Sorting and the Skill Premium: The Role of Nonhomothetic Housing Demand *

Robert Finlay †   Trevor C. Williams ‡

February 3, 2021

Abstract

Housing expenditure shares decline with income. Less skilled, and therefore poorer, households are more exposed to high housing costs and so choose to live and work in cheaper locations. Skilled households are less exposed and choose more expensive locations. This gap in location choices is ultimately determined by the size of the skill premium, and as the skill premium grows, spatial sorting by skill intensifies. We estimate nonhomothetic preferences over housing using consumption microdata and embed these preferences in a model of spatial sorting with heterogeneous workers. We show analytically that nonhomothetic preferences create a tight link between the skill premium and sorting. Our quantitative model finds that the rising skill premium explains 22% of the increase in spatial sorting by skill observed between 1980 and 2010. Our results highlight a previously unexplored connection between rising income inequality and the spatial distribution of skill and point to the importance of incorporating quantitatively realistic nonhomotheticities in spatial models.

*We thank Costas Arkolakis, Lorenzo Caliendo, Fabian Eckert, Brian Greaney, Sam Kortum, Giuseppe Moscarini, Michael Peters, Vítor Possebom, Jaehee Song, Conor Walsh, and seminar participants at Yale, the Young Economists' Symposium, and the Urban Economics Association PhD Student Workshop for their helpful comments, feedback, and suggestions. We also thank the Yale Department of Economics for financial support in obtaining restricted-access PSID data.

†Yale University. robert.finlay@yale.edu

‡Yale University. trevor.c.williams@yale.edu

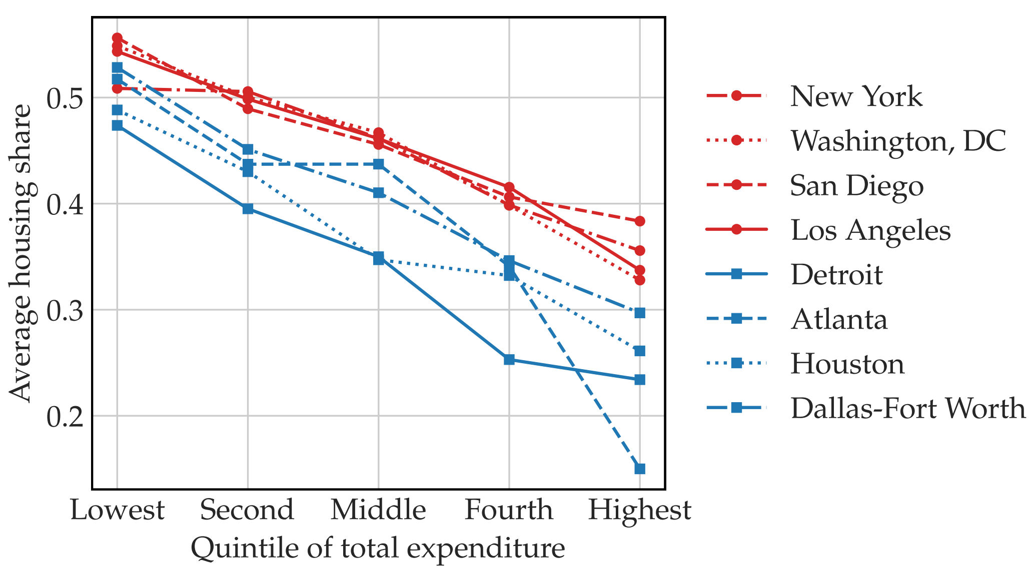
1 Introduction

The gap in wages between workers with and without a college degree—the skill premium—has grown rapidly since 1980 (Acemoglu and Autor 2011). So too has spatial sorting: workers increasingly self-select into different cities on the basis of skill (Berry and Glaeser 2005).

In this paper, we propose and quantify a novel mechanism by which the rising skill premium has caused the increase in spatial sorting. We show income-inelastic housing demand—a first-order feature of the data—translates income inequality at the aggregate level into spatial sorting. Less skilled households, with lower incomes, devote a large share of their budgets to housing. Because they are more sensitive to housing costs, they sort into cheaper locations. By contrast, skilled households are relatively insensitive to housing costs and sort into more expensive locations. Higher inequality increases this difference in sensitivity to housing costs across skill groups and causes ever more divergent location choices. We estimate nonhomothetic housing demand using consumption microdata and show the interaction of nonhomotheticity and the skill premium explains just over a fifth of the increase in spatial sorting observed between 1980 and 2010.

Figure 1 introduces data on housing consumption from the Consumer Expenditure Survey (CEX). We split the national sample into five equally sized quintiles based on total expenditure. Within each bin, we compute the average housing share by city, where the housing share is housing expenditure divided by total expenditure. Housing shares decline with total expenditure, and tend to be higher in more expensive cities at every expenditure level. We formalize these insights

Figure 1: Housing Share by Location and Expenditure



Source: CEX, 2006-2017, renters only. Housing share is defined as housing expenditure divided by total expenditure. The sample is divided into five groups ordered by total expenditure. We rank all cities identified in the CEX according to the BEA's regional price index for housing services. The plot features the four lowest-price metropolitan areas (squares) and the four-highest price areas (circles) in the data.

by estimating nonhomothetic constant elasticity of substitution (NHCES) preferences over housing and nonhousing consumption. Our estimation procedure uses expenditure rather than income data, instruments for expenditure to deal with measurement error, and controls for variation in prices across locations. We show all three features are important for consistently estimating the key preference parameters. The third, in particular, is necessary because high-income households tend to live in expensive housing markets, so that expenditure and price are positively correlated. We find housing demand is moderately income inelastic — for a household in the middle of the expenditure distribution, a 10% increase in total expenditure causes a 2.4% decrease in the housing expenditure share.

To study sorting, we embed nonhomothetic preferences into a standard spatial model. Heterogeneous households with NHCES preferences trade off wages, amenities, and housing costs. Our model is tractable and allows us to prove our main result analytically — increases in the skill premium causes increases in spatial sorting. When preferences are nonhomothetic, the skill premium creates a wedge between the ideal price indices of skilled and unskilled households. Less skilled, and hence lower-income, households are endogenously more sensitive to housing costs relative to wages. An increase in the skill premium increases this wedge in price indices and therefore causes location choices to diverge across skill groups.

To isolate the role of the rising skill premium in the increase in spatial sorting since 1980, we build a quantitative model with heterogeneity in productivities, amenities, and housing costs across locations. Our model is flexible enough to match sorting patterns generated by wages and amenities, as well as those generated by the nonhomothetic preferences we focus on: even so, we show changes in the skill premium only cause changes in sorting when preferences are nonhomothetic. We study the effect of the rising skill premium by scaling the relative productivity of skilled workers in every location so that the skill premium, which is endogenous in the model, remains fixed at its 1980 level. We find that in the absence of the rising skill premium, sorting increases 22% less than it did in the data. Our model explains the remaining 78% of the observed increase as the result of location-specific shocks to productivities, amenities, and housing costs. Finally, we extend our model to highlight other implications of nonhomothetic preferences: we show that housing subsidies targeted towards unskilled households can dampen spatial sorting by skill, that the welfare benefits of skill-neutral productivity shocks are skewed toward skilled households, and that nonhomotheticity amplifies the welfare effect of skill-biased productivity shocks in the presence of agglomeration externalities.

Our estimated housing-demand system is comparable to Albouy, Ehrlich, and Liu (2016). They also estimate NHCES preferences over housing and nonhousing consumption but exploit variation in expenditure across cities and over time rather than at the household level. Also related is Davis and Ortalo-Magné (2011), who find median housing shares vary little across cities and argue this finding is evidence for Cobb-Douglas preferences. Our paper offers an alternative explanation — roughly constant expenditure shares can also be rationalized by offsetting price and income effects.

By incorporating nonhomothetic preferences into a spatial model (see also Schmidheiny (2006), Eckert and Peters (2018), and Handbury (2019)), we emphasize sorting on prices. In related work, Ganong and Shoag (2017) connect changes in housing-supply regulations to the end of regional wage convergence. Diamond (2016) shows endogenous amenities intensify sorting.[1] Whereas these two papers depend on idiosyncratic shocks — to local regulations and local productivities, respectively — we instead consider an aggregate shock that nevertheless has very different consequences across different locations. In this respect, we are similar to Eckert (2018), Giannone (2019), and Eckert, Ganapati, and Walsh (2020), who study urban-biased aggregate productivity shocks. However, in those papers, changes in sorting are driven by wages, whereas our mechanism operates through changes in ideal price indices.

Finally, we contribute to a literature linking the income distribution to the spatial distribution of households. Couture et al. (2019) and Fogli and Guerrieri (2019) study how shocks to the income distribution trigger a resorting of households across neighborhoods within a city, with amplification from endogenous amenities and human capital spillovers, respectively. Gyourko, Mayer, and Sinai (2013) use a stylized model to show how an increase in the number of high-income households can create "superstar" cities and explain diverging house prices. These papers exploit an extreme form of nonhomotheticity — a unit housing requirement — to link the income distribution to the spatial distribution of households. By contrast, our model uses NHCES preferences to flexibly accommodate any degree of nonhomotheticity, and nests both a unit housing requirement and Cobb-Douglas preferences (the standard nonhomothetic and homothetic choices, respectively, in the spatial literature). Estimating preference parameters in the microdata disciplines the model's predictions and allows us to examine how our results vary with the strength of nonhomotheticity.

The rest of the paper proceeds as follows. Section 2 estimates a model of nonhomothetic housing demand. Section 3 uses a simple model to show that when housing demand is income inelastic, changes in the skill premium cause changes in sorting. Section 4 calibrates a quantitative version of this model, and Section 5 uses the calibrated model to quantify the role of the skill premium in increasing sorting. Section 6 highlights extensions and implications of our model, and Section 7 concludes.

2 Estimating the Income Elasticity of Housing Demand

Estimating housing demand presents three main challenges. First, we require consumption data because the key ingredient of the model is the elasticity of housing expenditure with respect to total expenditure.[2] Second, OLS estimates are biased by measurement error in expenditure, so we require an instrument. Finally, and most importantly, the price of housing varies widely across space, and does so in a way that is correlated with household income. Therefore, we need

[1] Our baseline model abstracts from endogenous amenities. In Section 5, we discuss incorporating this feature and show our theoretical results are unchanged by this extension.

[2] We use the familiar term "income elasticity" as shorthand for "expenditure elasticity."

to control for variation in housing prices. As we show below, failing to do so would strongly bias our results toward homotheticity.

2.1 Data

Consumption microdata

We use data from the Panel Study of Income Dynamics (PSID). Since 2005, the PSID has collected information on essentially all consumption covered by the Consumer Expenditure Survey (CEX) (Andreski et al. 2014). Because price data (discussed below) are not available until 2008, we use the 2009-2017 surveys. Our baseline sample is restricted to renting households because they have a clear measure of housing consumption, but we also present similar results using homeowners.

The PSID has two important advantages. First, the restricted-access PSID has county identifiers for the entire sample. Mapping counties to metropolitan statistical areas (MSAs), we can link price data to about 90% of households in the PSID, the remaining 10% being rural households for which we lack housing price indices. By contrast, the CEX has geographic identifiers only for households in large cities, which represent less than half the sample. Second, the PSID follows the same households over time, so we can study how housing expenditure responds to changes in total expenditure within the same household.

Housing prices

Our primary data on housing prices come from the Metropolitan Regional Price Parities produced by the Bureau of Economics Analysis (BEA 2020). For every MSA in the US, the BEA estimates price indices for goods, housing, and all other services besides housing. The housing index is constructed using American Community Survey (ACS) data on rents. By explicitly controlling for differences in housing quality and quantity across regions, the index reflects the price per standardized unit of housing services.

2.2 Preferences

Households have nonhomothetic constant elasticity of substitution (NHCES) preferences over housing and a composite consumption good.[3] The utility U of a household consuming h units of housing and c units of the consumption good is implicitly defined by

$$1 = \Omega^{\frac{1}{\sigma}} \left(hU^{-(1+\epsilon)} \right)^{\frac{\sigma-1}{\sigma}} + (1-\Omega)^{\frac{1}{\sigma}} \left(cU^{-1} \right)^{\frac{\sigma-1}{\sigma}}, \tag{1}$$

[3]For a detailed discussion of NHCES preferences, see Comin, Lashkari, and Mestieri (2021).

where $\epsilon \geq -1$, $0 < \sigma < 1$, and $0 < \Omega < 1$ are parameters.[4] The household maximizes U subject to a standard budget constraint $ph + c \leq e$, where p is the price of housing, e is total expenditure, and the consumption good is the numéraire.

We are interested in seeing how the housing share, which we denote by $\eta \equiv \frac{ph}{e}$, varies with expenditure. NHCES preferences admit a straightforward Hicksian demand function. The log of the relative housing share satisfies

$$\log\left(\frac{\eta}{1-\eta}\right) = \log\left(\frac{\Omega}{1-\Omega}\right) + (1-\sigma)\log p + \epsilon(1-\sigma)\log U. \tag{2}$$

We can see σ determines the sensitivity of housing expenditure to prices, and ϵ then determines how housing expenditure varies with utility. In particular, given our assumption that $\sigma < 1$, $\epsilon < 0$ implies the housing expenditure share falls with utility, whereas $\epsilon > 0$ implies the opposite. Because utility is monotonically increasing in total expenditure, ϵ determines the sign of the income elasticity of housing expenditure. If $\epsilon < 0$, housing is a necessity and its expenditure share falls with total expenditure, whereas if $\epsilon > 0$, it is a luxury and its share rises with total expenditure.

To take equation (2) to the data, we substitute out unobservable utility U in (1).[5] The substitution yields an expression that implicitly defines η as function of expenditure, prices, and parameters:

$$\left(\frac{\eta}{\Omega}\right) = e^{\epsilon(1-\sigma)} p^{1-\sigma} \left(\frac{1-\eta}{1-\Omega}\right)^{1+\epsilon}. \tag{3}$$

NHCES preferences nest two specifications commonly used in the spatial literature. Cobb-Douglas preferences are obtained by taking $\epsilon = 0$ and $\sigma \to 1$ in (3). In this case, the expenditure share is constant, equal to

$$\eta = \Omega. \tag{4}$$

The opposite case, a unit housing requirement, is obtained by taking $\epsilon = -1$ and $\sigma \to 0$. Each household consumes Ω units of housing. In this case, the expenditure share is

$$\eta = \Omega\left(\frac{p}{e}\right). \tag{5}$$

The dashed and dotted lines in Figure 2 depict these two cases. For values of ϵ and σ between these two extremes, housing demand is income- and price-inelastic, but not perfectly so.

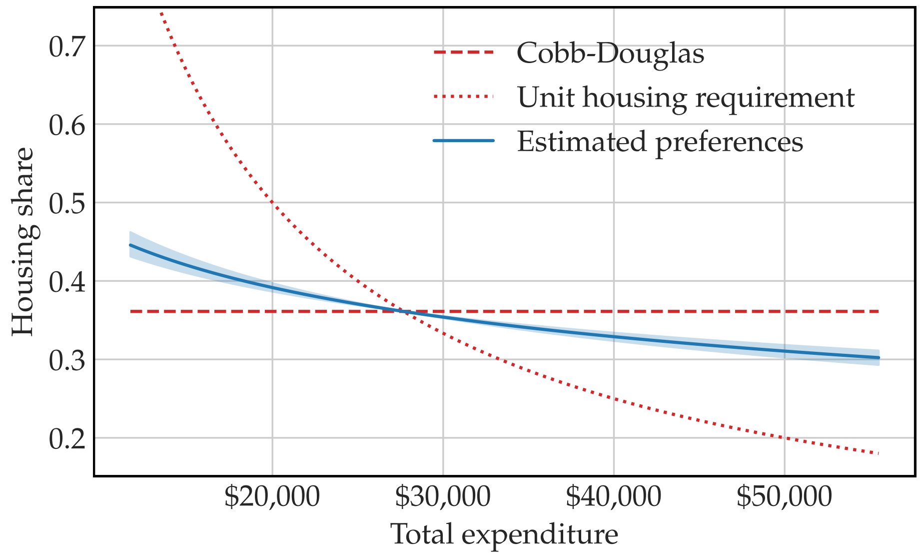
2.3 Estimation

We consider households indexed by i in locations n and years t. Housing prices vary by location and year and are denoted by p_{nt}, whereas the price of the consumption good is assumed not to vary across space and is normalized to 1. We interpret $\varpi \equiv \Omega(1-\Omega)^{-(1+\epsilon)}$ as an idiosyncratic

[4]The restriction $\sigma < 1$ implies housing demand is price-inelastic, which turns out to be the empirically relevant case. We impose $\sigma < 1$ purely for ease of exposition. NHCES preferences in general do allow $\sigma > 1$.

[5]From the Hicksian demand for the consumption good, $U = (1-\eta)^{\frac{1}{1-\sigma}} e(1-\Omega)^{\frac{-1}{1-\sigma}}$.

Figure 2: Housing Expenditure Shares



Notes: "Cobb-Douglas" and "Unit Housing Requirement" plot the preferences described by (4) and (5), respectively. "Estimated Preferences" plots (3) at the parameter values in Table 1, column (4). The shaded area represents a 99% confidence interval. In each case, the scale parameter Ω is chosen to match an expenditure share of 0.37 at the median total expenditure — this value is taken from our baseline sample.

shock to an individual household's taste for housing, so that (3) becomes

$$\eta_{int} = \varpi_{int} e_{int}^{\epsilon(1-\sigma)} p_{nt}^{1-\sigma} \left(1 - \eta_{int}\right)^{1+\epsilon}. \tag{6}$$

To build intuition for our estimation strategy, we log-linearize (6) around the median housing share $\bar{\eta}$ to obtain

$$\hat{\eta}_{int} = \left(\frac{1-\bar{\eta}}{1-\bar{\eta}+(\epsilon+1)\bar{\eta}}\right)\left(\hat{\varpi}_{int} + \epsilon(1-\sigma)\hat{e}_{int} + (1-\sigma)\hat{p}_{nt}\right), \tag{7}$$

where $\hat{x}$ denotes the log deviation of a variable x from its median. Equation (7) simplifies to

$$\hat{\eta}_{int} = \omega_{int} + \beta\hat{e}_{int} + \psi\hat{p}_{nt}, \tag{8}$$

where $\omega_{int} \equiv \left(\frac{1-\bar{\eta}}{1-\bar{\eta}+(\epsilon+1)\bar{\eta}}\right)\hat{\varpi}_{int}$ and β and ψ are defined analogously. Under the null of homothetic preferences, $\epsilon = \beta = 0$. We bring (8) to the data by modeling the demand shifter ω_{int} as a function of observables, year fixed effects, and an additive error. Formally,

$$\hat{\eta}_{int} = \omega_t + \omega' X_{int} + \beta\hat{e}_{int} + \psi\hat{p}_{nt} + \xi_{int}, \tag{9}$$

where X_{int} is a vector of the following demographic characteristics: the age, gender, and race of the household head; household size; and the number of earners in the household. We observe total expenditure e, the housing expenditure share η, and prices p. The error term ξ_{int} represents measurement error in expenditure and random shocks to housing demand, which are assumed to be uncorrelated with expenditure and prices conditional on the controls.

Table 1, columns (1) - (4), show the estimates of equation (9). Column (1) estimates equation (9) by OLS without controlling for price $\hat{p}_{nt}$. The point estimate indicates significant nonhomotheticity, but two sources of bias are evident. First, measurement error in expenditure is likely to bias $\hat{\beta}$ downwards.[6] Second, a positive correlation between prices and expenditure (reflecting the sorting of high-income households into-high price MSAs) will bias $\hat{\beta}$ upwards.

Column (2) addresses the first source of bias by instrumenting for log expenditure using log income. The exclusion restriction is that income is unrelated to the housing share, conditional on the true level of expenditure. Consistent with the argument above, $\hat{\beta}$ rises toward zero. This specification is comparable to those used in Davis and Ortalo-Magné (2011) and Aguiar and Bils

Table 1: Preference Estimates

Dependent variable: Log housing share

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) GMM	(5) 2SLS
Log expenditure	-0.248***	-0.133***	-0.238***		-0.259***
	(0.018)	(0.028)	(0.031)		(0.076)
Log price			0.365***		0.364***
			(0.031)		(0.067)
Implied ϵ			-0.642***	-0.679***	-0.713***
			(0.079)	(0.074)	(0.229)
Implied σ			0.571***	0.563***	0.583***
			(0.043)	(0.044)	(0.102)
Demographic controls	✓	✓	✓	✓	
Household FE					✓
Year FE	✓	✓	✓	✓	✓
R^2	0.09	0.08	0.15	-	0.13
First-stage F-statistic	-	1,196.4	883.0	-	98.1
N	10,129	10,129	8,923	8,923	7,017
No. of clusters	4,816	4,816	4,229	4,229	2,377

Source: PSID and BEA

Note: Renters only. Instrument is log family income. Demographic controls are bins for family size, number of earners, and sex, race, and age of household head. Standard errors clustered at household level. See Appendix A for further details of sample construction. Null hypotheses are $\epsilon = 0$ and $\sigma = 1$, respectively.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

[6]Because expenditure appears in the denominator of $\hat{\eta}$, the bias in $\hat{\beta}$ is not standard classical measurement error. See Appendix B for a short proof.

(2015), and similar to the results in those papers, it finds only a small amount of nonhomotheticity. However, column (3) additionally controls for MSA-level prices and the point estimate falls to -0.238, indicating significant nonhomotheticity. In column (3), we also use equation (7) to back out estimates for the structural parameters ϵ and σ.

Columns (2) and (3) show that omitting prices biases estimates toward homotheticity. This observation is consistent with high-income households sorting into high price locations — exactly the pattern our model predicts in the next section. This discussion also suggests how our finding of significant nonhomotheticity can be reconciled with the fact in Davis and Ortalo-Magné (2011) that median housing shares are roughly constant across cites.[7] The housing expenditure share in any location represents a mixture of income and price effects. If high-price locations disproportionately attract high-income households, these two forces will offset one another, leading to similar housing shares across cities.

Column (4) is our preferred specification. Here, we estimate the nonlinear equation (3) directly by GMM.[8] Similar to column (3), we instrument for expenditure using income, use price data, and allow Ω_{int} to vary with demographic characteristics and year. The implied $\hat{\epsilon}$ and $\hat{\sigma}$ are close to their values in column (4). Moreover, $\hat{\epsilon}$ is significantly different from 0 at the 1% level, allowing us to reject the null hypothesis of homothetic preferences. Because $\hat{\epsilon}$ is far from -1, we can also reject a unit housing requirement. The solid line in Figure 2 plots our estimated housing-demand function.

Instrumenting for expenditure using income addresses measurement error, but other threats to identification exist. Suppose ability is positively correlated with income and expenditure but negatively correlated with taste for housing. For example, a highly paid consultant who anticipates being out of town most of the week on business trips may choose a relatively small apartment since she will spend little time in it. In this case, we would observe a negative relationship between expenditure and the housing share in the cross section, but changes in the consultant's total expenditure might not prompt her to change her housing expenditure share.

To investigate this possibility, column (5) explores an alternative specification for the shifter ω_{int}. We suppose it can be written as a household fixed effect and an additive error term. Formally,

$$\hat{\eta}_{int} = \omega_i + \omega_t + \beta \hat{e}_{int} + \psi \hat{p}_{nt} + \xi_{int}. \tag{10}$$

In this specification, β is identified using only within-household variation in expenditure. It allows us to distinguish our explicitly nonhomothetic preferences from alternative models in which expenditure shares differ across different "types" of household, but are fixed for a given type over time (e.g. Diamond (2016)). The point estimate for β is very similar to our baseline linearized specification in column (3). The standard error of the implied structural parameter ϵ is large, but

[7] See Appendix B for a more detailed discussion of the relationship between our results and Davis and Ortalo-Magné (2011), and Appendix Table B.5 for other estimates of the income elasticity in the literature.

[8] In stating our preferences, we imposed $\epsilon > -1$ and $0 < \sigma < 1$. We do not impose these restrictions in our estimation procedure, but they are satisfied by the values obtained in column (4).

the point estimate is also similar to column (4). We conclude that preferences are nonhomothetic: even within a single household, changes in total expenditure cause changes in the housing expenditure share.

2.3.1 *Alternative Specifications*

We explore a number of alternative specifications and data sources in Appendix B.

In Table B.1, we present alternative specifications in the PSID. The results do not change when we control for local non-housing prices, when we use an alternative price index from Zillow, or when we use alternative instruments for expenditure. In particular, we consider instrumenting using education rather than income. This approach leverages the fact that education changes a household's permanent income. If our basline results just reflected sluggish adjustment of housing expenditure to temporary income shocks — because, for example, rental contracts can only be changed slowly — shocks to permanent income should not cause any changes in the housing expenditure share. But, in fact, we continue to find very similar results to our baseline specification in column (4).

We replicate our main results using the CEX. We find similar parameter estimates in Table B.2. We then extend our results to homeowners using the CEX, which has superior coverage of homeowners relative to the PSID. The estimated parameters look very similar when we include homeowners.

Finally, our choice of NHCES preferences imposes a particular functional form on the relationship between total expenditure and the housing expenditure share η. To assess the sensitivity of our results to this restriction, in Appendix B.2.4, we estimate a semiparametric model relating total expenditure to η. We continue to find the housing expenditure share falls as total expenditure rises, and the function estimated using this procedure is broadly similar to the one estimated in column (4) of Table 1.

3 Model

Having established that housing demand is income inelastic, we now explore the implications for spatial sorting by skill. We focus on how income-inelastic housing demand connects changes in the skill premium to changes in sorting. Starting with a simple model, we analytically characterize the relationship between the skill premium and sorting. Then we construct a more realistic quantitative model that we use in counterfactuals in Section 5.

3.1 Simple Model

Production and Wages

There are two types of household, skilled and unskilled, with types denoted by $i = s, u$. Households supply labor to tradable-goods producers in their home location, denoted by n. Firms are perfectly competitive and produce using skilled and unskilled labor according to the function

$$F_n(l_{sn}, l_{un}) = z_n(A \cdot l_{sn} + l_{un}). \tag{11}$$

Skilled and unskilled labor are perfect substitutes, and their relative productivities do not vary across locations. This assumption implies that the skill premium is exogenous and equal to A in every location. We therefore refer to A as the skill premium. Households do not save, so wages w_{in} are exactly equal to expenditure e_{in}. Expenditures and wages satisfy

$$e_{sn} = w_{sn} = z_n A \tag{12}$$

$$e_{un} = w_{un} = z_n. \tag{13}$$

Housing Supply

Housing is supplied perfectly elastically at exogenous price p_n. Housing is produced using consumption goods.

Location Choice and Preferences

We first describe the problem of a household in a given location, then turn to the household's choice of location. Households have NHCES preferences as in (1). The utility of a household of type i in location n, denoted by v_{in}, is

$$\begin{aligned} v_{in} &= \max_{c,h} \quad U \\ \text{s.t} \quad 1 &= \Omega^{\frac{1}{\sigma}} \left(hU^{-(1+\epsilon)} \right)^{\frac{\sigma-1}{\sigma}} + (1-\Omega)^{\frac{1}{\sigma}} \left(cU^{-1} \right)^{\frac{\sigma-1}{\sigma}}, \\ e_{in} &= c + p_n h. \end{aligned} \tag{14}$$

The solution to this problem yields housing expenditure shares η_{in}. If $\epsilon < 0$, as we estimated in Section 2, η_{in} is a decreasing function of total expenditure e_{in}.

We close the model by assuming location n's share of total employment of type i is an isoelastic function of utility v_{in} given by

$$l_{in} = \frac{v_{in}^{\theta} B_n}{\sum_m v_{im}^{\theta} B_m} L_i, \tag{15}$$

where L_i is the exogenous national population of households of type i.[9] We refer to B_n as the amenity value of location n and θ as the migration elasticity. Amenities do not differ by type i, an assumption we relax later.

[9]One microfoundation of the employment shares (15), common in quantitative spatial models, is that each household draws an n-vector of idiosyncratic location preference shocks from independent Fréchet distributions with scale B_n and shape θ (Allen and Arkolakis 2014; Redding 2016). We do not assume a particular microfoundation.

Equilibrium

Given parameters $\epsilon, \sigma, \Omega, \theta$, location-specific fundamentals (z_n, B_n, p_n), the skill premium A, and labor supplies (L_u, L_s), an equilibrium is a vector of populations l_{in}, wages w_{in}, and total expenditures e_{in} satisfying (12), (13), (14) and (15).

3.2 Analytical Results

We now characterize spatial sorting by skill.[10] We define sorting in terms of the log skill ratio in each location, which we denote by s_n and define as the log ratio of skilled to unskilled households in location n:

$$s_n = \log\left(\frac{l_{sn}}{l_{un}}\right).$$

Our preferred measure of sorting, which we denote by S, is the variance of the log skill ratio,

$$S = \text{Var}\left(s_n\right).$$

S is zero when skilled workers are distributed in proportion to unskilled workers across space, and rises as they become more clustered. Additionally, S is invariant to proportional increases in the number of skilled workers in all locations. This invariance property is desirable because the number of skilled workers in the US has grown relative to the number of unskilled workers since 1980. In Section 5, we also report results for other measures of sorting.

The Determinants of Sorting

Equation (15) yields a simple expression for s_n in terms of utilities v_{in}:

$$s_n = \zeta + \theta \log\left(\frac{v_{sn}}{v_{un}}\right) \tag{16}$$

where ζ is a function of fundamentals that does not vary across locations. Note B_n is absent: because amenities do not differ by type, they do not drive sorting.

To relate v_{in} to wages and prices, consider the ideal price indices, P_{in}, which satisfy

$$v_{in} = \frac{w_{in}}{P_{in}}, \tag{17}$$

where we are exploiting the fact that in the simple model wages are equal to expenditure. Substituting (17) into (16) and using (12) and (13) to replace wages with productivities yields

$$s_n = \zeta + \theta \log A - \theta \log\left(\frac{P_{sn}}{P_{un}}\right). \tag{18}$$

[10] As in Section 2, we assume $\sigma \in (0,1)$. Proofs of all the statements in this section can be found in Appendix C.

This expression clarifies the sources of sorting in our model. Wages do not cause sorting conditional on the price indices, because the ratio of skilled to unskilled wages is constant across locations. Instead, sorting is only a result of differences in the ideal price indices.

To see how these price indices depend on the wages and prices, we use expressions for P_{un} and P_{sn} implied by our NHCES preferences:

$$P_{un} = \left((1 - \Omega) + \Omega \left(\frac{z_n}{P_{un}} \right)^{\epsilon(1-\sigma)} p_n^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{19}$$

$$P_{sn} = \left((1 - \Omega) + \Omega A^{\epsilon(1-\sigma)} \left(\frac{z_n}{P_{sn}} \right)^{\epsilon(1-\sigma)} p_n^{1-\sigma} \right)^{\frac{1}{1-\sigma}}. \tag{20}$$

These price indices resemble ordinary CES price indices, except the weight placed on housing is a function of real income as long as $\epsilon \neq 0$. In particular, if $\epsilon < 0$ — as we found in Section 2 — this weight is decreasing in real income. Moreover, the housing weight for skilled workers is always lower than for unskilled workers and shrinks as A grows. Inspection of (19) and (20) shows each can be written as a function $P_i(c_n)$, where $c_n \equiv z_n^{\epsilon} p_n$ is defined as productivity-adjusted housing cost. Intuitively, z_n^{ϵ} should appear in c_n because a higher income lowers the burden of higher house prices when housing demand is income inelastic.

We consider the implications for sorting, starting with the homothetic case. In this case, $\epsilon = 0$ and $P_u(c) = P_s(c)$ for all c. Inspection of (18) then shows s_n does not depend on n. Skilled and unskilled workers are distributed in proportion to one another in every location and $S = 0$.

When preferences are nonhomothetic and $\epsilon < 0$, $P_u(c)$ is a steeper function of c than $P_s(c)$. As c grows, the wedge between unskilled and skilled price indices grows and high c locations look increasingly unattractive to unskilled workers. Lemma 1 formalizes this argument.

Lemma 1 *Suppose housing demand is income inelastic so that $\epsilon < 0$. Then, $\log P_u(c) - \log P_s(c)$ is a strictly increasing function of productivity-adjusted housing cost c. Equation (18) then implies the skill ratio s_n is a strictly increasing function of c.*

Panel (a) of Figure 3 illustrates this mechanism. The ideal price index for unskilled households is a steep function of c, whereas the ideal price index for skilled households is flatter. By (18), the skill ratio s_n in Panel (b) is just an affine transformation of the gap between the dotted and solid lines in Panel (a).

Sorting and Changes in the Skill Premium

So far, we have focused on the level of sorting. Now we turn to changes in sorting caused by changes in the skill premium. Proposition 1 states our main result by studying a small increase in the skill premium, $d \log A > 0$.

Proposition 1 *Suppose housing demand is income-inelastic so that $\epsilon < 0$. Consider an increase in the skill premium, $d \log A > 0$. Then, ds_n is a strictly increasing function of s_n. Sorting rises, $dS > 0$. If, instead, $\epsilon = 0$, then $ds_n = 0$ for all n and $dS = 0$.*

Equations (19) and (20) show skilled and unskilled ideal price indices differ only because of the skill premium A. As A rises, skilled households place less weight on housing costs and become more willing to live in locations with a high productivity-adjusted housing cost c. This flattening of the ideal price index is illustrated by the dashed line in Panel (a) of Figure 3, which increases the skill premium relative to the solid line. The gap between P_u and P_s grows and (18) tells us s_n must then become a steeper function of c, as shown by the dashed line in Panel (b). Skilled households, newly insensitive to housing costs, flee cheap locations toward the left of Panel (b), and instead cluster in expensive ones on the right. The higher skill premium reinforces pre-existing patterns of sorting and so S rises.

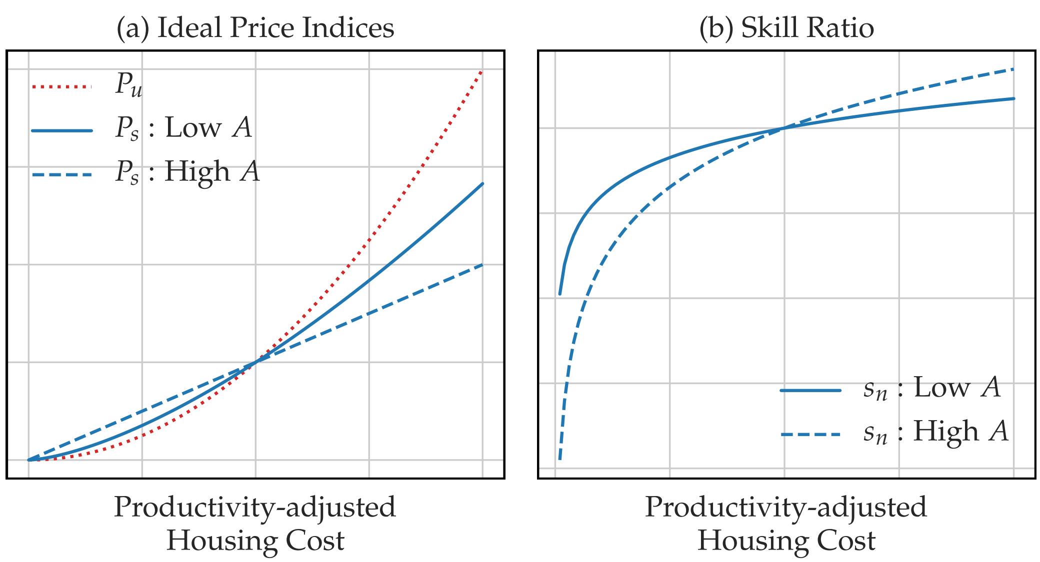
3.3 Quantitative Model

To take the model to the data, we enrich it on several dimensions.

The simple model deliberately shut down sorting based on wages. We relax this assumption by replacing (11) with a CES production function,

$$F_n(l_s, l_u) = Zz_n \left((Aa_n l_s)^{\frac{\rho-1}{\rho}} + l_u^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}, \tag{21}$$

Figure 3: Ideal Prices Indices and the Skill Ratio



Note: The solid lines in Panel (a) plot the price indices defined by (19) and (20) as functions of productivity-adjusted housing cost. The dashed line plots (20) again but uses a higher value of the skill premium A. The solid line in Panel (b) plots the log skill ratio s_n given by (18), corresponding to the solid lines in Panel (a). The dashed line plots s_n but uses the value for P_s given by the dashed line in Panel (a).

where ρ is the elasticity of substitution between skilled and unskilled labor. Implied wages are

$$w_{un} = Zz_n l_{un}^{\frac{-1}{\rho}} \left((Aa_n l_{sn})^{\frac{\rho-1}{\rho}} + l_{un}^{\frac{\rho-1}{\rho}} \right)^{\frac{1}{\rho-1}} \tag{22}$$

$$w_{sn} = (Aa_n)^{\frac{\rho-1}{\rho}} Zz_n l_{sn}^{\frac{-1}{\rho}} \left((Aa_n l_{sn})^{\frac{\rho-1}{\rho}} + l_{un}^{\frac{\rho-1}{\rho}} \right)^{\frac{1}{\rho-1}}. \tag{23}$$

As in the simple model, the quantitative model contains a location-specific productivity shock z_n and an economy-wide skill bias A. We additionally allow for location-specific skill bias using the shifter a_n, so that skilled households may have a comparative advantage working in San Francisco relative to, say, Detroit. The economy-wide productivity shifter Z is for notational convenience when we conduct counterfactuals.

We allow amenities B_{in} to differ by skill, so that (15) becomes

$$l_{in} = \frac{v_{in}^{\theta} B_{in}}{\sum_m v_{im}^{\theta} B_{im}} L_i. \tag{24}$$

With these modifications, our model can capture changes in sorting driven by location-specific changes in a location's attractiveness to skilled versus unskilled households, through both wages and amenities. But as in the simple model, our focus remains on changes in overall sorting driven by location-neutral changes in A.

In the simple model, the price of housing p_n was exogenous. In reality, increases in A push skilled households toward expensive cities, putting upward pressure on housing costs and crowding out unskilled households. The quantitative model captures this feedback to house prices by including inelastic housing supply as in Hsieh and Moretti (2019). The price of housing in location n is given by

$$p_n = \Pi_n (HD_n)^{\gamma_n}, \tag{25}$$

where HD_n is (physical) housing demand in n and Π_n is an exogenous price shifter. γ_n, the inverse elasticity of housing supply, is allowed to vary by location to reflect different physical or regulatory constraints on building. Housing demand is the sum of housing consumption by both types of households:

$$HD_n = p_n^{-1} \sum_i \eta_{in} e_{in} l_{in}. \tag{26}$$

Finally, we allow for progressive taxation following Heathcote, Storesletten, and Violante (2017). Expenditure equals after tax income, defined as

$$e_{in} = \lambda w_{in}^{1-\tau}. \tag{27}$$

τ determines the progressivity of the tax system and λ is chosen so that the government budget balances. We introduce this feature because the utility-maximization problem described in (14) depends on total expenditure e_{in}, but in the data used to calibrate the model, we observe only income w_{in}. By explicitly modeling the tax system, we account for an important wedge between income and expenditure.[11]

Equilibrium

Given parameters $\epsilon, \sigma, \Omega, \theta, \rho, \tau, \lambda, \{\gamma_n\}$, location-specific fundamentals $(a_n, z_n, B_{un}, B_{sn}, \Pi_n)$ for all n, aggregate fundamentals (Z, A), and labor supplies (L_u, L_s), an equilibrium is a vector of populations l_{in}, wages w_{in}, total expenditures e_{in}, expenditure shares η_{in}, housing demands HD_n, and prices p_n satisfying equations (14), (23), (24), (25), (26), and (27).

A Neutrality Result

We conclude this section by extending part of Proposition 1 to the quantitative model. Crucially, although our quantitative model generates rich patterns of sorting based on location-specific skill biases a_n and amenities B_{in}, homotheticity shuts down any relationship between A and sorting. Proposition 2 formalizes this point.

Proposition 2 *Suppose $\epsilon = 0$ so that preferences are homothetic. Then, S, the level of sorting, does not depend on A.*

See Appendix C.3 for a proof. To gain intuition for this result, return to the expression for the log skill ratio s_n derived in the simple model. In the quantitative model with $\epsilon = 0$, (16) is modified to

$$s_n = \zeta + \theta \log\left(\frac{e_{sn}}{e_{un}}\right) + \log\left(\frac{B_{sn}}{B_{un}}\right). \tag{28}$$

Using (27) to replace expenditures with wages, and then using (23) to replace wages with productivities and the skill ratio, we obtain

$$s_n = \tilde{\zeta} + \tilde{\theta}_1 \log a_n + \tilde{\theta}_2 \log\left(\frac{B_{sn}}{B_{un}}\right), \tag{29}$$

where $\tilde{\theta}_1$ and $\tilde{\theta}_2$ are functions of θ, τ and ρ. The striking feature of (29) is that s_n is entirely determined by location-specific fundamentals a_n and B_{in}, up to the additive constant $\tilde{\zeta}$. Changes in A have no impact on the distribution of skill ratios, and thus no impact on sorting, when preferences are homothetic. Proposition 2 is useful because it implies any changes in sorting observed in our quantitative model are ultimately the result of nonhomothetic housing demand.

[11] Another obvious source of differences between total expenditure and income is saving. However, our model is static and does not have a role for saving. In our counterfactual, we only consider shocks to the permanent incomes of households. A standard model of consumption and saving would predict the elasticity of consumption expenditure to permanent income shocks to be 1 — but see Straub (2019) for evidence for an elasticity other than 1.

Table 2: Parameters

Parameter	Value	Source
ϵ	-0.68	PSID
σ	0.56	PSID
ρ	3.85	Card (2009)
τ	0.17	PSID
$\{\gamma_n\}$	0.78[a]	Census
θ	3.89	Indirect inference

[a] Employment-weighted mean

4 Calibration

Section 2 estimated nonhomothetic preferences over housing consumption, and Section 3 embedded these preferences in a simple model to make our key theoretical point — increases in the skill premium cause increases in spatial sorting. To determine the importance of this force in explaining trends in sorting since 1980, we now calibrate the quantitative model. The crucial preference parameters — ϵ and σ — are set at the values obtained in Section 2. The scale parameter Ω is not identified separately from the scale of prices, so we normalize it to match the aggregate housing share in each year. The calibration of the remaining parameters is discussed in detail below: the elasticity of substitution ρ is calibrated from the literature; we derive estimating equations from the model for the tax-progressivity parameter τ and the housing-supply elasticities γ_n; and we calibrate the migration elasticity θ by targeting literature estimates. The results of this exercise are summarized in Table 2.

4.1 Data

Location-level information on wages, rents and employment are from IPUMS (Ruggles et al. 2020). We use the 5% population samples of the 1980, 1990, and 2000 decennial censuses and the 3% population sample from the 2009-2011 ACS. We have a balanced panel of 269 locations: 219 MSAs and the 50 non-metropolitan portions of states. Our census sample is made up of prime-age adults who report strong labor-force attachment. See Appendix A for more details.

4.2 Parameters

Elasticity of substitution The production side of the model is standard and we externally calibrate $\rho = 3.85$ to match Card (2009).[12] That paper estimates the elasticity of substitution between workers of different skill groups at the MSA level using immigration as an instrument for labor-supply changes. The elasticity is larger than canonical estimates from Katz and Murphy (1992) and Acemoglu and Autor (2011), who report values close to 1.6. However, Katz and Murphy (1992) estimate an aggregate production function fitted to time-series data, whereas Card (2009) estimates a

[12] See Table 5, column (7), in Card (2009) for the negative inverse elasticity of -0.26.

city-level production function fitted to cross-sectional data. Studies estimating the elasticity of substitution at the city level tend to find values between 3 and 5 (Bound et al. 2004; Beaudry, Doms, and Lewis 2010; Baum-Snow, Freedman, and Pavan 2018; Eckert, Ganapati, and Walsh 2020). [13]

Tax system To calibrate the progressivity parameter τ, we follow Heathcote, Storesletten, and Violante (2017). From (27), log post-tax income for individual i in location n and year t is equal to

$$\log y_{it} = \log \lambda_t + (1-\tau)\log w_{it}. \tag{30}$$

Regressing log post-tax income on log pre-tax income and a year fixed effect in the PSID for 1980, 1990, 2000, and 2010 yields $\hat{\tau} = 0.174$, close to the value of 0.181 reported by Heathcote, Storesletten, and Violante (2017) for 1978-2006.

Housing-supply elasticities The housing-supply equation (25) is specified in terms of the physical quantity of housing, HD_n, which is not observed. To obtain an estimating equation, rewrite (25) with HD_n expressed in terms of price and expenditure:

$$p_n = \tilde{\Pi}_n \left(\sum_i R_{in} l_{in} \right)^{\chi_n},$$

where $\chi_n = \frac{\gamma_n}{1+\gamma_n}$, $\tilde{\Pi}_n = \Pi_n^{\frac{1}{1+\gamma_n}}$, and $R_{in} = \eta_{in} e_{in}$ is the rental expenditure of households of type i in location n. Taking logs and differencing over time yields an equation linear in χ_n

$$\Delta \log p_n = \Delta \log \tilde{\Pi}_n + \chi_n \Delta \log \left(\sum_i R_{in} l_{in} \right).$$

Following Saiz (2010), we parameterize χ_n as a function of geographical and regulatory constraints, $\chi_n = \chi + \chi_L UNAVAL_n + \chi_R WRLURI_n$. $UNAVAL_n$ is a measure of geographic constraints from Saiz (2010) and $WRLURI_n$ is the Wharton Residential Land Use Regulation Index developed by Gyourko, Saiz, and Summers (2008). Substituting the expression χ_n into (4.2) yields an estimating equation for χ, χ_L and χ_R:

$$\Delta \log p_n = \Delta \log \tilde{\Pi}_n + (\chi + \chi_L UNAVAL_n + \chi_R WRLURI_n)\, \Delta \log \left(\sum_i R_{in} l_{in} \right).$$

We use data on rents and wages from the 1980 and 2010 censuses to construct $\Delta \log p_n$ using (3). Because rents and employment are endogenous, we follow Diamond (2016) and use Bartik shocks, as well as their interactions with $UNAVAL_n$ and $WRLURI_n$, as instruments for labor demand. We set $\gamma_n = \frac{\chi_n}{1-\chi_n}$. The employment-weighted average of the γ_n obtained using this procedure is 0.78, very close to the value of 0.77 reported by Saiz (2010).

[13] An exception is Diamond (2016), who estimates an elasticity close to 1.6 in line with the time-series results.

Migration elasticity We match the long-run employment elasticity reported by Hornbeck and Moretti (2019), who use a measure of TFP recovered from plant-level data as an instrument for labor-demand shocks. They find an elasticity of 2.76. We solve for the value of θ that produces the same response in our model. This procedure implies $\theta = 3.89$. Although not strictly comparable because of the difference in utility functions and the presence of progressive taxation, this number is similar to the values usually reported in the literature. For example, Hsieh and Moretti (2019) calibrate θ to be 3.33.

4.3 Fundamentals

We now turn to the fundamentals of the quantitative model: the location-specific productivity, amenity and housing-supply shifters $\left(a_n^t, z_n^t, B_{un}^t, B_{sn}^t, \Pi_n^t\right)$ and the aggregate productivity parameters A^t and Z^t. Note we have added a time superscript because we allow all fundamentals to vary by year. We obtain these fundamentals for each year $t \in \{1980, 1990, 2000, 2010\}$ by inverting the model so that it exactly matches census data on wages and employment by skill and MSA and average rental expenditures by MSA.[14]

5 The Skill Premium and Sorting

How did the increase in the aggregate skill premium alter the spatial distribution of skill 1980-2010? To answer this question, we perform a counterfactual experiment using the quantitative model developed in Section 3 and calibrated in Section 4. For each census year between 1980 and 2010, we solve for the spatial distributions of skilled and unskilled workers that would have occurred had skill premium remained constant at its 1980 level. In the implementation, we choose values for aggregate productivity Z^t and aggregate skill bias A^t such that (i) average skilled wages grow at the same rate as average unskilled wages, and (ii) average unskilled wages grow at the same rate as they did in the data. All location-specific fundamentals — productivities, skill biases, amenities, and housing-supply shifters — evolve as they did in the data. Because only Z^t and A^t are changed, the difference between the data and the model represents a *location-neutral* shock. This difference identifies the causal effect of the rising aggregate skill premium.

5.1 Main Results

Figure 4 shows our main result. In the absence of a rising national skill premium, sorting only rises by 25.5%, whereas in the data it rose by 32.6%. We conclude that without the increase in the skill premium, sorting would have risen by 7.1 percentage points less — 22% of the overall increase between 1980 and 2010. The shaded area in Figure 4 shows this difference. Our model explains

[14] In each year, A^t and Z^t are not separately identified from the scale of a_n^t and z_n^t. We normalize the mean of a_n^t and z_n^t to each be 1 in every year, which has no impact on our results.

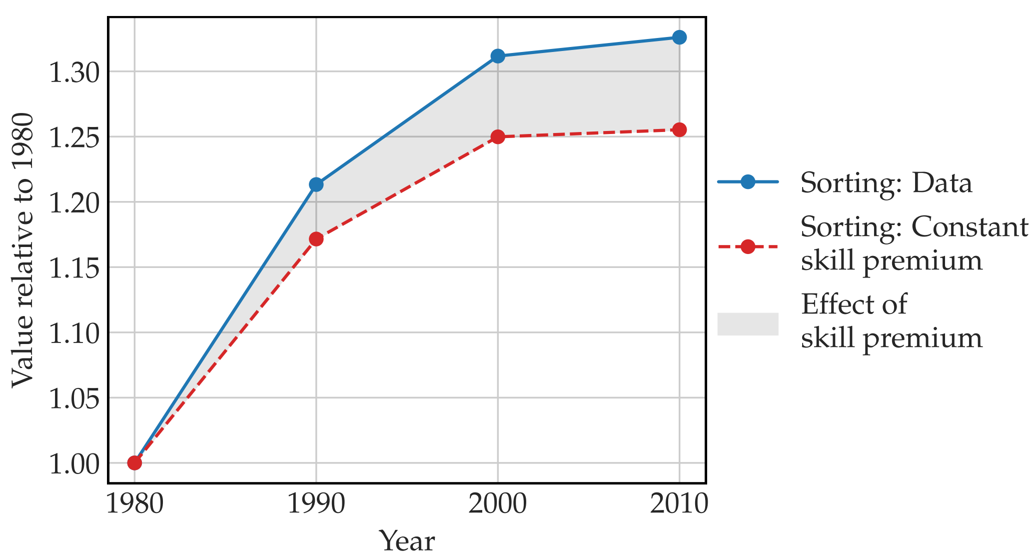
the remaining 78% as the result of idiosyncratic amenity, productivity, and housing-supply shocks such as those highlighted by Diamond (2016) or Ganong and Shoag (2017).

Figure 5 explores the model mechanism. In the simple model in Section 3, the key driver of sorting was the productivity-adjusted cost of housing, c_n. There, we showed that increases in the skill premium make skilled households less sensitive to housing costs, thereby pushing them toward high c_n locations. Because these locations are already relatively skill intensive, sorting rises. Figure 5 shows each of these steps in the quantitative model. Panel (a) plots the causal effect of the higher skill premium in 2010 against the productivity-adjusted housing cost in 2010. As in the simple model, the higher skill premium encouraged skilled workers to move toward high cost locations. Panel (b) translates the skill-housing cost relationship into a statement about sorting by plotting the causal effect of the higher skill premium against city-level skill ratios in 2010. The positive slope of the regression line shows skill-intensive locations generally gained skilled workers as a result of the rising skill premium, and so sorting increased. The relationship in (b) is noisy because in the quantitative model, c_n is only one determinant of sorting, alongside wages and amenities.

The Role of Nonhomotheticity

Here, we consider how the effect of the skill premium depends on ϵ. Recall that given a price elasticity parameter σ, ϵ determines the income elasticity of housing demand. When $\epsilon = 0$, housing expenditure shares do not depend on income. When ϵ becomes more negative, housing is

Figure 4: Sorting since 1980



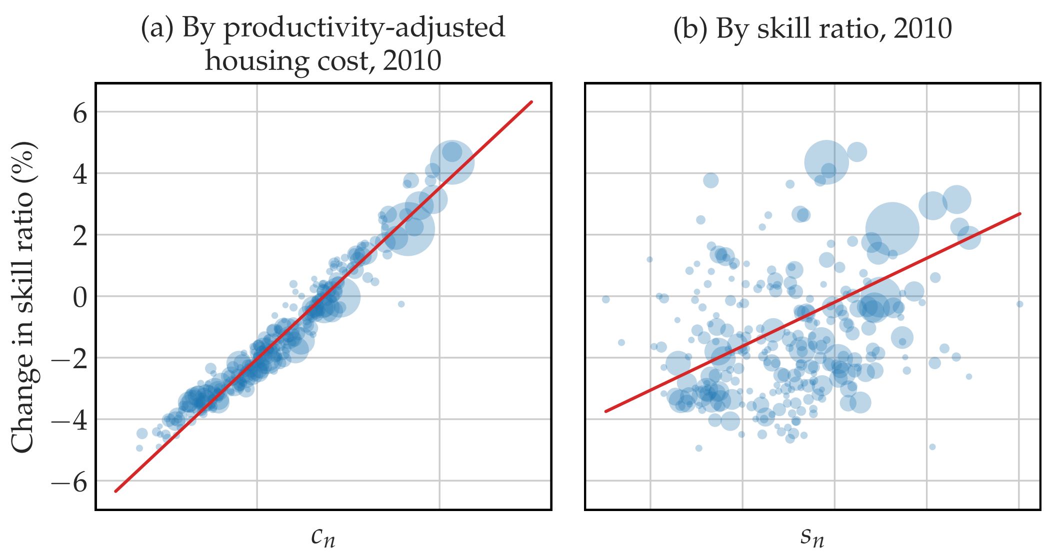
Note: "Data" calculates sorting using census data on employment by education level, 1980-2010. "Model" shows the level of sorting in an economy with the same fundamentals as in the data, except that aggregate productivities are changed to eliminate the increase in the skill premium. Shaded area shows the causal effect of the rising skill premium. Sorting is defined as the variance of the log skill ratio across cities, weighted by 1980 employment.

more of a necessity and housing expenditure shares decline more steeply with income. Figure 6 plots the results of our main counterfactual exercise for values of ϵ between -1 and 0. The effect of the rising skill premium is monotonically decreasing in ϵ and falls to 0 when $\epsilon = 0$, in line with Proposition 2. At the other extreme of $\epsilon = -1$, the rising skill premium accounts for 37% of the increase in sorting observed in the data. Our estimated value of -0.679, shown by the vertical line, lies between these two cases. We additionally investigate how assuming a unit housing requirement would alter our results. Recall from Section 2 that a unit housing requirement corresponds to $\epsilon = -1$ and $\sigma = 0$. We find the share explained rises to 30%, a substantial increase over our baseline. Overall, the results here show the importance of carefully estimating nonhomothetic housing demand for gauging the strength of the skill premium as a driver of sorting.

5.2 Extensions and Robustness

Below, we discuss several extensions and robustness checks.[15] Appendix E provides more details.

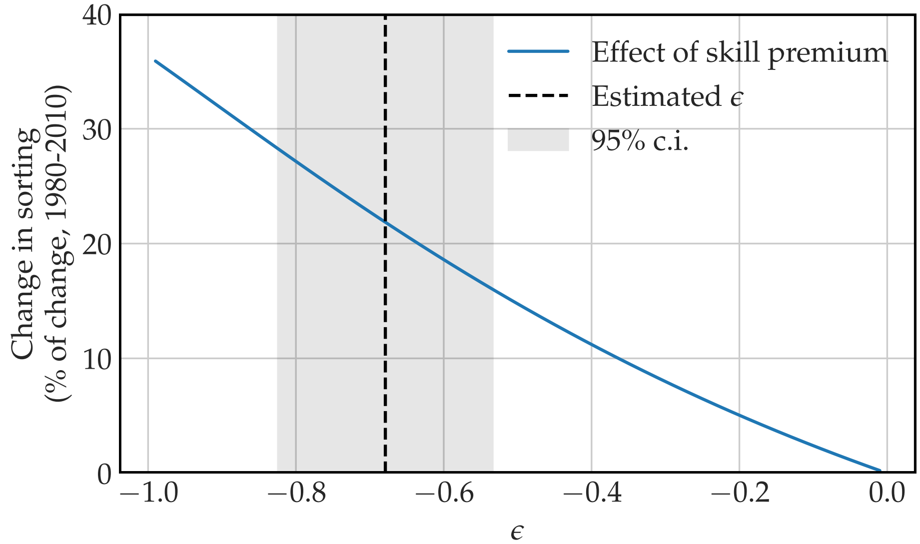
Figure 5: City Level Effects of the Skill Premium



Note: Panels (a) and (b) plot the causal effect of the increase in the skill premium in each MSA in 2010 against the productivity-adjusted housing cost in 2010 and log skill ratio in 2010, respectively. The causal effect of the skill premium is defined as the difference between the economy with all fundamentals changing as in the data, and the economy with the same fundamentals except for aggregate productivities Z and A, which are changed to eliminate the increase in the skill premium 1980-2010. The dots are sized proportionally to 1980 employment. The solid line is the regression line.

[15]Whenever we alter a parameter or add a new feature to the model, we re-calibrate the model following the same steps as in Section 4.

Figure 6: The Role of Nonhomotheticity



Note: The solid line plots the share of the increase in sorting between 1980 and 2010 caused by the increase in the skill premium for 30 different values of ϵ between -1 and 0. For each value of ϵ, we repeat the counterfactual exercise described above, holding all other parameters at the values chosen in Section 4. The dashed line shows the estimated value of ϵ from Table 1, column (4). The shaded area shows a 95% confidence interval for this point estimate.

Alternative Measures of Sorting

So far, we have focused on the variance of the log skill ratio as a measure of sorting. Here, we consider three alternative measures. The change in each measure between 1980 and 2010 is reported in Table 3, along with the effect of the skill premium in each case. We relegate the mathematical details to Appendix E.1. Column (1) reports our baseline measure, the variance of the log skill ratio. Column (2) shows the Theil index. The Theil index is a measure of the unevenness of the distribution of some characteristic across different units — in our case, the distribution of skill ratios across MSAs. The index is zero when the skill ratio is constant across locations, and a higher Theil index corresponds to more intense sorting. The dissimilarity index in column (3) measures how differently two groups are distributed over space and has recently been used by Fogli and Guerrieri (2019) to measure within-MSA segregation by income. Finally, in column (4), we report the difference between the 90th and 10th percentiles of the log skill ratio distribution. All measures of sorting have increased since 1980, and for columns (2) and (4), the effect of the skill premium is quite similar to our baseline result. For the dissimilarity index, we find a somewhat lower value.

Counterfactual Implementation

In our baseline counterfactual, the values A^t and Z^t are chosen to (i) fix the skill premium at its 1980 level and (ii) match the growth of average unskilled wages from the data. As an alternative, we modify (ii) to match the growth of average wages (pooling unskilled and skilled together). Although the implied sequence of (A^t, Z^t) is somewhat different, the counterfactual result is virtually unchanged at 22%.

Preference Specification

We investigate the sensitivity of our findings to our parameterization of utility by replacing the NHCES function used above with Price Independent Generalized Linear (PIGL) preferences, a leading case of nonhomothetic preferences (Boppart 2014; Eckert and Peters 2018). PIGL preferences imply a log-linear demand system that maps directly to the linearized specifications estimated in Section 2, allowing us to read off the key parameters of these preferences from column (2) of Table 1.[16] After recalibrating the quantitative model, we find the skill premium explains 19% of the increase in sorting since 1980, comparable to our baseline result. We conclude our findings are not sensitive to the parametrization of utility.

Choice of Elasticity of Substitution in Production

We next turn to the production side of our model. We consider a lower value of ρ equal to 1.6, taken from Acemoglu and Autor (2011). Using this value of ρ, the share of the increase in sorting explained by the rising skill premium falls to 13%. The fact that the share explained falls is intuitive — when skilled and unskilled labor are not close substitutes, the influx of skilled workers into expensive cities is dampened by falling skilled wages. Stronger congestion in production implies the impact of the rising skill premium is limited. Our quantitative results are therefore

Table 3: Alternative Measures of Sorting

	(1) Var. log skill ratio	(2) Theil	(3) Dissimilarity	(4) 90-10
Change, 1980-2010				
Data	32.6%	34.5%	19.0%	8.28 pp.
Model	25.5%	27.3%	16.4%	6.71 pp.
Effect of skill premium (% of observed change)	21.7	20.8	13.5	19.0

Note: Each column reports the change in sorting in the data and in the economy in which all fundamentals change as in the data, apart from the aggregate productivity parameters A and Z. The aggregate parameters A and Z are changed to eliminate the observed increase in the skill premium, 1980-2010. Column (1) is our preferred measure of sorting, while columns (2)-(4) present alternative measures of sorting. The final row reports the difference between the data and the model economy, which measures the causal effect of the rising skill premium on each measure of sorting. See Appendix E.1 for details of each measure of sorting.

[16] See Appendix E.2 for more details.

fairly sensitive to changes in this parameter. Nevertheless, as we argued in Section 4, city-level estimates of the elasticity of substitution are consistently higher than aggregate estimates like the one used here, and our baseline value of $\rho = 3.85$ is around the middle of such estimates.

Endogenous Amenities

Diamond (2016) allows amenities to respond endogenously to the local skill mix and concludes skilled workers value amenities more, causing sorting. In Appendix E.3, we extend our model to incorporate endogenous amenities; here, we summarize the key points. As in Diamond (2016), endogenous amenities in our model take the form $B_{in} = b_{in} \left(\frac{l_{sn}}{l_{un}}\right)^{\beta_i}$. First, in the simple model of Section 3, endogenous amenities amplify the effect of the skill premium on spatial sorting in the presence of nonhomothetic housing demand. An increase in the skill premium makes skilled households less sensitive to housing costs and encourages movement toward more expensive cities, just as in our baseline model. Then, amenities endogenously rise in expensive cities, encouraging further skilled in-migration. Second, we extend the neutrality result in Proposition 2 to endogenous amenities. If housing demand is homothetic, changes in the skill premium continue to have no effect on sorting. This result emerges because even in the presence of endogenous amenities, the log skill ratio s_n is entirely determined by location-specific fundamentals: with endogenous amenities, (28) becomes

$$s_n = \zeta + \theta \log\left(\frac{e_{sn}}{e_{un}}\right) + (\beta_s - \beta_u)s_n + \log\left(\frac{b_{sn}}{b_{un}}\right),$$

implying

$$s_n = \check{\zeta} + \check{\theta}_1 \log a_n + \check{\theta}_2 \log\left(\frac{b_{sn}}{b_{un}}\right)$$

for constants $\check{\zeta}, \check{\theta}_1, \check{\theta}_2$ In summary, endogenous amenities are likely to amplify the effect of the skill premium on spatial sorting when preferences are nonhomothetic, but they do not create an independent link between the skill premium and spatial sorting.

6 Extensions

6.1 Housing subsidies

In Section 5 we showed that increases in the skill premium since 1980 have increased spatial sorting by skill via nonhomothetic housing demand. In this subsection, we ask whether a simple policy intervention — a housing subsidy offered to unskilled households — might be able to reverse this trend.

To understand why such a subsidy would alter sorting patterns, we return to the simple model introduced in Section 3. There, we showed the ideal price indices are key drivers of sorting by skill. Now, we introduce a proportional subsidy s for unskilled households — the ideal price indices become

$$P_{un} = \left((1-\Omega) + \Omega(1-s)^{1-\sigma} \left(\frac{z_n}{P_{un}} \right)^{\epsilon(1-\sigma)} p_n^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$
$$P_{sn} = \left((1-\Omega) + \Omega A^{\epsilon(1-\sigma)} \left(\frac{z_n}{P_{sn}} \right)^{\epsilon(1-\sigma)} p_n^{1-\sigma} \right)^{\frac{1}{1-\sigma}}.$$

As in Section 3, P_{in} is the ideal price index of type i in location n, z_n and p_n are the productivity and housing price of location n, and A is the skill premium. We can see $(1-s)$ plays a similar role in P_{un} as the skill premium does in P_{sn}. A higher subsidy s reduces the sensitivity of P_{un} to prices p_n and so encourages unskilled households to move toward more expensive locations. Because skilled households disproportionately chose to live in such locations, we expect a higher subsidy to decrease sorting. This implication is not shared by Cobb-Douglas preferences — in this case a housing subsidy scales P_{un} proportionally in all locations and does not alter location choices.

We quantify the impact of such a subsidy using our calibrated model. Starting from 2010 fundamentals, we introduce a 15% subsidy.[17] We find the variance of the log skill ratio — our key measure of sorting — falls from 0.169 to 0.159, just over 24% of the increase since 1980. In Section 5 we found the rising skill premium accounted for 22% of this increase. Thus, a housing subsidy targeted to unskilled households can more than reverse the increase in sorting caused by the rising skill premium.

6.2 The incidence of local productivity shocks

We have shown nonhomothetic preferences introduce location bias to skill-specific shocks. The converse is also true: location-specific shocks are skill biased.

Consider a shock to z_n, the skill-neutral labor productivity in region n. The equivalent variation (EV) of the shock (as a percentage of initial expenditure) is

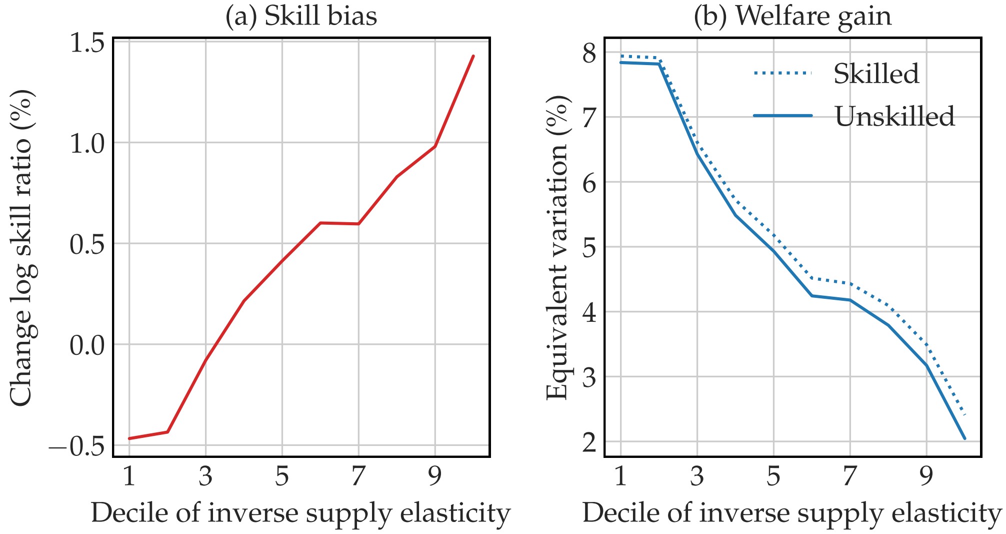
$$EV_{in} = \hat{e}_{in} - \eta_{in}\hat{p}_n$$

where hats denote log changes. Income growth is partially offset by rising house prices. Under homotheticity, all households are equally exposed to rising prices because they have the same housing share. By contrast, in our model, unskilled households are relatively more exposed to rising prices.

We quantify how the incidence of productivity shocks varies with γ_n, the inverse housing-supply elasticity. Starting from 2010 fundamentals, we shock z_n for a single location by 10% and recompute the new equilibrium. We repeat the exercise one location at a time. The left panel of Figure 7 plots binned average changes in the log skill ratio between the baseline and each counterfactual for deciles of γ_n. Inelastic housing markets become more skill intensive following

[17] We increase the shifter λ in the tax system (27) to ensure the government budget balances.

Figure 7: Incidence of Local Productivity Shocks



Source: We shock the skill-neutral productivity z_n by 10%, location by location. The horizontal axis is the decile of the inverse supply elasticity (higher means more inelastic). The left panel shows the change in the log skill ratio. The right panel shows equivalent variation, as a percentage of initial expenditure.

the shock. The right panel of Figure 7 plots, for each type, binned welfare changes against γ_n. In the top decile, the EV for skilled households is nearly 20% higher than for unskilled, 2.4% to 2.0%.

6.3 Agglomeration externalities and the spatial multiplier

Finally, we extend the model to incorporate productivity spillovers as in Fajgelbaum and Gaubert (2020). Agglomeration forces imply the spatial distribution of skills directly affects productivity and the income distribution. We show that because nonhomotheticity fuels spatial segregation by skill, it amplifies skill-biased productivity shocks.

We parametrize agglomeration as isoelastic functions of employment by type. The productivity terms in (21) become

$$z_n = \overline{z}_n l_{sn}^{\phi^{z,s}} l_{un}^{\phi^{z,u}} \tag{31}$$

$$a_n = \overline{a}_n l_{sn}^{\phi^{a,s}} l_{un}^{\phi^{a,u}}. \tag{32}$$

Calibrating to match Fajgelbaum and Gaubert (2020) implies $\phi^{z,u} = 0.003$, $\phi^{z,s} = 0.044$, $\phi^{a,u} = 0.017$, $\phi^{a,s} = 0.009$, and $\rho = 1.6$.[18]

Starting from 2010 fundamentals, we shock the aggregate skill bias A by 10% and compute the elasticity of the aggregate skill premium with respect to A. In a model without space, this elasticity would be exactly $1 - \rho^{-1} = 0.375$. We first consider a model with Cobb-Douglas preferences.

[18]We recalibrate ρ, the elasticity of substitution, because it has a different quantitative interpretation in models with and without spillovers.

In this case, the elasticity rises 12% to 0.42. The elasticity rises because the shock to A disproportionately raises wages in skill-intensive locations. However, the elasticity of sorting S to A is only 0.004, indicating changes in sorting do not play an important role in this amplification.

Now we turn to the model with estimated NHCES preferences. The elasticity of the skill premium with respect to A now rises 18% to 0.44, whereas the elasticity of sorting with respect to A rises to 0.07. Spatial amplification of the shock is stronger because the higher skill premium encourages more intense clustering of skilled households in a few cities. Sorting feeds back to wages through agglomeration externalities and leads to a higher aggregate skill premium.

7 Conclusion

When housing demand is income-inelastic, the skill premium causes spatial sorting because skilled and unskilled households face different ideal price indices in the same location. Skilled households have low housing shares and are insensitive to high house prices. The opposite is true for unskilled households. The growth in the skill premium since 1980 has amplified the cost-of-living wedge, causing skilled households to flock to expensive cities and unskilled households to flee to cheap cities. Our model attributes a fifth of the observed increase in spatial sorting to the growth of the skill premium.

Our contribution is twofold. First, we infer the extent of nonhomotheticity directly from consumption microdata. The estimates differ from the knife-edge cases commonly used. Preferences in which the housing share is independent of expenditure, such as Cobb-Douglas, sever the link between the skill premium and sorting. A unit housing requirement, in which the housing share declines one-for-one with expenditure, overstates it. Second, our theory connects sorting to a key macro aggregate, the skill premium. Nonhomotheticity makes aggregate shocks non-neutral, skewing the mapping between wages and welfare and helping to rationalize the observed disparity of skill across space.

References

Acemoglu, Daron, and David Autor. 2011. "Skills, tasks and technologies: Implications for employment and earnings". In *Handbook of Labor Economics*, 4:1043–1171.

Aguiar, Mark, and Mark Bils. 2015. "Has consumption inequality mirrored income inequality?" *American Economic Review* 105 (9): 2725–56.

Albouy, David, Gabriel Ehrlich, and Yingyi Liu. 2016. "Housing demand, cost-of-living inequality, and the affordability crisis". Working Paper.

Allen, Treb, and Costas Arkolakis. 2014. "Trade and the topography of the spatial economy". *The Quarterly Journal of Economics* 129 (3): 1085–1140.

Andreski, Patricia, Geng Li, Mehmet Zahid Samancioglu, and Robert Schoeni. 2014. "Estimates of annual consumption expenditures and its major components in the PSID in comparison to the CE". *American Economic Review* 104 (5): 132–35.

Aten, Bettina, and Eric Figueroa. 2019. "Estimating price levels for housing rents in the Regional Price Parities". *Survey of Current Business* 99 (6): 1–12.

Attanasio, Orazio P, Renata Bottazzi, Hamish W Low, Lars Nesheim, and Matthew Wakefield. 2012. "Modelling the demand for housing over the life cycle". *Review of Economic Dynamics* 15 (1): 1–18.

Attanasio, Orazio P., and Luigi Pistaferri. 2016. "Consumption Inequality". *Journal of Economic Perspectives* 30 (2): 3–28.

Baum-Snow, Nathaniel, Matthew Freedman, and Ronni Pavan. 2018. "Why has urban inequality increased?" *American Economic Journal: Applied Economics* 10 (4): 1–42.

BEA. 2020. *Metropolitan Area Regional Price Parities, Services: Rent*. Dataset.

Beaudry, Paul, Mark Doms, and Ethan Lewis. 2010. "Should the personal computer be considered a technological revolution? Evidence from US metropolitan areas". *Journal of Political Economy* 118 (5): 988–1036.

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

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten. 2016. "Consumption inequality and family labor supply". *American Economic Review* 106 (2): 387–435.

Boppart, Timo. 2014. "Structural change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences". *Econometrica* 82 (6): 2167–2196.

Bound, John, Jeffrey Groen, Gabor Kezdi, and Sarah Turner. 2004. "Trade in university training: cross-state variation in the production and stock of college-educated labor". *Journal of Econometrics* 121 (1-2): 143–173.

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

Chetverikov, Denis, Dongwoo Kim, and Daniel Wilhelm. 2018. "Nonparametric instrumental-variable estimation". *The Stata Journal* 18 (4): 937–950.

Comin, Diego A, Danial Lashkari, and Martí Mestieri. 2021. "Structural change with long-run income and price effects". *Econometrica* 89 (1): 311–374.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst. 2019. "Income growth and the distributional effects of urban spatial sorting". Working paper.

Davis, Morris A, and François Ortalo-Magné. 2011. "Household expenditures, wages, rents". *Review of Economic Dynamics* 14 (2): 248–261.

Diamond, Rebecca. 2016. "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000". *American Economic Review* 106 (3): 479–524.

Eckert, Fabian. 2018. "Growing apart: Tradable services and the fragmentation of the US economy". Working Paper.

Eckert, Fabian, Sharat Ganapati, and Conor Walsh. 2020. "Skilled scalable services: the new urban bias in recent economic growth". Working Paper.

Eckert, Fabian, and Michael Peters. 2018. "Spatial structural change". Working Paper.

Fajgelbaum, Pablo D, and Cecile Gaubert. 2020. "Optimal spatial policies, geography, and sorting". *The Quarterly Journal of Economics* 135 (2): 959–1036.

Fogli, Alessandra, and Veronica Guerrieri. 2019. "The end of the American Dream? Inequality and segregation in US cities". Working Paper.

Ganong, Peter, and Daniel Shoag. 2017. "Why has regional income convergence in the US declined?" *Journal of Urban Economics* 102:76–90.

Giannone, Elisa. 2019. "Skilled-biased technical change and regional convergence". Working Paper.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai. 2013. "Superstar cities". *American Economic Journal: Economic Policy* 5 (4): 167–99.

Gyourko, Joseph, Albert Saiz, and Anita Summers. 2008. "A new measure of the local regulatory environment for housing markets: the Wharton Residential Land Use Regulatory Index". *Urban Studies* 45 (3): 693–729.

Handbury, Jessie. 2019. "Are poor cities cheap for everyone? Non-homotheticity and the cost of living across u.s cities". Working Paper.

Hansen, Julia L, John P Formby, W James Smith, et al. 1996. "The income elasticity of demand for housing: evidence from concentration curves". *Journal of Urban Economics* 39 (2): 173–192.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante. 2017. "Optimal tax progressivity: an analytical framework". *The Quarterly Journal of Economics* 132 (4): 1693–1754.

Hornbeck, Richard, and Enrico Moretti. 2019. "Who benefits from productivity growth? Direct and indirect effects of local TFP growth on wages, rents, and inequality". Working Paper.

Hsieh, Chang-Tai, and Enrico Moretti. 2019. "Housing constraints and spatial misallocation". *American Economic Journal: Macroeconomics* 11 (2): 1–39.

Ioannides, Yannis M, Jeffrey E Zabel, et al. 2008. "Interactions, neighborhood selection and housing demand". *Journal of Urban Economics* 63 (1): 229.

Katz, Arnold. 2017. "Imputing rents to owner-occupied housing by directly modelling their distribution". Working paper.

Katz, Lawrence F, and Kevin M Murphy. 1992. "Changes in relative wages, 1963–1987: supply and demand factors". *The Quarterly Journal of Economics* 107 (1): 35–78.

Larsen, Erling Røed. 2014. "The Engel curve of owner-occupied housing consumption". *Journal of Applied Economics* 17 (2): 325–352.

Lewbel, Arthur, and Krishna Pendakur. 2009. "Tricks with Hicks: The EASI demand system". *American Economic Review* 99 (3): 827–63.

Poterba, James, and Todd Sinai. 2008. "Tax expenditures for owner-occupied housing: Deductions for property taxes and mortgage interest and the exclusion of imputed rental income". *American Economic Review* 98 (2): 84–89.

Real Personal Income and Regional Price Parities. 2020. Tech. rep. Bureau of Economic Analysis.

Redding, Stephen J. 2016. "Goods trade, factor mobility and welfare". *Journal of International Economics* 101:148–167.

Rosenthal, Stuart S. 2014. "Are private markets and filtering a viable source of low-income housing? Estimates from a "repeat income" model". *American Economic Review* 104 (2): 687–706.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. *IPUMS USA: Version 10.0*. Minneapolis, MN.

Saiz, Albert. 2010. "The geographic determinants of housing supply". *The Quarterly Journal of Economics* 125 (3): 1253–1296.

Schmidheiny, Kurt. 2006. "Income segregation from local income taxation when households differ in both preferences and incomes". *Regional Science and Urban Economics* 36 (2): 270–299.

Straub, Ludwig. 2019. "Consumption, savings, and the distribution of permanent income". Working Paper.

Zabel, Jeffrey E. 2004. "The demand for housing services". *Journal of Housing Economics* 13 (1): 16–35.

A Data

A.1 PSID

The primary consumption microdata comes from the Panel Study of Income Dynamics. The PSID is administered biannually, with about 9,000 households responding in each wave. It included a consumption module starting in 1999 and added several categories in 2005. The survey now covers about 70% of spending in the national accounts (Blundell, Pistaferri, and Saporta-Eksten 2016). Total expenditure is computed as the sum of all reported consumption categories: rent, food, utilities, telephone and internet, automobile expenses (including car loans, down payments, lease payments, insurance, repairs, gas, and parking), other transportation expenses, education, childcare, healthcare, home repairs, furniture, computers (2017 only), clothing, travel, and recreation. The PSID imputes a small number of observations to handle invalid responses. To match the definition in IPUMS, housing expenditure is equal to rent plus utilities. Homeowners were not asked to estimate the rental value of their home until 2017, so we restrict attention to renters and analyze homeowners with the CEX.

We select our sample according to the following criteria. We cut respondents in the top and bottom 1% of the pre-tax income distribution in each year to guard against serious misreporting errors. We also select households in which the head is prime-age (25-55, inclusive) and attached to the labor force (head or spouse reports usually working at least 35 hours per week). The controls included in the regressions are dummies for family size bins, number of earners, age bins, sex of household head, race of household head, and year. We use PSID sample weights in all regressions.

Using the restricted access county identifiers, we are able to assign local prices to 88% of households in the PSID sample. The missing 12% is by construction: the BEA estimates price indices only for MSAs, not for micropolitan statistical areas or rural areas. MSAs, in turn, cover only about 86% of the US population. Because the price data is available only starting in 2008, we use the 2009-2017 waves of the PSID (note that each survey asks about consumption in the previous year).

A.2 Regional Price Parities

The BEA has published metropolitan price indices since 2008 (Aten and Figueroa 2019). In this section we discuss the construction of the rental index; for more information about the price parities for other consumption components, we refer the interested reader to *Real Personal Income and Regional Price Parities* (2020).

Using ACS data, the BEA estimates a standard hedonic regression model of the form

$$\log r_{in} = \alpha_n + X_{in}'\beta + \varepsilon_{in} \tag{33}$$

where r is rent, α_n is an MSA dummy and X is the set of observable dwelling characteristics available in the ACS: the type of structure interacted with the number of bedrooms, the total number of rooms, building age, a rural dummy, and a dummy for whether utilities are included in the contract rent. The log price indices are given by $\{\alpha_n\}$.

A.3 CEX

We append the 2006-2017 Consumer Expenditure Surveys (CEX) together and annualize at the household level. We define rental expenditure as actual rent paid for renters (`rendwe`) and self reported rental-equivalent (`renteqvx`) for owners. As in PSID, we add utilities `util` to be consistent with the data available in the Census. To solicit rental equivalent, homeowners are asked "If someone were to rent your home today, how much do you think it would rent for monthly, unfurnished and without utilities?" We define total consumption expenditure as equal to total reported expenditure `totexp` *less* retirement and pension savings `retpen`, cash contributions `cashco`, miscellaneous outlays `misc` (which includes mortgage principal), and life and personal insurance `lifins`. We apply the exact same sample selection criteria and controls in the CEX as in the PSID (see Section A.1). We use CEX sample weights in all regressions.

In 2006, the CEX added more detailed geographic identifiers in the variable `psu`. The primary sampling unit, i.e. the MSA of residence, is available for a subset of households. The CEX identifies twenty-four large MSAs, which cover about 45% of households in the survey. In regressions with price data, we include data starting in 2008.

A.4 Census

We use the 5% public use samples from the 1980, 1990, and 2000 Censuses. For the final period of data, we use the 2009-2011 American Community Survey, a 3% sample. For convenience we refer to this as the "2010 data." IPUMS attempts to concord geographic units across years, although complete concordance is not possible because of data availability and disclosure rules. We classify MSAs according to the variable `metarea`. We produce a balanced panel using the following rule: if an MSA appears in all four years, then it is kept. If an MSA does not appear in all four years, then we assign all individuals in that MSA across all years to a residual state category. For example, Charlottesville, VA appears in 1980, 2000, and 2010, but not in 1990. Therefore we assign all individuals in Charlottesville in every year to "Virginia." This procedure gives us 219 MSAs (including Washington, D.C.) and 50 residual state categories, for a total of 269 regions. The share of national employment which can be assigned to an MSA, rather than a state residual, is 70% in 1980, 72% in 1990, and 75% in 2000 and 2010.

A worker is considered skilled if she or he has completed at least a four year college degree according to the variable `educ`. By this metric, the national fraction of workers who are skilled is 22.5% in 1980, 26.5% in 1990, 30.2% in 2000, and 35.7% in 2010.

We compute wages and employment for each region, skill level, and year. Wages are from the IPUMS variable `incwage`. To be included in the wage and employment sample, workers must be between 25 and 55 years old, inclusive; not have any business or farm income; work at least

40 weeks per year and 35 hours per week; and earn at least one-half the federal minimum wage. For each region, skill level, and year, we average rents paid by renting households whose head is included in the wage and employment sample. Rents are from the IPUMS variable `rentgrs`. Wages and rents are adjusted to 2000 real values using BLS' Non-Shelter CPI.

A.4.1 Reconciling income and expenditure in the data

The necessary model inputs are (1) average total expenditures by skill, location, and year, denoted by e_{int}; and (2) average housing expenditures by skill, location, and year, denoted by r_{int}. Rents and wages in the data, which we denote $\tilde{r}_{int}$ and $\tilde{w}_{int}$ do not exactly correspond to their model counterparts because of differences in savings and household composition. In addition, we only observe rents paid by renting households, who tend to have lower incomes and total expenditures than non-renting households. We now describe the steps to address these discrepancies in the Census data.

Income Compute post-tax income as $\tilde{y}_{int} = \lambda_t \tilde{w}_{int}^{1-\tau}$, where λ_t is chosen to balance the budget. We assume that the elasticity of expenditure to permanent post-tax income is unity, so from the household's budget constraint it is immediate that expenditure is equal to permanent post-tax income, $e_{int} = \tilde{y}_{int}$. Households save in the data, but savings wash out in the aggregate since we focus on permanent income.

We could relax this assumption following Straub (2019). Suppose that expenditure were given by $e_{int} = \bar{c}_{int}\tilde{y}_{int}^{\phi}$. If $\phi < 1$, expenditure would be nonhomothetic in permanent income. Qualitatively, this feature would increase the strength of our sorting mechanism. If consumption were less important for high-income households (relative to income), then they would be less sensitive to the price of local housing consumption (again, relative to income). We do not pursue a quantitative treatment of nonhomothetic total expenditure, which would require a dynamic quantitative spatial model beyond the scope of our paper.

Rent In the model, households are unitary decisionmakers. In the data, rent may be apportioned among multiple earners or dependents. To reconcile the two, we multiply rents in the data by a scalar, κ, such that the average rental share in the Census data is equal to the average rental share in the CEX, in each year. Specifically, κ_t is chosen to fit

$$\frac{\sum_i \sum_n \frac{\kappa_t \tilde{r}_{int}}{e_{int}} l_{int}}{\sum_i \sum_n l_{int}} = \overline{\eta}_t$$

where e, r, and l are for renters only. The final step is to impute rents by homeowners in the Census. This can be backed out given the data on renters and our estimated preferences. From (3), the expenditures of any two groups a and b in the same MSA satisfy

$$\frac{r_a}{r_b} = \left(\frac{e_a}{e_b}\right)^{-\epsilon\sigma} \left(\frac{e_a - r_a}{e_b - r_b}\right)^{1+\epsilon}$$

We compute r_{int}, rental expenditure by all households, as the solution to this equation given data on rental expenditure by renters, total expenditure by renters, and total expenditure by all households.

A.4.2 Bartik instruments

In order to obtain instrumental variables for labor demand, we construct Bartik shift-share variables. The share is a region's industrial composition in 1980, and the shift is change in average wages nationwide (excluding the region itself).

We use the industry categories in the Census variable `ind1990`. Harmonizing the industries with our own crosswalk yields 208 industries which are consistently defined over all four periods. We drop individuals who cannot be classified into any industry ($\approx 0.3\%$ of workers) or who are in the military ($\approx 0.9\%$ of workers).

B Estimation

We first describe how measurement error biases OLS estimates of the log-linearized estimating equation (9). We then describe alternative specifications to estimate the preferences in Section 2.

B.1 Measurement error

Recall that the log-linearized estimating equation is

$$\hat{\eta}_{int} = \omega_t + \omega' X_{int} + \beta \hat{e}_{int} + \gamma \hat{p}_{nt} + \xi_{int}$$

We address measurement error in expenditure in the following way. First, partialling out observable demographics and prices, write the reduced-form relationship between expenditure shares and total expenditure as

$$\eta = \beta^0 e + \xi \tag{34}$$

where we have suppressed subscripts and hats for notational convenience. Expenditure and rental expenditure are measured with error: $\tilde{e} \equiv e + v^e$, $\tilde{r} \equiv r + v^r$, and $\tilde{\eta} \equiv \tilde{r} - \tilde{e}$. v^e and v^r are assumed to be uncorrelated with e, r, and ξ.

The OLS estimate of β^0 is asymptotically

$$\begin{aligned}\widehat{\beta^0}_{OLS} &= \frac{\mathrm{cov}(\tilde{\eta}, \tilde{e})}{\mathrm{var}(\tilde{e})} \\ &= \frac{\beta^0 \sigma_e^2 + \sigma_{v^r, v^e} - \sigma_{v^e}^2}{\sigma_e^2 + \sigma_{v^e}^2}\end{aligned}$$

The attenuation bias $\sigma_e^2 / (\sigma_e^2 + \sigma_{v^e}^2)$ is familiar from classical measurement error. There are two additional sources of bias: (1) measurement error in expenditure appears on both the left- and right-hand sides of (34) and (2) measurement errors in expenditure and rent are mechanically correlated. The direction of the bias is ambiguous, although in practice the OLS estimate is lower than the 2SLS estimate.

B.2 Alternative specifications in PSID

We present several alternative specifications in Table B.1, still using our baseline sample of renters in the PSID.

B.2.1 Non-housing prices

In the baseline model, we assume that the price of non-housing consumption, C, is equalized across space. Here, we relax that assumption. Let non-housing consumption be a Cobb-Douglas aggregate of goods, with share ϱ, and non-housing services, with share $1 - \varrho$. This modification implies that equation (3) becomes

$$\eta_{int} = \Omega_{int} \left(\frac{e_{int}}{P_{nt}} \right)^{\epsilon(1-\sigma)} \left(\frac{p_{nt}}{P_{nt}} \right)^{1-\sigma} (1 - \eta_{int})^{1+\epsilon} . \tag{35}$$

where $P_{nt} = p_{Gnt}^{\varrho} p_{Snt}^{1-\varrho}$ is the local price index of goods and non-housing services.

We measure each index using the BEA's Metropolitan Regional Price Parities and set $\varrho = 0.51$ to replicate the BEA expenditure weights. We estimate (35) by GMM in column (1) of Table B.1. The estimated coefficients are not significantly different from the baseline values because prices of goods and non-housing services do not vary too much across space, relative to housing prices.

B.2.2 Alternative measures of price

In column (2) we show that the estimates are not sensitive to our choice of price index. Instead of the BEA Regional Price Parities, we use a Rent Index from Zillow, a real estate analytics firm. Zillow estimates the market price per square foot of rented units in most metropolitan areas starting in 2011. Despite the difference in datasets, the estimated coefficients are reassuringly close to the baseline values.

To study sorting *across* cities in a tractable way, we have assumed that prices are equalized

Table B.1: Preferences, alternative specifications (PSID)

Dependent variable: Log housing share

	(1) GMM Local prices	(2) GMM Zillow	(3) GMM Zillow	(4) 2SLS Fixed effects	(5) 2SLS Fixed effects	(6) GMM Alt. IV	(7) GMM Alt. IV
Log expenditure				-0.222***	-0.183***		
				(0.030)	(0.029)		
Implied ϵ	-0.650***	-0.705***	-0.890***			-0.594***	-0.626***
	(0.077)	(0.096)	(0.101)			(0.096)	(0.114)
Implied σ	0.534***	0.623***	0.707***			0.578***	0.587***
	(0.051)	(0.047)	(0.033)			(0.053)	(0.056)
MSA FE				✓			
State + Urban FE					✓		
Demographic controls	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
Excluded IV	log expenditure	log expenditure	log expenditure	log expenditure	log expenditure	lagged log expenditure	education
R^2	-	-	-	0.12	0.10	-	-
First-stage F-statistic	-	-	-	802.1	984.2	-	-
N	8,923	7,052	7,133	8,879	8,915	5,392	8,772
No. of clusters	4,229	3,640	3,674	4,201	4,224	2,743	4,148

Source: PSID, BEA, and Zillow.

Note: Column (1) extends the model to incorporate variation in local non-housing prices. Column (2) uses Zillow metropolitan price data while column (3) uses Zillow county price data. Column (4) uses MSA fixed effects. Column (5) uses state fixed effects and a dummy for living in an urban area. Columns (6) and (7) instrument log expenditure with lagged log expenditure and education, respectively. Standard errors clustered at the household level.

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

within each MSA. However, urban models typically feature within-city price heterogeneity.[19] We relax the assumption of common prices by assigning households to a *county-level* rather than a *metropolitan-level* price.

We use the county price index from Zillow. The estimated parameters in column (3) imply more nonhomotheticity than our MSA-level results. Incorporating more granular price data, for example at the neighborhood or Census tract level, is a worthwhile avenue for future research.

Columns (4) and (5) dispense with price data altogether. Instead, we re-estimate the linearized equation (9), replacing p with location fixed effects. The advantage of this strategy is that we can be agnostic on the correct measure of housing prices; the disadvantage is that we can no longer identify the price elasticity. Column (5) uses MSA fixed effects. The estimated coefficient on log expenditure is nearly identical to the baseline estimate. Alternatively, column (5) uses state fixed effects plus a dummy equal to one if the household lives in an urban area. We include this robustness check because it can be done using the public-access PSID.

B.2.3 Alternative instruments

In columns (6) and (7) we consider alternative instruments for log expenditure. Column (6) uses the log of lagged expenditure, which is valid as long as measurement error in expenditure is not serially correlated. Column (7) instead uses education as an instrument. We assign households into four bins based on the household head's level of education: less than high school graduate, high school graduate, some college, and college graduate or higher.[20] The exclusion restriction is that, conditional on the true level of expenditure and observable demographics like age and family size, education does not determine the housing share. The coefficient estimates are stable with these alternative instruments.

B.2.4 Semi-parametric model

We consider a semi-parametric of the form

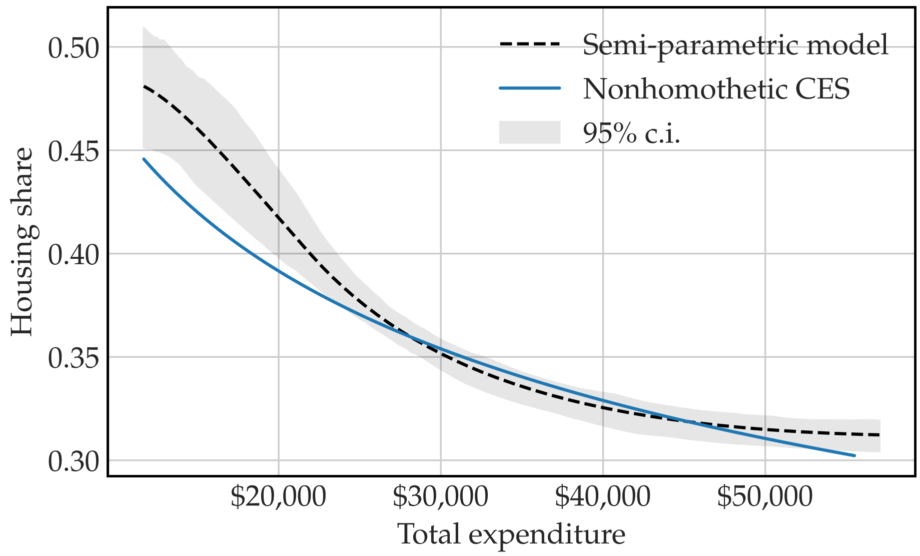
$$\log \eta_{int} = \omega_t + \omega' X_{int} + f(\log e_{int}) + \psi \log p_{nt} + \xi_{int}. \tag{36}$$

Note that this is very similar to (9), except that the model is no longer restricted to be log-linear in total expenditure. We estimate the unknown function f by nonparametric IV, using income as an instrument for total expenditure as in Section 2. We use the Stata package `npiv` developed by Chetverikov, Kim, and Wilhelm (2018), although we do not impose the monotonicity restriction suggested by that paper. In Figure B.1 we plot the estimated values of η (evaluated at the average values of the controls in (36)), along with a bootstrapped 95 % confidence interval. For comparison, we also show our baseline NHCES preferences, estimated in column (4) of Table 1. The two function are similar, but it is apparent that NHCES preferences slightly understate the slope of η

[19] Differentiated neighborhood quality or commuting costs will generate price dispersion within a city.

[20] Results are similar if we use years of education.

Figure B.1: Housing Expenditure Shares: Semi-parametric Model



Notes: The dashed line shows the predicted housing share as a function of total expenditure for the semi-parametric model (36), estimated following Chetverikov, Kim, and Wilhelm (2018). The solid line shows the housing share predicted by the NHCES model (6), estimated in column (4) of Table 1.

for low levels of expenditure while overstating the slope for high levels of expenditure. Despite this, we conclude from Figure B.1 that NHCES preferences provide a reasonable fit to the data.

B.3 Consumer Expenditure Survey (CEX)

In this section we present additional results from the Consumer Expenditure Survey (CEX). Reassuringly, all findings are close to our main results.

In the first column of Table B.2, we re-estimate our baseline specification in the CEX. The estimated expenditure elasticity is slightly higher, but the difference is not statistically significant.

B.3.1 Homeowners

Thus far we have focused on renting households because we do not observe expenditure on owner-occupied housing. In this section we explore whether our results extend to homeowners too. An appropriate measure of housing expenditure by homeowners is *rent equivalent*. Respondents are asked: "If someone were to rent your home today, how much do you think it would rent for monthly, unfurnished and without utilities?" The PSID consumption module did not elicit rent equivalent until 2017, but rent equivalent is available in all recent waves of the CEX. Therefore we use the CEX to study homeowners.

Column (2) of Table B.2 pools renting and owning households together. The estimate is consistent with significant nonhomotheticity. Restricting attention only to owners (column (3)) yields even stronger nonhomotheticity than the baseline estimate for renters.

Table B.2: Preferences (CEX)

Dependent variable: Log housing share

	(1) GMM Baseline	(2) GMM Pooled	(3) GMM Owners	(4) GMM Owners, incl second	(5) GMM Pooled, oop	(6) GMM Owners, oop
ϵ	-0.598	-0.507	-0.650	-0.656	-0.621	-0.487
	(0.073)	(0.040)	(0.040)	(0.038)	(0.052)	(0.076)
σ	0.812	0.802	0.780	0.783	0.767	0.761
	(0.031)	(0.018)	(0.021)	(0.020)	(0.023)	(0.030)
Demographic controls	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
N	2,503	6,719	4,216	4,506	6,719	4,216

Source: CEX and BEA

Note: Column (1) replicates our baseline specification of Table 1 column (5), using the CEX. Column (2) adds homeowners, measuring housing expenditure by self-reported rental equivalent. Column (3) restricts to homeowners only. Column (4) includes households who own second homes. Columns (5) and (6) measures housing expenditure as out-of-pocket expenses, defined as mortgage interest, property taxes, insurance, maintenance, and repairs; mortgage principle is excluded. Instrument is log household income. Robust standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

In columns (5) and (6), we use an alternative measure of housing expenditure for homeowners, *out-of-pocket expenses*. We define out-of-pocket expenses as the sum of mortgage interest, property tax, insurance, maintenance, and repairs. We omit payments on mortgage principal since these payments are savings, not consumption. Out-of-pocket expenses reflect the user cost of housing, which is equal to the rental value of the house in equilibrium. The estimates are close to our baseline results.

In our main analysis of homeowners, we restrict our sample to households who own a single home, which includes 94% of homeowners in the CEX. The reason is that expenditure on second homes does not reflect the local cost of living, but rather is a luxury more akin to recreation or vacations. That said, it is possible that second homes are a substitute for primary homes in expensive markets: for example, a household could live in a small house in the city and maintain a larger house in the country. Including second homeowners in column (4) leaves our results virtually unchanged.

B.3.2 Imputing rents from home values

Because data on rent equivalent is (until very recently) unavailable in the PSID, the standard approach has been to impute rents as a constant fraction of self-reported home value, generally six percent (Attanasio and Pistaferri 2016; Straub 2019). We argue that this is not an appropriate strategy. The six percent figure is from Poterba and Sinai (2008), who compute the user cost of housing with data from the Survey of Consumer Finances. Poterba and Sinai (2008) document considerable variation in the user cost across different types of homeowners, with a mean of six percent. We provide further evidence on heterogeneity in rent-to-value ratios from the CEX in Figure B.2, panel (a). The median is 7.8% with a mean of 8.5%. Crucially, there is a clear negative relationship: more valuable properties have systematically lower rent-to-value ratios. The same pattern appears in the Residential Financial Survey, used by the BEA to impute rents in the national accounts (Katz 2017). Imputing rent as a constant fraction of home value tends to deflate the housing shares of households with low home values, obscuring nonhomotheticity in the data.

B.4 Comparison to Davis and Ortalo-Magné (2011)

We reconcile our paper with Davis and Ortalo-Magné (2011), who argue in favor of Cobb-Douglas preferences over housing and non-housing consumption.

First, Davis and Ortalo-Magné (2011) show that median rent-to-income ratios were roughly constant across time and space between 1980 and 2000. We replicate and extend this result in Table B.3. The mean is constant from 1980-2000 but increases after 2000, suggesting that the main findings in Davis and Ortalo-Magné (2011) may be sensitive to the period studied.

Second, constant housing shares over space are necessary but not sufficient to conclude that preferences are Cobb-Douglas. Cross-city comparisons of housing shares reflect both income and

Table B.3: Median housing share across cities

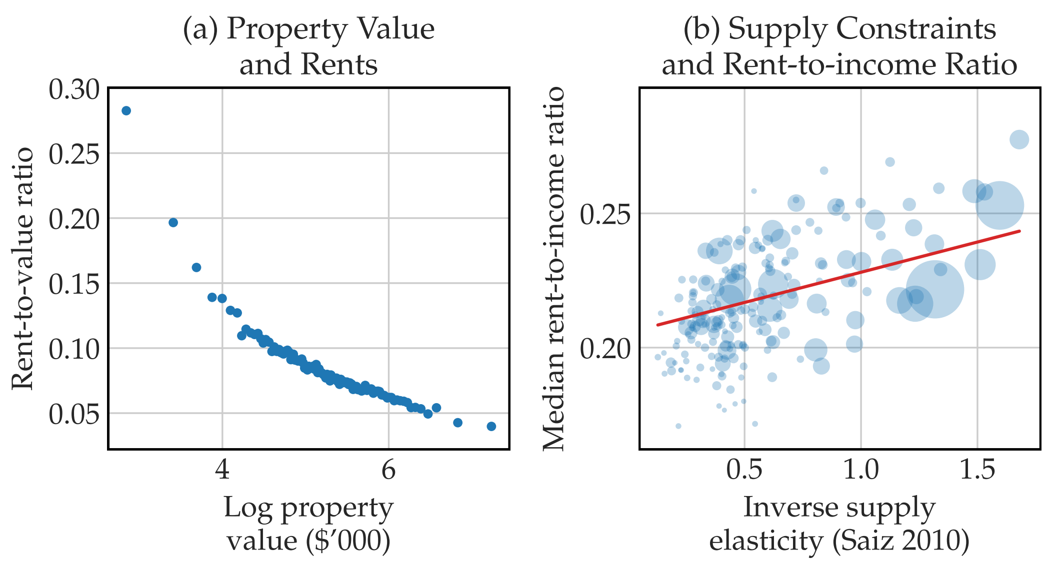
	DOM sample		Full sample	
	Mean	Std. dev.	Mean	Std. dev.
1980	0.22	0.02	0.22	0.02
1990	0.23	0.02	0.22	0.02
2000	0.22	0.02	0.22	0.02
2010	0.27	0.03	0.26	0.03

Source: Census. These are unweighted statistics of the median rent-to-income ratio across 50 US metropolitan areas ("DOM sample") and 219 US metropolitan areas ("Full sample")

price effects. Households in expensive cities have a higher housing share than households in inexpensive cities at every level of income; but the composition of expensive cities is tilted toward high-income households, who tend to have lower housing shares. In Figure B.2, panel (b), we plot the median rent-to-income ratio and the inverse housing supply elasticity from Saiz (2010), a reduced-form price shifter. Expenditure is positively correlated with prices at the city level, with a correlation of 0.59. Cobb-Douglas preferences imply zero correlation.

More generally, preferences are not identified from aggregate data on housing shares. Turning to the microdata, Davis and Ortalo-Magné (2011) report that the elasticity of rent (net of utilities) to total expenditure in the CEX is 0.99. Aguiar and Bils (2015) report an elasticity of 0.92 with the same data. This is the elasticity *unconditional on prices*, and hence not appropriate for a model of sorting across cities. We replicate and extend Davis and Ortalo-Magné (2011) in Table B.4. To

Figure B.2: Additional Figures



Notes: (a) CEX, 2006-2017. We compute the average ratio of rent equivalent to property value for 100 property value bins. (b) Census data from 2000 and Saiz (2010). Size of marker indicates city size. Solid line is a regression of median rent-to-income ratio on inverse supply elasticity.

facilitate a fair comparison, we estimate the elasticity of rent (net of utilities) to total expenditure in the CEX.[21] As in Davis and Ortalo-Magné (2011), the unconditional expenditure elasticity is close to unity. Crucially, controlling for local prices—whether by directly including prices as in column (2), or by using MSA fixed effects as in column (3)—confirms that preferences are significantly nonhomothetic. In columns (4) and (5) we repeat the exercise for utilities, which we model as part of housing.

B.5 Income elasticities from the literature

Table B.5 summarizes estimates of the income elasticity of housing demand from the literature. Controlling for local prices, using expenditure on the right hand side, and accounting for measurement error with an IV are all key in obtaining a consistent estimate of the elasticity.

Table B.4: Separate rent and utilities

	Log net rent			Log utilities	
	(1)	(2)	(3)	(4)	(5)
Log expenditure	0.982***	0.824***	0.844***	0.541***	0.639***
	(0.025)	(0.026)	(0.026)	(0.029)	(0.037)
Log price		0.576***			
		(0.033)			
MSA FE			✓		✓
Demographic controls	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
First stage F-stat.	2,404	1,137	1,305	2,382	1,295
R^2	0.38	0.49	0.50	0.37	0.44
N	5,803	2,501	3,029	5,795	3,022

Source: CEX and BEA
Note: Controls are bins for age, race, and gender of household head; household size; and number of earners in household. Instrument is log household income. Robust standard errors in parentheses.
* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses.

[21] Our baseline specifications define housing costs as rent plus utilities.

Table B.5: Income elasticities in the literature

Paper	Elasticity	Sample	Local prices?	Expenditure?	IV?
Rosenthal (2014)[a]	-0.88	Renters	✓		
Ioannides, Zabel, et al. (2008)[b]	-0.79	Owners	✓		
Hansen, Formby, Smith, et al. (1996)[c]	-0.73	Renters			
Larsen (2014)[d]	-0.67	Owners			
Zabel (2004)[e]	-0.52	Owners	✓		
Albouy, Ehrlich, and Liu (2016)[f]	-0.28	Renters	✓		
Lewbel and Pendakur (2009)[g]	-0.28	Renters		✓	✓
Attanasio et al. (2012)[h]	-0.22	Both		✓	
Aguiar and Bils (2015)[i]	-0.08	Both		✓	✓
Davis and Ortalo-Magné (2011)[j]	-0.01	Both		✓	✓
Paper benchmark[k]	-0.24	Renters	✓	✓	✓

[a] American Housing Survey, 1985-2011. Table 5, column 1.
[b] American Housing Survey, 1985-1993. Table 5, column 1.
[c] American Housing Survey, 1989. Table 5, column 2, last row.
[d] Norweigian Rental Survey and Consumer Expenditure Survey, 2007. Table 2, row 5.
[e] American Housing Survey, 2001. Table 3, row 3.
[f] US Census, 1970-2014. Table 1, column 3.
[g] Canadian Family Expenditure Surveys, 1969-1996. Median uncompensated elasticity computed using authors' replication file following their Appendix VII.1.
[h] British Household Panel Survey, 1991-2002. Table 4, panel B. Estimates for high- and low-education groups are averaged with weights one-third and two-thirds, respectively.
[i] US CEX, 1980-2010. Table 2, column 1.
[j] US CEX, 1982-2003. Text, page 253.
[k] PSID and BEA, 2009-2017.

C Theory

C.1 Proof of Lemma 1

To derive the ideal price indices in (19) and (20), substitute the expression for P_{in} in (17) into the Hicksian demand function (2) to obtain an expression in terms of η_{in}

$$\frac{\eta_{in}}{1-\eta_{in}} = \Omega p_n^{1-\sigma} \left(\frac{e_{in}}{P_{in}}\right)^{\epsilon(1-\sigma)}.$$

Substituting this into the expression for η_{in} in (3) and rearranging yields

$$P_{in}^{1-\sigma} = \left((1-\Omega) + \Omega p_n^{1-\sigma} \left(\frac{e_{in}}{P_{in}}\right)^{\epsilon(1-\sigma)}\right).$$

Replacing expenditures with productivities following (12) and (13) yields the ideal price indices (19) and (20), reproduced below

$$P_{un} = \left((1-\Omega) + \Omega \left(\frac{z_n}{P_{un}}\right)^{\epsilon(1-\sigma)} p_n^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$
$$P_{sn} = \left((1-\Omega) + \Omega A^{\epsilon(1-\sigma)} \left(\frac{z_n}{P_{sn}}\right)^{\epsilon(1-\sigma)} p_n^{1-\sigma}\right)^{\frac{1}{1-\sigma}}.$$

Defining $c_n = z_n^{\epsilon} p_n$, we define the functions $P_u(c)$ and $P_s(c)$ implictly

$$P_u(c) = \left((1-\Omega) + \Omega \left(\frac{c}{P_u(c)^{\epsilon}}\right)^{(1-\sigma)}\right)^{\frac{1}{1-\sigma}} \tag{37}$$
$$P_s(c) = \left((1-\Omega) + A^{\epsilon(1-\sigma)}\Omega \left(\frac{c}{P_s(c)^{\epsilon}}\right)^{(1-\sigma)}\right)^{\frac{1}{1-\sigma}}. \tag{38}$$

Clearly $P_i(c_n) = P_{in}$.

To prove Lemma 1, we establish that $\log P_u(c) - \log P_s(c)$ is strictly increasing in c. This is equivalent to showing that

$$\delta_u(c) > \delta_s(c)$$

where $\delta_i(c)$ is the elasticity of P_i with respect to c. Differentiating (19) and (20) and rearranging yields

$$\delta_i(c) = \frac{\eta_i(c)}{1+\epsilon\eta_i(c)}$$

where η_i is the housing expenditure share of type i when facing productivity adjusted housing cost c. δ_i is clearly a strictly increasing function of $\eta_i(c)$. Whenever $\epsilon < 0$, so that housing demand is income inelastic, $\eta_u(c) > \eta_s(c)$ because the expenditure of the skilled worker is always higher.

Therefore $\delta_u > \delta_s$ and $\log P_u(c) - \log P_s(c)$ is strictly increasing in c. Lemma 1 follows.

C.2 Proof of Proposition 1

We first assume $\epsilon < 0$. P_{un} is unaffected by changes in A, so

$$ds_n = -\theta d \log P_{sn} d\zeta$$

from (18). Inspection of (38) shows that A^ϵ appears isomorphically to c, and so

$$ds_n = -\theta\epsilon\delta_s(c_n) d \log A + d\zeta$$

where δ_s is the elasticity of P_s with respect to productivity adjusted housing costs. In Appendix C.1 we showed that δ_s is a strictly increasing function of c. Since $\epsilon < 0$, this implies that ds_n is also strictly increasing in c. Since by Lemma 1 s_n is strictly increasing in c, we then have that ds_n is strictly increasing in s_n. Now we turn to sorting S, defined as the variance of s_n. We prove this statement for the weighted variance with positive (and fixed) weights ω_n which sum to 1 since we will weight by 1980 employment in our empirical application. By definition

$$S = \sum_n \omega_n \left(s_n - \bar{s}\right)^2$$

$$\bar{s} = \sum_n \omega_n s_n.$$

Differentiating

$$dS = 2 \sum_n \omega_n \left(ds_n - d\bar{s}\right)\left(s_n - \bar{s}\right) = 2Cov(ds_n, s_n).$$

Since ds_n is a strictly increasing function of s_n, this covariance is positive and so $dS > 0$. This completes the proof for the case of $\epsilon < 0$. When $\epsilon = 0$, $P_{sn} = P_{un}$ and (18) then implies that $ds_n = 0$ for all n. $dS = 0$ follows.

C.3 Proof of Proposition 2

We start by taking logs of (24)

$$\log l_{in} = \theta \log v_{in} + \log B_{in} - \log U_i$$

where U_i is just the denominator in (24), divided by L_i. We difference this across types in the same location n and use the definition of the log skill ratio s_n

$$s_n = \theta \log\left(\frac{v_{sn}}{v_{un}}\right) + \log\left(\frac{B_{sn}}{B_{un}}\right) - \log\left(\frac{U_s}{U_u}\right).$$

Now when $\epsilon = 0$ preferences are homothetic, and v_{in} is given by

$$v_{in} = e_{in} \left((1 - \Omega) + \Omega p_n^{1-\sigma} \right)^{\frac{-1}{1-\sigma}}.$$

This implies that the ratio v_{sn}/v_{un} is just the ratio of expenditures. Therefore

$$s_n = \theta \log \left(\frac{e_{sn}}{e_{un}} \right) + \log \left(\frac{B_{sn}}{B_{un}} \right) - \log \left(\frac{U_s}{U_u} \right).$$

Now use (27) to replace expenditures with wages

$$s_n = \theta(1 - \tau) \log \left(\frac{w_{sn}}{w_{un}} \right) + \log \left(\frac{B_{sn}}{B_{un}} \right) - \log \left(\frac{U_s}{U_u} \right).$$

Next, we replace wages with productivities and labor supplies, using (23), and rearrange

$$(1 + \theta(1 - \tau)\rho^{-1}) s_n = \theta(1 - \tau) \log a_n + \log \left(\frac{B_{sn}}{B_{un}} \right) - \log \left(\frac{U_s}{U_u} \right).$$

Differencing this equation between any two locations shows that the difference between s_n and s_m for any two locations n and m depends only on location specific fundamentals and not on A. So changes in A have no effect on the variance of s_n, i.e on sorting S.

D Calibration

Tax system

We use data from the 1981/91/2001/11 waves of the PSID (each containing summary information on the *prior* year's income). Using the same sample restrictions as in section 2, we run the PSID data through the NBER's TAXSIM program. For each household, pre-tax income is computed as adjusted gross income minus Social Security transfers. Post-tax income is computed as pre-tax income minus federal and state taxes (including payroll taxes) plus Social Security transfers. We estimate (30) by pooled OLS over the four periods. Our estimated $\hat{\tau}$ is 0.174 (robust s.e. 0.003). The R^2 of the regression is 0.98, suggesting that, despite its parsimony, a log-linear tax equation is a good approximation to the actual tax system in the United States. Our estimate is quite close to Heathcote, Storesletten, and Violante (2017), who estimate $\hat{\tau} = 0.181$ despite constructing a household sample with slightly different inclusion criteria.

Housing Supply Elasticities

Our estimating equation is

$$\Delta \log p_n = \Delta \log \tilde{\Pi}_n + (\chi + \chi_L UNAVAL_n + \chi_R WRLURI_n) \Delta \log \left(\sum_i R_{in} l_{in} \right). \tag{39}$$

We bring this to the data using changes between 1980 and 2010. Saiz (2010) reports values of land unavailability $UNAVAL_n$ and regulatory constrains $WRLURI_n$ for a subset of MSAs. After dropping those for which these measures are missing, we are left with 193 MSAs. We use Census data on rents and incomes to construct model-consistent prices p_n using (3), setting $\Omega = 1$ throughout. We use Census data on rents and employment to construct housing expenditure $\sum_i R_{in} l_{in}$ for each MSA. Finally we use the Bartik shifter Z_{int} (and its interactions with $UNAVAL_n$ and $WRLURI_n$) described in the text as an instrument for housing expenditure. Table D.1 reports the result of estimating (39) by 2SLS. For the 193 locations with complete data, we then define

$$\gamma_n = \frac{\chi_n}{1 - \chi_n}$$

where

$$\chi_n = \chi + \chi_L UNAVAL_n + \chi_R WRLURI_n.$$

This procedure results in a negative value for one MSA. We replace this with a zero. Of the remaining locations, 50 are the nonmetro portions of states and 26 are MSAs for which $UNAVAL_n$ and $WRLURI_n$ are not available. For the 26 MSAs, we define γ_n to be the median among the 193 MSAs with complete information. For the 50 state residuals, we set γ_n to the lowest value among the 193 MSAs with complete information, on the assumption that supply is likely to be more elastic in nonmetro areas.

Migration elasticity

We estimate θ by requiring our model to match the results of Hornbeck and Moretti (2019). That paper estimates the causal effect of TFP shocks between 1980 and 1990 on employment and wages. We mimic their setting by shutting down all shocks other than shocks to productivity and then repeating their regressions using the output of our model. Our target is the ratio of the effect on employment to the effect on wages by 2010 — the long run elasticity of employment to wages. This implies a target of $4.03/1.46 = 2.76$ (see Table 2, Column (3) of Hornbeck and Moretti (2019)). Formally we proceed as follows:

(i) Guess θ

Table D.1: Housing Supply Elasticity Estimates
Dependent variable: Log price change, 1980-2010

χ	-0.177
	(0.208)
χ_L	0.363
	(0.106)
χ_R	0.526
	(0.118)

Source: Census. Robust standard errors in parentheses.

(ii) Invert the model in 1980 and 1990 to obtain fundamentals $\left(A^t_{in}, B^t_{in}\right)_{i,n}, \left(\Pi^t_n\right)_n, \left(L^t_i\right)_i$ for $t = 1980, 1990$.

(iii) Solve the model with fundamentals $\left(A^{90}_{in}, B^{80}_{in}\right)_{i,n}, \left(\Pi^{80}_n\right)_n, \left(L^{80}_i\right)_i$ to obtain $\left(\hat{l}^{90}_{in}, \hat{w}^{90}_{in}\right)_{i.n}$.

(iv) Define $L^{80}_n = \sum_i l^{80}_{in}$, $W^{80}_n = \sum_i l^{80}_{in} w^{80}_{in} / \sum_i l^{80}_{in}$ and $\log Z^{80}_n = \sum_i l^{80}_{in} \log A^{80}_{in} / \sum_i l^{80}_{in}$ and likewise for $\hat{L}^{90}_n$, $\hat{W}^{90}_n$ and $\log \hat{Z}^{90}_n$

(v) Estimate the models below by OLS, weighting by 1980 employment:

$$\log \hat{L}^{90}_n - \log L^{80}_n = \pi^L \left(\hat{Z}^{90}_n - Z^{80}_n\right) + v^L_n$$
$$\log \hat{W}^{90}_n - \log W^{80}_n = \pi^W \left(\hat{Z}^{90}_n - Z^{80}_n\right) + v^W_n.$$

Note that the fact that we only study changes between 1980 and 1990 is innocuous, because our model has no transitional dynamics.

(vi) Calculate π^L / π^W.

(vii) Update θ until π^L / π^W converges to the target value.

This procedure yields $\theta = 3.89$.

E Counterfactual

E.1 Alternative Measures of Sorting

Here we define the alternative measures of sorting discussed in subsection 5.2.

The Theil index for a non-negative variable x with weights ω_n is defined as

$$T = \sum_i \omega_n \left(\frac{x_n}{\bar{x}}\right) \log \left(\frac{x_n}{\bar{x}}\right)$$

where $\bar{x}$ is the weighted average of x_n. We use this as a measure of sorting by setting $x_n = \exp\left(s_n\right)$, where s_n is the log-skill ratio, and weight by 1980 employment.

The dissimilarity index D for two populations u and s, spread over geographical units indexed by n is given by

$$D = \frac{1}{2} \sum_n | \left(\frac{l_{sn}}{L_s}\right) - \left(\frac{l_{un}}{L_u}\right) |$$

Note that employment weights are already implicit in this expression.

E.2 Alternative Parametrization of Preferences

We recalibrate our model to Price Independent Generalized Linear (PIGL) utility, a leading case of nonhomothetic preferences (Boppart 2014; Eckert and Peters 2018). PIGL admits a closed form for the indirect utility function (14),

$$v_{in} = \frac{1}{\varepsilon}(e_{in}^{\varepsilon} - 1) - \frac{\nu}{\varsigma}(p_n^{-\varsigma} - 1)$$

for parameters $0 < \varepsilon < \varsigma < 1$ and $\nu > 0$. By Roy's identity, the housing share is

$$\eta_{in} = \nu e_{in}^{-\varepsilon} p_n^{\varsigma} \tag{40}$$

Taking logs, adding a time subscript, and interpreting the scalar ν into an idiosyncratic household demand shifter ν_{int}, (40) is equivalent to the estimating equation (7) for NHCES utility. The income elasticity is ε and the price elasticity is ς, which correspond to β and ψ, respectively, in (8). We can therefore read the parameters directly off Table 1, setting $\varepsilon = 0.24$ and $\varsigma = 0.37$. After recalibrating the full model we find that the skill premium explains 19% of the increase in sorting since 1980, comparable to our baseline results. More generally, (8) is a first order approximation to any demand system. We conclude that our findings are not sensitive to the parametrization of utility.

E.3 Endogenous Amenities

Diamond (2016) shows the importance of endogenous amenities for understanding the location choices of skilled versus unskilled workers. In this subsection we consider how our results might change in the presence of endogenous amenities.

We start by incorporating them into the simple model described in Section 3. Following Diamond (2016) we model amenities as

$$B_{in} = b_{in} \left(\frac{l_{sn}}{l_{un}} \right)^{\beta_s}. \tag{41}$$

That is, for both types amenities depend on the skill ratio, but different types may value them differently — this is captured by β_s. In the context of our simple model, we do not allow exogenous differences in amenities across types, and so we impose $b_{sn} = b_{un} = b_n$. Diamond (2016) shows that skilled households value endogenous amenities more than unskilled households, implying $\beta_s > \beta_u$. We also impose $\beta_s - \beta_u < 1$ to avoid endogenous amenities so strong that they cause perfect sorting (i.e a situation in which skilled and unskilled workers inhabit totally different locations).

It is helpful to compare two economies with the same fundamentals — one without endogenous amenities, whose variables are denoted by $\bar{x}$, and one with endogenous amenities, whose variables are denoted by $\tilde{x}$. In the economy with endogenous amenities (18) becomes

$$\tilde{s}_n = \tilde{\zeta} - \theta \left(\log \tilde{P}_{sn} - \log \tilde{P}_{un} \right) + (\beta_s - \beta_u)\, \tilde{s}_n. \tag{42}$$

Notice that in our model the ideal price indices are independent of the presence of endogenous amenities, and so $\bar{P}_{in} = \tilde{P}_{in}$. This implies

$$\bar{s}_n = \bar{\zeta} - \theta \left(\log \tilde{P}_{sn} - \log \tilde{P}_{un} \right)$$

and therefore

$$\tilde{s}_n = (1 - (\beta_s - \beta_u))^{-1} \left(\bar{s}_n + (\tilde{\zeta} - \bar{\zeta}) \right). \tag{43}$$

That is, skill ratios in the economy with endogenous amenities are simply an affine transformation of skill ratios in the economy without endogenous amenities. In particular given our assumption on β_s and β_u, the slope of $\tilde{s}_n$ with respect to $\bar{s}_n$ is above one. This leads to the first result of this section

Proposition 3 *Suppose $\beta_s > \beta_u$. If $\epsilon < 0$, sorting is higher in the presence of endogenous amenities, i.e $\tilde{S} > \bar{S}$. If instead $\epsilon = 0$ then $\tilde{S} = \bar{S} = 0$.*

This follows directly from observing that $\tilde{s}_n$ is an affine transformation of $\bar{s}_n$ with a coefficient on $\bar{s}_n$ above 1. Proposition 3 tells us that endogenous amenities amplify the effects of nonhomothetic housing demand. Nonhomothetic housing demand ensures that high price locations have a higher skill ratio. Endogenous amenities then encourage even more skilled workers to locate there. But it is important to note that when $\epsilon = 0$, there is no sorting even with endogenous amenities, showing that they do not create an independent motive for sorting in our model, but rather amplify existing ones.

We now proceed to our next result, concerning the effect of an increase in the skill premium, $d \log A > 0$. Differentiating (42) yields

$$d\tilde{s}_n = (1 - (\beta_s - \beta_u))^{-1} \left(d\tilde{\zeta} - \theta \left(d \log \tilde{P}_{sn} - \log \tilde{P}_{un} \right) \right).$$

Again substituting out prices using the economy without endogenous amenities, we obtain

$$d\tilde{s}_n = (1 - (\beta_s - \beta_u))^{-1} \left(d\bar{s}_n + \left(d\tilde{\zeta} - d\bar{\zeta} \right) \right).$$

Now $d\tilde{s}_n$ is an affine function of $d\bar{s}_n$ with a coefficient on $d\bar{s}_n$ above 1. Following the same steps as above, we obtain the result.

Proposition 4 *Suppose $\beta_s > \beta_u$. If $\epsilon < 0$, sorting increases more when amenities are endogenous. Formally, $d\tilde{S} > d\bar{S}$. When $\epsilon = 0$ then $d\tilde{S} = d\bar{S} = 0$*

Proposition 4 shows that endogenous amenities amplify the mechanism we focus on in this paper — diverging incomes causing diverging sensitivities to housing costs and thus diverging location choices – but do not independently link the skill premium to spatial sorting.

Finally, we extend Proposition 2 to a richer environment with endogenous amenities. We drop the assumption that $b_{sn} = b_{un}$. Adding endogenous amenities does not change the derivation presented in the proof of Proposition 2, so we start from

$$(1 + \theta(1-\tau)\rho^{-1})s_n = \theta(1-\tau)\log a_n + \log\left(\frac{B_{sn}}{B_{un}}\right) - \log\left(\frac{U_s}{U_u}\right).$$

Inserting our definition of B_{in} and rearranging, we obtain

$$(1 + \theta(1-\tau)\rho^{-1} - (\beta_s - \beta_u))s_n = \theta(1-\tau)\log a_n + \log\left(\frac{\bar{b}_{sn}}{\bar{b}_{un}}\right) - \log\left(\frac{U_s}{U_u}\right).$$

Following exactly the same steps as in Proposition 2, we obtain our final result:

Proposition 5 *Suppose $\beta_i \neq 0$. Suppose also $\epsilon = 1$ so that preferences are homothetic. Then changes in aggregate skill-bias A have no effect on sorting S.*

Proposition 5 tells us that even in the quantitative model, if preferences are homothetic then endogenous amenities do not independently link the skill premium to spatial sorting.

E.4 Agglomeration Externalities

The production function in Fajgelbaum and Gaubert (2020) is

$$y_n = ((\tilde{z}_{un} l_{un})^{\frac{\rho-1}{\rho}} + (\tilde{z}_{sn} l_{sn})^{\frac{\rho-1}{\rho}})^{\frac{\rho}{\rho-1}}.$$

This is equivalent to our framework with

$$\begin{aligned}
\tilde{z}_{un} &= Z z_n \\
\tilde{z}_{sn} &= Z z_n A a_n
\end{aligned}$$

Fajgelbaum and Gaubert (2020) parametrize the productivities as $\tilde{z}_{un} = \tilde{Z}_{un} l_{sn}^{\gamma_{su}} l_{un}^{\gamma_{uu}}$ and $\tilde{z}_{sn} = \tilde{Z}_{sn} l_{sn}^{\gamma_{ss}} l_{un}^{\gamma_{us}}$. Matching coefficients to (31) and (32) gives

$$\begin{aligned}
\phi^{z,s} &= \gamma_{su} \\
\phi^{z,u} &= \gamma_{uu} \\
\phi^{a,s} &= \gamma_{ss} - \gamma_{su} \\
\phi^{a,u} &= \gamma_{us} - \gamma_{uu}
\end{aligned}$$

where $(\gamma_{su}, \gamma_{uu}, \gamma_{ss}, \gamma_{us}) = (0.044, 0.003, 0.053, 0.02)$ from Fajgelbaum and Gaubert (2020).