Distributional Effects of Uneven Regional Growth," Working Paper.

GREANEY, B., A. PARKHOMENKO, AND S. VAN NIEUWERBURGH (2024): "Dynamic Urban Economics," Working Paper.

GREENWALD, D. L. (2018): "The Mortgage Credit Channel of Macroeconomic Transmission," Working Paper.

GREENWALD, D. L., AND A. GUREN (2021): "Do Credit Conditions Move House Prices?," Working Paper 29391, National Bureau of Economic Research.

GREENWALD, D. L., M. LEOMBRONI, H. LUSTIG, AND S. VAN NIEUWERBURGH (2021): "Financial and Total Wealth Inequality with Declining Interest Rates," Working Paper 28613, National Bureau of Economic Research.

GUERRIERI, V., D. HARTLEY, AND E. HURST (2013): "Endogenous gentrification and housing price dynamics," *Journal of Public Economics*, 100(C), 45–60.

GYOURKO, J., C. MAYER, AND T. SINAI (2013): "Superstar Cities," *American Economic Journal: Economic Policy*, 5(4), 167–99.

HENDERSON, J. V. (1974): "The Sizes and Types of Cities," *American Economic Review*, 64(4), 640–56.

HENDERSON, J. V., AND Y. M. IOANNIDES (1983): "A Model of Housing Tenure Choice," *The American Economic Review*, 73(1), 98–113.

HERKENHOFF, K. F., L. E. OHANIAN, AND E. C. PRESCOTT (2018): "Tarnishing the Golden and Empire States: Land-Use Restrictions and the U.S. Economic Slowdown," *Journal of Monetary Economics*, 93, 89 – 109.

HILBER, C., AND T. M. TURNER (2014): "The Mortgage Interest Deduction and its Impact on Homeownership Decisions," *The Review of Economics and Statistics*, 96(4), 618–637.

HILBER, C. A. L., AND A. MENSE (2021): "Why have house prices risen so much more than rents in superstar cities?," Working Paper.

Hilber, C. A. L., and T. M. Turner (2024): "Land use regulation, homeownership and wealth inequality," Discussion paper, Centre for Economic Performance, LSE.

Hornbeck, R., and E. Moretti (2024): "Estimating Who Benefits from Productivity Growth: Local and Distant Effects of City Productivity Growth on Wages, Rents, and Inequality," *The Review of Economics and Statistics*, 106(3), 587–607.

Howard, G. (2020): "The migration accelerator: Labor mobility, housing, and demand," *American Economic Journal: Macroeconomics*, 12(4), 147–179.

Howard, G., and J. Liebersohn (2023): "Regional Divergence and House Prices," *Review of Economic Dynamics*, 49, 312–350.

Hsieh, C.-T., and E. Moretti (2019): "Housing Constraints and Spatial Misallocation," *American Economic Journal: Macroeconomics*, 11(2), 1–39.

Imrohoroğlu, A., K. Matoba, and S. Tüzel (2018): "Proposition 13: An Equilibrium Analysis," *American Economic Journal: Macroeconomics*, 10(2), 24–51.

Kaplan, G., K. Mitman, and G. L. Violante (2020): "The Housing Boom and Bust: Model Meets Evidence," *Journal of Political Economy*, 128(9), 3285 – 3345.

Kaplan, G., and S. Schulhofer-Wohl (2010): "Interstate Migration Has Fallen Less Than You Think: Consequences of Hot Deck Imputation in the Current Population Survey," NBER Working Paper 16536.

Karlman, M. (2022): "Why have house prices diverged? The role of the interest rate," Working Paper.

Katz, L., and K. M. Murphy (1992): "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107(1), 35–78.

Kennan, J., and J. R. Walker (2011): "The effect of expected income on individual migration decisions," *Econometrica*, 79(1), 211–251.

Kiel, K., and J. Zabel (1999): "The Accuracy of Owner-Provided House Values: The 1978–1991 American Housing Survey," *Real Estate Economics*, 27(2), 263–298.

Lindley, J., and S. Machin (2014): "Spatial changes in labour market inequality," *Journal of Urban Economics*, 79(C), 121–138.

Lutz, C., and B. Sand (2023): "Highly Disaggregated Land Unavailability," Working paper.

Lyons, R. C., A. Shertzer, R. Gray, and D. N. Agorastos (2024): "The Price of Housing in the United States, 1890-2006," Working Paper 32593, National Bureau of Economic Research.

Mabille, P. (2023): "The Missing Home Buyers: Regional Heterogeneity and Credit Contractions," *The Review of Financial Studies*, 36(7), 2756–2796.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018): "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 108(12), 3855–90.

Moretti, E. (2011): "Local Labor Markets," *Handbook of Labor Economics*, 4B, 1237–1313.

——— (2012): *The New Geography of Jobs*. Houghton Mifflin Harcourt, New York.

——— (2013): "Real Wage Inequality," *American Economic Journal: Applied Economics*, 5(1), 65–103.

Oswald, F. (2019): "The effect of homeownership on the option value of regional migration," *Quantitative Economics*, 10(4), 1453–1493.

Parkhomenko, A. (2023): "Local Causes and Aggregate Implications of Land Use Regulation," *Journal of Urban Economics*, 138, 103605.

Piazzesi, M., and M. Schneider (2016): "Housing and Macroeconomics," vol. 2, chap. Chapter 19, pp. 1547–1640. Elsevier.

Piketty, T., and E. Saez (2003): "Income Inequality in the United States, 1913–1998," *The Quarterly Journal of Economics*, 118(1), 1–41.

Roback, J. (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90(6), 1257–1278.

Rosen, S. (1979): "Wage-based indexes of urban quality of life," *Current issues in urban economics*, pp. 74–104.

Rossi-Hansberg, E., P.-D. Sarte, and F. Schwartzman (2019): "Cognitive Hubs and Spatial Redistribution," Working Paper 26267, National Bureau of Economic Research.

Ruggles, S., S. Flood, S. Foster, R. Goeken, J. Pacas, M. Schouweiler, and M. Sobek (2021): "Integrated Public Use Microdata Series, Current Population Survey: v. 11.0 [dataset]," .

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

Schubert, G. (2021): "House Price Contagion and U.S. City Migration Networks," Working paper.

Sommer, K., and P. Sullivan (2018): "Implications of US Tax Policy for House Prices, Rents, and Homeownership," *American Economic Review*, 108(2), 241–74.

Tolbert, C. M., and M. Sizer (1996): "U.S. Commuting Zones and Labor Market Areas: A 1990 Update," Staff Reports 278812, United States Department of Agriculture, Economic Research Service.

Truffa, S. (2017): "On the Geography of Inequality: Labor Sorting and Place-Based Policies in General Equilibrium," Working Paper.

Van Nieuwerburgh, S., and P.-O. Weill (2010): "Why Has House Price Dispersion Gone Up?," *Review of Economic Studies*, 77(4), 1567–1606.

Online Appendix

Homeownership, Polarization, and Inequality

Andrii Parkhomenko

A Appendix: Data

A.1 Locations

Empirical analysis is done at the level of U.S. states (48 mainland states plus the District of Columbia) and commuting zones (CZs). CZ definitions follow Tolbert and Sizer (1996); there are 741 CZs available for all years used in this study. CZ borders may change over time and, to guarantee comparability over time, I use the crosswalk by Eckert, Gvirtz, Liang, and Peters (2020). To ensure that I have enough observations to compute CZ-level aggregates using micro data, such as wage or house price indices, I only keep CZs with at least 2,000 observations per sample year (for 5% Census/ACS samples, this means keeping CZs with population of about 40,000 or above). This reduces the number of CZs to 465.

A.2 Wages

Wage data for each CZ and year comes the Census data from 1980, 1990, and 2000, and the 5-year American Community Survey (ACS) samples from 2006–2010 and 2015–2019.[47]

Sample cleaning. I exclude observations who live in group quarters, are younger than 25 and older than 64 years old, worked less than 26 weeks last year and less than 35 hours last week, work in the government or the military, and had non-positive wage or total income last year. Also excluded are observations with reported annual wage and salary income equivalent to less than half the minimum federal hourly wage.

Wages. To construct price-wage and rent-wage ratios, I compute median wages for each CZ and year.

A.3 Rent and House Price Indices

Hedonic rent and price indices are constructed both at the state and at the CZ level. I use rent and price data from the Census data from 1980, 1990, and 2000, and the 5-year ACS samples from 2006–2010 and 2015–2019.

Sample cleaning. I keep only household heads to ensure that the analysis is conducted at the household level. I exclude observations who live in group quarters; live in farm

[47]Tabulated by the IPUMS-USA (Ruggles, Flood, Foster, Goeken, Pacas, Schouweiler, and Sobek, 2021).

houses, mobile homes, trailers, boats, tents, etc.; are younger than 25 and older than 64 years old; and live in a dwelling that has no information on the year of construction.

Hedonic rent and price indices. To construct rent and price indices, I use self-reported rents and home values (variables `RENT` and `VALUEH` in the IPUMS-USA).[48] I estimate the following regression,

$$\ln \mathbf{q}_{n,it} = \beta_0 + \boldsymbol{\beta}_1 \mathcal{X}_{n,it} + \varphi_{it} + \varepsilon_{n,it}, \tag{30}$$

where $\mathbf{q}_{n,it} \in \{r_{n,it}, p_{n,it}\}$ is either the rent or the house value reported by household n in state i and year t, while $\mathcal{X}_{n,it}$ is a vector of controls that includes the number of rooms in the dwelling, the number of units in the structure (e.g., single-family detached, 2-family building), and the year of construction. Parameter φ_{it} is a location-year fixed effect. The rent or price index, $\mathbf{Q}_{it} \in \{R_{it}, P_{it}\}$, represents the rent or price after controlling for the observable characteristics listed before and idiosyncratic effects, and is given by $\mathbf{Q}_{it} \equiv \exp(\beta_0 + \varphi_{it})$. In the empirical analysis, I use either the price index, P_{it}, the price-rent index, P_{it}/R_{it}, or the ratio of either price index or rent index to median wages.

A.4 Employment Polarization

For the ease of comparison with the literature on labor market polarization, I follow Autor and Dorn (2013) in defining employment polarization. In particular, I use the skill percentile ranking of 3-digit normalized occupations. The percentile ranking was constructed in the aforementioned paper using 1980 wages at the national level. It assigns each occupation into a percentile bin, and a worker who holds an occupation that belongs to bin k is interpreted to be in the k-th percentile of the skill distribution. I compute the skill distribution at the CZ level using the Census and the ACS data, and using the same sample cleaning criteria as for constructing wage indices. In order to ensure that occupations are comparable between 1980 and 2015–2019, I use the crosswalk of occupations between 1980 and 2005 from Autor and Dorn (2013) and supplement it by constructing a crosswalk between the 2005 and the 2010 definitions of occupations (2010 definitions are used in the 2019 5-year sample).

A.5 Income Inequality

I estimate income inequality at the CZ level in 1980, 1990, and 2000 using the Census data, and in 2006–2010 and 2015–2019 using 5-year ACS samples. I focus on two measures of inequality: the variance of log wages and the Gini coefficient. Each measure is estimated for either annual or hourly wage income, where wage income can be either the reported

[48] Self-reported prices have been widely used in the literature in cases when other measures were not available. Kiel and Zabel (1999) show that, while self-reported prices are 3-8% higher than actual prices, the size of the bias does not depend on observable owners' characteristics and location. Thus, for the purposes of comparison across metro areas, self-reported prices are good proxies for market prices.

monetary income or income adjusted for the effects of gender, race, industry, occupation, education, and age, as described below.

Sample cleaning. I exclude observations who live in group quarters, younger than 25 and older than 64 years old, worked less than 26 weeks last year and less than 35 hours last week, work in the government or the military, and had non-positive wage or total income last year. Also excluded are observations with reported annual wage and salary income equivalent to less than half the minimum federal hourly wage.

Adjusted wage income. To construct adjusted wage income, I estimate

$$\ln w_{n,it} = \beta_0 + \boldsymbol{\beta}_{1,t} \mathcal{X}_{n,it} + \varepsilon_{n,it}, \tag{31}$$

where $w_{n,it}$ is the wage income, either annual or hourly, of individual n in commuting zone i and year t. $\mathcal{X}_{n,it}$ is a vector of controls that includes dummies for gender, race, 2-digit industry, 2-digit occupation, college, and 5-year age groups. The adjusted wage income is then given by

$$\tilde{w}_{n,it} \equiv \exp\left(\beta_0 + \varepsilon_{n,it}\right). \tag{32}$$

A.6 Interstate Migration Data

The data on interstate migration comes from the Annual Social and Economic Supplements (ASEC) of the Current Population Survey (CPS). I use the data for years 2001–2019.[49]

Sample cleaning. I first clean the sample by only keeping household heads to ensure that the level of observation is a household. Then I exclude observations in group quarters; younger than 18 years old; those who live in mobile homes, trailers, boats, tents, etc.; government and military employees; those who did not report location of residence last year or resided in a foreign country; those with non-positive or missing total income.

Definition of a migrant. A migrant is a household who reported that their state of residence last year was different from the state of residence at the time of the survey (variable `MIGSTA1` in the IPUMS-CPS).

Imputed migration. Kaplan and Schulhofer-Wohl (2010) show that imputations of missing data by the Census Bureau significantly bias estimated interstate migration rates. Thus, I drop observations with imputed migration status using the procedure in the aforementioned paper. In particular, I categorize an observation as having imputed migration status when the migration status of the individual was imputed, when the state of residence last year was imputed, when the migration status was not allocated, or when the migration status was allocated from another household member and the status of that member was imputed or inferred from yet another member whose status was imputed.

Moving reasons. The CPS also asks about the reason for moving (variable `WHYMOVE` in the IPUMS-CPS). In particular, the questionnaire asks what was the "main reason for

[49] Tabulated by the IPUMS-CPS (Flood, King, Rodgers, Ruggles, and Warren, 2020).

Table A.1: Summary statistics for interstate migration in the CPS

	Number of households	*Percentage of total*
Full sample	1,265,832	100
Moved between states		
all observations	22,518	1.8
non-imputed observations	15,815	1.2
Reason for moving between states (non-imputed observations)	*Number of households*	*Percentage of interstate migrants*
Job-related	8,097	51.2
Family-related	3,659	23.1
Housing-related	1,920	12.1
Other	2,139	13.5

Notes: The table reports the number of households (i.e., the number of observations in the cleaned sample) in the 2001–2019 CPS by their migration status and the reason for moving.

moving to this house (apartment)." Respondents can choose from 20 distinct reasons which can be grouped into job-related reasons, family-related reasons, housing-related reasons, and other reasons. Housing-related reasons combine the following answers: "wanted to own home, not rent," "wanted new or better housing," "wanted better neighborhood," "for cheaper housing," "other housing reason."[50]

Summary statistics. Table A.1 summarizes the data. It shows that 15,815 non-imputed observations (1.2% of the sample) moved across states. Among these, 1,920 (12.1% of interstate migrants) reported having moved for housing-related reasons.

A.7 Household Income (For Migration Analysis)

The data on household income for migration analysis comes from the Annual Social and Economic Supplements (ASEC) of the Current Population Survey (CPS).

Sample cleaning. I first clean the sample by only keeping household heads to ensure that the level of observation is a household. Then I exclude observations in group quarters; younger than 18 years old; those who live in mobile homes, trailers, boats, tents, etc.; government and military employees; those who did not report location of residence last year or resided in a foreign country; those with non-positive or missing total income.

Definition of household income. Household income is the total monetary income during the previous calendar year of all adult household members (variable `HHINCOME` in

[50] Even though the answer "wanted to own home, not rent" explicitly indicates that the primary reason for relocation was the desire to become a homeowner, other reasons may also be related with the transition to ownership. For this reason, I focus on all housing-related reasons in the empirical analysis.

the IPUMS-CPS).

Income distribution at the migration origin. Household income is reported for the previous year. This allows me to understand how the position of a household in the income distribution in the state of origin affects the likelihood of migration.

A.8 Additional Controls

In polarization and inequality regressions in Section 2.3.1, I include additional controls in columns (3) and (6) of Tables 2 and 4. These controls include the 1980 number of workers, as a share of total CZ employment, who work in manufacturing industries, who are women, who have a college degree, and who were born abroad. I also include a state dummy after merging several states such that the number of CZs in each state or group of states is at least five. In particular, I merge Connecticut, New Jersey, and New York; Maine and Rhode Island; Delaware and Pennsylvania; Arizona and Nevada; Maine, New Hampshire, and Vermont; North Dakota and South Dakota; Colorado and Utah; Idaho, Montana, and Wyoming; Maryland and Virginia.

B Appendix: Additional Empirical Results

B.1 Rents

The paper focuses on the evolution of prices between 1980 and 2019. In this section, I examine the evolution of the hedonic rent index and its ratio to wages. Table B.1 shows that rents actually increased more slowly in large cities. While this result may seem unexpected, the data from another study of long-run rent growth by Lyons, Shertzer, Gray, and Agorastos (2024) shows the same pattern. Regressing the log difference of 1980–2005 changes in their rent index on log 1980 city population from the Census, I find a negative coefficient of −0.074 which is nearly identical as the one in my table above.[51]

B.2 Prices within CZs

In Table 1, I showed that prices, price-wage, and price-rent ratios increased more in large CZs. To examine how price and rent growth in different neighborhoods within CZs relates to CZ size, I compute , price-wage, and price-rent ratios for each Public-Use Microdata Area (PUMA) within every CZ.[52] Then I look at the relationship between CZ size and price growth at the 10th, 25th, 50th, 75th, and 90th percentiles of the price distribution at the

[51]Their data ends in 2006 but the last year for which the data is available for all cities in their sample is 2005.

[52]PUMA is the smallest geographical unit for which individual-level Census and ACS data is publicly available. Largest CZs consist of over 100 PUMAs.

Table B.1: Relationship between city size and rents

	(1) Log rent chg.	(2) Log r/w chg.
Log initial population	-0.0671***	-0.0699***
	(0.00977)	(0.00820)
R-squared	0.0700	0.0970

The table shows the results from first-difference OLS regressions for the 1980–2019 period. Columns (1) and (2) report the coefficients from the regression of the change in log housing rents and rent-wage ratios on log CZ population in 1980. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

PUMA level within each CZ.[53] Table B.3 shows that price-rent ratios everywhere from the 10th to the 90th percentile of the within-CZ distribution grew more in big cities. Prices and price-wage ratios from the 25th to the 90th percentile grew more in large CZs, whereas the relationship between size and price growth is not statistically significant at the 10th percentile. That is, large CZs not only experienced faster price growth on average, but the growth of prices was faster in nearly all neighborhoods of large CZs.

B.3 Polarization

Next, I study whether the results in Table 2 are robust to using other thresholds to split employment into low-, middle-, and high-skilled groups. Tables B.4 and B.5 show the results when the groups are split at the 10th and the 90th wage percentiles, and the 33rd and the 67th percentiles. In both tables, the majority of coefficients remain negative and statistically significant, although significance becomes weaker when additional controls are used. Nonetheless, these findings suggest that the results presented in the main text are not the artifact of using the 20th and the 80th wage percentiles to define skill groups.

In Table B.6, I show that the findings in Table 2 are robust to using shorter time intervals. I estimate stacked regressions with two periods, 1980–2000 and 2000–2019, and year fixed effects. Coefficients remain negative and significant, and have comparable magnitudes.

B.4 Inequality

In the main text, I investigated the relationship between the growth in housing prices and the change in income inequality, measured as the variance of log annual wages. In Table B.7, I show that the findings presented in the main text hold when inequality is measured as the Gini coefficient of annual wages. In Table B.8, I demonstrate that the results are

[53] In particular, I calculate the growth between the n-th percentile in 1980 and the n-th percentile in 2019. While the PUMA at the n-th percentile may change between 1980 and 2019, this calculation allows to look at how prices change in segments of the housing market with the same *relative* prices in each year.

robust to using hourly, not annual wages. Finally, in Table B.9 I show that results do not change much when I use shorter time intervals and estimate a stacked regression with two periods, 1980–2000 and 2000–2019, and year fixed effects.

B.5 Migration

In Figure 1 in the main text I show that middle-income households are more likely to move for housing-related reasons from expensive to affordable states than low- or high-income households. In Figure B.1, I show that this hump-shaped relationship between income and the probability of moving only holds for housing-related migration and not for other migration reasons. Panel A shows that for job-related reasons, the relationship is somewhat U-shaped, though it is not statistically significant at the 95% level in most cases. This suggests that middle-income households may be less likely to move to a more affordable location when their relocation is related to a job. This could occur, for example, when the move is accompanied with a pay rise that compensates for an increase in housing costs. Panels B and C show that there is no statistically significant relationship between the probability of moving for family or other reasons and the income level.

B.6 Homeownership

Table B.2 shows the relationship between the growth in the price index, price-rent, and price-wage ratios and the change in the homeownership rate between 1980 and 2019. The relationships are not stastistically significant.

Table B.2: Relationship between homeownership and housing costs

	(1)	(2)	(3)
Housing costs change	0.394*	2.307*	0.641
	(0.225)	(1.375)	(0.898)
R-squared	0.0120	0.0110	0.00200

Notes: The table shows the results from first-difference OLS regressions for the 1980–2019 period. Column (1) reports the coefficient from a regression of the change in log housing price index on the change in 100× the homeownerhip rate. Columns (2) and (3) report results for price-rent and price-wage ratios, respectively. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

Table B.3: Relationships between city size, polarization, inequality, and prices (by percentile)

Panel A: 10th percentile

	(1) Log price chg.	(2) Log p/w chg.	(3) Log p/r chg.
Log initial population	-0.00185	-0.00464	0.0800***
	(0.0126)	(0.0106)	(0.0103)
R-squared	0	0.00100	0.113

Panel B: 25th percentile

	(1) Log price chg.	(2) Log p/w chg.	(3) Log p/r chg.
Log initial population	0.0176	0.0148	0.0921***
	(0.0112)	(0.00937)	(0.00940)
R-squared	0.00700	0.00700	0.152

Panel C: 50th percentile

	(1) Log price chg.	(2) Log p/w chg.	(3) Log p/r chg.
Log initial population	0.0315***	0.0287***	0.0852***
	(0.0118)	(0.00954)	(0.0102)
R-squared	0.0250	0.0300	0.123

Panel D: 75th percentile

	(1) Log price chg.	(2) Log p/w chg.	(3) Log p/r chg.
Log initial population	0.0487***	0.0459***	0.0709***
	(0.0124)	(0.0103)	(0.0129)
R-squared	0.0520	0.0620	0.0670

Panel E: 90th percentile

	(1) Log price chg.	(2) Log p/w chg.	(3) Log p/r chg.
Log initial population	0.0848***	0.0820***	0.0713***
	(0.0132)	(0.0112)	(0.0140)
R-squared	0.123	0.148	0.0620

The table shows the results from first-difference OLS regressions for the 1980–2019 period. In panel A, column (1) reports the coefficient from the regression of the change in the log of the 10th percentile of the house price distribution across PUMAs within each CZ on log CZ population in 1980. Columns (2) to (5) report the coefficients for log price-wage ratios, rents, rent-wage ratio, and price-rent ratios. Panels B to E repeat the analysis for the 25th, the 50th, the 75th, and the 90th percentiles. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

Table B.4: Change in the middle-skilled share and house price growth, 10/90 split

Panel A: Prices

	(1)	(2)	(3)	(4)	(5)	(6)
Price change	-0.374***	-0.0782	-0.234	-1.327***	-0.933**	-2.263***
	(0.122)	(0.113)	(0.147)	(0.412)	(0.401)	(0.874)
Log initial population		-1.267***	-0.992***		-1.099***	-1.206***
		(0.106)	(0.134)		(0.142)	(0.174)
Mean of dependent variable	-2.335	-2.335	-2.335	-2.335	-2.335	-2.335
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.0170	0.233	0.520			
1st-stage F-statistic				53.66	52.61	17.98

Panel B: Price-rent ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-rent ratio change	-7.247***	-4.368***	-1.302	-12.47***	-10.51**	-26.72*
	(0.771)	(0.902)	(1.132)	(3.431)	(4.207)	(14.42)
Log initial population		-0.991***	-0.929***		-0.582*	-0.184
		(0.111)	(0.141)		(0.308)	(0.458)
Mean of dependent variable	-2.335	-2.335	-2.335	-2.335	-2.335	-2.335
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.170	0.282	0.519			
1st-stage F-statistic				22.74	15.53	4.319

Panel C: Price-wage ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-wage ratio change	-2.829***	-1.783***	-1.909***	-4.433***	-3.060**	-7.857***
	(0.436)	(0.491)	(0.663)	(1.271)	(1.218)	(2.944)
Log initial population		-1.203***	-0.981***		-1.146***	-1.024***
		(0.104)	(0.131)		(0.121)	(0.136)
Mean of dependent variable	-2.335	-2.335	-2.335	-2.335	-2.335	-2.335
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.0580	0.254	0.528			
1st-stage F-statistic				73.42	68.89	21.85

Notes: The table shows the results from first-difference regressions for the 1980–2019 period. Panel A shows results for the house price index, panel B shows results for price-rent ratios, and panel C shows results for price-wage ratios. Column (1) reports the results from the OLS regression of the change in 100× the middle-skilled share on the change in prices, where the middle-skilled group is defined as occupations between the 10th and the 90th wage percentile. Column (2) includes initial CZ population as a control. Column (3) adds manufacturing share, female share, college share, foreign-born share, and state dummy as additional controls. Columns (4)–(6) report the results from 2SLS estimation. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

Table B.5: Change in the middle-skilled share and house price growth, 33/67 split

Panel A: Prices

	(1)	(2)	(3)	(4)	(5)	(6)
Price change	-0.624***	-0.366***	-0.891***	-2.060***	-1.763***	-2.317**
	(0.145)	(0.134)	(0.184)	(0.436)	(0.424)	(0.954)
Log initial population		-1.104***	-0.638***		-0.830***	-0.789***
		(0.109)	(0.149)		(0.162)	(0.185)
Mean of dependent variable	2.560	2.560	2.560	2.560	2.560	2.560
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.0440	0.199	0.476			
1st-stage F-statistic				53.66	52.61	17.98

Panel B: Price-rent ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-rent ratio change	-7.110***	-4.579***	-2.691**	-19.37***	-19.86***	-27.36*
	(0.835)	(0.928)	(1.244)	(4.209)	(5.514)	(15.51)
Log initial population		-0.871***	-0.465***		0.147	0.257
		(0.119)	(0.155)		(0.413)	(0.508)
Mean of dependent variable	2.560	2.560	2.560	2.560	2.560	2.560
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.155	0.237	0.448			
1st-stage F-statistic				22.74	15.53	4.319

Panel C: Price-wage ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-wage ratio change	-2.768***	-1.815***	-3.085***	-6.885***	-5.784***	-8.045**
	(0.524)	(0.503)	(0.734)	(1.388)	(1.310)	(3.324)
Log initial population		-1.096***	-0.567***		-0.920***	-0.603***
		(0.107)	(0.150)		(0.140)	(0.155)
Mean of dependent variable	2.560	2.560	2.560	2.560	2.560	2.560
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.0520	0.206	0.466			
1st-stage F-statistic				73.42	68.89	21.85

Notes: The table shows the results from first-difference regressions for the 1980–2019 period. Panel A shows results for the house price index, panel B shows results for price-rent ratios, and panel C shows results for price-wage ratios. Column (1) reports the results from the OLS regression of the change in 100× the middle-skilled share on the change in prices, where the middle-skilled group is defined as occupations between the 33rd and the 67th wage percentile. Column (2) includes initial CZ population as a control. Column (3) adds manufacturing share, female share, college share, foreign-born share, and state dummy as additional controls. Columns (4)–(6) report the results from 2SLS estimation. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

Table B.6: Change in the middle-skilled share and house price growth, 20-year intervals

Panel A: Prices

	(1)	(2)	(3)	(4)	(5)	(6)
Price change	-1.116***	-0.654***	-1.227***	-6.099***	-4.975***	-6.226**
	(0.264)	(0.235)	(0.252)	(1.265)	(1.231)	(2.607)
Log initial population		-0.911***	-0.782***		-0.688***	-0.906***
		(0.0727)	(0.104)		(0.107)	(0.139)
Mean of dependent variable	-1.430	-1.430	-1.430	-1.430	-1.430	-1.430
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.318	0.417	0.522			
1st-stage F-statistic				49.70	44.65	11.08

Panel B: Price-rent ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-rent ratio change	-6.905***	-4.824***	-3.946***	-25.57***	-25.09***	-47.02
	(0.714)	(0.710)	(0.756)	(5.695)	(7.323)	(36.41)
Log initial population		-0.777***	-0.681***		-0.0710	0.0944
		(0.0732)	(0.105)		(0.280)	(0.705)
Mean of dependent variable	-1.430	-1.430	-1.430	-1.430	-1.430	-1.430
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.378	0.445	0.526			
1st-stage F-statistic				22.68	13.93	1.631

Panel C: Price-wage ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-wage ratio change	-3.933***	-2.910***	-3.117***	-11.59***	-9.295***	-13.98**
	(0.530)	(0.522)	(0.544)	(2.078)	(1.991)	(5.588)
Log initial population		-0.881***	-0.786***		-0.740***	-0.906***
		(0.0731)	(0.105)		(0.0939)	(0.141)
Mean of dependent variable	-1.430	-1.430	-1.430	-1.430	-1.430	-1.430
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.338	0.431	0.525			
1st-stage F-statistic				75.41	66.79	12.91

Notes: The table shows the results from regressions for the 1980–2000 and 2000–2019 periods. All regressions include a dummy for the 2000–2019 period. Panel A shows results for the house price index, panel B shows results for price-rent ratios, and panel C shows results for price-wage ratios. Column (1) reports the results from the OLS regression of the change in 100× the middle-skilled share on the change in prices. Column (2) includes initial CZ population as a control. Column (3) adds manufacturing share, female share, college share, foreign-born share, and state dummy as additional controls. Columns (4)–(6) report the results from 2SLS estimation. The number of observations in each regression is 930 (465 CZs times 2 periods). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

Table B.7: Change in income inequality and house price growth, Gini coefficient

Panel A: Prices

	(1)	(2)	(3)	(4)	(5)	(6)
Price change	0.785***	0.515***	0.679***	1.446***	1.073***	2.095***
	(0.110)	(0.0838)	(0.105)	(0.294)	(0.266)	(0.667)
Log initial population		1.152***	0.920***		1.043***	1.069***
		(0.0745)	(0.102)		(0.0994)	(0.134)
Mean of dependent variable	6.063	6.063	6.063	6.063	6.063	6.063
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.120	0.412	0.577			
1st-stage F-statistic				53.66	52.61	17.98

Panel B: Price-rent ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-rent ratio change	6.301***	3.298***	2.778***	13.59***	12.08***	24.73**
	(0.603)	(0.614)	(0.874)	(2.802)	(3.338)	(11.73)
Log initial population		1.034***	0.767***		0.448*	0.124
		(0.0843)	(0.104)		(0.247)	(0.391)
Mean of dependent variable	6.063	6.063	6.063	6.063	6.063	6.063
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.210	0.409	0.555			
1st-stage F-statistic				22.74	15.53	4.319

Panel C: Price-wage ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-wage ratio change	3.709***	2.724***	2.538***	4.832***	3.518***	7.271***
	(0.439)	(0.348)	(0.484)	(0.922)	(0.795)	(2.312)
Log initial population		1.133***	0.867***		1.098***	0.901***
		(0.0729)	(0.103)		(0.0809)	(0.115)
Mean of dependent variable	6.063	6.063	6.063	6.063	6.063	6.063
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.162	0.447	0.571			
1st-stage F-statistic				73.42	68.89	21.85

Notes: The table shows the results from first-difference regressions for the 1980–2019 period. Panel A shows results for the house price index, panel B shows results for price-rent ratios, and panel C shows results for price-wage ratios. Column (1) reports the results from the OLS regression of the change in 100× the Gini coefficient of annual wages on the change in prices. Column (2) includes initial CZ population as a control. Column (3) adds manufacturing share, female share, college share, foreign-born share, and state dummy as additional controls. Columns (4)–(6) report the results from 2SLS estimation. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

Table B.8: Change in income inequality and house price growth, hourly wages

Panel A: Prices

	(1)	(2)	(3)	(4)	(5)	(6)
Price change	2.185***	1.682***	1.486***	2.156***	1.363***	2.909***
	(0.236)	(0.187)	(0.213)	(0.466)	(0.403)	(0.999)
Log initial population		2.150***	1.511***		2.212***	1.661***
		(0.136)	(0.184)		(0.153)	(0.213)
Mean of dependent variable	9.598	9.598	9.598	9.598	9.598	9.598
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.258	0.541	0.677			
1st-stage F-statistic				53.66	52.61	17.98

Panel B: Price-rent ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-rent ratio change	10.12***	3.621***	5.176***	20.26***	15.35***	34.35**
	(1.483)	(1.219)	(1.391)	(4.999)	(5.533)	(17.41)
Log initial population		2.239***	1.203***		1.458***	0.349
		(0.156)	(0.206)		(0.399)	(0.577)
Mean of dependent variable	9.598	9.598	9.598	9.598	9.598	9.598
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.151	0.410	0.643			
1st-stage F-statistic				22.74	15.53	4.319

Panel C: Price-wage ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-wage ratio change	6.898***	4.932***	3.605***	7.202***	4.470***	10.10***
	(1.124)	(0.792)	(1.000)	(1.701)	(1.407)	(3.726)
Log initial population		2.262***	1.381***		2.282***	1.428***
		(0.142)	(0.197)		(0.153)	(0.205)
Mean of dependent variable	9.598	9.598	9.598	9.598	9.598	9.598
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.156	0.471	0.647			
1st-stage F-statistic				73.42	68.89	21.85

Notes: The table shows the results from first-difference regressions for the 1980–2019 period. Panel A shows results for the house price index, panel B shows results for price-rent ratios, and panel C shows results for price-wage ratios. Column (1) reports the results from the OLS regression of the change in 100× the variance of log hourly wages on the change in prices. Column (2) includes initial CZ population as a control. Column (3) adds manufacturing share, female share, college share, foreign-born share, and state dummy as additional controls. Columns (4)–(6) report the results from 2SLS estimation. The number of observations in each regression is 465 (the number of CZs). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

Table B.9: Change in income inequality and house price growth, 20-year intervals

Panel A: Prices

	(1)	(2)	(3)	(4)	(5)	(6)
Price change	2.975***	2.347***	2.108***	4.943***	2.976***	3.774
	(0.387)	(0.311)	(0.372)	(1.211)	(1.138)	(2.578)
Log initial population		1.238***	1.095***		1.205***	1.136***
		(0.0928)	(0.135)		(0.104)	(0.151)
Mean of dependent variable	5.012	5.012	5.012	5.012	5.012	5.012
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.111	0.284	0.355			
1st-stage F-statistic				49.70	44.65	11.08

Panel B: Price-rent ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-rent ratio change	7.787***	4.573***	5.321***	20.73***	15.00**	28.50
	(0.988)	(0.832)	(0.946)	(5.580)	(6.351)	(26.05)
Log initial population		1.199***	0.948***		0.836***	0.530
		(0.0954)	(0.135)		(0.235)	(0.503)
Mean of dependent variable	5.012	5.012	5.012	5.012	5.012	5.012
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.0980	0.250	0.348			
1st-stage F-statistic				22.68	13.93	1.631

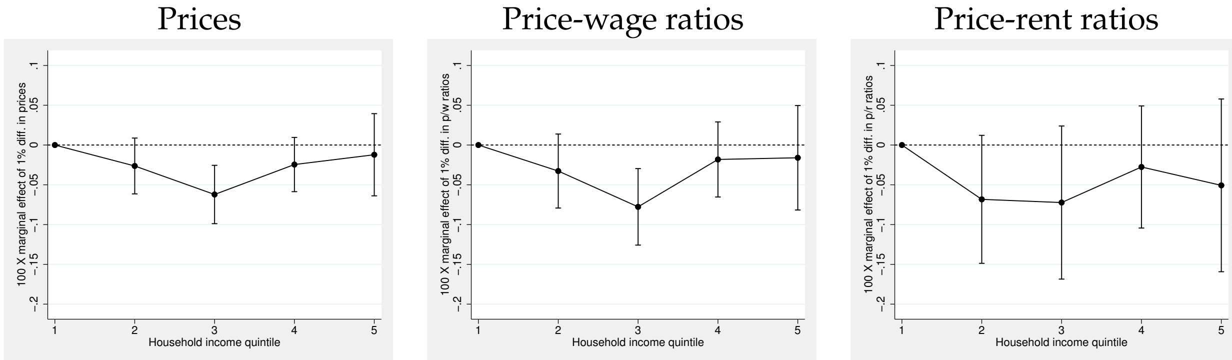
Panel C: Price-wage ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Price-wage ratio change	4.990***	3.502***	2.389***	9.394***	5.559**	8.474
	(0.839)	(0.732)	(0.814)	(2.350)	(2.161)	(5.970)
Log initial population		1.281***	1.070***		1.236***	1.137***
		(0.0978)	(0.140)		(0.105)	(0.160)
Mean of dependent variable	5.012	5.012	5.012	5.012	5.012	5.012
Model	OLS	OLS	OLS	2SLS	2SLS	2SLS
Additional controls	No	No	Yes	No	No	Yes
R-squared	0.0610	0.247	0.329			
1st-stage F-statistic				75.41	66.79	12.91

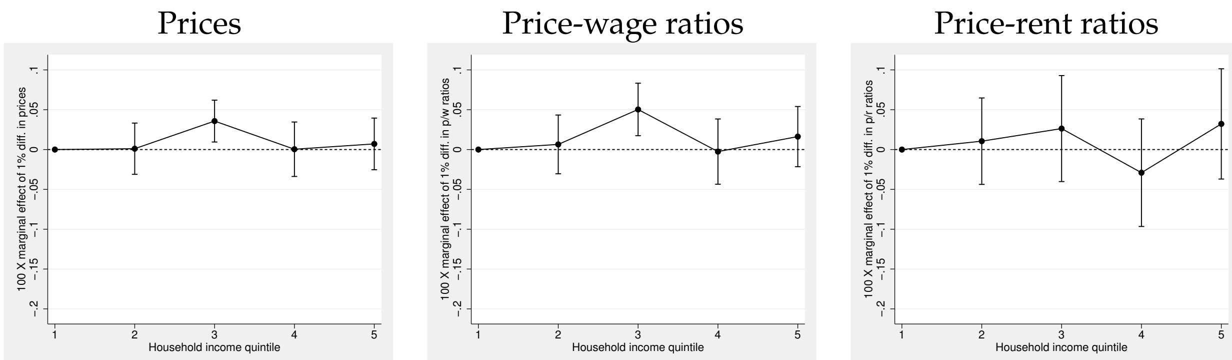
Notes: The table shows the results from regressions for the 1980–2000 and 2000–2019 periods. All regressions include a dummy for the 2000–2019 period. Panel A shows results for the house price index, panel B shows results for price-rent ratios, and panel C shows results for price-wage ratios. Column (1) reports the results from the OLS regression of the change in 100× the variance of log annual wages on the change in prices. Column (2) includes initial CZ population as a control. Column (3) adds manufacturing share, female share, college share, foreign-born share, and state dummy as additional controls. Columns (4)–(6) report the results from 2SLS estimation. The number of observations in each regression is 930 (465 CZs times 2 periods). Robust standard errors are reported in parentheses. *, **, and *** indicate 10%, 5%, and 1% significance levels.

Figure B.1: Marginal effects on migration for non-housing reasons

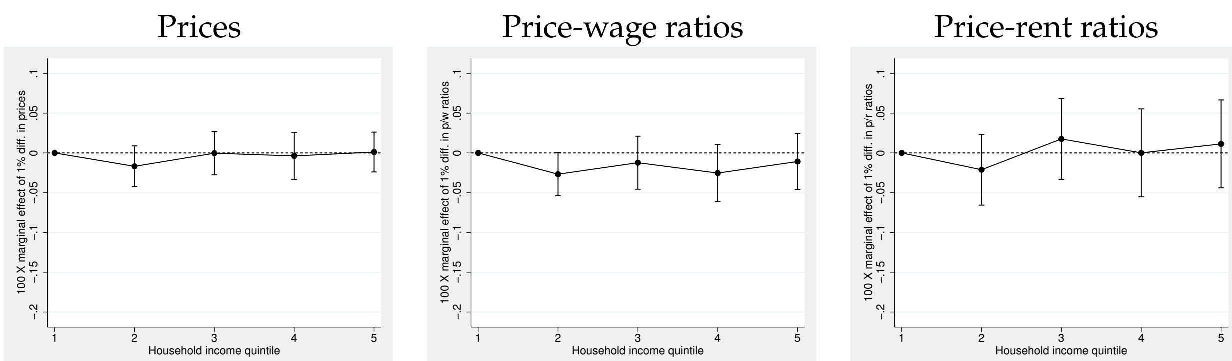
Panel A: Job-related reasons



Panel B: Family-related reasons



Panel C: Other reasons



Notes: Panel A shows 100× marginal effects on the probability of moving for job-related reasons. The left plot shows the marginal effect of the log ratio of house prices in the location of origin to the prices in the location of destination for each quintile in the household income distribution at the location of origin. The center plot shows the marginal effects of price-wage ratios, and the right plot shows the marginal effects of price-rent ratios. Panel B shows the results for family-related migration, and panel C shows the results for migration for other reasons. Marginal effects are estimated at the average observation from coefficients δ_2^q from regression (3) on the sample of 15,815 interstate migrants. Vertical bars represent the 95% confidence interval. Standard errors of marginal effects are computed using the Delta method. Standard errors of the underlying logit regression are clustered by the state of origin.

C Appendix: Derivations and Proofs

C.1 Proof of Lemma 1

Based on the expression for optimal housing consumption of owners (10), there are homeowners who buy houses larger than the minimum size only if income earned by the most skilled worker ($s = 1$) in the city is high enough. When the PTI constraint does not bind ($\lambda > \gamma/\delta_i$), the wage of the most skilled worker must satisfy $w_i(s = 1) > \delta_i p_i \bar{h}/\gamma$. When the PTI constraint binds ($\lambda \leq \gamma/\delta_i$), the wage of the most skilled worker must satisfy $w_i(s = 1) > p_i \bar{h}/\lambda$. Thus, the necessary condition for homeownership is

$$\bar{h} < \min\left\{\frac{\gamma}{\delta_i}, \lambda\right\} \frac{w_i(s=1)}{p_i} \tag{33}$$

However, even if this necessary condition is satisfied, households may find it suboptimal to own houses if prices are too high relative to rents. For a high enough income, ownership is preferred if the indirect utility of owning (11) is greater or equal to the utility of renting (6). In case of a non-binding PTI constraint, we have

$$\frac{w_i(s)}{\delta_i p_i^{\gamma}} \left(\frac{\delta_i - \gamma}{1-\gamma}\right)^{1-\gamma} \geq \frac{w_i(s)}{r_i^{\gamma}}. \tag{34}$$

When the PTI constraint binds, we have

$$\frac{w_i(s)}{p_i^{\gamma}} \left(\frac{1 - \delta_i \lambda}{1-\gamma}\right)^{1-\gamma} \left(\frac{\lambda}{\gamma}\right)^{\gamma} \geq \frac{w_i(s)}{r_i^{\gamma}}. \tag{35}$$

These two conditions can be rewritten in terms of the price-rent ratio as

$$\frac{p_i}{r_i} \leq \begin{cases} \left(\frac{\delta_i - \gamma}{1-\gamma}\right)^{\frac{1-\gamma}{\gamma}} \frac{1}{\delta_i^{1/\gamma}} & \text{if } \lambda > \frac{\gamma}{\delta_i}, \\ \left(\frac{1-\delta_i\lambda}{1-\gamma}\right)^{\frac{1-\gamma}{\gamma}} \frac{\lambda}{\gamma} & \text{if } \lambda \leq \frac{\gamma}{\delta_i}. \end{cases} \tag{36}$$

C.2 Proof of Lemma 2

Parts (a) and (b)

Case 1: PTI constraint does not bind. First, since $\bar{h} > 0$, there is $s' > 0$ such that $w_i(s') = \delta_i p_i \bar{h}$. This means that $v_i^O(w_i(s), p_i) = -\infty$ for all $s \leq s'$. At the same time, because renter's housing consumption does not have a minimal level and because $w_i(s) > 0$ for all s, we have $v_i^R(w_i(s), r_i) > 0$ for any $s \leq s'$. Thus, $v_i^R(w_i(s), r_i) > v_i^O(w_i(s), p_i)$ for all $s \leq s'$.

Next, let $v_i^{MO}(w_i(s), p_i)$ be the utility derived from consuming $h = \bar{h}$ units of housing and $v_i^{FO}(w_i(s), p_i)$ be the utility derived from consuming $h \geq \bar{h}$ units (*MO* and *FO* for marginal and full owners). Let s_i^{**} be such that $w_i(s_i^{**}) = \delta_i p_i \bar{h}/\gamma$. Note that at this skill level it is optimal to switch to become a full owner but still consume $h = \bar{h}$. As a result,

$$v_i^{FO}(w_i(s_i^{**}), p_i) = v_i^{MO}(w_i(s_i^{**}), p_i).$$

Consuming $h > \bar{h}$ when $s < s_i^{**}$ is suboptimal; hence, $v_i^{MO}(w_i(s), p_i) > v_i^{FO}(w_i(s), p_i)$ for all $s \in (s', s_i^{**})$. At the same time, $v_i^{FO}(w_i(s), p_i) > v_i^{MO}(w_i(s), p_i)$ for all $s \in (s_i^{**}, \infty)$ because in this interval $h_i^{FO}(s) > \bar{h}$ and an individual would not maximize lifetime utility by choosing house size $\bar{h}$. Also, since the condition (14) holds, we have $v_i^{FO}(w_i(s_i^{**}), p_i) > v_i^{R}(w_i(s_i^{**}), r_i)$.

Third, note that the wage threshold for full ownership is higher than the threshold for marginal ownership. As a consequence, since $v_i^{MO}(w_i(s), p_i)$ is a continuous function of s, there exists $s'' < s_i^{**}$ such that $v_i^{MO}(w_i(s''), p_i) > v_i^{R}(w_i(s''), r_i)$.

Fourth, recall that $\Phi(s)$ has full support on $s \in (0, \infty)$ and that both $v_i^{R}(w_i(s), r_i)$ and $v_i^{MO}(w_i(s), p_i)$ are continuous functions of s. This means that there exists $s^* \in (s', s'')$ such that

$$v_i^{R}(w_i(s_i^{*}), r_i) = v_i^{MO}(w_i(s_i^{*}), p_i).$$

Finally, since $s_i^* < s''$ and $s_i^{**} > s''$, we have $s_i^* < s_i^{**}$

Case 2: PTI constraint binds. If the PTI constraint binds, the proof is identical, except that there are no marginal owners and $s_i^* = s_i^{**}$ ■

Part (c). The full-ownership threshold is given by $w_i(s_i^{**}) = \delta_i p_i \bar{h}/\min\{\gamma/\delta_i, \lambda\}$. Since, $w_i(s) = A_i a(s)$, the skill threshold s_i^{**} is

$$a(s_i^{**}) = \frac{p_i \bar{h}}{A_i \min\{\gamma/\delta_i, \lambda\}}. \tag{37}$$

Because $a(s)$ is a strictly increasing function, s_i^{**} is decreasing in A_i.

The marginal-ownership threshold s_i^* is implicitly defined by the equivalence between the value of renting and the value of owning:

$$\Phi = \frac{w_i(s_i^*)}{r_i^{\gamma}} - \left(\frac{w_i(s_i^*) - \delta_i p_i \bar{h}}{1-\gamma}\right)^{1-\gamma} \left(\frac{\bar{h}}{\gamma}\right)^{\gamma} = 0. \tag{38}$$

Using the implicit function theorem, we can determine the relationship between s_i^* and A_i as $\partial s_i^*/\partial A_i = -(\partial\Phi/\partial A_i)/\left(\partial\Phi/\partial s_i^*\right)$. We have

$$\frac{\partial \Phi}{\partial A_i} = \frac{a_i(s_i^*)}{r_i^{\gamma}} - a_i(s_i^*)\left(\frac{1-\gamma}{\gamma}\frac{\bar{h}}{w_i(s_i^*) - \delta_i p_i \bar{h}}\right)^{\gamma}, \tag{39}$$

and

$$\frac{\partial \Phi}{\partial s_i^*} = \frac{A_i a_i'(s_i^*)}{r_i^\gamma} - A_i a_i'(s_i^*) \left(\frac{1-\gamma}{\gamma} \frac{\bar{h}}{w_i(s_i^*) - \delta_i p_i \bar{h}} \right)^\gamma . \tag{40}$$

Therefore,

$$\frac{\partial s_i^*}{\partial A_i} = -\frac{a_i(s_i^*)}{A_i a_i'(s_i^*)}. \tag{41}$$

Since all of the components of the ratio on the right-hand side are positive, $\partial s_i^* / \partial A_i < 0$, i.e., the skill level required to buy a house falls when a city has higher productivity ■

C.3 Proof of Proposition 1

By Lemma 2, higher price-wage ratio in city 1 implies that the skill threshold to become an owner is higher in this city, i.e., $s_1^* > s_2^*$. Define the low-skilled households as those with $s \leq s'$ and high-skilled households as those with $s \geq s''$, and let $s' < s_2^*$ and $s'' > s_1^*$. The remainder of households are middle-skilled.

Equilibrium skill shares. Denote $v_i^R(s) \equiv v_i^R(w_i(s), r_i)$ and $v_i^O(s) \equiv v_i^O(w_i(s), p_i)$. Optimal location and tenure choices imply that the measure of low-skilled workers in city 1 is

$$\mathcal{L}_1 = \int_0^{s'} \left(1 + \left(\frac{v_2^R(s)}{v_1^R(s)} \right)^\epsilon \right)^{-1} \mathrm{d}\Phi(s) = \int_0^{s'} \left(1 + \left(\frac{A_2 X_2}{A_1 X_1} \left(\frac{r_1}{r_2} \right)^\gamma \right)^\epsilon \right)^{-1} \mathrm{d}\Phi(s) = \frac{1}{1+\mathbf{r}} \mathbf{\Phi}_0^{s'},$$

where $\mathbf{r} \equiv \left(\frac{A_2 X_2}{A_1 X_1} \left(\frac{r_1}{r_2} \right)^\gamma \right)^\epsilon$ and $\mathbf{\Phi}_0^{s'} \equiv \int_0^{s'} \mathrm{d}\Phi(s)$. Similarly, the number of the high-skilled is given by

$$\mathcal{H}_1 = \int_{s''}^1 \left(1 + \left(\frac{v_2^O(s)}{v_1^O(s)} \right)^\epsilon \right)^{-1} \mathrm{d}\Phi(s) = \frac{1}{1+\mathbf{p}} \mathbf{\Phi}_{s''}^1,$$

where $\mathbf{p} \equiv \left(\frac{A_2 X_2}{A_1 X_1} \left(\frac{p_1}{p_2} \right)^\gamma \right)^\epsilon$. Finally, the measure of the middle-skilled is

$$\begin{aligned} \mathcal{M}_1 &= \int_{s'}^{s_2^*} \left(1 + \left(\frac{v_2^R(s)}{v_1^R(s)} \right)^\epsilon \right)^{-1} \mathrm{d}\Phi(s) + \int_{s_2^*}^{s_1^*} \left(1 + \left(\frac{v_2^O(s)}{v_1^R(s)} \right)^\epsilon \right)^{-1} \mathrm{d}\Phi(s) + \int_{s_1^*}^{s''} \left(1 + \left(\frac{v_2^O(s)}{v_1^O(s)} \right)^\epsilon \right)^{-1} \mathrm{d}\Phi(s) \\ &= \frac{1}{1+\mathbf{r}} \mathbf{\Phi}_{s'}^{s_2^*} + \frac{1}{1+\mathbf{q}} \mathbf{\Phi}_{s_2^*}^{s_1^*} + \frac{1}{1+\mathbf{p}} \mathbf{\Phi}_{s_1^*}^{s''}, \end{aligned} \tag{42}$$

where $\mathbf{q} \equiv \left(\frac{A_2 X_2}{A_1 X_1} \left(\frac{r_1}{p_2} \right)^\gamma \frac{\lambda^\gamma}{\gamma^\gamma (1-\gamma)^{1-\gamma}} \right)^\epsilon$.

In city 2, the number of the low-skilled is

$$\mathcal{L}_2 = \frac{\mathbf{r}}{1+\mathbf{r}} \mathbf{\Phi}_0^{s'}, \tag{43}$$

the number of the high-skilled is

$$\mathcal{H}_2 = \frac{\mathbf{p}}{1+\mathbf{p}} \mathbf{\Phi}_{s''}^{1}, \tag{44}$$

while the number of the middle-skilled households is

$$\mathcal{M}_2 = \frac{\mathbf{r}}{1+\mathbf{r}} \mathbf{\Phi}_{s'}^{s_2^*} + \frac{\mathbf{q}}{1+\mathbf{q}} \mathbf{\Phi}_{s_2^*}^{s_1^*} + \frac{\mathbf{p}}{1+\mathbf{p}} \mathbf{\Phi}_{s_1^*}^{s''}. \tag{45}$$

City 1 has higher s'-low-skilled share. The goal is to show that

$$n_1^L(s') = \frac{\mathcal{L}_1}{\mathcal{L}_1 + \mathcal{M}_1 + \mathcal{H}_1} > n_2^L(s') = \frac{\mathcal{L}_2}{\mathcal{L}_2 + \mathcal{M}_2 + \mathcal{H}_2}, \tag{46}$$

or, equivalently, that

$$\frac{\mathcal{M}_2}{\mathcal{L}_2} + \frac{\mathcal{H}_2}{\mathcal{L}_2} > \frac{\mathcal{M}_1}{\mathcal{L}_1} + \frac{\mathcal{H}_1}{\mathcal{L}_1}. \tag{47}$$

Using the expressions for skill shares derived above, we can rewrite this inequality as

$$\frac{\mathbf{\Phi}_{s'}^{s_2^*}}{\mathbf{\Phi}_0^{s'}} + \frac{\mathbf{q}}{\mathbf{r}} \frac{1+\mathbf{r}}{1+\mathbf{q}} \frac{\mathbf{\Phi}_{s_2^*}^{s_1^*}}{\mathbf{\Phi}_0^{s'}} + \frac{\mathbf{p}}{\mathbf{r}} \frac{1+\mathbf{r}}{1+\mathbf{p}} \frac{\mathbf{\Phi}_{s_1^*}^{s''}}{\mathbf{\Phi}_0^{s'}} + \frac{\mathbf{p}}{\mathbf{r}} \frac{1+\mathbf{r}}{1+\mathbf{p}} \frac{\mathbf{\Phi}_{s''}^{1}}{\mathbf{\Phi}_0^{s'}} > \frac{\mathbf{\Phi}_{s'}^{s_2^*}}{\mathbf{\Phi}_0^{s'}} + \frac{1+\mathbf{r}}{1+\mathbf{q}} \frac{\mathbf{\Phi}_{s_2^*}^{s_1^*}}{\mathbf{\Phi}_0^{s'}} + \frac{1+\mathbf{r}}{1+\mathbf{p}} \frac{\mathbf{\Phi}_{s_1^*}^{s''}}{\mathbf{\Phi}_0^{s'}} + \frac{1+\mathbf{r}}{1+\mathbf{p}} \frac{\mathbf{\Phi}_{s''}^{1}}{\mathbf{\Phi}_0^{s'}}. \tag{48}$$

Canceling common terms and combining, this inequality can be simplified to

$$\left(\frac{\mathbf{q}}{\mathbf{r}} - 1\right) \frac{1+\mathbf{r}}{1+\mathbf{q}} \frac{\mathbf{\Phi}_{s_2^*}^{s_1^*}}{\mathbf{\Phi}_0^{s'}} + \left(\frac{\mathbf{p}}{\mathbf{r}} - 1\right) \frac{1+\mathbf{r}}{1+\mathbf{p}} \left(\frac{\mathbf{\Phi}_{s_1^*}^{s''}}{\mathbf{\Phi}_0^{s'}} + \frac{\mathbf{\Phi}_{s''}^{1}}{\mathbf{\Phi}_0^{s'}}\right) > 0. \tag{49}$$

First, let us examine the ratio $\mathbf{q}/\mathbf{r}$:

$$\frac{\mathbf{q}}{\mathbf{r}} = \frac{\left(\frac{A_2 X_2}{A_1 X_1} \left(\frac{r_1}{p_2}\right)^{\gamma} \frac{\lambda^{\gamma}}{\gamma^{\gamma}(1-\gamma)^{1-\gamma}}\right)^{\epsilon}}{\left(\frac{A_2 X_2}{A_1 X_1} \left(\frac{r_1}{r_2}\right)^{\gamma}\right)^{\epsilon}} = \Lambda^{\epsilon} \left(\frac{p_2}{r_2}\right)^{-\gamma\epsilon}, \tag{50}$$

where $\Lambda \equiv \frac{\lambda^{\gamma}}{\gamma^{\gamma}(1-\gamma)^{1-\gamma}}$. Note that $\mathbf{q}/\mathbf{r} \geq 1$ if and only if $p_2/r_2 \leq \Lambda^{1/\gamma}$. The latter inequality is exactly the same as the condition (14) in Lemma 1 which, I assumed, must hold. Thus, $\mathbf{q}/\mathbf{r} \geq 1$. Next, let us examine the ratio $\mathbf{p}/\mathbf{r}$:

$$\frac{\mathbf{p}}{\mathbf{r}} = \frac{\left(\frac{A_2 X_2}{A_1 X_1} \left(\frac{p_1}{p_2}\right)^{\gamma}\right)^{\epsilon}}{\left(\frac{A_2 X_2}{A_1 X_1} \left(\frac{r_1}{r_2}\right)^{\gamma}\right)^{\epsilon}} = \left(\frac{p_1/r_1}{p_2/r_2}\right)^{\gamma\epsilon}. \tag{51}$$

Since I assumed that the price-rent ratio is higher in city 1, $\mathbf{p}/\mathbf{r} > 1$.

Therefore, since $\mathbf{q}/\mathbf{r} \geq 1$ and $\mathbf{p}/\mathbf{r} > 1$, and because all other object in inequality (49) are positive, the inequality holds. This proves that the low-skilled employment share is higher in city 1 ■

City 1 has higher s''-high-skilled share. The objective is to show that, under certain conditions,

$$n_1^H(s'') = \frac{\mathcal{H}_1}{\mathcal{L}_1 + \mathcal{M}_1 + \mathcal{H}_1} > n_1^H(s'') = \frac{\mathcal{H}_2}{\mathcal{L}_2 + \mathcal{M}_2 + \mathcal{H}_2}, \tag{52}$$

or, equivalently, that

$$\frac{\mathcal{M}_2}{\mathcal{H}_2} + \frac{\mathcal{L}_2}{\mathcal{H}_2} > \frac{\mathcal{M}_1}{\mathcal{H}_1} + \frac{\mathcal{L}_1}{\mathcal{H}_1}. \tag{53}$$

Using the expressions for skill shares derived above, we can rewrite this inequality as

$$\frac{\mathbf{r}}{\mathbf{p}}\frac{1+\mathbf{p}}{1+\mathbf{r}}\frac{\boldsymbol{\Phi}_{s'}^{s_2^*}}{\boldsymbol{\Phi}_{s''}^{1}} + \frac{\mathbf{q}}{\mathbf{p}}\frac{1+\mathbf{p}}{1+\mathbf{q}}\frac{\boldsymbol{\Phi}_{s_2^*}^{s_1^*}}{\boldsymbol{\Phi}_{s''}^{1}} + \frac{\boldsymbol{\Phi}_{s_1^*}^{s''}}{\boldsymbol{\Phi}_{s''}^{1}} + \frac{\mathbf{r}}{\mathbf{p}}\frac{1+\mathbf{p}}{1+\mathbf{r}}\frac{\boldsymbol{\Phi}_{0}^{s'}}{\boldsymbol{\Phi}_{s''}^{1}} > \frac{1+\mathbf{p}}{1+\mathbf{r}}\frac{\boldsymbol{\Phi}_{s'}^{s_2^*}}{\boldsymbol{\Phi}_{s''}^{1}} + \frac{1+\mathbf{p}}{1+\mathbf{q}}\frac{\boldsymbol{\Phi}_{s_2^*}^{s_1^*}}{\boldsymbol{\Phi}_{s''}^{1}} + \frac{\boldsymbol{\Phi}_{s_1^*}^{s''}}{\boldsymbol{\Phi}_{s''}^{1}} + \frac{1+\mathbf{p}}{1+\mathbf{r}}\frac{\boldsymbol{\Phi}_{0}^{s'}}{\boldsymbol{\Phi}_{s''}^{1}}. \tag{54}$$

Canceling common terms and combining, this inequality can be simplified to

$$\left(\frac{\mathbf{r}}{\mathbf{p}} - 1\right)\frac{1+\mathbf{p}}{1+\mathbf{r}}\left(\frac{\boldsymbol{\Phi}_{s'}^{s_2^*}}{\boldsymbol{\Phi}_{s''}^{1}} + \frac{\boldsymbol{\Phi}_{0}^{s'}}{\boldsymbol{\Phi}_{s''}^{1}}\right) + \left(\frac{\mathbf{q}}{\mathbf{p}} - 1\right)\frac{1+\mathbf{p}}{1+\mathbf{q}}\frac{\boldsymbol{\Phi}_{s_2^*}^{s_1^*}}{\boldsymbol{\Phi}_{s''}^{1}} > 0. \tag{55}$$

First, note that because $\mathbf{p}/\mathbf{r} > 1$, as shown above, $\mathbf{r}/\mathbf{p} < 1$. Next, let us examine the ratio $\mathbf{q}/\mathbf{p}$:

$$\frac{\mathbf{q}}{\mathbf{p}} = \frac{\left(\frac{A_2X_2}{A_1X_1}\left(\frac{r_1}{p_2}\right)^{\gamma}\Lambda\right)^{\epsilon}}{\left(\frac{A_2X_2}{A_1X_1}\left(\frac{p_1}{p_2}\right)^{\gamma}\right)^{\epsilon}} = \Lambda^{\epsilon}\left(\frac{p_1}{r_1}\right)^{-\gamma\epsilon}. \tag{56}$$

Note that $\mathbf{q}/\mathbf{p} \geq 1$ if and only if $p_1/r_1 \leq \Lambda^{1/\gamma}$. The latter inequality is exactly the same as the condition (14) in Lemma 1 which, I assumed, must hold. Thus, $\mathbf{q}/\mathbf{p} \geq 1$.

With $\mathbf{r}/\mathbf{p} < 1$ and $\mathbf{q}/\mathbf{p} \geq 1$, inequality (55) may not hold. However, note that $\mathbf{r}/\mathbf{p}$ approaches 1 when the difference in price-rent ratios between the two cities shrinks. Hence, it is possible to define a threshold $\mathcal{B} > 1$ such that the expression on the left-hand side of inequality (55) is equal to zero. If the ratio of price-rent ratios is smaller than $\mathcal{B}$, i.e., $\frac{p_1/r_1}{p_2/r_2} < \mathcal{B}$, then inequality (55) holds. In this case, city 1 has a larger high-skilled employment share ■

C.4 Proof of Proposition 2

Scenario 1: all households are renters. If at least one of the conditions of Lemma 1 does not hold, all households choose to rent in each city. As a result, the threshold s_i^* is not defined. Following the same steps as in the proof of Proposition 2 (Section C.3), one can show that the number of low, middle, and high-income workers in city 1 is

$$\mathcal{L}_1 = \frac{1}{1+\mathbf{r}}\mathbf{\Phi}_0^{s'}, \quad \mathcal{M}_1 = \frac{1}{1+\mathbf{r}}\mathbf{\Phi}_{s'}^{s''} \quad \mathcal{H}_1 = \frac{1}{1+\mathbf{r}}\mathbf{\Phi}_{s''}^{1}, \tag{57}$$

where $\mathbf{\Phi}$ and $\mathbf{r}$ were defined in Section C.3, and the latter variable does not depend on the skill level. Similarly, in city 2

$$\mathcal{L}_2 = \frac{\mathbf{r}}{1+\mathbf{r}}\mathbf{\Phi}_0^{s'}, \quad \mathcal{M}_2 = \frac{\mathbf{r}}{1+\mathbf{r}}\mathbf{\Phi}_{s'}^{s''} \quad \mathcal{H}_2 = \frac{\mathbf{r}}{1+\mathbf{r}}\mathbf{\Phi}_{s''}^{1}, \tag{58}$$

Using previous expressions and the fact that $\mathbf{\Phi}_0^{s'} + \mathbf{\Phi}_{s'}^{s''} + \mathbf{\Phi}_{s''}^{1} = 1$, the low-skilled employment share in both city 1 and city 2 is

$$\frac{\mathcal{L}_i}{\mathcal{L}_i + \mathcal{M}_i + \mathcal{H}_i} = \mathbf{\Phi}_0^{s'}, \tag{59}$$

and the high-skilled share in both cities is

$$\frac{\mathcal{H}_i}{\mathcal{L}_i + \mathcal{M}_i + \mathcal{H}_i} = \mathbf{\Phi}_{s''}^{1}. \tag{60}$$

Scenario 2: all households are owners. If both conditions of Lemma 1 are satisfied and the minimum size of an owner-occupied property is zero, $\bar{h} = 0$, then all households in all cities choose to be homeowners. Using the same steps as for Scenario 1 and replacing $\mathbf{r}$ with $\mathbf{p}$, one can show that in this case there are no differences in high or low-skilled employment shares between the two cities.

Therefore, when there are no differences in housing tenure choices within and across cities, both cities have exactly the same low and high-skilled employment shares ■

D Appendix: Calibration

D.1 Housing Expenditure Shares

For the time being, I will assume that $\gamma/\delta_i < \lambda$ for all i and all time periods and will confirm this assumption after obtaining the values of δ_i. To calibrate γ, δ_i, and $\tilde{\delta}_i$, I first take the share of income renters and owners spend on shelter using the Consumer Expenditure Survey (CEX) data. These shares are 0.204 in 1980 and 0.255 in 2019 for renters, and 0.146 in 1980 and 0.157 in 2019 for homeowners.[54] The value of γ in each year is equal to the

[54]The earliest available data is 1989. I use the 1989–1993 five-year average for 1980, the 1996-2000 average for 2000, and the 2015–2019 average for 2019. The renters' expenditure shares are close to 0.24 estimated by Davis and Ortalo-Magné (2011), a standard value for the renters' share used in the literature. The owners' expenditure shares are also consistent with Davis and Van Nieuwerburgh (2015) who report that, based on

renters' expenditure share in each year.

To obtain the value of δ_i, I proceed in two steps. First, I specify $\tilde{\delta}_i = \hat{\delta} + \delta_i$. In other words, I assume that, even though homeowners' expenses δ_i can be different from real estate managers' expenses $\tilde{\delta}_i$, local variation in these two parameters comes from the same source (e.g., differences in property taxes or depreciation rates). Second, I pin down the managers' expenses $\tilde{\delta}_i$ from the observed price-rent ratios and the discount factor (see equation 25). Third, to make sure that the nationwide housing expenditure share of owners is the same in the model and the data, I specify $\delta_i = \bar{\delta} + \delta_i^O$ and impose that the weighted-average of local components of δ_i is zero, i.e., $\sum_{i \in \mathcal{I}} \delta_i^O N_i = 0$, and obtain $\bar{\delta}$ by setting the nationwide expenditure share of owners equal to the weighted average of γ/δ_i in each year. This allows me to identify $\bar{\delta}$ and then, using the differences in price-rent ratios across cities, I can obtain $\hat{\delta}$.

Dividing the values of γ by δ_i and using the value of $\lambda = 0.308$ (see Section D.2 below) confirms that $\gamma/\delta_i < \lambda$ for all i in all years. Note that the housing expenditure share of marginal owners may still be higher than γ/δ_i, all the way up to λ.

D.2 PTI Constraint

I follow the empirical evidence presented in Greenwald (2018) and assume that the payment-to-income (PTI) constraint is 0.5. Recall that the model is calibrated to the data on individuals aged 25–64 (i.e., 40-year interval), while the vast majority of mortgage contracts are underwritten for 30 years. Since the model has a 40-year life cycle and households buy houses in the first period, the 0.5 PTI constraint would apply to the first-period income and households would pay off their house before the life cycle ends. Therefore, the value of the PTI constraint must be adjusted in two ways. First, I must incorporate the fact that in the data household income tends to be lower at the beginning of the life cycle. Second, I must take into account the fact that household income streams arrive for 40 years, while mortgage payments last for only 30 years. The adjusted PTI constraint can be written as

$$\lambda = \text{PTI} \times 30 \times \frac{w_1}{\sum_{t=1}^{40} w_t}. \tag{61}$$

I find that the average annual wage income of an individual in the 25–29 age group constitutes a fraction 0.0205 of the lifetime income from 25 to 64 years old from the ACS data. Therefore, the PTI constraint applicable to the quantitative model is $\lambda = 0.5 \times 30 \times 0.0205 = 0.308$.

the NIPA data, housing accounts for about 17% of household expenditures of owners and renters combined.

E Appendix: Model Extensions

E.1 Multiple Locations within a City

In this section, I construct a quantitative model with multiple locations within a city. The model is the same as described in Sections 3 and 4, except that each city i consists of two neighborhoods, $\ell \in \{L, H\}$. The neighborhoods differ by land-augmented housing productivity $\phi_{i\ell}$, land supply $\Lambda_{i\ell}$, ownership expenditures $\gamma_{i\ell}$ and $\tilde{\gamma}_{i\ell}$, as well as residential amenities, $X_{i\ell}$.[55] The city is still a single labor market, and productivity and skill dispersion parameters are the same in all neighborhoods.

In equilibrium, neighborhoods have different supply of each skill $N_{i\ell}(s)$, rents $r_{i\ell}$, and prices, $p_{i\ell}$. They also differ in homeownership skill thresholds $s^*_{i\ell}$ and $s^{**}_{i\ell}$, allowing for differences in the ability of local residents to buy a house. After having chosen which city to live in, but before making housing tenure and consumption choices, workers receive a neighborhood preference shock drawn from the Fréchet distribution with scale parameter ρ. The probability that a worker chooses to live in neighborhood ℓ, conditional on living in city i, is given by

$$\pi_{\ell|i}(s) = \frac{v_{i\ell}\left(w_i(s), p_{i\ell}, r_{i\ell}\right)^{\rho}}{\sum_{\ell' \in \{L,H\}} v_{i\ell'}\left(w_i(s), p_{i\ell'}, r_{i\ell'}\right)^{\rho}}. \tag{62}$$

The probability of choosing to live in city i is then given by

$$\pi_i(s) = \frac{\tilde{v}_i^{\epsilon}}{\sum_{j \in \mathcal{I}} \tilde{v}_j^{\epsilon}}, \tag{63}$$

where $\tilde{v}_i \equiv \left[\sum_{\ell' \in \{L,H\}} v_{i\ell'}\left(w_i(s), p_{i\ell'}, r_{i\ell'}\right)^{\rho}\right]^{1/\rho}$ is the expected value of living in city i before observing neighborhood-specific preference shocks.

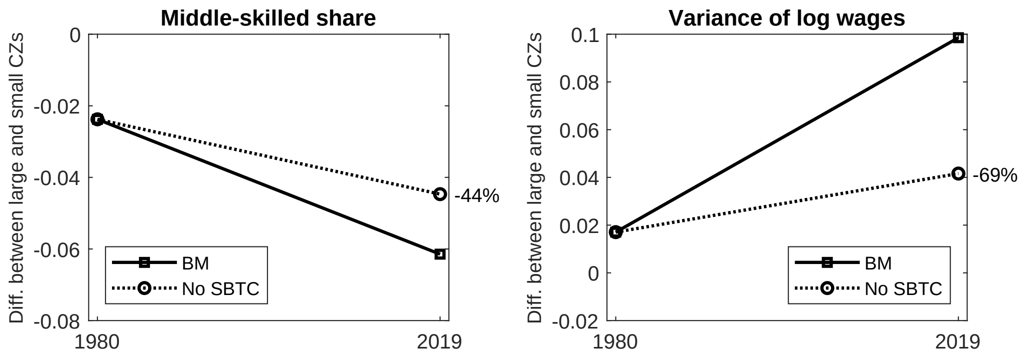
In the quantitative model, the $\ell = L$ neighborhood represents PUMAs with below-the-median prices and the $\ell = H$ neighborhood represents PUMAs with above-the-median prices. I calibrate the model as in Section 4, with two differences. First, I calibrate $\phi_{i\ell}\Lambda_{i\ell}$ and $\tilde{\gamma}_{i\ell}$ to match the 25th percentiles of the within-CZ distributions of price-wage and price-rent ratios for $\ell = L$, and the 75th percentiles for $\ell = H$.[56] This means that I split neighborhoods

[55] For simplicity, I do not model commuting. Differences in $X_{i\ell}$ can represent differences in access to jobs, among other things.

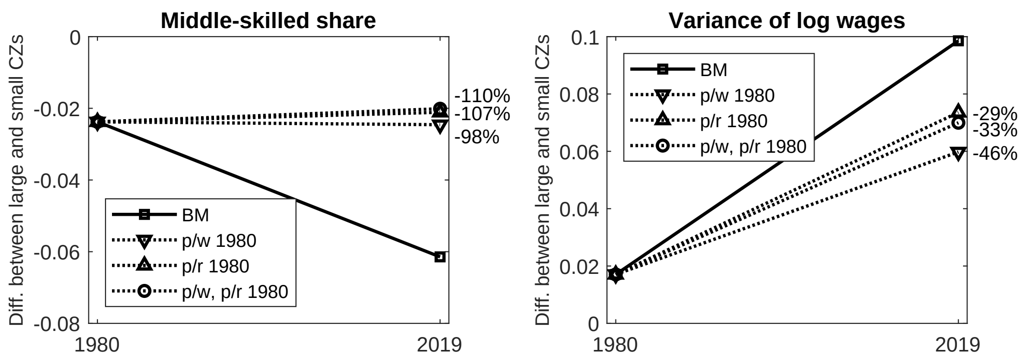
[56] Separate calibration of $\phi_{i\ell}\Lambda_{i\ell}$ in each neighborhood means that the model will generate stronger price increases in the H neighborhood of large CZs by construction. In particular, prices in the L neighborhood in large CZs went up by 565% from 1980 to 2019, while prices in the H neighborhood went up by 660%. To check if the model can generate higher price growth in the H neighborhood without relying on $\phi_{i\ell}\Lambda_{i\ell}$, I calibrate an alternative model where I adjust $\phi_{iH}\Lambda_{iH}$ such that $\phi_{iH}\Lambda_{iH}$ changes by the same percentage as $\phi_{iL}\Lambda_{iL}$ between 1980 and 2019. This lowers the price growth in the H neighborhood to 634% but it is still much larger than the 565% in the L neighborhood. In other words, most of the differential in price growth between top and bottom half of neighborhoods in large CZs is driven by rising income inequality and sorting of the high-income households into the top neighborhoods, and not by $\phi_{i\ell}\Lambda_{i\ell}$.

Figure E.1: Counterfactual Results (multiple locations within a city)

Panel A: Role of SBTC



Panel B: Role of Price-Wage and Price-Rent Ratios



Notes: The left figure in panel A shows the difference between the share of middle-skilled employment in large CZs and the share in small CZs in the benchmark economy (solid line) and in the counterfactual (dotted line) where SBTC is shut down. The right figure in panel A shows the difference between the variance of log wages in large CZs and the variance in small CZs for the benchmark and the counterfactual economies. The numbers to the right of each panel report the percentage change from the benchmark to the counterfactual economy in the gap between large and small CZs. Panel B shows results for the counterfactual where the levels of price-wage and price-rent ratios are fixed at their 1980 levels. The numbers to the right of each panel report the percentage change from the benchmark to the counterfactual economy in the gap between large and small CZs.

into top and bottom halves of the distribution of prices, and therefore I calibrate $X_{i\ell}$ such that the share of city residents in each neighborhood is 1/2. For comparability with the main model, I keep the value of ϵ at 6.3. The value of ρ is calibrated as follows. Baum-Snow and Han (2024) study a nested location choice model, similar to the one in this model extension, and estimate the scale parameter of the Fréchet distribution of shocks across cities at 3.9 and the parameter of the distribution of shocks within cities at 8.5. The ratio between the two parameters is $8.5/3.9 = 2.18$, hence $\rho = 2.18\epsilon = 13.3$.

Then, I run the same set of counterfactuals as in Section 5, shutting down SBTC and keeping prices constant relative to wages and rents. The results are shown in Figure E.1. Shutting down SBTC lowers the difference in polarization between large and small cities by 44% and the difference in the increase in the variance of log wages by 69%, similar to the results from the main model. Allowing for SBTC but keeping price-wage and price-rent ratios at the 1980 level reduces the gap in polarization between large and small cities by 98–110% (compared to 89–96% in the main model) and the gap in the rise in inequality by 29–46% (compared to 27–40% in the main model). In other words, allowing for multiple locations within cities does not change the conclusion that both SBTC and greater price growth in big cities contributed to greater polarization and faster rise in inequality in those cities.

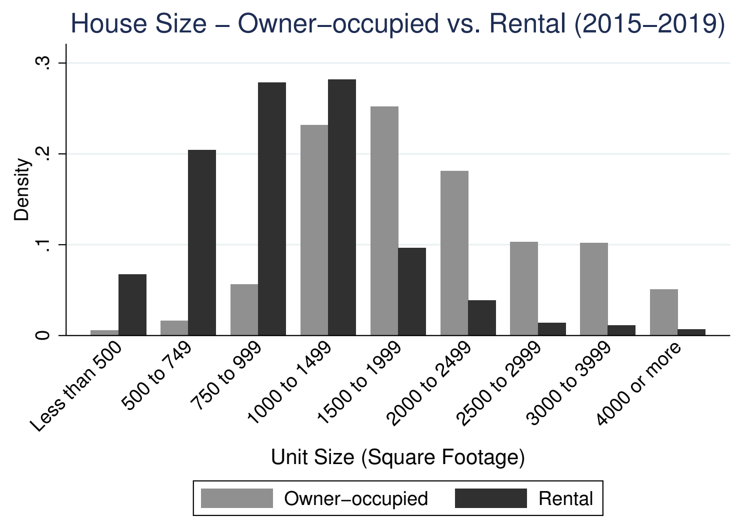
Why are the results larger in the model with multiple locations within a city? First, as shown in Table B.3, prices in big cities increase faster than in small cities not only on average but in nearly all neighborhoods. This means that if a household is priced out of homeownership in a large expensive city, it is highly unlikely that this city will have attractive neighborhoods where homeownership is within reach. Second, while the PTI constraint is rarely binding in the model with no heterogeneity within cities, in this version of the model it is binding in expensive neighborhoods of large CZs. This further reduces attractiveness of those neighborhoods for middle-skilled workers and encourages them to move to smaller cities.

In theory, workers could move to more affordable neighborhoods when prices in other neighborhoods go up.[57] However, all neighborhoods within a city belong to the same labor market, and positive labor demand or labor supply shocks should raise prices in all neighborhoods. In rare cases when neighborhoods have low prices in an expensive city, often this is because they either have poor amenities or poor access to jobs.

[57]And, to some extent, they do. In the main model, the threshold for being a homeowner in large CZ in 2019 is $s^* = 0.57$. In the model with 2 neighborhoods, the threshold is $s^*_H = 0.70$ in the H neighborhood and $s^*_L = 0.52$ in the L neighborhood. This means that, while in the economy without different neighborhoods workers with skills $s \in [0.52, 0.57)$ must rent in large CZs, in the economy with two neighborhoods, they can buy in the cheaper L neighborhood. In other words, allowing for different neighborhoods means that some households can buy a house in cheaper neighborhoods of large CZs, whereas without neighborhood heterogeneity those same households would have to move to another CZ in order to buy a house.

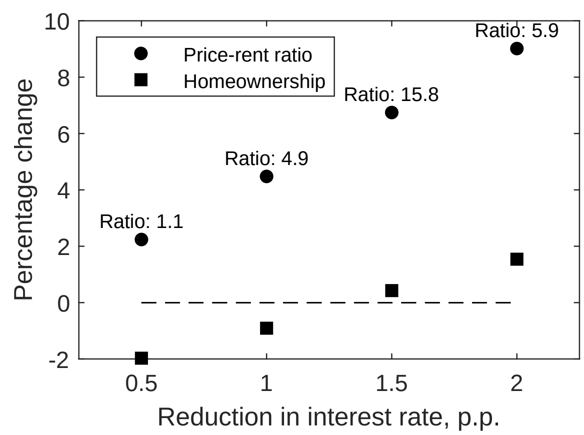
F Appendix: Additional Figures and Tables

Figure F.1: Distribution of sizes of rental and owner-occupied housing units



Notes: This figure shows the distribution of sizes of rental and owner-occupied housing units from the pooled 2015–2019 American Housing Survey samples.

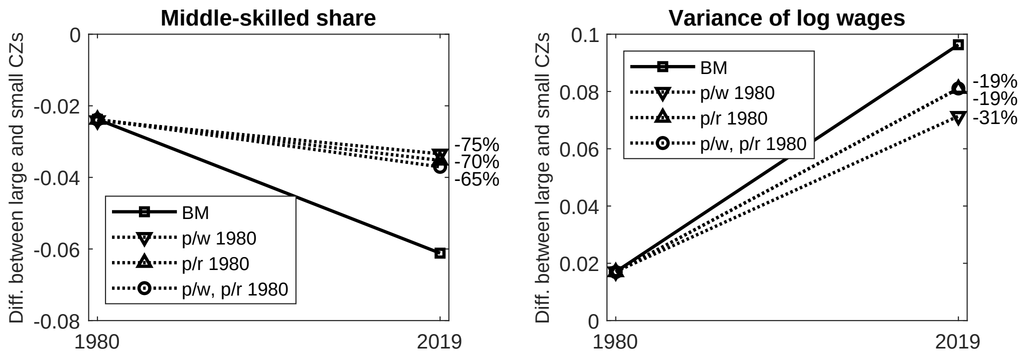
Figure F.2: Response of price-rent ratios and homeownership to lower interest rates



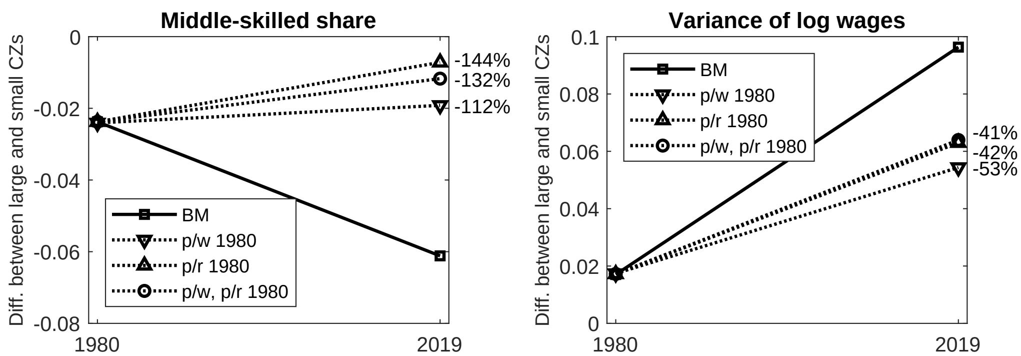
Notes: This figure shows percentage changes in the price-rent ratio and the homeownership rate in response to 0.5, 1, 1.5, and 2 percentage point reductions in the interest rate, as well as the absolute value of the ratio of the two variables. The reduction in the interest rate is engineered via changes in δ_i and $\tilde{\delta}_i$. In particular, I first assume that δ_i and $\tilde{\delta}_i$ include the annual expenditures on mortgage interest in principal using the 2015–2019 average real mortgage interest rate of 2.44% (30-year mortgage interest rate minus annual CPI inflation from the FRED database). Then I lower δ_i and $\tilde{\delta}_i$ by the amount that would result from 0.5, 1, 1.5, and 2 percentage point reductions in the interest rate, and solve the equilibrium of the model with updated δ_i and $\tilde{\delta}_i$, and compare price-rent ratios and homeownership rates to the benchmark model.

Figure F.3: Robustness and sensitivity: labor supply elasticity

Panel A: Labor supply elasticity = 2



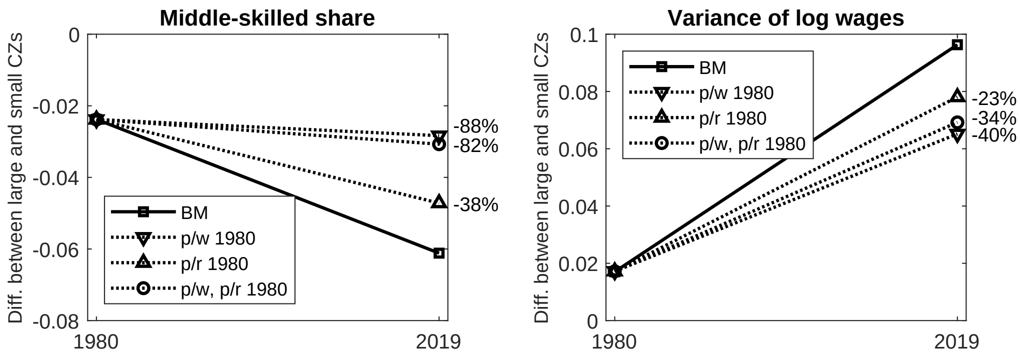
Panel B: Labor supply elasticity = 4



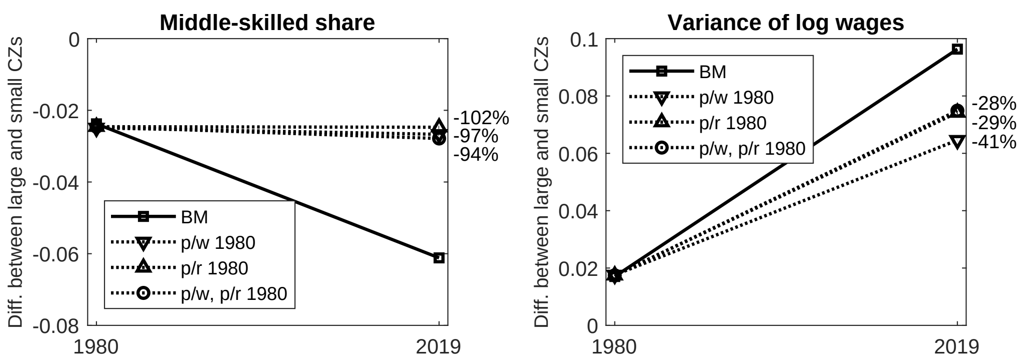
Notes: The left figure in panel A shows the difference between the share of middle-skilled employment in large CZs and the share in small CZs in the benchmark economy (solid line) and in the counterfactual (dotted line) where the scale parameter of the Fréchet distribution ϵ was calibrated to labor supply elasticity of 2. The right figure shows the difference between the variance of log wages in large CZs and the variance in small CZs for the benchmark and the counterfactual economies. Panel B shows the results in counterfactuals where ϵ was calibrated to labor supply elasticity of 4. The numbers to the right of each panel report the percentage change from the benchmark to the counterfactual economy in the gap between large and small CZs.

Figure F.4: Robustness and sensitivity: rent redistribution

Panel A: Local redistribution



Panel B: Nationwide redistribution



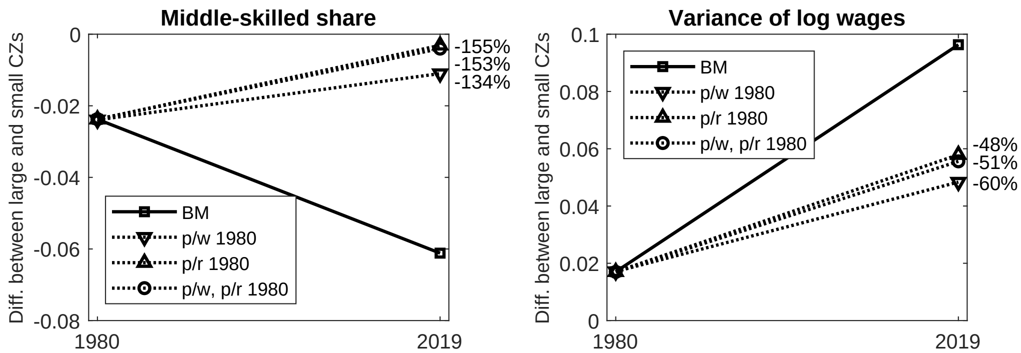
Notes: The left figure in panel A shows the difference between the share of middle-skilled employment in large CZs and the share in small CZs in the benchmark economy (solid line) and in the counterfactual (dotted line) where price-wage and price-rent ratios are kept constant at their 1980 levels, and local rent revenues are redistributed proportionally. The right figure shows the difference between the variance of log wages in large CZs and the variance in small CZs for the benchmark and the counterfactual economies. Panel B shows the results in counterfactuals where nationwide rent revenues are redistributed proportionally among full homeowners. The numbers to the right of each panel report the percentage change from the benchmark to the counterfactual economy in the gap between large and small CZs.

Table F.1: Comparison of welfare gains with and without rent redistribution

	Bench-mark	Const. p/w	Const. p/r	Const. $p/w, p/r$
Welfare without redistribution (main model)	100	112.4	98.6	112.4
Welfare with local rent redistribution	100	114.3	104.2	113.0
Welfare with nationwide rent redistribution	100	114.4	104.4	112.3

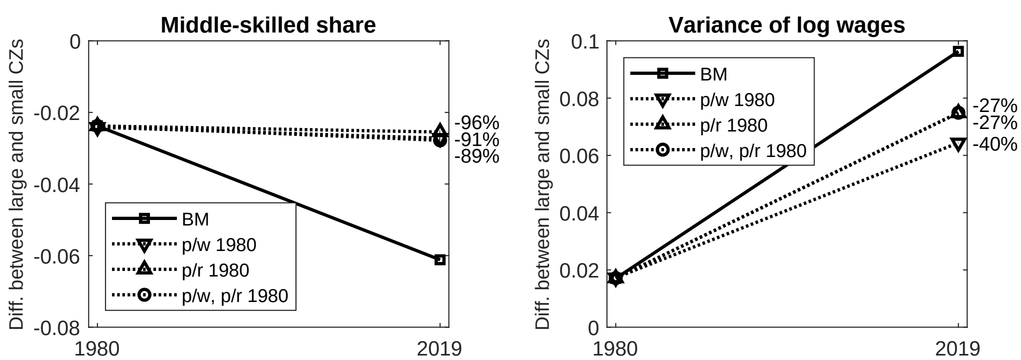
Notes: The table shows welfare gains with and without rent redistribution.

Figure F.5: Robustness and sensitivity: tighter PTI constraint



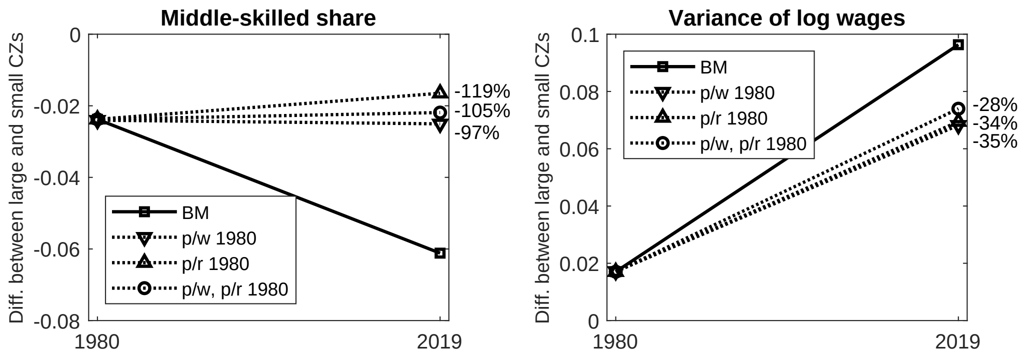
Notes: The left figure shows the difference between the share of middle-skilled employment in large CZs and the share in small CZs in the benchmark economy (solid line) and in the counterfactual (dotted line) with a PTI constraint of 1/3. The right figure shows the difference between the variance of log wages in large CZs and the variance in small CZs for the benchmark and the counterfactual economies. The numbers to the right of each panel report the percentage change from the benchmark to the counterfactual economy in the gap between large and small CZs.

Figure F.6: Robustness and sensitivity: no agglomeration externalities



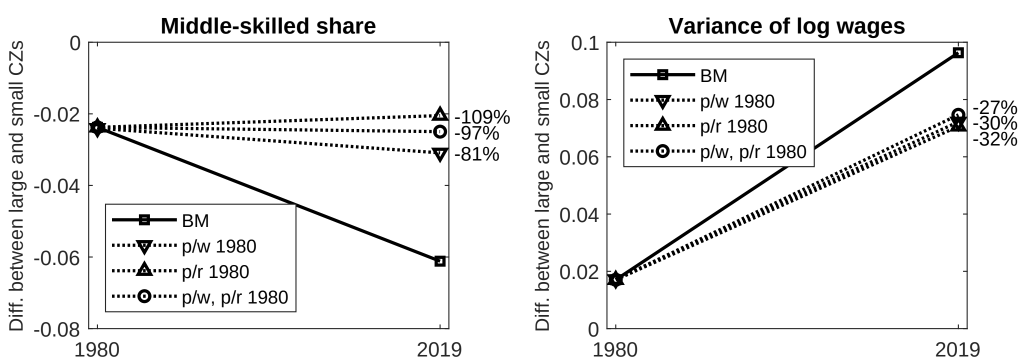
Notes: The left figure shows the difference between the share of middle-skilled employment in large CZs and the share in small CZs in the benchmark economy (solid line) and in the counterfactual (dotted line) where price-wage and price-rent ratios are kept constant at their 1980 levels, and local labor productivity does not depend on local employment, i.e., $\theta = 0$. The right figure shows the difference between the variance of log wages in large CZs and the variance in small CZs for the benchmark and the counterfactual economies. The numbers to the right of each panel report the percentage change from the benchmark to the counterfactual economy in the gap between large and small CZs.

Figure F.7: Robustness and sensitivity: fixed housing expenditure share



Notes: The left figure shows the difference between the share of middle-skilled employment in large CZs and the share in small CZs in the benchmark economy (solid line) and in the counterfactual (dotted line) where price-wage and price-rent ratios are kept constant at their 1980 levels, and the housing expenditure share parameter γ is constant at the 1980 level. The right figure shows the difference between the variance of log wages in large CZs and the variance in small CZs for the benchmark and the counterfactual economies. The numbers to the right of each panel report the percentage change from the benchmark to the counterfactual economy in the gap between large and small CZs.

Figure F.8: Robustness and sensitivity: fixed minimum owner-occupied size



Notes: The left figure shows the difference between the share of middle-skilled employment in large CZs and the share in small CZs in the benchmark economy (solid line) and in the counterfactual (dotted line) where price-wage and price-rent ratios are kept constant at their 1980 levels, and the minimum owner-occupied size $\bar{h}$ is constant at the 1980 level. The right figure shows the difference between the variance of log wages in large CZs and the variance in small CZs for the benchmark and the counterfactual economies. The numbers to the right of each panel report the percentage change from the benchmark to the counterfactual economy in the gap between large and small CZs.