SORTING TO EXPENSIVE CITIES*

Cécile Gaubert† Frédéric Robert-Nicoud‡

March 2025

Abstract

We propose a spatial equilibrium model with heterogeneous households holding general non-homothetic preferences over tradable goods and housing. In equilibrium, desirable and productive locations command high housing prices. So long as housing is a necessity, these locations are disproportionately inhabited by high-income earners who are relatively less affected by high housing prices. We clarify how this source of sorting complements other potential sorting forces in spatial equilibrium models, namely, comparative advantage in production and heterogeneous preferences for locations. We show how to measure changes in welfare inequality across income groups in a theoretically-consistent way when housing is a necessity, extending the approach popular in models with homothetic preferences. We use our framework to track the evolution of welfare inequality between college and non-college graduates in the United States between 1980 and 2020. We find that, accounting for change in prices, it has risen by more than nominal wage inequality, even as college graduates increasingly sort into cities with expensive housing over this time period.

*For comments, suggestions, and discussions, we are grateful in particular to Don Davis, Jonathan Dingel, Gilles Duranton, J. Peter Neary, Rowan Shi, Kohei Takeda, and participants to seminars and conferences at Bern, Bocconi, CURE, Duisburg, Geneva, HEC Paris, NYU-AD, OSUS, SecGo, Singapore Management University, and the Urban Economics Association. We thank Matthew Tauzer for excellent research assistance. Cécile Gaubert acknowledges support from NSF CAREER grant #1941917.

†UC Berkeley, NBER and CEPR. cecile.gaubert@berkeley.edu

‡GSEM (Université de Genève) and CEPR. Frederic.Robert-Nicoud@unige.ch

1 Introduction

Cities differ dramatically in size, productivity, and cost of living. They also vary in composition: larger, more productive, and more expensive cities typically attract a disproportionate share of high-skill workers and high-income earners — that is, they feature spatial sorting (Diamond and Gaubert, 2022). While a growing literature relies on quantitative spatial models to study the uneven distribution of economic activity in space, these models typically assume homogeneous agents with homothetic preferences over housing and consumption (Redding and Rossi-Hansberg, 2017). These assumptions facilitate tractability and welfare analysis and make these models well suited to analyze quantitative differences across cities, such as their relative sizes. However, they cannot, by construction, explain why cities differ in their composition.

In this paper, we propose a spatial model where heterogeneous households hold general non-homothetic preferences for housing. The model is well-suited to analyze spatial sorting, as it makes consumption patterns—and thus location choices— income-dependent. Higher-ability and higher-income workers, who allocate a smaller share of their income to housing, are more willing to pay the premium to live in attractive, high-cost cities, leading to endogenous spatial sorting. Our framework has three main advantages. First, the assumption of non-homotheticity is realistic. Second, it provides a microfounded driver of sorting. This is in contrast to a common alternative in the literature that relies on a standard homothetic model but assume exogenous differences in preferences for urban amenities across worker groups. While useful for describing observed patterns, such models assume preference heterogeneity rather than deriving spatial sorting from fundamental economic behavior. Third, our approach preserves a common preference structure across agents, making welfare comparisons meaningful.

We first characterize the theoretical properties of a spatial model with general non-homothetic preferences. We then show how the model allows to track welfare inequality between skill groups using a convenient sufficient statistics approach. This extends the popular methods in the quantitative spatial literature to settings with non-homothetic preferences. A key sufficient statistic for welfare in our framework is the housing expenditure share, which plays no role in models with homothetic preferences. Finally, we use our framework to quantify the evolution of welfare inequality between college and high-school graduates in the United States from 1980 to 2020, accounting for changes in spatial sorting. Moretti (2013), using a framework with homothetic preferences, finds that real wages grew less unequal than nominal wages because college graduates increasingly clustered in expensive cities. In contrast, when we incorporate income-dependent housing expenditure shares, we find that welfare inequality rose by more than nominal wage differences suggest. Although low-wage workers are under-represented in expensive cities, those who do live there bear a heavier burden from high housing costs compared to high-income workers.

Our framework is motivated by four established stylized facts:

1. **Housing is more expensive in large cities**. For instance, dwelling rent increase with respect to urban population density with an elasticity of about 0.15 – as reported in the meta-analysis of Ahlfeldt and Pietrostefani (2019).

2. **Housing demand is income-inelastic**. The expenditure share of housing falls with household income.[1] In addition, housing is the largest single component of household consumption, making it determinant in location choices. Specifically, housing costs represent a median of 28% of a household's expenditures and up to 62% for the bottom decile of renters by income based on French data (Combes, Duranton, and Gobillon, 2019).

3. **The urban wage premium increases with workers' skill**. Wages rise with city size, and this urban wage premium is larger for more skilled workers.[2] Davis and Dingel (2020) report that the elasticity of the urban skill premium with respect to city population size is 0.3 in the United States.

4. **Higher incomes and highly educated workers are over-represented in larger cities** in the U.S. (Glaeser and Mare, 2001; Baum-Snow and Pavan, 2013) and beyond.[3] This pattern holds within narrowly defined skilled groups: the elasticity of PhD graduates to city population is 1.14 in the United States (Davis and Dingel, 2020).

Fact 2 is the foundation for our core assumption: we take household preferences to be non-homothetic, with housing as a necessity. Combined with Fact 1, this will generate a natural mechanism for sorting: high-income households self-select into larger, more expensive cities. Alone, however, Facts 1 and 2 would imply that the urban wage premium should *decrease* with city size (Black, Kolesnikova, and Taylor, 2009). In spatial equilibrium, differences in

[1] See e.g. Aguiar and Bils (2015); Albouy, Ehrlich, and Liu (2016); Finlay and Williams (2022); Ganong and Shoag (2017); Glaeser, Kahn, and Rappaport (2008).

[2] See e.g. Baum-Snow and Pavan (2013); Baum-Snow, Freedman, and Pavan (2018); Davis and Dingel (2019); Card, Rothstein, and Yi (2023).

[3] The fact was established in particular for France (Combes, Duranton, and Gobillon, 2008) , Britain (D'Costa and Overman, 2014), as well as in the developing world (Dingel, Miscio, and Davis, 2021). In contrast to this evidence of sorting on observed individual characteristics, the evidence of sorting on unobserved ability is more mixed (Baum-Snow and Pavan, 2012; Combes, Duranton, and Gobillon, 2008; Dauth, Findeisen, Moretti, and Suedekum, 2022; De La Roca and Puga, 2017).

the burden of housing costs should offset differences in skill returns across locations. Since housing costs affect a smaller share of the expenditure of high-income households than that of low-income households, standard compensating differential logic suggests that skill premia should be lower in expensive cities. Therefore, Fact 3 suggests an additional force at play: high-ability workers have a comparative advantage in productive cities. This comparative advantage together with cost-of-living effects drive spatial sorting (Fact 4).[4]

The paper is organized as follows. We first lay out our spatial sorting model and establish conditions for equilibrium existence and uniqueness (Section 2). Cities differ in exogenous amenities and productivity. Workers are endowed with heterogeneous ability, which may be relatively more valuable in some cities than others, and have general, potentially non-homothetic preferences over a freely traded consumption good and housing. They also have idiosyncratic preferences for locations, generating imperfect sorting within skill. We do not impose specific functional forms on preferences or technology. We establish that the equilibrium exists under general technical conditions. Moreover, we show that the equilibrium is unique if there is a single type of workers, whether preferences are homothetic or not, or if utility is homogeneous in income, whether workers have homogeneous ability or not. In contrast, the interaction between non-homothetic preferences and heterogeneity in worker types may generate multiple equilibria. In this case, we provide sufficient conditions for uniqueness that can be evaluated in quantitative applications.

We then characterize theoretical conditions under which spatial sorting emerges (Section 3). We highlight how these conditions relate to existing frameworks in the literature. A key insight is that seemingly innocuous functional form choices can strongly influence sorting outcomes. We clarify that both the sub-utility of consumption and the distribution of locational preference shocks affect sorting. In fact, it is the combination of these two that drive theoretical results, rather than each of them separately. As a result, the a priori uncontroversial choice of the distribution of idiosyncratic shocks for locations can be as consequential for spatial sorting as choosing between a homothetic and a non-homothetic utility function. We next examine how city fundamentals impact worker composition. In equilibrium, cities with better consumption or production amenities exhibit higher housing rents, consistent with classic urban theory (Rosen, 1979; Roback, 1982). We establish conditions under which higher-ability workers sort into these cities. Non-homothetic preferences alone yield such sorting, since even in the absence of comparative advantage and even if household agree on the quality of local amenities, higher income workers can more easily afford the higher cost

[4] A third driver of sorting may be taste-based, driven by heterogeneous preferences for local amenities or local public goods (Kuminoff, Smith, and Timmins, 2013; Diamond, 2016; Fajgelbaum and Gaubert, 2020; Diamond and Gaubert, 2022). We discuss this source of sorting briefly in Section 3.1.

of living in attractive places. In addition, sorting also arises if high-type workers have what we define as a "strong comparative advantage" in productive cities. In this case, the two sorting mechanisms reinforce each other.

Third, we study the normative implications of the model (Section 4). In quantitative spatial models with homogeneous agents, welfare changes are often measured using a "sufficient statistics" approach, which computes the representative welfare effects of economic shocks based on model elasticities and a small set of key statistics from the data. We extend this approach to our model with heterogeneous agents and non-homothetic preferences to measure the distributional consequences of shocks. We show that relative welfare changes across any pair of skill types in the economy can be expressed as a function of three statistics that are usually readily available in the data: (i) relative change in nominal incomes, (ii) relative changes in location choices, and (iii) relative change in housing expenditure shares. The first two elements are standard in models with homothetic preferences, but the third component is novel to our framework. It captures how local housing price changes affect workers differently based on income, reflecting the unequal burden of housing costs. Intuitively, measuring changes in expenditure share on a necessity is informative about welfare change (Atkin, Faber, Fally, and Gonzalez-Navarro, 2023). Finally, to enable meaningful interpersonal welfare comparisons in the context of non-homothetic preferences, we convert utility changes into money-metric terms.

We apply this framework to measure how welfare inequality between college and high-school graduates evolved in the United States between 1980 and 2020 (Section 5). Our approach differs from Moretti (2013)'s original contribution on this question in two ways. First, we account for heterogeneous housing expenditure shares. College graduates spend a substantially smaller share of income on housing than high-school graduates, which alters the welfare effects of housing price changes. Second, we incorporate the fact that sorting is not random – newcomers attracted to a city following a shock are a selected group that has lower intrinsic preference for that location than incumbent residents. This leads to a selection correction when computing welfare changes. If nominal income inequality has increased by 25% between 1980 and 2020, welfare inequality has increased by 27% to 31%, depending on the specification. In the context of the housing affordability crisis, taking into account the disproportionate burden that housing cost place on lower-income households – even though they sort less into expensive cities than high-income ones – sheds a strikingly different light on welfare inequality.

This paper contributes to the Economic Geography literature by studying the properties of spatial models with non-homothetic preferences, thereby enriching the quantitative spatial toolbox pioneered by Ahlfeldt, Redding, Sturm, and Wolf (2015) and Allen and Arkolakis

(2014). Within quantitative spatial frameworks, a literature posits different homothetic preferences for different groups to generate sorting across cities (Diamond, 2016; Giannone, 2019; Fajgelbaum and Gaubert, 2020) or sorting within cities across neighborhoods (Tsivanidis, 2019; Almagro and Domínguez-Iino, 2020), different from our approach. In contrast, some recent applied quantitative work has incorporated income-based sorting in their analysis of sorting across regions or cities (Finlay and Williams, 2022; Fretz, Parchet, and Robert-Nicoud, 2021; Takeda, 2022) or within cities (Couture, Gaubert, Handbury, and Hurst, 2023; Hoelzlein, 2023; Gaigné, Koster, Moizeau, and Thisse, 2022; Tsivanidis, 2019). Also related, Handbury (2021) measures the prevalence and welfare incidence of heterogeneous prices over space. Our work complements these papers by deriving general positive and normative theoretical results. Our research also connects with earlier (non-quantitative) theories of systems of cities, specifically those featuring heterogeneous agents making heterogeneous location choices in equilibrium (Behrens, Duranton, and Robert-Nicoud, 2014; Davis and Dingel, 2019). In these papers, sorting follows from the popular assumption of unit housing consumption, an extreme form of non-homotheticity, which we relax. Finally our welfare analysis relates to the recent work of Baqaee and Burstein (2022), Jaravel (2021), Jaravel and Lashkari (2022) and Oberfield (2023) that study the measurement of welfare inequality with income-dependent preferences in non-spatial contexts.

2 An Elementary Spatial Sorting Model With Non-Homothetic Preferences

2.1 Setup

Consider an economy populated by a continuum of workers of heterogeneous skills $t, s \in \mathcal{T}$. The set of skills is discrete, and there is a fixed aggregate supply L^t of workers of skill t. Workers make a discrete choice of where to live among $n, m \in \mathcal{C} \equiv \{1, ..., N\}$ locations. A worker of type t supplies one unit of labor where they live, facing wage w_n^t in location n and a rental cost of housing r_n.

Preferences Preferences are the same for all workers. They consume a freely traded good c (taken as the numeraire), housing h, and directly derive utility from the amenities A_n of the city where they live. Worker ω has idiosyncratic preferences for locations $\{\varepsilon_n^\omega\}_{n\in\mathcal{C}}$, iid across workers. In line with extant work, we assume that utility is separable in its idiosyncratic component, ε_n^i, and its common component, $\tilde{U}\left(c, h, A_n\right)$ and restrict the analysis to distributions of idiosyncratic shocks that result in Luce aggregation. That is, there exists an increasing function $g(.)$ such that we can write the fraction of workers of type t choosing to live and work in location n as:

$$\lambda_n^t = \frac{V_n^t}{\sum_{m\in\mathcal{C}} V_m^t},$$

where $V_n^t \equiv g\left(\tilde{V}_n^t\right)$ in a monotonic transform of the common component of indirect utility in location n:

$$\tilde{V}_n^t = \max_{c,h} \tilde{U}\left(c, h, A_n\right)$$
$$\text{such that } c + r_n h = w_n^t.$$

In particular, under the assumption popular in the quantitative spatial literature that utility is $\tilde{U}\varepsilon$, i.e. multiplicative and separable in idiosyncratic shocks, and that idiosyncratic shocks are distributed Frechet with shape parameter $\kappa > 1$, we have $g\left(\tilde{V}_n^t\right) = \left(\tilde{V}_n^t\right)^\kappa$. Location choice probabilities become:

$$\lambda_n^t = \frac{\left(\tilde{V}_n^t\right)^\kappa}{\sum_{m\in\mathcal{C}}\left(\tilde{V}_m^t\right)^\kappa}. \tag{1}$$

If instead utility is $\tilde{U} + \varepsilon$ and idiosyncratic shocks follow the standard Gumbel distribution, an assumption popular in IO that is also used in the urban economics literature (see e.g., Almagro and Domínguez-Iino 2020; Diamond 2016; Giannone 2019), then $g\left(\tilde{V}_n^t\right) = \exp \tilde{V}_n^t$ and choice probabilities are:

$$\lambda_n^t = \frac{\exp \tilde{V}_n^t}{\sum_{m\in\mathcal{C}} \exp \tilde{V}_m^t}. \tag{2}$$

With some abuse of terminology, we henceforth refer to V_n^t as the (common component of) indirect utility of type t in city n, even though it is a combination of the indirect utility proper, $\tilde{V}_n^t$, and of the functional form chosen for the distribution of the idiosyncratic preferences shocks.

The Marshallian demand for housing in city n for skill t is denoted h_n^t :

$$h_n^t \equiv h\left(w_n^t, r_n, A_n\right).$$

We assume it is continuously differentiable in all its arguments. Furthermore, we take housing to be a normal good so that h_n^t is increasing in its first argument and decreasing in the second. Our object of interest in this paper is the case where preferences are non-homothetic in (c, h). Specifically, we will focus on the case where housing is a necessity, so that the expenditure share on housing is non-decreasing in income, consistent with empirical evidence.[5] Denoting:

$$s_n^t = \frac{r_n h_n^t}{w_n^t} \tag{3}$$

the housing expenditure shares, we assume in what follows that

$$\frac{\partial s_n^t}{\partial w_n^t} \leq 0. \tag{4}$$

Popular functional forms Though we work with minimalist assumptions throughout, and most of our results do not depend on parametric assumptions, applied work requires using specific functional forms. It is therefore useful to summarize how popular functional forms for preferences fit in our framework.

Table 1 does so. The first column lists Stone-Geary and the acronyms of five other classes of non-homothetic preferences, several of which are well-known (see the Table caption for the meaning of acronyms and for references). The second column reports the indirect utility function V as a function of the vector of prices (p, r, w) and amenities A. We choose V to be multiplicatively separable in amenities as we discuss in Section 3.1. Parameter $\nu \in \{0, 1\}$ governs the (non-)homotheticity of each class of preferences. In all cases, preferences are homothetic if $\nu = 0$ and non-homothetic if $\nu = 1$. Under the assumption $\nu = 1$, the third column imposes restrictions on the structural parameters that ensure that the inequality in equation (4) holds (housing is a necessity).

Most applications in urban economics use either of the following polar preferences over housing and the consumption good: homothetic Cobb-Douglas (in which case the price- and income-elasticity of housing demand are unitary), or unit housing consumption (in which households consume exactly one unit of housing).[6] Recent exceptions include Albouy, Ehrlich, and Liu (2016) and Finlay and Williams (2022), who use NH-CES, Eckert and

[5] In contrast, the Economic Geography literature usually models housing consumption assuming Cobb-Douglas utility, so that expenditure share on housing is constant across households and cities ($\frac{\partial s_n^t}{\partial w_n^t} = 0$). This latter assumption is convenient, together with the assumption of homogeneous workers, to reflect the aggregate patterns reported by Davis and Ortalo-Magné (2011): the median of the distribution of expenditure share on housing is stable across cities. However, this aggregate empirical pattern masks heterogeneity in consumption patterns of housing across incomes within cities – poorer households spend more on housing as a share of their income – as well as heterogeneity in income distribution across cities. These cannot be rationalized with Cobb-Douglas preferences and instead call for non-homothetic preferences, motivating the current study.

[6] These polar cases are limiting cases of some of the functional forms listed in Table 2. Homothetic Cobb-Douglas preferences correspond to limit case $\nu = 0$ in all cases of Table 2, with the exception of CRIE and NH-CES. The case of unit housing consumption corresponds to the limiting case $\psi \to 0$ and the normalization $\underline{h} = 1$ if preferences are Stone-Geary, and to the limiting case $\psi \to 0, \iota \to 1$ if preferences are PIGL.

Table 1: Non-homothetic preferences: Functional forms

Preferences	Indirect utility V	Structural parameters
Stone-Geary	$V = A\frac{w-\nu \underline{h} r}{p^{1-\psi}r^{\psi}}$	$\psi \in (0,1)$, $\underline{h} > 0$
PIGL	$V = A\left[\left(\frac{w}{p^{1-\psi}r^{\psi}}\right)^{\iota} - \frac{\nu}{\gamma}\left(\frac{r}{p}\right)^{\gamma\iota}\right]^{1/\iota}$	$\iota, \psi \in (0,1)$, $\gamma \geq 1$
PIGLOG	$V = A\left(\frac{r}{p}\right)^{\nu}\ln\left(\frac{w}{p^{1-\psi}r^{\psi}}\right)$	$\psi \in (0,1)$
CRIE	$V = A\left[\left(\frac{w}{r}\right)^{-\psi} + \left(\frac{w}{p}\right)^{-\psi+\nu(\psi-\mu)}\right]^{-1}$	$\psi > \mu$
NH-CES	$w^{1-\sigma} = \left(\frac{V}{A}\right)^{1-\sigma}\left[\left(\frac{V}{A}\right)^{\nu(\mu-\psi)}r^{1-\sigma} + p^{1-\sigma}\right]$	$\sigma \in (0,1)$, $\psi > \mu > \sigma - 1$
HCD	$0 = \left(\frac{V}{A}\right)^{\mu}\ln\left(\frac{r}{w}\right) + \left(\frac{V}{A}\right)^{\mu+\nu(\psi-\mu)}\ln\left(\frac{p}{w}\right)$	$\psi > \mu$

Notes: PIGL stands for "Price Independent Generalized Linearity" (Muellbauer, 1975), PIGLOG stands for "Price Invariant Generalized Logarithmic" (Deaton and Muellbauer, 1980), CRIE stands for "Constant Relative Income Elasticity" (Caron, Fally, and Markusen, 2014; Eeckhout, Pinheiro, and Schmidheiny, 2014; Fieler, 2011; Hanoch, 1975), NH-CES stands for "Non Homothetic Constant Elasticity of Substitution" (Albouy, Ehrlich, and Liu, 2016; Comin, Lashkari, and Mestieri, 2021; Hanoch, 1975; Matsuyama, 2019; Sato, 1977), and HCD stands for "Heterothetic Cobb-Douglas" (Bohr, Mestieri, and Robert-Nicoud, 2024). HCD preferences are a limiting case of NH-CES preferences. The indicator variable $\nu \in \{0,1\}$ governs the (non-)homotheticity of each class of preferences. In all cases, preferences are homothetic if $\nu = 0$ and non-homothetic if $\nu = 1$.

Peters (2022) and Fan, Peters, and Zilibotti (2023), who use PIGL, and Fretz, Parchet, and Robert-Nicoud (2021) and Tsivanidis (2019), who use Stone-Geary.

Production The traded good is produced in each city by perfectly competitive producers using labor under constant returns to scale. Worker skills are perfect substitutes in production, but they may differ in their complementarity with city productivity T_n. Workers of type t who work in city n provide $\ell_n^t \equiv \ell\left(T_n, t\right)$ effective units or labor, with

$$\frac{\partial \ell\left(T_n, t\right)}{\partial T_n} \geq 0, \quad \frac{\partial \ell\left(T_n, t\right)}{\partial t} \geq 0.$$

Higher-t workers are more productive everywhere, and all worker types are more productive in high-T cities. We leave $\ell\left(T_n, t\right)$ general for now, but will explore assumptions on the complementarity between T_n and t later on. For simplicity, we model here vertical differentiation of worker skills (ability) without modeling their horizontal differentiation (between different sectors or occupation for instance). Therefore, results have to be understood as characterizing spatial sorting on ability within a given sector or occupation. Traded output is simply:

$$Y_n = \sum_{t \in \mathcal{T}} L^t \lambda_n^t \ell_n^t.$$

Housing Housing is produced by perfect competitors in each market combining land and the traded good under constant returns to scale. We parameterize the aggregate supply function for housing in city n so that both the level of housing supply and the elasticity of housing supply can vary across markets (Saiz, 2010). They are captured respectively by parameters H_n and η_n in the following production function:

$$\mathcal{H}_n = H_n \, k\left(\frac{r_n}{p}; \eta_n\right), \tag{5}$$

where we assume:

$$\frac{\partial k}{\partial r} > 0, \quad \frac{\partial k}{\partial \eta} > 0, \qquad \frac{\partial^2 \ln k}{\partial \eta \partial \ln r} \geq 0, \tag{6}$$

This latter assumption insures that the elasticity of housing supply is increasing in η, and we may refer to η as the housing supply elasticity for short when we compare two locations with different η's.

Taking stock, cities are differentiated in four dimensions: productivity T_n, amenities A_n, land endowment H_n, and housing supply elasticity η_n. We now analyze an equilibrium of this economy, with the goal of understanding how these city primitives impact the sorting of heterogeneous skills across locations.

2.2 Equilibrium

An equilibrium of this economy is a set of prices $\left(\{r_n\}_n, \{w_n^t\}_{n,t}\right)$, a distribution of types across space $\{\lambda_n^t\}_{n,t}$, and consumption profiles by type and city $\{c_n^t, h_n^t\}_{n,t}$ such that the following conditions hold:

(i) Workers are paid their marginal product, so that $\forall n, t$:

$$w_n^t = \ell\left(T_n, t\right). \tag{7}$$

(ii) Workers choose location n, and consumption c and h to maximize utility under their budget constraint, given prices. In particular, $\forall n, t$:

$$\lambda_n^t = \frac{V_n^t}{\sum_{m \in \mathcal{C}} V_m^t}, \tag{8}$$

$$\text{where } V_n^t = V\left(w_n^t, r_n; A_n\right). \tag{9}$$

(iii) The housing market clears in each city n:

$$H_n \, k \left(\frac{r_n}{p} ; \eta_n \right) = \sum_{t \in \mathcal{T}} L^t \lambda_n^t h_n^t. \tag{10}$$

The consumption good market then clears by Walras' law. Note that after substitution of equations (7)-(9) into equation (10), the set of N equations (10) pins down the N unknown $\{r_n\}_n$ that fully summarize equilibrium conditions. All other endogenous variables are explicit functions of the vector $\{r_n\}_n$.

Our first result concerns the existence of such an equilibrium:

Proposition 1. ***Existence.*** *An equilibrium of the economy always exists.*

Proof. The proof relies on Brouwer's fixed point theorem, and is provided in Appendix A. □

The next result establishes sufficient conditions for the equilibrium to be unique. Throughout the paper, it will be convenient to use the city n-type t elasticity notation:

$$\mathcal{E}_n^t \left(x, y \right) \equiv \frac{\partial \ln x_n^t}{\partial \ln y_n^t}.$$

We also denote by $\mathcal{E}_n \left(x, y \right)$, i.e. without a superscript, elasticities of variables that are independent from worker-type.

Proposition 2. ***Uniqueness.*** *The spatial equilibrium is unique if either of the following conditions holds:*

(i) there is only one type of workers;

(ii) preferences are homothetic and multiplicatively separable in amenities;

(iii) preferences and the production function for housing jointly satisfy:

$$\inf_n \mathcal{E}_n \left(k, r \right) + \inf_{n,t} \left| \mathcal{E}_n^t \left(h, r \right) \right| > \sup_{n,t} \left| \mathcal{E}_n^t \left(V, r \right) - \overline{\mathcal{E}^t \left(V, r \right)} \right|, \tag{11}$$

where we have defined $\overline{\mathcal{E}^t \left(V, r \right)} \equiv \sum_{i=1}^N \lambda_i^t \mathcal{E}_i^t \left(V, r \right).$

Proof. See Appendix A. □

Interestingly, the threat to uniqueness comes from the *interaction* between non-homothetic preferences and type heterogeneity. Taken in isolation, neither is sufficient to lead to multiple equilibria, as conditions (i) and (ii) highlight.

To understand the threat to equilibrium uniqueness when preferences are non-homothetic and workers are of heterogeneous type, start from a reference equilibrium and consider the impact of an increase in the the price of the traded good on the housing markets of all locations. Relative housing prices decrease everywhere and, holding constant the distribution of workers across locations, this equilibrium perturbation increases demand for housing in all locations. However, what then happens to the distribution of population across cities is ambiguous. The relative attractiveness of cities changes following a housing price decrease everywhere, and these relative changes in attractiveness may lead some types to migrate out of some cities, given income effects in location choice. If out-migration is large enough to decrease total demand for housing in some city n, then the possibility of multiple equilibria arises. In particular, the demand system does not exhibit the gross substitutes property in this case, as an increase in the price of the traded good leads to a *decrease* in excess demand for housing in n. Several equilibria may coexist in principle: for instance, a reference equilibrium alongside an equilibrium with higher traded-good price and higher housing costs everywhere except in n, where they are lower.

When the economy features several workers types, a sufficient condition for uniqueness is given by equation (11). Consider first the left-hand side of this inequality. When the housing supply elasticity $\mathcal{E}_n(k, r)$ is high, a pervasive decrease in housing rents yields strong reductions in housing supply. Similarly, a high price-elasticity of housing demand also leads to a large increase in excess demand when housing prices decrease. Both channels lead to large responses in the excess demand for housing, which makes the gross substitutes property more likely to hold. Consider next the right-hand side of the inequality, which captures the equilibrium location choice response of workers. If the elasticity of indirect utility with respect to housing prices is heterogeneous across locations, a proportional decrease in housing costs everywhere leads to net out-migrations in some locations. If this effect is strong enough, it can lead to a net decrease in housing demand in some locations, violating the gross substitutes property. Overall, uniqueness is all the more guaranteed as supply and demand for housing respond strongly to prices, or as the relative attractiveness of cities does not change too much when housing prices change in the same proportion everywhere.[7] In quantitative applications, the sufficient condition in equation (11) can be checked.

We assume from now on that model parameters are such that the equilibrium is unique, and turn to studying how heterogeneous skill types sort into different locations.

[7] We show in Appendix A.1 that Condition (11) is akin to a single-crossing condition in the space of housing prices in the two-city case.

3 Sorting in Spatial Equilibrium

In this section, we first give conditions under which the equilibrium exhibits sorting – i.e., location choices are heterogeneous across skills (Section 3.1). We then characterize the sorting equilibrium. We establish in particular that spatial sorting is shaped by the distribution of equilibrium housing prices. We characterize these equilibrium prices in Section 3.2, before establishing in Section 3.3 which skills sort into which type of cities.

3.1 Conditions for Spatial Sorting

An equilibrium with sorting is defined as follows:

Definition 1. Sorting. Sorting occurs when equilibrium location choices differ by skill type:

$$\exists n \in \{1, ..., N\}, \quad \exists s \neq t \in \mathcal{T}, \qquad \text{such that } \lambda_n^t \neq \lambda_n^s.$$

In contrast, a no-sorting equilibrium is characterized by:

$$\forall m, n \in \mathcal{C}, \quad \forall s, t \in \mathcal{T}: \qquad \left(\frac{\lambda_n^t}{\lambda_n^s}\right)\left(\frac{\lambda_m^t}{\lambda_m^s}\right)^{-1} = 1. \tag{12}$$

We lay out and discuss two conditions, which jointly will rule out sorting.

Condition 1. The productivity advantage of cities is skill-neutral. That is, for all $n \in \{1, ..., N\}$ and all $t \in \mathcal{T}$, $\frac{\partial^2 \ln \ell_n^t}{\partial T_n \partial t} = 0$.

Note that in this case, we can re-normalize t and T without loss of generality and write $w_n^t = T_n t$. This condition shuts down the possibility of a given skill enjoying comparative advantage in production in some location.

Condition 2. The indirect utility V is homogeneous in income. That is, there exists a function P, common across types, and a real number $\kappa > 0$ such that $V_n^t = \left[\frac{w_n^t}{P(r_n, A_n)}\right]^\kappa$ for all $n \in \mathcal{C}$ and all $t \in \mathcal{T}$.[8]

This condition crucially requires that the common component of preferences be *homothetic* as all types face the same price index $P(r_n, A_n)$. However, the condition is more stringent, as it is in effect a joint condition on the *representation* chosen for these homothetic preferences, $\tilde{U}$, and on the distribution of idiosyncratic shocks $\{\varepsilon_n\}$. Specifically, Condition 2 can fail when the common component of preferences $\tilde{U}$ is homothetic. Assume

[8] If $V(w, r, A)$ is homogeneous in income, then $\forall x > 0$, $V(xw, r, A) = x^\kappa V(w, r, A)$ and, in particular, $V(w, r, A) = w^\kappa V(1, r, A)$. Denoting $P(r_n, A_n) = V(1, r, A)^{-\frac{1}{\kappa}}$, we get $V_n^t = \left[\frac{w_n^t}{P(r_n, A_n)}\right]^\kappa$.

for instance that utility is $\tilde{U}(c,h;A_n)\,\varepsilon_n^{\omega}$ where $\tilde{U} = 1 + c^{\alpha}h^{1-\alpha}$ and shocks are distributed Frechet, with shape parameter κ. $\tilde{U}$ represents homothetic preferences, but condition 2 fails, as $\left(1 + \frac{w_n^t}{r_n^{1-\alpha}}\right)^{\kappa}$ is not homogeneous in income. What matters ultimately is not the separate properties of $\tilde{U}$ and of the distribution of $\{\varepsilon_n\}$, but only the properties of the resulting probability choice λ. Specifically, the property of location choices that matter for sorting is that, given Condition 2, location choices are given by:

$$\forall m,n \in \mathcal{C}, \quad \forall s,t \in \mathcal{T}: \qquad \lambda_n^t = \frac{\left[\frac{w_n^t}{P(r_n,A_n)}\right]^{\kappa}}{\sum_{m\in\mathcal{C}}\left[\frac{w_m^t}{P(r_m,A_m)}\right]^{\kappa}}, \tag{13}$$

so that location choice are homogeneous of degree zero in wages.

We are now ready to establish necessary conditions under which spatial equilibria exhibit sorting, generically (that is, for arbitrary distributions of city primitives):

Proposition 3. ***Sorting.***

(i) (Necessary condition) If there is sorting, then at most one of Conditions 1 and 2 holds.

(ii) (Sufficient condition) If exactly one of the two Conditions 1 or 2 holds, then sorting obtains.

Proof. See Appendix A.3. □

A few comments are in order. First, Proposition 3 characterizes a necessary condition for sorting, as well as a more stringent sufficient condition. To see why the necessary condition is not sufficient, consider the following case, which shows that even if neither Condition 1 nor Condition 2 hold, the equilibrium can still have no sorting. Assume that $\ell_n^t = w_n^t = t + T_n$, so the productivity advantage of cities is unskilled-biased ($\frac{\partial^2 \ln \ell_n^t}{\partial T_n \partial t} < 0$) and Condition 1 fails. In addition, assume that utility is $\exp(c)\,\varepsilon_n^{\omega}$ with unit housing consumption, so that $c = w - r$ (which generates housing housing expenditure shares $s_n^t = \frac{r_n}{w_n^t}$ that are decreasing with income). Assume that ε_n^{ω} are distributed Frechet with shape parameter κ, so that $V_n^t = \left[\frac{\exp(w_n^t)}{\exp(r_n)}\right]^{\kappa}$. Under these assumptions, Condition 2 fails. However, it is easy to see that $\lambda_n^t = \frac{\exp\kappa(T_n - r_n)}{\sum_{m\in\mathcal{C}}\exp\kappa(T_m - r_m)}$ does not depend on t, hence the spatial equilibrium features no sorting. In this example, the comparative advantage in production that the unskilled enjoy in high-productivity cities is exactly offset by the differential burden that housing prices places on them, so that location choices are ultimately independent from skill. In knife-edge cases where neither Conditions 1 nor condition 2 hold, but they exactly offset each other, there can still be no sorting.

Second, the representations commonly used in quantitative spatial equilibrium models feature a single type of agents, with Cobb-Douglas preferences between traded good and housing, multiplicative amenities and Frechet idiosyncratic shocks (see Redding and Rossi-Hansberg, 2017). Once extended to several agent types, these baseline assumptions allow for sorting across locations to be driven by comparative advantage only (i.e., heterogeneous nominal wage premium of groups across locations). To allow for sources of sorting other than these nominal wage effects, the most popular approach in the literature has been to keep this general Cobb-Douglas multiplicative structure within groups, but allow the parameters that govern preferences to vary by group. The alternative we explore in this paper is to maintain the assumption of preferences that are common across all groups, but to allow for non-homotheticity in these preferences. Doing so offers a microfounded model where groups with different incomes make systematically different choices.

Before studying the consequences of our alternative approach, we discuss here how various papers in the literature have generated the failure of Condition 2 to model sorting. In what follows, we assume for simplicity that Condition 1 holds. First, some papers assume that different groups have different common components of utility. In order to illustrate the consequences of this assumption in a simple way, assume that preferences are Cobb-Douglas, but with different housing shares $\mu^t \neq \mu^s$ for some types $s \neq t$. In this case, Condition 2 is violated, and households types that spend a higher share of income on housing are disproportionately drawn into cheaper cities:

$$\forall m, n \in \mathcal{C} \quad \forall s, t \in \mathcal{T}: \qquad \left(\frac{\lambda_n^t}{\lambda_n^s}\right)\left(\frac{\lambda_m^t}{\lambda_m^s}\right)^{-1} = \left(\frac{r_n}{r_m}\right)^{-\kappa\left(\mu^t-\mu^s\right)}.$$

It is also often assumed that different types have ex-ante heterogeneous preferences for amenities (as in Almagro and Domínguez-Iino, 2020; Diamond, 2016; Fajgelbaum and Gaubert, 2020). To illustrate this channel, assume that $V_n^t = A_n^t \frac{w_n^t}{r_n^\mu}$, where the notation A_n^t highlights that the valuation of amenities differ across groups. Then, households types are disproportionately drawn into cities with amenities that they disproportionately care about:

$$\forall m, n \in \mathcal{C} \quad \forall s, t \in \mathcal{T}: \qquad \left(\frac{\lambda_n^t}{\lambda_n^s}\right)\left(\frac{\lambda_m^t}{\lambda_m^s}\right)^{-1} = \left(\frac{A_n^t}{A_n^s}\right)\left(\frac{A_m^t}{A_m^s}\right)^{-1}.$$

Alternatively, some authors assume that the common component of utility is homothetic and identical for all skills, but sorting is obtained through the distribution of shocks. For instance, if shocks are distributed Frechet, but the distribution of shocks differ by skill with a Frechet parameter $\kappa^t \neq \kappa^s$ for some types $s \neq t$ (as in e.g. Tsivanidis, 2019), then Condition

2 is violated. It then follows that:

$$\forall m,n \in \mathcal{C} \quad \forall s,t \in \mathcal{T}: \qquad \left(\frac{\lambda_n^t}{\lambda_n^s}\right)\left(\frac{\lambda_m^t}{\lambda_m^s}\right)^{-1} = \left(\frac{\frac{T_n}{P(r_n,A_n)}}{\frac{T_m}{P(r_m,A_m)}}\right)^{\kappa^t-\kappa^s},$$

and high-κ types disproportionately sort into cities displaying high productivity, desirable amenities, and cheap housing, as their sensitivity to these differences in local attractiveness is higher. Relatedly, assume that the common component of utility is a homothetic function, but consider additive logit shocks as in equation 2. In this case, Condition 2 also fails. Specifically, so long as Condition 1 holds, relative choice probabilities are given by:

$$\left(\frac{\lambda_n^t}{\lambda_n^s}\right)\left(\frac{\lambda_m^t}{\lambda_m^s}\right)^{-1} = \exp\left[\left(\frac{T_n}{P\left(r_n,A_n\right)} - \frac{T_m}{P\left(r_m,A_m\right)}\right)(t-s)\right],$$

which generically differs from unity. The reason why this seemingly innocuous change of functional forms has a direct impact on sorting is that, as discussed above, what matters for the properties of sorting is the combined feature of the representation of preferences and the distribution of shocks. With additive logit shocks, the migration elasticity $\mathcal{E}_n^t\left(\frac{\lambda}{1-\lambda},\tilde{V}\right)$ is city- and type-specific and equal to V_n^t, and thus is increasing in type t, given n. Through this mechanism, more mobile types (that is, higher utility/higher-t type in this case) are disproportionately represented in more attractive cities.

Our analysis on the role of identical but non-homothetic preferences in generating sorting complements this literature. A common definition of preferences is particularly desirable in that it allows for transparent welfare comparison across individuals, as well as for welfare accounting in the context of changing groups composition – e.g., the context of the stark rise in college education in developed economies.

We now move on characterizing the sorting equilibrium, i.e. which skills sort to which type of cities, in our model with income-inelastic preferences for housing.

3.2 Equilibrium Housing Prices

We first make a detour through housing rents and establish that, under quite general conditions, desirable and productive cities command relatively high housing prices in spatial equilibrium. However, we will see that this result may not hold when the quality of some local amenities directly reduce housing consumption. The following assumption rules out this case:

Assumption 1. *Housing and amenities are q-complements in the sense that*

$$\forall n \in \mathcal{C}, \quad \forall t \in \mathcal{T}: \qquad \frac{\partial h_n^t}{\partial A_n} \geq 0.$$

Note that Assumption 1 is warranted in particular in the case that is ubiquitous in the literature where $U(c, h, A_n) = U(u(c, h), A_n)$ for some function u, i.e., when consumption and amenities are separable. In this case, $\partial h_n^t / \partial A_n = 0$.[9]

The following proposition summarizes how equilibrium housing rents reflect city characteristics:

Proposition 4. *"**Rosen-Roback**." Consider two cities $m, n \in \mathcal{C}$. If $T_n > T_m$, $H_n < H_m$ or $\eta_n < \eta_m$, all else equal, then city n is more expensive than city m in equilibrium ($r_n > r_m$). In addition, under Assumption 1, if $A_n > A_m$ all else equal, then city n is more expensive than city m in equilibrium ($r_n > r_m$).*

Proof. See Appendix A.3. □

Note that the effect of a higher amenity A_n on the equilibrium housing price r_n is ambiguous if Assumption 1 is violated. Indeed, when housing and amenities are q-substitutes, higher amenities reduce housing demand of locals. This force pushes housing prices down. At the same time, higher amenities attract more households to the city, pushing prices up, so that the net effect of amenities on housing prices is ambiguous. However, Assumption 1 is strong for our purposes: Proposition 4 holds so long as local *aggregate* demand for housing in the city, $\sum_t L^t \lambda_n^t h_n^t$, is non-decreasing in A_n, a condition that may hold even if Assumption 1 (that pertains to individual housing demand h_n^t) is violated.

3.3 City Primitives and Sorting

We now characterize spatial sorting, that is, we establish the effects of cities primitives $\{A_n, T_n, H_n, \eta_n\}$ on the *composition* of their labor force. We restrict the analysis to a demand system separable in amenities and consumption, in line with the literature.[10] Specifically,

[9] The case $\frac{\partial h_n^t}{\partial A_n} > 0$ corresponds to a situation in which natural amenities are a complement to housing services – e.g., a soft climate that makes a large garden more enjoyable, or nice views that make a spacious living room with large windows more desirable. The case $\frac{\partial h_n^t}{\partial A_n} < 0$ corresponds to cultural or natural amenities that make hanging out outside more attractive than staying home watching Netflix, or if the variety and quality of local restaurants makes you use your kitchen less frequently.

[10] Arntz, Brull, and Lipowski (2023) use a stated-preference experiment to elicit preferences for urban amenities. They do not find any significant differences among skill groups. We are unaware of other empirical results measuring substitution patterns between local amenities and consumption. We therefore make the assumption, pervasive in the literature, that preferences are separable in these two terms. Relaxing this assumption would lead to a typology of cases that are ancillary to the main point of the paper.

we assume from now on that:

Assumption 2. *Preferences are separable in amenities and consumption. Location choices are given by:*

$$\lambda_n^t = \frac{\left[v\left(w_n^t, r_n\right) A_n\right]^\kappa}{\sum_{m \in \mathcal{C}} \left[v\left(w_m^t, r_m\right) A_m\right]^\kappa}. \tag{14}$$

In particular, Assumption 1 holds under Assumption 2, since $\frac{\partial h_n^t}{\partial A_n} = 0$. This formulation allows for two distinct drivers of sorting: comparative advantage (Condition 1 may fail), and income effects (Condition 2 may fail). The assumption that preferences are separable in amenities and real wages v conveniently emphasizes the role of non-homothetic preferences (embedded in v) on sorting: heterogeneous amenities do drive sorting in equilibrium, but only indirectly – through their capitalization in housing prices.

As is intuitive, the direction of sorting depends on which worker type gets a higher benefit from city productivity and/or is more burdened by local cost of living. We detail here assumptions on primitives that drive these two forces, before characterizing sorting under these assumptions. First, we assume that high-skill workers have a comparative advantage in cities with high productivity, as it is the more empirically relevant case. :

$$\frac{\partial \mathcal{E}_n^t\left(\ell, T\right)}{\partial t} > 0. \tag{15}$$

We will see below that a stronger form of comparative advantage may sometimes be required. We say that high-skill workers have a *strong comparative advantage* in cities with high productivity when:

$$\frac{\partial \mathcal{E}_n^t\left(v, T\right)}{\partial t} > 0. \tag{16}$$

Strong comparative advantage means a high-skill worker is comparatively more productive than a low-skill in a high productivity city, not just in income terms, as in expression (15), but also in utility terms, as in expression (16).[11] Second, income-based sorting is driven by how concave is utility. This is because concavity in utility makes differences in real income across locations more salient for low-skill workers than for high-skill workers. To fix ideas, we derive results under the following assumption:

Assumption 3. *The elasticity of indirect utility with respect to income is non-increasing in*

[11] The cardinality of utility matters in discrete choice models of the sort we adopt here since the magnitude of utility differences affects choice probabilities (McFadden, 1974).

income and therefore in worker type by $\partial w_n^t/\partial t > 0$:

$$\frac{\partial^2 \left(\ln v_n^t\right)}{\partial \left(\ln w_n^t\right)^2} \leq 0 \quad \Rightarrow \quad \frac{\partial \mathcal{E}_n^t \left(v, w\right)}{\partial t} \leq 0. \tag{17}$$

Note that under Assumption 3, strong comparative advantage (16) implies comparative advantage (15). We show in Appendix B that all the functional forms for utility listed in Table 2 satisfy Assumption 3.[12] We are now ready to characterize sorting patterns in this economy.

Proposition 5. ***Sorting Patterns****. Consider two cities $m, n \in \mathcal{C}$, and two worker types $s, t \in \mathcal{T}$ with $s < t$. In equilibrium, city n disproportionately attracts high-t workers relative to city m ($\frac{\lambda_n^t}{\lambda_n^s} > \frac{\lambda_m^t}{\lambda_m^s}$):*

(i) If $H_n < H_m$ or $\eta_n < \eta_m$, all else equal.

(ii) Under Assumption 1, if $A_n > A_m$, all else equal.

(ii) Under Assumption 3, if $T_n > T_m$ and if high-ability workers have a strong comparative advantage in productive cities, all else equal.

Proof. See Appendix A.3. □

Two comments are in order. The first part of the proof shows that any city characteristic that increases local housing prices, all else equal, increases the relative supply of high-skill workers there. In that sense, *high-skill workers sort into expensive cities*. In particular, this mechanism leads high-skill workers to sort into productive cities (which pay high wages and tend to be more expensive in equilibrium). In addition, their comparative advantage in production also leads high-skill workers to sort to high-productivity places. Though seemingly intuitive, this result is more subtle, as the decreasing marginal utility of income dampens the effect of comparative advantage on the relative attractiveness of cities. As the proposition shows, this type of comparative advantage has to be stronger than the effect of decreasing marginal utility of income to generate this sorting. Condition (16) provides a sufficient condition for the comparative advantage effect to dominate decreasing marginal utility of income, and drive sorting of high-skill workers to high-productivity locations. With this correction, non-homothetic preferences and comparative advantage do reinforce each other in shaping the sorting of high skill into high productivity cities. The role played by

[12] Note also that $\frac{\partial^2 v_n^t}{\partial (w_n^t)^2} \leq 0$ and $\mathcal{E}_n^t\left(v, w\right) \geq 1$ together imply $\frac{\partial \mathcal{E}_n^t (v,w)}{\partial t} \leq 0$. That is, the log-log marginal utility of income $\mathcal{E}_n^t\left(v, w\right)$ is decreasing in income if (i) the marginal utility of income $\frac{\partial v_n^t}{\partial w_n^t}$ is decreasing in income and (ii) if $\mathcal{E}_n^t\left(v, w\right) \geq 1$ holds, as is the case for the functional forms that we consider in Appendix B. This result follows from $\frac{\partial \mathcal{E}(v,w)}{\partial w} = \frac{\partial^2 v}{\partial^2 w} + \frac{\partial v}{\partial w}\left[1 - \mathcal{E}\left(v, w\right)\right]$.

the concavity of preferences does not arise in quantitative models that use preferences that are unit-elastic in income (typically Cobb-Douglas).

4 Welfare Inequality

We have focused so far on positive questions, and have discussed in particular how city characteristics shape the spatial sorting of heterogeneous types of workers. Another class of questions urban economists are interested in is normative in nature. What are the distributional consequences of spatial shocks? To make progress on these issues, we propose results on the measurement of welfare and welfare inequality between types, in our spatial model with heterogeneous agents and non-homothetic preferences.

4.1 Welfare Changes: Sufficient Statistics Approach

Models with homogeneous agents typically rely on a "sufficient statistics" approach for welfare, which allows to compute representative welfare changes following an economic shock, given available statistics. We extend this approach to our model with heterogeneous agents and non-homothetic preferences, to measure the distributional consequences, across heterogeneous groups, of given economic shocks. We assume that indirect utility takes the form in (1), maintaining Assumption 2. That is, the indirect utility of a worker ω of type t in location m is $v\left(w_m^t, r_m\right) A_m \epsilon_m^t(\omega)$, where $\epsilon_m^t(\omega)$ are distributed i.i.d. Frechet with shape parameter $\kappa > 1$. In particular, expected utility of type t is equal to $\Gamma\left(1-\frac{1}{\kappa}\right)$ times $\mathbb{V}^t$, where Γ is the gamma function and:

$$\mathbb{V}^t = \left[\sum_{m \in \mathcal{C}} \left[v\left(w_m^t, r_m\right) A_m\right]^{\kappa}\right]^{\frac{1}{\kappa}}. \tag{18}$$

Compared to the literature, we relax the assumption that $v(.)$ is Cobb-Douglas, and consider instead functions that exhibit income effects. We use the hat notation to compute relative changes in variables, $\hat{x} \equiv x'/x$, where x' denotes the value of x after an economic shock or a policy change.

We first look into how expected utility changes for different types following a shock. Given the properties of the Frechet distribution, expected utility of type t changes as follows, for any location n:

$$\hat{\mathbb{V}}^t = \hat{A}_n \frac{\hat{w}_n^t}{\hat{P}_n^t} \left(\hat{\lambda}_n^t\right)^{-\frac{1}{\kappa}}, \tag{19}$$

where $P_n^t \equiv w_n^t / v_n^t = P\left(w_n^t, r_n; \Theta\right)$ is a price index for type t in city n. It depends on wages and prices (which can a priori be observed) and on the vector Θ that summarizes the parameters of the demand function (which can a priori be estimated). This price index is generically both city-specific, because housing prices differ across space, and type-specific, because of non-homothetic preferences.

In turn, change in expected utility inequality between types s and t is given by:

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \frac{\hat{w}_n^t}{\hat{w}_n^s} \left(\frac{\hat{P}_n^t}{\hat{P}_n^s} \right)^{-1} \left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s} \right)^{-\frac{1}{\kappa}}, \text{for any } n \in \mathcal{C}. \tag{20}$$

Expression (20) makes clear that simple measures of changes in relative real income in a given city, such as $\frac{\hat{w}_n^t}{\hat{w}_n^s} \left(\frac{\hat{P}_n^t}{\hat{P}_n^s} \right)^{-1}$, do not deliver a theoretically consistent measure of change in expected utility between types. This is because, in spatial equilibrium, the economy responds to shocks in two ways: through a change in prices, captured by this real income measure, but also through a reallocation of population, moving out of (relatively) adversely affected locations. In addition to real income changes, one must account for a selection effect in how population responds to shock: those newly attracted to a location following a positive shock have typically lower idiosyncratic preference for that location than incumbents. Therefore, larger inflows signal larger selection effects, which lower average utility. This is controlled for through the term $\left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s} \right)^{-\frac{1}{\kappa}}$, to obtain a theoretically grounded measure of the change in utility inequality.

Note that if the first and last terms in the right hand side of equation (20), $\frac{\hat{w}_n^t}{\hat{w}_n^s}$ and $\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s}$, are readily observable empirically, changes in type-specific price indices $\frac{\hat{P}_n^t}{\hat{P}_n^s}$ are model-based constructs and are not directly observable. In order to make progress, then, one has to impose a functional form for utility and hence for price indices, as in Handbury (2021). We consider several popular classes of non-homothetic preferences, and we propose below a new convenient result for distributional analysis in these classes of models: in each case, the change in expected utility inequality in equation (20) depends on a small set of simple statistics, including – crucially – changes in the expenditure shares on housing, s_n^t. These housing expenditure shares are key statistics that measure the distributional effect of a shock, as they are informative about relative welfare changes when preferences are non-homothetic. The key insight here is that these statistics exploit the mapping between Engel curves and welfare, as in Aguiar and Bils (2015) and Atkin, Faber, Fally, and Gonzalez-Navarro (2023).

Specifically, we show that in all functional forms we consider, the type-t-specific price index can be written as

$$P_n^t = P_n S_n^t, \tag{21}$$

the product of a component common to all incomes ($P_n \equiv P(p, r_n)$, where P is homogeneous of degree one in p and r_n) and a correction for income effects (S_n^t) which only depends on the expenditure share on housing.[13] Consider the following Stone-Geary preferences as an example:

$$v\left(w_n^t, r_n\right) = \frac{w_n^t - \underline{h} r_n}{p^{1-\psi}\left(r_n\right)^{\psi}}, \tag{22}$$

where $\underline{h}$ can be interpreted as the subsistence level of housing if it is positive (so that housing is a necessity) and $\psi \in (0,1)$. This expression encompasses Cobb-Douglas as a special case when $\underline{h} = 0$. This attractive feature and its simplicity make this functional form a popular one – see e.g. Tsivanidis (2019) and Gaigné, Koster, Moizeau, and Thisse (2022) for applications in spatial contexts. Using Roy's identity, the expenditure share of housing is $s_n^t = \psi + (1-\psi)\frac{\underline{h} r_n}{w_n^t}$, so that in this case, we get:

$$v\left(w_n^t, r_n\right) = \frac{w_n^t}{P_n S_n^t}, \text{with } P_n = p^{1-\psi}\left(r_n\right)^{\psi} \text{ and } S_n^t = \frac{1-\psi}{1-s_n^t}.$$

[14] Therefore, relative changes in indirect utility V only depend on prices via the housing expenditure share term S_n^t. Specifically, the relative change in expected utility between types s and t is :

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \frac{\widehat{w_n^t}}{\widehat{w_n^s}} \frac{\widehat{1-s_n^t}}{\widehat{1-s_n^s}} \left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s}\right)^{-\frac{1}{\kappa}}, \tag{23}$$

where n is any arbitrary city. That is, changes in nominal incomes, housing expenditure shares and skill composition of any location n, as well as an estimate of the migration elasticity κ, are sufficient to compute changes in expected utility inequality between any pair of types. Thus, any change in expected utility inequality can be decomposed into changes in nominal earnings (which are equal to changes in real earning inequality if preferences are homothetic), a correction for location changes, and a correction for non-homothetic preferences, as captured by heterogeneous changes in expenditure shares.

This logic generalizes to the six alternative functional forms for utility that are listed in Table 1. To establish this result, we define the composite demand parameters $\alpha, \beta, \delta, a, b > 0$ that combine structural parameters of each demand system, as described in Columns 3 to 7

[13] This decomposition is not unique; we normalize the homothetic component consistently across equilibria to ensure comparability.

[14] By the same token, in the case of non-homothetic CES preferences that we also use in the application of Section 5 below, we show in Appendix section B.5 that P_n and S_n^t obey

$$P_n = p \text{ and } S_n^t = \left(1 - s_n^t\right)^{-1/(1-\sigma)}.$$

We also provide two alternative decompositions in this appendix section.

of Table 2. These composite parameters $(a, b, \alpha, \beta, \delta)$ are the elasticities needed to implement welfare computations.[15]

Table 2: Non-homothetic preferences: Definition of composite parameters.

Preferences	Indirect utility V	Composite parameters				
		α	β	δ	a	b
Stone-Geary	$V = A\frac{w-\nu \underline{h} r}{p^{1-\psi}r^{\psi}}$	1	0	1	0	1
PIGL	$V = A\left[\left(\frac{w}{p^{1-\psi}r^{\psi}}\right)^{\iota} - \frac{\nu}{\gamma}\left(\frac{r}{p}\right)^{\gamma\iota}\right]^{1/\iota}$	1	0	$\frac{1}{\iota}$	ψ	$\frac{1}{\gamma}$
PIGLOG	$V = A\left(\frac{r}{p}\right)^{\nu}\ln\left(\frac{w}{p^{1-\psi}r^{\psi}}\right)$	0	0	1	0	$\frac{1}{\psi}$
CRIE	$V = A\left[\left(\frac{w}{r}\right)^{-\psi} + \left(\frac{w}{p}\right)^{-\psi+\nu(\psi-\mu)}\right]^{-1}$	1	$\frac{1}{\mu}$	$\frac{1}{\mu}$	$\frac{\psi}{\psi-\mu}$	$-\frac{\psi-\mu}{\mu}$
NH-CES	$w^{1-\sigma} = \left(\frac{V}{A}\right)^{1-\sigma}\left[\left(\frac{V}{A}\right)^{\nu(\mu-\psi)} r^{1-\sigma} + p^{1-\sigma}\right]$	1	0	$\frac{1}{1-\sigma}$	0	1
HCD	$0 = \left(\frac{V}{A}\right)^{\psi+\nu(\mu-\psi)}\ln\left(\frac{r}{w}\right) + \left(\frac{V}{A}\right)^{\psi}\ln\left(\frac{p}{w}\right)$	0	$\frac{1}{\psi-\mu}$	$\frac{1}{\psi-\mu}$	0	1

Notes: PIGL stands for "Price Independent Generalized Linearity" (Muellbauer, 1975), PIGLOG stands for "Price Invariant Generalized Logarithmic" (Deaton and Muellbauer, 1980), CRIE stands for "Constant Relative Income Elasticity" (Caron, Fally, and Markusen, 2014; Eeckhout, Pinheiro, and Schmidheiny, 2014; Fieler, 2011; Hanoch, 1975), NH-CES stands for "Non Homothetic Constant Elasticity of Substitution" (Albouy, Ehrlich, and Liu, 2016; Comin, Lashkari, and Mestieri, 2021; Hanoch, 1975; Matsuyama, 2019; Sato, 1977), and HCD stands for "Heterothetic Cobb-Douglas" (Bohr, Mestieri, and Robert-Nicoud, 2024). HCD preferences are a limiting case of NH-CES preferences. The indicator variable $\nu \in \{0, 1\}$ governs the (non-)homotheticity of each class of preferences. In all cases, preferences are homothetic if $\nu = 0$ and non-homothetic if $\nu = 1$.

Second, we define S_n^t, a monotonically increasing transformation of the housing expenditure share s_n^t:

$$S_n^t \equiv \frac{\left(s_n^t - a\right)^{\beta}}{\left[1 - b\left(s_n^t - a\right)\right]^{\delta}}. \tag{24}$$

With these notations, the following result obtains:

Proposition 6. *Relative Welfare Changes.* *Under all of the functional forms for V considered in Table 2, the relative change in expected utility between any pair of types can simply be expressed as:*

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \left(\frac{\widehat{w_n^t}}{\widehat{w_n^s}}\right)^{\alpha}\left(\frac{\widehat{S_n^t}}{\widehat{S_n^s}}\right)^{-1}\left(\frac{\widehat{\lambda_n^t}}{\widehat{\lambda_n^s}}\right)^{-\frac{1}{\kappa}}, \textit{for any } n \in \mathcal{C}. \tag{25}$$

[15] Appendix B provides the details.

Proof. Appendix B derives this result for each class of non-homothetic preferences. □

Proposition 6 establishes that changes in nominal incomes and populations by skill in any location n, pre- and post-housing expenditure shares, and relative migration changes, together with knowledge of some structural parameters of the model, are sufficient to compute changes in expected utility inequality between any pair of types. Note that the formula does not necessarily requires estimates of all structural parameters of the utility function, in line with the "hat algebra" approach more generally. Consider the NH-CES case as an example. The formula requires an estimate for the elasticity of substitution σ between housing and other goods, but not of the parameters governing the income elasticity of demand μ, ψ. Similarly, the approach does not require estimating any of the structural parameters of the Stone-Geary utility function. In either case, the effect of the "superfluous" parameters on equilibrium outcomes are subsumed in the (potentially observable) expenditure shares.

Finally, we note that one can augment the model to allow for type-specific local amenities, as in Diamond (2016) and Fajgelbaum and Gaubert (2020). In this case maintaining the assumption of multiplicatively separable amenities, equation (25) becomes

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \frac{\widehat{A_n^t}}{\widehat{A_n^s}} \left(\frac{\widehat{w_n^t}}{\widehat{w_n^s}}\right)^{\alpha} \left(\frac{\widehat{S_n^t}}{\widehat{S_n^s}}\right)^{-1} \left(\frac{\widehat{\lambda_n^t}}{\widehat{\lambda_n^s}}\right)^{-\frac{1}{\kappa}}. \tag{26}$$

We return to that case in the application of Section 5 below.

4.2 Money-Metric Welfare Changes

Proposition 6 takes a first step to extend the "exact hat algebra" approach to welfare analysis, common in trade, economic geography, and other fields, to a setting with general non-homothetic preferences. In the context of non-homothetic preferences, however, changes in utility are not money-metric changes. An additional step is thus required to convert changes in utility into interpretable changes. We now show how to map these changes in utils into an equivalent variation-type measure. As will become clear, $\mathcal{E}_n^t(v, w)$ is an important variable for our analysis, and we henceforth refer to this object as the "log-log marginal utility of income." (the marginal utility of income, expressed in elasticity terms).

Specifically, we interpret the expected utility function (18) as the utility function of a representative type-t agent. We measure the impact of a shock on welfare of type t by asking what should be her proportional increase in income in the period prior to the shock to reach the well-being of the period posterior to the shock? This offers a natural unit-less measure of money-metric welfare change between the two periods, which we write $\hat{y}^t$. It is an equivalent variation concept, expressed in percentage terms. Formally, $\hat{y}^t$ is implicitly defined as:

$$\left[\sum_{m\in\mathcal{C}} v\left(w_m^t \hat{y}^t, r_m\right)^\kappa (A_m)^\kappa\right]^{\frac{1}{\kappa}} = \left[\sum_{m\in\mathcal{C}} v\left(w_m^{\prime t}, r_m^{\prime}\right)^\kappa (A_m^{\prime})^\kappa\right]^{\frac{1}{\kappa}}. \tag{27}$$

We consider small shocks to the economy. Using a first order approximation, we establish the following result.[16]

Proposition 7. ***Equivalent Variation****.The equivalent variation (in proportional terms) is related to the relative change in expected utility as follows:*

$$\hat{y}^t = \left(\hat{\mathbb{V}}^t\right)^{1/\bar{\varepsilon}^t}, \tag{28}$$

where $\bar{\varepsilon}^t \equiv \sum_n \lambda_n^t \mathcal{E}_n^t(v, w)$ *is the average log-log marginal utility of income across all cities.*

Proof. See Appendix B.7. □

The change in income that is equivalent to the change in the economic environment caused by a given shock is given by the corresponding proportional change in utility, converted in money-metric unit (converting 1% util increase into 1% dollar increase) by the term $\bar{\varepsilon}^t$. When utility is homothetic and unit-elastic in income, it is clear that $\bar{\varepsilon}^t = 1$, so that $\hat{y}^t = \hat{\mathbb{V}}^t$. Hence, expected utility can be used directly in this case to measure welfare change in an appropriate metric. Here, in contrast, the log-log marginal utility of income is not constant, and varies across individuals and choices. McFadden (1999) and Karlstrom (2000) discuss how to compute equivalent variation in discrete choice models with income effects, which is generally intractable and computationally intensive. The Frechet formulation, combined with a generic non-homothetic common utility, strikes a balance between allowing for realistic income effects and preserving some tractability for welfare computations.

We use $\frac{\hat{y}^t}{\hat{y}^s}$ as a money metric measure of change in welfare inequality. It measures the difference in the proportional income growth equivalent to a given shock for type t versus type s. Using equation (28) and considering two arbitrary types t, s, this change in welfare inequality following a small shock can be expressed as follows:

$$\frac{\hat{y}^t}{\hat{y}^s} = \frac{\left(\hat{\mathbb{V}}^t\right)^{1/\bar{\varepsilon}^t}}{\left(\hat{\mathbb{V}}^s\right)^{1/\bar{\varepsilon}^s}}. \tag{29}$$

[16] We note that Proposition 7 does not require Assumption 2 (that is, V does not have to be multiplicatively separable in v and A). The equivalent variation term captures both change in prices and changes in amenities.

Expression (29) tells us that the measure of change in expected utility inequality using sufficient statistics, $\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s}$, in equation (25) has to be corrected in two ways to deliver a meaningful measure of change in inequality between types. First, it has to be corrected by the average log-log marginal utility of income, $\bar{\varepsilon}^t$ to convert a change in utils into a change in income. The second correction accounts for the fact that the average log-log marginal utility of income differs across types. That is, using only expected utility inequality $\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s}$ generally leads to a biased measure of change in welfare because $\bar{\varepsilon}^t, \bar{\varepsilon}^s \neq 1$, and because $\bar{\varepsilon}^t \neq \bar{\varepsilon}^s$. To conclude, we highlight that taking into account these corrections for different marginal utility of incomes leads to less simple expressions that our initial welfare inequality measure (25). To see this, develop the change in inequality from (28) fully, using equations (19) and (21). The change in inequality is:

$$\frac{\hat{y}^t}{\hat{y}^s} = \frac{\left[\left(\widehat{w_n^t}\right)^{\alpha}\left(\widehat{S_n^t}\right)^{-1}\left(\widehat{\lambda_n^t}\right)^{-\frac{1}{\kappa}}\right]^{1/\bar{\varepsilon}^t}}{\left[\left(\widehat{w_n^s}\right)^{\alpha}\left(\widehat{S_n^s}\right)^{-1}\left(\widehat{\lambda_n^s}\right)^{-\frac{1}{\kappa}}\right]^{1/\bar{\varepsilon}^s}}\left(\frac{\widehat{A_n}}{\widehat{P_n}}\right)^{\frac{1}{\bar{\varepsilon}^t}-\frac{1}{\bar{\varepsilon}^s}}. \tag{30}$$

The first term in the right-hand side corrects relative utility changes in equation (25) using the equivalent-variation correction. Since it is income-, hence type-specific, the correction is done separately in the numerator and the denominator. The second term corresponds to elements that are common to both types and enter multiplicatively their utility, hence disappeared from equation (25). In contrast, when expressing inequality in equivalent-variation term, the homothetic components of utility common to both types also enter the computation of inequality. They are valued differently by the two types, in money-metric terms. This second correction is moot if changes in local amenities are fully capitalized in changes in the homothetic component of local prices (and hence$\widehat{A_n} = \widehat{P_n}$). By contrast, if changes in local amenities are not fully capitalized in changes in the homothetic component of local prices, and if type-t corresponds to a higher-income type in that it has a lower marginal utility of income than type-s , then this correction increases the money-metric correction whenever amenities increase by more than the homothetic component of local prices (i.e., if $\widehat{A_n} > \widehat{P_n}$).

5 Measuring Changes In Welfare Inequality: USA 1980-2020

In this section, we evaluate the model-consistent money-metric change in welfare inequality between high- and low-skill workers in the U.S., in the last four decades, using the results from Section 4. In doing so we revisit results from a literature that has measured these changes taking into account the heterogeneous spatial sorting of high- and low-skill workers (Moretti, 2013; Diamond, 2016). We show that accounting for income effects in housing consumption changes these measures significantly.

We maintain Assumption 2 throughout this section, so that we can write $V_n^t = (v_n^t A_n)^\kappa$, where $v_n^t = v(w_n^t, r_n)$ and $\kappa > 1$ governs the elasticity of the spatial labor supply.

5.1 Data

To operationalize the formulas of Section 4, we need data on nominal wages, location choices of high- vs low-skill workers, and – crucially – expenditure on housing over time. We describe these here in turn and refer the reader to Appendix B for more details.

Raw data Our data collection and definitions follow Moretti (2013) as closely as possible. Data for our key variables come from the Decennial Census for 1980 to 2000 and the 5-year American Community Survey (ACS) sample for 2010 and 2020 (Ruggles et al., 2024). A location in the analysis is one of 303 Metropolitan Statistical Areas (MSAs) and "rest-of-state" regions. The sample includes US-born, full-time workers aged 25 to 59 who live in private households. We define individuals with 4+ years of higher education as "high skill" and individuals with a high school degree as "low skill."

Sufficient statistics First, from the dataset above, we directly read the share $\lambda_{n,y}^t$ of workers of type t living in location n in year y. Second, we use an adjusted measure of wages to control for demographic changes among worker types over time. To that end, we use the same controls as in Moretti (2013): sex, dummies for race, and a cubic in work experience. The adjustment is performed separately for 1980-2000 and 2010-2020 periods to maintain consistency with Moretti (2013), who analyzes the earlier period. For each period, we regress wages on these demographic controls and use the residuals, holding demographic characteristics constant at their period-specific means. We compute $\left\{w_{n,y}^t\right\}_{t,n,y}$ as the median adjusted wage for each city-year-skill. Housing expenditure shares $\left\{s_{n,y}^t\right\}_{t,n,y}$ are calculated as the ratio of annual gross rent to household income, where again we use the median within a city-year-skill group. Our measure shows that in 2020 the median share of income spent on housing was 20% for high-skill individuals and 24% for low-skill individuals. These shares vary across cities from 11% to 24% for high-skill households, and from 16% to 31% for low-skill ones. Moreover, from 1980 to 2020, housing expenditure shares rose relatively more for low-skill individuals, increasing by 5 percentage points against 2 percentage point increase for high-skill individuals. Much of the increase in spending shares occurred after 2000. Summary statistics for these housing shares are detailed in Appendix Table OA.1.

5.2 Model Parameters

In addition to this data, implementing the formulas of Section 4 requires estimates of the elasticity of population to wages κ and estimates of the average log-log marginal utility of income by group, $\left\{\bar{\varepsilon}^L, \bar{\varepsilon}^H\right\}$.

Non-homotheticity and marginal utility of income We apply equation (25) considering two alternative specifications of non-homothetic demand: Stone-Geary and Non-Homothetic CES. In each case, we review here how we calibrate the average log-log marginal utility of income for each group, $\left\{\bar{\varepsilon}^L, \bar{\varepsilon}^H\right\}$.

In the Stone-Geary case, recall from equation (22) that two parameters, $\psi \in (0,1)$ and $\underline{h} \geq 0$, characterize the indirect sub-utility function: $v_n^t = \frac{w_n^t - \underline{h} r_n}{r_n^{\psi}}$. We show in Appendix B.1 that the log-log marginal utility of income for type t in location n is given by:

$$\mathcal{E}_{n,SG}^t\left(v,w\right) = \frac{1-\psi}{1-s_n^t},$$

which is larger than one given that $s_n^t \geq \psi$ (here we add the subscript "SG" to emphasize that the computation of this elasticity is specific to the choice of functional form for preferences, and to avoid confusion with the computation of $\mathcal{E}_{n,NHCES}^t$ below). We calibrate parameter ψ noting that this parameter relates to the expenditure share on housing as follows: $s_n^t = \psi + (1-\psi)\frac{\underline{h} r_n}{w_n^t}$. In particular, ψ corresponds to the housing expenditure share for infinitely high incomes. We calibrate ψ to match the 10th percentile of the expenditure share distribution of high skills as the actual minimum of the housing expenditure share $\min_{n,t} s_n^t$ is very prone to measurement error. This yields $\psi = 0.1$. We then compute the distribution of $\left\{\mathcal{E}_{n,SG}^t\left(v,w\right)\right\}_{n,t}$ given that of $\left\{s_n^t\right\}_{n,t}$ using the median housing expenditure share for each combination of city and type and finally compute the mean $\bar{\varepsilon}_{SG}^t \equiv \sum_n \lambda_n^t \mathcal{E}_{n,SG}^t\left(v,w\right)$. We report the results in Table 3. We also report the median rather than the mean value to assess the stability of the results.

In the Non-Homothetic CES case, v_n^t is implicitly defined by

$$\left(w_n^t\right)^{1-\sigma} = \left(v_n^t\right)^{1-\sigma}\left[\left(v_n^t\right)^{\mu-\psi}\left(r_n\right)^{1-\sigma} + p^{1-\sigma}\right].$$

Housing demand is price-inelastic if $\sigma \in (0,1)$ and income-inelastic if $\sigma - 1 \leq \mu - \psi < 0$. Appendix B.5 provides details. The log-log marginal utility of income for type t in location n is given by:

$$\mathcal{E}_{n,NHCES}^t\left(v,w\right) = \frac{1-\sigma}{1-\sigma-\left(\psi-\mu\right)s_n^t},$$

Table 3: Log-log marginal utility of a dollar (1980)

	Mean		Median	
	Low Skill	High Skill	Low Skill	High Skill
Homothetic	1.000	1.000	1.000	1.000
Nominal	1.000	1.000	1.000	1.000
Stone-Geary	1.104	1.094	1.096	1.094
NH-CES	1.132	1.126	1.127	1.126

which is strictly larger than one if $\mu < \psi$. We set $\sigma = 0.53$ and $\psi - \mu = 0.306$ based on Finlay and Williams (2022). As for the Stone-Geary case, we then compute the distribution of $\left\{\mathcal{E}^t_{n,NHCES}(v,w)\right\}_{n,t}$ using the median housing expenditure share for each combination of city and type, and report $\bar{\varepsilon}^t_{NHCES} \equiv \sum_n \lambda^t_n \mathcal{E}^t_{n,NHCES}(v,w)$ in Table 3. We also report the median of $\mathcal{E}^t_{n,NHCES}(v,w)$ for robustness.

Two intermediate results are worth emphasizing. First, the computed values of the log-log marginal utility of income are very stable whether we use the median value of $\mathcal{E}^t_n$ instead of the weighted average (a comparison across column pairs). Second, these computed values are remarkably similar across types and across functional forms, around 1.1 for both high-skill and low-skill renters. By inspection of equation (30), an immediate, quantitative implication of this finding is that the elasticity of inequality to changes in the *levels* of local amenities (and the common component of price indices) is two orders of magnitude smaller than its responsiveness to changes in wages, expenditure shares, or migration patterns.[17] For this reason, we can neglect the last term in Equation (30) when computing changes in inequality between groups in terms of equivalent variation.

Migration elasticity Implementing the formulas of Section 4 also requires an estimate of the migration elasticity, κ. Taking logs of equation (14) and differentiating the equation between types (we denote high-skill households with H and low-skill households with L) and

[17]The change in money-metric inequality in Equation (30) responds with an elasticity greater than $\min\left\{\frac{1}{\bar{\varepsilon}^L}, \frac{1}{\bar{\varepsilon}^H}\right\}$ to the change in inequality in terms of $\left(\widehat{w^t_n}\right)^{\alpha}\left(\widehat{S^t_n}\right)^{-1}\left(\widehat{\lambda^t_n}\right)^{-\frac{1}{\kappa}}$ (first term in (30)), whereas it responds to changes in amenities with elasticity $\frac{\bar{\varepsilon}^L - \bar{\varepsilon}^H}{\bar{\varepsilon}^L \bar{\varepsilon}^H}$ (second term in (30)). Given the mean values in Table 3 computed in the Stone-Geary case, the ratio of the first elasticity to the second is at least $\frac{1.094}{1.104-1.094} = 109.4$ (the same ratio in the NH-CES case is at least $\frac{1.127}{1.132-1.127} = 225.4$).

over time by decade (as symbolized by the operator Δ) yields our estimating equation for κ:

$$\Delta \ln \lambda^H_{n,y} - \Delta \ln \lambda^L_{n,y} = \alpha + \kappa \left(\Delta \ln \frac{w^H_{n,y}}{P^H_{n,y}} - \Delta \ln \frac{w^L_{n,y}}{P^L_{n,y}} \right) + \kappa \underbrace{\left(\Delta \ln A^H_{n,y} - \Delta \ln A^L_{n,y} \right)}_{\text{unobserved}} + \epsilon_{n,y}, \tag{31}$$

where $\lambda^t_{n,y}$ is the population share of type t in MSA n and year y, $w^t_{n,y}$ is the nominal wage, $P^t_{n,y}$ is the type-specific price index, $A^t_{n,y}$ is type-specific amenities and the term $\epsilon_{n,y}$ captures measurement error.

One may worry that OLS estimates of κ using equation (31) are biased due to reverse causality, due e.g. to housing prices responding to population changes, or omitted variables (amenity changes). In addition to double differencing our estimating equation to control for city-level shocks that are common to both types and for city-level characteristics that are constant over time, we follow common practice in the literature and instrument for relative real wage changes using a shift-share instrument differenced between types and over time. Our instrument predicts local changes in real wages by fixing type- and city-specific employment shares by industries to their 1980 levels and using national leave-one-out wage growth by industry. The identifying assumption is that relative 1980 employment shares in a city are uncorrelated with future changes in relative amenities (Goldsmith-Pinkham, Sorkin, and Swift, 2020).

We estimate κ for the homothetic case – in which case both types face the same price index $P^H_{n,y} = P^L_{n,y}$ – as well as for Stone-Geary and NH-CES specifications of non-homothetic preferences. In these latter cases, we use the decomposition in equation (21) and the result in Proposition 6 to replace relative changes in city-type-specific price indices, which we do not observe, by relative changes in city-type-specific expenditure shares, which we do.[18] Results are presented in Table 4. Our preferred specifications using the IV estimator yield an estimate of the migration elasticity of 1.87 in the homothetic case, 1.82 in the non-homothetic Stone Geary case, and 1.76 in the non-homothetic CES case. Our estimates

[18] Specifically, building on the derivations in Appendix B.1 for Stone-Geary preferences, we use:

$$\frac{\hat{P}^H_n}{\hat{P}^L_n} = \left(\frac{\widehat{1 - s^H_n}}{\widehat{1 - s^L_n}} \right)^{-1}.$$

By the same token, building on the derivations in Appendix B.5 for NH-CES preferences, we use:

$$\frac{\hat{P}^H_n}{\hat{P}^L_n} = \left(\frac{\widehat{1 - s^H_n}}{\widehat{1 - s^L_n}} \right)^{-1/(1-\sigma)},$$

where we use the estimates from Finlay and Williams (2022) to set $\sigma = 0.53$.

of decadal migration elasticities are within the range $[0.6, 4]$ of estimates in the literature (Diamond, 2016; Notowidigdo, 2020; Suarez Serrato and Zidar, 2016).[19]

Table 4: Estimation of Migration Elasticity κ

	Homothetic preferences		Stone-Geary		NH-CES	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\left(\ln \frac{w_n^H}{P_n^H} - \ln \frac{w_n^L}{P_n^L}\right)$	0.50***	1.87***	0.46***	1.82***	0.34***	1.76***
	(0.13)	(0.42)	(0.13)	(0.42)	(0.10)	(0.41)
Constant	0.00		-0.01		0.00	
	(0.01)		(0.01)		(0.01)	
Observations	876	876	876	876	876	876
F-Stat		627.50		499.92		307.85

Dependent variable: $\Delta \ln \lambda_n^H - \Delta \ln \lambda_n^L$ (relative labor supply changes). IVs are shift-shares.
Standard errors clustered at the MSA level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Regressions are weighted using 1980 MSA population.

One can interpret the IV coefficient in the second column as the labor supply elasticity with respect to nominal wages, and the IV coefficient in the fourth and sixth columns as the labor supply elasticity with respect to real wages when preferences are non-homothetic (Stone-Geary or NH-CES). Although the difference between these coefficients is not statistically significant, we note the coefficient on nominal wages is higher than on real wages. This means that omitting skill-specific prices from the regression (as we do under the homothetic assumption) yields an upward bias in the elasticity. This suggests that, in the cities where nominal wage ratios w_n^H/w_n^L have gone up the most and hence attracted the largest fractions of high-skill households, real wage ratios v_n^H/v_n^L have increased by even more, i.e. relative prices for high skill have increased *less* than for low skill. This is consistent with housing being a necessity: when house price rise in these skill-intensive cities, lower-skill are more sensitive to the corresponding price change.

5.3 Welfare Inequality 1980-2020

With this data at hand, we are ready to compute the model-consistent change in welfare inequality from 1980 to 2000 and 2000 to 2020 between high- and low-skill households, revisiting the seminal work of Moretti (2013). Our approach differs from that in Moretti

[19] Our estimates in the homothetic case are close to the upper bound of those of Diamond (2016), Notowidigdo (2020), and Suarez Serrato and Zidar (2016), which vary between 2 and 4, and above those in Goldsmith-Pinkham, Sorkin, and Swift (2020) and Autor, Dorn, and Hanson (2013). Goldsmith-Pinkham, Sorkin, and Swift (2020) report an inverse labor elasticity in the range $[1.28, 1.76]$ (and hence an elasticity in the range $[0.6, 0.8]$).

(2013) in two ways. First, our framework allows us to go from real wages to expected utility by taking into account the revealed preferences of households across locations, as we demonstrate in Section 4. Second, our non-homothetic preferences allow for price indices to be income specific, while Moretti (2013) imposes homothetic preferences.

We start by computing changes in nominal inequality as a benchmark. We report the findings in the first row of the left panel of Table 5. We find that the skill premium has risen by 19% between 1980 and 2000 (the time period covered by Moretti, 2013), by 5% between 2000 and 2020, and overall by 25% between 1980 and 2020. For the sake of comparison, we report the change in real-wage inequality computed using the methodology in Moretti (2013). Specifically, this corresponds to the difference between the average change in real wage for high-skill and low-skill, weighted by their respective populations across cities. Real wages are measured by skill-specific nominal wages deflated by a common price index for low- and high-skill workers:

$$\begin{aligned}\ln \frac{\hat{W}^H}{\hat{W}^L} = & \left[\sum_{n\in\mathcal{C}} \omega^H_{n,y'} \ln\left(\frac{w^H_{n,y'}}{P_{n,y'}}\right) - \sum_{n\in\mathcal{C}} \omega^H_{n,y} \ln\left(\frac{w^H_{n,y}}{P_{n,y}}\right)\right] \\ & - \left[\sum_{n\in\mathcal{C}} \omega^L_{n,y'} \ln\left(\frac{w^L_{n,y'}}{P_{n,y'}}\right) - \sum_{n\in\mathcal{C}} \omega^L_{n,y} \ln\left(\frac{w^L_{n,y}}{P_{n,y}}\right)\right]. \end{aligned} \tag{32}$$

According to this measure, real wage inequality has increased by 16% in the period 1980-2000 and 24% in the period 1980-2020, less than the 19%- and 25% corresponding increases in nominal wage inequality.

Table 5: Welfare Results

		Relative Welfare Change			Relative Equivalent Variation		
		1980–2000	2000–2020	1980–2020	1980–2000	2000–2020	1980–2020
Nominal	$\frac{\widehat{w}^H}{\widehat{w}^L}$	1.19	1.05	1.25	1.19	1.05	1.25
Real (Moretti, 2013)	$\frac{\hat{W}^H}{\hat{W}^L}$ in eq. (32)	1.16	1.07	1.24	1.16	1.07	1.24
Nominal + sorting	$\frac{\widehat{w}^H}{\widehat{w}^L}\left(\frac{\widehat{\lambda}^H}{\widehat{\lambda}^L}\right)^{-\frac{1}{\kappa}}$	1.15	1.06	1.23	1.15	1.06	1.23
Homothetic + sorting	$\frac{\widehat{w}^H}{\widehat{w}^L}\frac{\widehat{P}}{\widehat{P}}\left(\frac{\widehat{\lambda}^H}{\widehat{\lambda}^L}\right)^{-\frac{1}{\kappa}}$	1.15	1.06	1.23	1.15	1.06	1.23
Stone-Geary + sorting	$\frac{\widehat{w}^H}{\widehat{w}^L}\frac{\widehat{1-s^H}}{\widehat{1-s^L}}\left(\frac{\widehat{\lambda}^H}{\widehat{\lambda}^L}\right)^{-\frac{1}{\kappa}}$	1.17	1.09	1.28	1.16	1.09	1.27
NH-CES + sorting	$\frac{\widehat{w^t_n}}{\widehat{w^s_n}}\left(\frac{\widehat{1-s^H_n}}{\widehat{1-s^L_n}}\right)^{1/(1-\sigma)}\left(\frac{\hat{\lambda}^H_n}{\hat{\lambda}^L_n}\right)^{-\frac{1}{\kappa}}$	1.19	1.12	1.34	1.17	1.10	1.31

Note: Our estimate of a 19% change in nominal wage inequality over 1980–2020 corresponds to Table 5, Column 8 of Moretti (2013), where the log wage difference is 0.18, implying exp(0.18) = 1.20, or 20%. The slight discrepancy is due to rounding. When we replicate his specification, we obtain a coefficient of 0.177, so exp(0.177) = 1.19, or 19%.

Next, we implement our approach. Starting with nominal wage inequality, our first step is to correct for idiosyncratic location preferences and differential sorting patterns. Results are reported in the third row of Table 5. We find a relative increase in inequality of 23% between 1980 and 2020, which is slightly lower than the 25%-increase in nominal inequality and the

24%-increase in real inequality. We report in the fourth row of the table the model-consistent welfare changes in equation (25) under the assumption that preferences are homothetic. In this case, high- and low-skill households face the same changes in local prices, and hence measured changes in inequality are identical to those of the previous specification (in the row above).

We then allow for non-homothetic housing demand. We compute the change in utility inequality from equation (25) in the Stone-Geary and NH-CES cases, which we recall here for convenience:

$$\frac{\hat{\mathbb{V}}^t_{SG}}{\hat{\mathbb{V}}^s_{SG}} = \frac{\widehat{w}^t_n}{\widehat{w}^s_n}\frac{\widehat{1-s^t_n}}{\widehat{1-s^s_n}}\left(\frac{\hat{\lambda}^t_n}{\hat{\lambda}^s_n}\right)^{-\frac{1}{\kappa}}, \qquad \frac{\hat{\mathbb{V}}^t_{NHCES}}{\hat{\mathbb{V}}^s_{NHCES}} = \frac{\widehat{w}^t_n}{\widehat{w}^s_n}\left(\frac{\widehat{1-s^t_n}}{\widehat{1-s^s_n}}\right)^{\frac{1}{1-\sigma}}\left(\frac{\hat{\lambda}^t_n}{\hat{\lambda}^s_n}\right)^{-\frac{1}{\kappa}}. \tag{33}$$

Given our calibrated $\sigma = 0.53$, changes in welfare inequality put a geometric weight on changes in expenditure shares that is about twice as large if preferences are NH-CES than if preferences are Stone-Geary (i.e., $\frac{1}{1-0.53} \approx 2$). In the model, the formulas in equation (33) hold exactly across all locations; in the data, the resulting measure differs by location, which we interpret as measurement error. We compute the formula for each city and report the geometric average as our main estimate of $\frac{\hat{\mathbb{V}}^H}{\hat{\mathbb{V}}^L}$, weighted by 1980 population shares.[20] We report the results in the fifth and sixth rows of Table 5. The main takeaway from this exercise is that welfare inequality between high- and low-skill households has risen by *more* than nominal inequality once we account for non-homothetic preferences. Rather than increasing by 23% under homothetic preferences, inequality has risen by 28% between 1980 and 2020 if we assume that preferences are Stone-Geary, and by 34% if preferences are NH-CES.

Our third and final modification is to convert these changes in "utils" into money-metric changes, using the average log-log marginal utility of income for each worker type. Specifically, we use equation (30) for high- and low-skill households, treating the contribution of the final term in the expression as negligible (i.e., $\simeq 1$) as discussed above. This yields for each specification of preferences $\zeta = SG, NHCES$:

$$\ln\frac{\hat{y}^H_\zeta}{\hat{y}^L_\zeta} = \frac{1}{\bar{\varepsilon}^H_\zeta}\ln\hat{\mathbb{V}}^H_\zeta - \frac{1}{\bar{\varepsilon}^L_\zeta}\ln\hat{\mathbb{V}}^L_\zeta. \tag{34}$$

The rightmost panel of Table 5 applies this measure on the relative welfare changes in the left panel using the estimates of the log-log marginal utility of incomes for renters that we report in Table 3. Results are qualitatively similar to those in the left panel, but a bit more

[20]Results are almost identical when they are non-weighted or weighted by low-skill or high-skill population instead.

muted. We find that money-metric welfare inequality has increased by 27% if we assume that preferences are Stone-Geary, and by 31% if preferences are NH-CES.

We conclude that accounting for non-homothetic preferences considerably changes conclusions on welfare inequality. High-skill households have increasingly flocked into cities with high and fast-growing housing costs, but they did so in a large part because their purchasing power is less affected by housing costs than low-skill households. As a result, welfare inequality has risen by more than nominal inequality, not less.

6 Conclusion

This paper develops a spatial equilibrium model where agents have heterogeneous incomes and non-homothetic consumption preferences. In this framework, the relative price of traded vs. non-traded goods varies across locations, and because agents share identical but non-homothetic preferences, they value locations differently. This generates spatial sorting as a direct consequence of income-inelastic housing demand. We embed this force in a broader model where comparative advantage in production further shapes location choices. Our results show that under general conditions, high-income agents systematically sort into expensive cities, driven both by their lower housing expenditure share and their productivity advantage in large cities.

Beyond explaining sorting, the model has implications for welfare inequality measurement in a spatial equilibrium context. We extend standard welfare measurement approaches in economic geography to account for non-homothetic preferences. Applying this framework to the United States (1980–2020), we show that while high-skill workers increasingly relocated to high-cost cities, they did so in part because their purchasing power is less affected by housing costs than low-skill workers. As a result, welfare inequality rose by 27-31%—outpacing the increase in nominal wage inequality.

Our model abstracts from endogenous agglomeration economies, amenities, and trade costs to highlight the role of housing demand in driving sorting. While some of our results could be adapted to specific functional forms that incorporate endogenous agglomeration and amenities, introducing heterogeneous trade costs poses significant analytical challenges. Existing methods Allen, Arkolakis, and Takahashi (2020) do not apply when agents have heterogeneous incomes and non-homothetic preferences. Addressing this limitation is an important direction for future research.

References

Aguiar, M. and M. Bils (2015, September). Has Consumption Inequality Mirrored Income Inequality? *American Economic Review 105*(9), 2725–2756.

Ahlfeldt, G. M. and E. Pietrostefani (2019, may). The economic effects of density: A synthesis. *Journal of Urban Economics 111*, 93–107.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Albouy, D., G. Ehrlich, and Y. Liu (2016, nov). Housing demand, cost-of-living inequality, and the affordability crisis. Technical report.

Allen, T. and C. Arkolakis (2014). Trade and the topography of the spatial economy. *Quarterly Journal of Economics 1085*, 1139.

Allen, T., C. Arkolakis, and Y. Takahashi (2020, feb). Universal gravity. *Journal of Political Economy 128*(2), 393–433.

Almagro, M. and T. Domínguez-Iino (2020). Location Sorting and Endogenous Amenities: Evidence from Amsterdam.

Arntz, M., E. Brull, and C. Lipowski (2023, May). Do preferences for urban amenities differ by skill? *Journal of Economic Geography 23*(3), 541–576.

Atkin, D., B. Faber, T. Fally, and M. Gonzalez-Navarro (2023, sep). Measuring welfare and inequality with incomplete price information. *The Quarterly Journal of Economics*.

Autor, D. H., D. Dorn, and G. H. Hanson (2013, October). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review 103*(6), 2121–2168.

Baqaee, D. R. and A. Burstein (2022, nov). Welfare and output with income effects and taste shocks. *The Quarterly Journal of Economics 138*(2), 769–834.

Baum-Snow, N., M. Freedman, and R. Pavan (2018, oct). Why has urban inequality increased? *American Economic Journal: Applied Economics 10*(4), 1–42.

Baum-Snow, N. and R. Pavan (2012). Understanding the city size wage gap. *The Review of economic studies 79*(1), 88–127.

Baum-Snow, N. and R. Pavan (2013). Inequality and city size. *Review of Economics and Statistics 95*(5), 1535–1548.

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

Black, D., N. Kolesnikova, and L. Taylor (2009, jan). Earnings functions when wages and prices vary by location. *Journal of Labor Economics 27*(1), 21–47.

Bohr, C. E., M. Mestieri, and F. Robert-Nicoud (2024). Economic geography and structural change. *CEPR Discussion Paper 19766*.

Boppart, T. (2014, nov). Structural change and the kaldor facts in a growth model with relative price effects and non-gorman preferences. *Econometrica 82*(6), 2167–2196.

Card, D., J. Rothstein, and M. Yi (2023). Location, location, location. *NBER Working Paper 31587*.

Caron, J., T. Fally, and J. R. Markusen (2014, may). International trade puzzles: A solution linking production and preferences. *The Quarterly Journal of Economics 129*(3), 1501–1552.

Combes, P.-P., G. Duranton, and L. Gobillon (2008, March). Spatial wage disparities: Sorting matters! *Journal of Urban Economics 63*(2), 723–742.

Combes, P.-P., G. Duranton, and L. Gobillon (2019, July). The Costs of Agglomeration: House and Land Prices in French Cities. *The Review of Economic Studies 86*(4), 1556–1589.

Comin, D., D. Lashkari, and M. Mestieri (2021). Structural change with long-run income and price effects. *Econometrica 89*(1), 311–374.

Couture, V., C. Gaubert, J. Handbury, and E. Hurst (2023, apr). Income growth and the distributional effects of urban spatial sorting. *Review of Economic Studies*.

Dauth, W., S. Findeisen, E. Moretti, and J. Suedekum (2022, feb). Matching in cities. *Journal of the European Economic Association 20*(4), 1478–1521.

Davis, D. R. and J. I. Dingel (2019, January). A Spatial Knowledge Economy. *American Economic Review 109*(1), 153–170.

Davis, D. R. and J. I. Dingel (2020, March). The comparative advantage of cities. *Journal of International Economics 123*, 103291.

Davis, M. A. and F. Ortalo-Magné (2011, apr). Household expenditures, wages, rents. *Review of Economic Dynamics 14*(2), 248–261.

D'Costa, S. and H. G. Overman (2014, sep). The urban wage growth premium: Sorting or learning? *Regional Science and Urban Economics 48*, 168–179.

De La Roca, J. and D. Puga (2017). Learning by working in big cities. *The Review of Economic Studies 84*(1), 106–142.

Deaton, A. and J. Muellbauer (1980, June). An almost ideal demand system. *American Economic Review 70*(3), 312–326.

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

Diamond, R. and C. Gaubert (2022). Spatial sorting and inequality. *Annual Review of Economics 14*, 795–819.

Dingel, J. I., A. Miscio, and D. R. Davis (2021, sep). Cities, lights, and skills in developing economies. *Journal of Urban Economics 125*, 103174.

Eckert, F. and M. Peters (2022, sep). Spatial structural change. Technical report.

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

Fajgelbaum, P. and C. Gaubert (2020). Optimal Spatial Policies, Geography and Sorting. *Quarterly Journal of Economics*.

Fan, T., M. Peters, and F. Zilibotti (2023). Growing like india—the unequal effects of service-led growth. *Econometrica 91*(4), 1457–1494.

Fieler, A. C. (2011). Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation. *Econometrica 79*(4), 1069–1101.

Finlay, R. and T. C. Williams (2022). Sorting and the Skill Premium: The Role of Nonhomothetic Housing Demand.

Fretz, S., R. Parchet, and F. Robert-Nicoud (2021, sep). Highways, market access and spatial sorting. *The Economic Journal 132*(643), 1011–1036.

Gaigné, C., H. R. Koster, F. Moizeau, and J.-F. Thisse (2022, mar). Who lives where in the city? amenities, commuting and income sorting. *Journal of Urban Economics 128*, 103394.

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

Giannone, E. (2019). Skill-Biased Technical Change and Regional Convergence.

Glaeser, E. L., M. E. Kahn, and J. Rappaport (2008, January). Why do the poor live in cities? the role of public transportation. *Journal of Urban Economics 63*(1), 1–24.

Glaeser, E. L. and D. C. Mare (2001). Cities and skills. *Journal of Labor Economics 19*(2), 316–342.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020, August). Bartik Instruments: What, When, Why, and How. *American Economic Review 110*(8), 2586–2624.

Handbury, J. (2021). Are poor cities cheap for everyone? non-homotheticity and the cost of living across u.s. cities. *Econometrica 89*(6), 2679–2715.

Hanoch, G. (1975, May). Production and demand models with direct or indirect implicit additivity. *Econometrica 43*(3), 395–419.

Hoelzlein, M. (2023). Two-sided sorting and spatial inequality in cities.

Jaravel, X. (2021, aug). Inflation inequality: Measurement, causes, and policy implications. *Annual Review of Economics 13*(1), 599–629.

Jaravel, X. and D. Lashkari (2022). Measuring growth in consumer welfare with income-dependent preferences. *CEPR Discussion Paper 16866*.

Karlstrom, A. (2000). Non-linear value functions in random utility econometrics. *Presented at 9th International Conference on Travel Behaviour Research, Queensland, Australia.*.

Kuminoff, N. V., V. K. Smith, and C. Timmins (2013, dec). The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature 51*(4), 1007–1062.

Matsuyama, K. (2019). Engel's law in the global economy: Demand-induced patterns of structural change, innovation, and trade. *Econometrica 87*(2), 497–528.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.

McFadden, D. (1999). *Computing willingness to pay in random utility models*, Chapter Trade, Theory and Econometrics, pp. 253–74. Routledge.

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

Muellbauer, J. (1975, oct). Aggregation, income distribution and consumer demand. *The Review of Economic Studies 42*(4), 525.

Notowidigdo, M. J. (2020, July). The incidence of local labor demand shocks. *Journal of Labor Economics 38*(3), 687–725.

Oberfield, E. (2023). Inequality and measured growth. *Mimeograph*.

Redding, S. J. and E. A. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*(1).

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, 1257–1278.

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics 3*, 324–345.

Ruggles, S., S. Flood, M. Sobek, D. Backman, A. Chen, G. Cooper, S. Richards, R. Rodgers, and M. Schouweiler (2024). IPUMS USA: Version 15.0 [dataset]. Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D010.V15.0.

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

Sato, R. (1977, September). Homothetic and non-homothetic ces production functions. *American Economic Review 67*(4), 559–569.

Suarez Serrato, J. C. and O. Zidar (2016, September). Who benefits from state corporate tax cuts? a local labor markets approach with heterogeneous firms. *American Economic Review 106*(9), 2582–2624.

Takeda, K. (2022). The geography of structural transformation: Effects on inequality and mobility. *CEP Discussion Paper 1893*.

Tsivanidis, N. (2019). Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio.

Appendix

A Theory Appendix

A.1 Proof of Proposition 1 (Existence).

Set $p = 1$ by choice of numeraire. We use bold letters to denote vectors: let $\mathbf{A} \equiv \{A_n\}_{n\in\mathcal{C}}$ and $\ell^{\mathbf{t}} \equiv \{L^t\lambda_n^t\}_{n\in\mathcal{C}}$. Manipulating (10) yields

$$r_n k\left(r_n, \eta_n\right) = \frac{r_n \sum_{t\in\mathcal{T}} L^t \lambda_n^t h_n^t}{H_n}.$$

Let $K\left(r_n, \eta_n\right) \equiv r_n k\left(r_n, \eta_n\right)$. K is continuously increasing in both its arguments, and $K\left(0, \eta\right) = 0$. Let

$$E_n \equiv r_n \sum_{t\in\mathcal{T}} L^t \lambda_n^t \left(\mathbf{r}, \ell^{\mathbf{t}}, \mathbf{A}\right) h\left(r_n, \ell_n^t, A_n\right)$$

denote total expenditure on housing in n. Inverting the housing market clearing condition above yields

$$r_n = K^{-1}\left(\frac{E_n}{H_n}, \eta_n\right) \equiv f_n\left(\mathbf{r}\right). \tag{35}$$

An equilibrium of the economy is then a solution of $\mathbf{f}\left(\mathbf{r}\right) = \mathbf{r}$. The function $\mathbf{f}$ is continuous on $\mathbb{R}_+^{N_C}$ by inspection since all functions λ_n^t and h are continuously differentiable. The lower bound of E_n is obviously equal to zero. To find an upper bound, assume that the whole labor force settles in city n and that it spends all its income on housing; in this case the numerator inside the parenthesis of f in equation (35) is equal to $E_n^{\max} \equiv \sum_{t\in\mathcal{T}} L^t \ell_n^t$, which is exogenously given and finite, hence $K^{-1}\left(\frac{E_n^{\max}}{H_n}, \eta_n\right)$ is finite, too. The image of f is therefore:

$$S \equiv \prod_{n\in\mathcal{C}} \left[0, K^{-1}\left(\frac{E_n^{\max}}{H_n}, \eta_n\right)\right] \subset \mathbb{R}_+^{N_C},$$

which is compact. Since $f : S \to S$ is continuous and S is compact, equation (35) admits a fixed point by Brouwer's fixed point theorem. QED.

A.2 Proof of Proposition 2 (Uniqueness).

Denote $\hat{Z}\left(\mathbf{r}\right)$ the adjusted aggregate excess demand function for housing as a function of prices (rewriting equation (10)), normalizing the price of the traded good to 1:

$$\hat{Z}_n\left(\mathbf{r}\right) = \frac{\sum_{t\in\mathcal{T}} L^t \lambda_n^t h_n^t}{H_n k\left(\frac{r_n}{p}, \eta_n\right)} - 1 \qquad \forall n \in \mathcal{C}.$$

The proof makes use of the index theorem . In step 1, we compute the Jacobian of the adjusted excess demand function, $D\hat{Z}(\mathbf{r})$, and compute the corresponding index of the Jacobian at equilibrium prices $\mathbf{r}^*$, where the index is defined as:

$$Index(\mathbf{r}^*) = sgn\left(Det\left[-D\hat{Z}(\mathbf{r}^*)\right]\right)$$

In step 2, we derive conditions under which that this index is necessarily positive at any equilibrium. Using the index theorem, we conclude that under these condition, the equilibrium is also unique.

After some algebra, the Jacobian of $\hat{Z}(\mathbf{r})$ is given by:

$$\frac{\partial \hat{Z}_n}{\partial r_i} = -\sum_{t\in\mathcal{T}} L^t\,\lambda_n^t h_n^t\left[H_n k\left(\frac{r_n}{p},\eta_n\right)\right]^{-1}\lambda_i^t\frac{\partial \log V_i^t}{\partial r_i} \qquad \text{for i}\neq n$$

$$\frac{\partial \hat{Z}_n}{\partial r_n} = \sum_{t\in\mathcal{T}} L^t\lambda_n^t h_n^t\left[H_n k\left(\frac{r_n}{p},\eta_n\right)\right]^{-1}\left[\frac{\partial \log V_n^t}{\partial r_n}\left(1-\lambda_n^t\right)+\frac{\partial \log h_n^t}{\partial r_n}-\frac{1}{r_n}\mathcal{E}_n(k,r)\right]. \tag{36}$$

We next show conditions under which $-D\hat{Z}(\mathbf{r})$ is an M-matrix, so that all its eigenvalues are positive hence, in particular, $sgn\left(Det\left[-D\hat{Z}(\mathbf{r}^*)\right]\right) = +1$. To do so, we use the following characterization of M-matrices: a matrix M with all its off-diagonal elements less than or equal to 0 (i.e., a Z-matrix) is an M-matrix if it is is semi-positive. That is, there exists some vector $x > 0$ such that $Mx > 0$.

Note that for $i \neq n$, it is clear from equation (36) that $-\frac{\partial \hat{Z}_n}{\partial r_i} \leq 0$, since $\frac{\partial \log V_i^t}{\partial r_i} \leq 0$, and the other terms are positive. Therefore, $-D\hat{Z}(\mathbf{r})$ is a Z-matrix. We next establish conditions under which it is an M-matrix.

We first consider the case $N_T = 1$, namely, there is only one type t in the economy. In this case, we compute the elements of $-\left(D\hat{Z}(\mathbf{r})\right)^T x$ for the positive vector $x = \left\{\frac{H_n^0 k\left(\frac{r_n}{p},\eta_n\right)}{h_n^t}\right\}_n$, leading to:

$$-\sum_n \frac{\partial \hat{Z}_n}{\partial r_i}x_n = L^t\lambda_i^t\left(\frac{1}{r_n}\mathcal{E}_n(k,r)-\frac{\partial \ln h_i^t}{\partial \ln r_i}\right) > 0.$$

Therefore, when $N_T = 1$, $-\left(D\hat{Z}(\mathbf{r})\right)^T$ is an M-matrix (and so is $-\left(D\hat{Z}(\mathbf{r})\right)$, since the transpose of an M-matrix is an M-matrix) and its eigenvalues are all positive. So is its determinant, so that:

$$Index(\mathbf{r}^*) = +1.$$

The index theorem, which states that $\sum_{r^* eq.} Index(r^*) = +1$, ensures that the equilibrium is unique.

Next, in the case $N_T > 1$, we apply this strategy to the positive vector $x = \{r_n\}_n$ and compute the elements of $-D\hat{Z}(\mathbf{r})\cdot x$, which are :

$$-\sum_i \frac{\partial \hat{Z}_n}{\partial r_i}x_i = \sum_{t\in\mathcal{T}} L^t\,\lambda_n^t h_n^t \left[H_n k\left(\frac{r_n}{p},\eta_n\right)\right]^{-1}\left(-\mathcal{E}_n^t(V,r)-\mathcal{E}_n^t(h,r)+\mathcal{E}_n(k,r)+\sum_{i=1}^N \lambda_i^t\mathcal{E}_i^t(V,r)\right).$$

A sufficient condition for this to be positive is that $\forall n$, $\forall t$:

$$\mathcal{E}_n(k,r)-\mathcal{E}_n^t(h,r)>\mathcal{E}_n^t(V,r)-\sum_{i=1}^N \lambda_i^t\mathcal{E}_i^t(V,r). \tag{37}$$

Notice that if V is homothetic in the sense of Condition 2, and amenities are multiplicatively separable in utility, then $\mathcal{E}_i^t(V,r)=\textit{constant}$ so that the sufficient condition become $\mathcal{E}_n(k,r)-\mathcal{E}_n^t(h,r)>0$, which is always satisfied when h is a normal good. That is, uniqueness is warranted under homothetic preferences and separable amenities.

If Condition 2 does not hold, the sufficient condition (37) requires that the heterogeneity in $\mathcal{E}_n^t(V,r)$ across cities, for each type, be not too large. A more stringent condition that ensures that equation (37) holds is simply:

$$\inf_n \mathcal{E}_n(k,r)+\inf_{n,t}\left|\mathcal{E}_n^t(h,r)\right|>\sup_{n,t}\left|\mathcal{E}_n^t(V,r)-\overline{\mathcal{E}^t(V,r)}\right|,$$

$$\overline{\mathcal{E}^t(V,r)}\equiv\sum_{i=1}^N \lambda_i^t\mathcal{E}_i^t(V,r)$$

which is the one reported in the proposition. When this condition holds, $-D\hat{Z}(\mathbf{r})$ is an M-matrix, hence its index is positive at all equilibrium r^*. By the index theorem, which states that $\sum_{r^*eq.} Index(r^*)=+1$, the uniqueness of the equilibrium is then warranted. Notice, equation (37) is also the condition that ensures that the excess demand system exhibits gross substitutes. Specifically, denoting $Z(\mathrm{r},p)$ the excess demand function, it is clear that $\frac{\partial Z_n}{\partial r_i}>0$ for $i\neq n$, but condition (37) ensures that, in addition, $\frac{\partial \hat{Z}_n}{\partial p}>0$.

A.3 Proof of Proposition 3 (Sorting).

(i) The contrapositive (Conditions 1 and 2⇒ No sorting) is immediate to verify, since in this case location choices are given by:

$$\lambda_n^t=\frac{\left[\frac{T_n}{P(r_n,A_n)}\right]^\kappa}{\sum_{m\in\mathcal{C}}\left[\frac{T_m}{P(r_m,A_m)}\right]^\kappa}.$$

Clearly, λ_n^t is independent of t, hence there is no spatial sorting by type.

(ii) Assume Condition 1 fails but Condition 2 holds. Then:

$$\frac{\partial^2}{\partial T_n\partial t}\ln\left(\frac{\lambda_n^t}{1-\lambda_n^t}\right)=\kappa\left(1-\lambda_n^t\right)\frac{\partial^2 \ln w_n^t}{\partial T_n\partial t},$$

which is different from 0 for some t, n given that Condition 1 is assumed to fail. Therefore, equation (12) cannot hold generically.

Finally, recast the no-sorting condition (12) as

$$\frac{\partial}{\partial \ln t} \ln \left(\frac{\lambda_n^t}{1-\lambda_n^t}\right) - \frac{\partial}{\partial \ln t} \ln \left(\frac{\lambda_m^t}{1-\lambda_m^t}\right) = 0$$

for all n, m, and assume that Condition 2 fails but Condition 1 holds, i.e. $w_n^t = T_n t$. Proceed by contradiction and assume that there is no sorting. Therefore:

$$0 = \frac{\partial}{\partial \ln t} \ln \left(\frac{\lambda_n^t}{1-\lambda_n^t}\right) - \frac{\partial}{\partial \ln t} \ln \left(\frac{\lambda_m^t}{1-\lambda_m^t}\right) = \frac{\partial \ln V_n^t}{\partial \ln t} - \frac{\partial \ln V_m^t}{\partial \ln t} = \frac{\partial \ln V_n^t}{\partial \ln w_n^t} - \frac{\partial \ln V_m^t}{\partial \ln w_m^t}.$$

In turn, if $\frac{\partial \ln V_n^t}{\partial \ln w_n^t} = \frac{\partial \ln V_m^t}{\partial \ln w_m^t}$ generically holds for any T_n, T_m , then $\frac{\partial \ln V}{\partial \ln w}$ is constant; hence V is homogeneous in income, a contradiction.

A.4 Proof of Proposition 4 ("Rosen-Roback").

Consider equation (10), $0 = \sum_{t \in \mathcal{T}} L^t \lambda_n^t h_n^t - H_n k\left(r_n, \eta_n\right)$. The right-hand side is increasing in A and T, and decreasing in H and η; it is also decreasing in r, so it has to be that $r_m < r_n$ if $A_n > A_m$ (respectively if $T_n > T_m$, $H_n < H_m$, $\eta_n < \eta_m$), all else equal.

A.5 Proof of Proposition 5 (Sorting Patterns).

By Definition 1, city n disproportionately attracts high-t workers relative to city m if and only if, for any $s, t \in \mathcal{T}$ with $s < t$:

$$1 < \frac{\lambda_n^t}{\lambda_m^t} \left(\frac{\lambda_n^s}{\lambda_m^s}\right)^{-1} = \left[\frac{v\left(r_n, w_n^t\right)}{v\left(r_m, w_m^t\right)}\right]^{\kappa} \left[\frac{v\left(r_n, w_n^s\right)}{v\left(r_m, w_m^s\right)}\right]^{-\kappa},$$

where the equality follows from equation (8). For further reference, observe also that the share of expenditure devoted to housing is equal to

$$s_n^t \equiv s\left(r_n, w_n^t\right) = -\frac{\mathcal{E}_n^t\left(V, r\right)}{\mathcal{E}_n^t\left(V, w\right)} \tag{38}$$

by Roy's identity. (***i***) and (***ii***) By assumption, $T_m = T_n$ so that $w_n^t = w_m^t = w^t$ for all t, and the inequality above obtains if and only if (taking logs):

$$0 < \int_{r_m}^{r_n} \left[\mathcal{E}^t\left(v, r\right) - \mathcal{E}^s\left(v, r\right)\right] \mathrm{d} \ln r.$$

Given equation (38), the right hand side of this expression is equal to

$$\int_{r_m}^{r_n} \left[-s\left(r, w_n^t\right) \mathcal{E}^t\left(v, w\right) + s\left(r, w_n^s\right) \mathcal{E}^s\left(v, w\right)\right] \mathrm{d} \ln r > \int_{r_m}^{r_n} \left[-s\left(r, w_n^t\right) + s\left(r, w_n^s\right)\right] \mathcal{E}^t\left(v, w\right) \mathrm{d} \ln r,$$

given $s < t$ and Assumption 3. Finally, the term in the square parenthesis is positive by equation (4) (housing is a necessity), and $\mathcal{E}_n^t(v, w) > 0$ by definition. Hence, this inequality holds if and only if $r_n > r_m$, which can arise only if either $A_n > A_m$, $H_n < H_m$, or, provided that the properties of function k obey equation (6), if $\eta_n < \eta_m$, by Proposition 4 (recall $T_m = T_n$ by assumption).

(***iii***) For this part, note first that $T_n > T_m$ implies $r_n > r_m$, everything else equal, by Proposition 4.Then,

$$\frac{v(r_n, w_n^t)}{v(r_m, w_n^t)}\left[\frac{v(r_n, w_m^s)}{v(r_m, w_m^s)}\right]^{-1} = \frac{v(r_n, w_n^t)}{v(r_n, w_m^t)}\left[\frac{v(r_n, w_n^s)}{v(r_n, w_m^s)}\right]^{-1}\frac{v(r_n, w_m^t)}{v(r_m, w_m^t)}\left[\frac{v(r_n, w_m^s)}{v(r_m, w_m^s)}\right]^{-1}.$$

This term is larger than unity because $\frac{v(r_n, w_n^t)}{v(r_n, w_m^t)}\left[\frac{v(r_n, w_n^s)}{v(r_n, w_m^s)}\right]^{-1} > 1$ by equation (16), and $\frac{v(r_n, w_m^t)}{v(r_m, w_m^t)}\left[\frac{v(r_n, w_m^s)}{v(r_m, w_m^s)}\right]^{-1} > 1$ by the first part of this Proposition.

B Appendix: Data

Sample This analysis uses decennial U.S. Census microdata from 1980-2000 and the 5-year American Community Survey (ACS) for 2010 and 2020 from IPUMS (Ruggles et al., 2024). The sample is restricted to individuals aged 26-59 in private households who worked at least 5 weeks in the previous year. The analysis excludes individuals enrolled in school and those born outside the United States.

Unit of Geography All geographic identifiers are based on the Office of Management and Budget's (OMB) Metropolitan Statistical Area definitions. To maintain consistent geographic definitions across the 1980-2020 period, we construct a crosswalk between Census METAREA codes (available through 2010) and 2013-based Metropolitan Statistical Area definitions. Using the 2010 Census sample, we observe both METAREA codes and 2013-based MSA codes for each individual. We calculate population-weighted concordance shares that map each METAREA to potentially multiple 2013 MSAs. Specifically, for each METAREA-2013 MSA pair, we compute the share of the METAREA's population that resides in each 2013 MSA. These shares are used as weights to allocate observations from the 2020 data to METAREA definitions. For 2013 MSAs that span multiple METAREAs, observations are probabilistically assigned based on these population shares, with sampling weights adjusted proportionally.

Education Education levels are coded into two categories: high school graduates and college graduates (those with 4+ years of college). Experience is calculated as age minus years of education minus 6, where years of education is assumed to be 16 for college graduates and 12 for high school graduates.

Wages Wages are constructed from annual wage income divided by the product of weeks worked and usual weekly hours. Weeks worked are imputed for categorical responses using interval midpoints.

Housing Costs Housing costs are measured using gross rent for renters. Housing cost shares are calculated as annual housing costs divided by household income, which we impute to the individual.

Local Prices Local price indices are constructed following Moretti (2013). The housing component uses rental costs for 2-3 bedroom units. The non-housing component uses local BLS CPI data. Housing expenditure shares are based on BLS Consumer Expenditure Survey weights for shelter plus utilities, ranging from 38.2% in 1980 to 37.7% in 2020.

Covariate Adjustment All wage and income measures are adjusted for demographic composition by regressing log outcomes on a cubic in experience, gender, and race indicators (White, Black, Asian, and other), separately for 1980-2000 and 2010-2020, nesting Moretti (2013). The adjusted values hold demographic characteristics constant at their period-specific means.

Bartik Instrument The Bartik labor demand instruments are constructed using local industry employment shares from 1980 interacted with national industry wage growth. We begin with individual-level Census data collapsed to MSA-year-industry-skill cells containing employment counts and average wages. Industries are harmonized using a balanced panel of 1990 Census industry codes following Autor, Dorn, and Hanson (2013). Industries with zero national employment in any sample year are dropped to maintain consistency across years.

For each MSA and skill group (college and non-college), we calculate 1980 employment shares across industries to capture the initial local industrial structure. National wage growth by industry and skill level is computed using a leave-one-out mean that excludes each MSA's own wages when calculating the national trend. Wage growth is measured as the change in log average wages between consecutive decades.

The instrument is then constructed by multiplying each MSA's 1980 industry employment shares by the subsequent national industry wage growth rates and summing across industries within MSA-skill cells. The final relative shock measure is calculated as the difference between the college and non-college predicted wage changes.

Online Appendix (not for publication)

A Additional Theory results

A.1 Sorting in the Two-City Case

We claim in Footnote 7 that, in the two-city case $n \in \{1,2\}$, the sufficient condition for uniqueness in equation (11) is akin to a single-crossing condition in the space of unit housing prices (the proof in Appendix A works with an arbitrary number of cities).

Totally differentiate equation (10) for city 1 with respect to r_1 and r_2 yields

$$0 = \sum_t \theta_1^t \left\{ \left[-\mathcal{E}_1(k,r) + \mathcal{E}_1^t(h,r) + \left(1-\lambda_1^t\right)\mathcal{E}_1^t(V,r) \right] \mathrm{d}\ln r_1 - \left(1-\lambda_1^t\right)\mathcal{E}_2^t(V,r)\,\mathrm{d}\ln r_2 \right\},$$

where $\theta_1^t \equiv L^t \lambda_1^t h_1^t \left(H_1^0 k_1\right)^{-1}$, $k_1 \equiv k(r_1, \eta_1)$, and $\sum_t \theta_1^t = 1$. Thus

$$\left.\frac{\mathrm{d}\ln r_2}{\mathrm{d}\ln r_1}\right|_{Z_1=0} = \frac{\sum_t \theta_1^t \left[\mathcal{E}_1(k,r) - \mathcal{E}_1^t(h,r) - \left(1-\lambda_1^t\right)\mathcal{E}_1^t(V,r)\right]}{-\sum_t \theta_1^t \left(1-\lambda_1^t\right)\mathcal{E}_2^t(V,r)},$$

which is positive by $\mathcal{E}_n^t(h,r), \mathcal{E}_n^t(V,r) < 0$. It follows that the locus for an equilibrium in the housing market for city 1 in the (r_1, r_2)-space is upward sloping. By the same token, totally differentiating (10) for city 2 yields:

$$\left.\frac{\mathrm{d}\ln r_2}{\mathrm{d}\ln r_1}\right|_{Z_2=0} = \frac{-\sum_t \theta_2^t \lambda_2^t \mathcal{E}_1^t(V,r)}{\sum_t \theta_2^t \left[\mathcal{E}_2(k,r) - \mathcal{E}_2^t(h,r) - \lambda_1^t \mathcal{E}_2^t(V,r)\right]} > 0.$$

The locus for an equilibrium in city 1 crosses the locus for an equilibrium in city 2 from below in the space of housing prices (r_1, r_2), where r_1 is on the horizontal axis and r_2 is on the vertical one if and only if , $\forall r, h, k$:

$$1 < \frac{\sum_t \theta_1^t \left[\mathcal{E}_1(k,r) - \mathcal{E}_1^t(h,r) - \left(1-\lambda_1^t\right)\mathcal{E}_1^t(V,r)\right]}{-\sum_t \theta_1^t \left(1-\lambda_1^t\right)\mathcal{E}_2^t(V,r)} \frac{\sum_t \theta_2^t \left[\mathcal{E}_2(k,r) - \mathcal{E}_2^t(h,r) - \lambda_1^t \mathcal{E}_2^t(V,r)\right]}{-\sum_t \theta_2^t \lambda_1^t \mathcal{E}_1^t(V,r)} \tag{39}$$

(and > 1 for the opposite ordering). Whenever $\mathcal{E}_n(k,r) - \mathcal{E}_n^t(h,r) > \left(1-\lambda_n^t\right)\left[\mathcal{E}_n^t(V,r) - \mathcal{E}_{m\neq n}^t(V,r)\right]$ holds for all n, t as in (11), we get that $\frac{\sum_t \theta_1^t \left[\mathcal{E}_1(k,r) - \mathcal{E}_1^t(h,r) - \left(1-\lambda_1^t\right)\mathcal{E}_1^t(V,r)\right]}{-\sum_t \theta_1^t \left(1-\lambda_1^t\right)\mathcal{E}_2^t(V,r)} > 1$ (and similarly for city 2), so that the single crossing property holds.

What is the economic intuition for the condition in equation (39)? Denote by (r_1^*, r_2^*) a pair of housing prices that satisfy $Z_1(r_1^*, r_2^*) = Z_2(r_1^*, r_2^*) = 0$ – that is, both housing markets are in equilibrium. Heuristically, if the conditioning equation (39) holds, then, whenever excess demand in either or both housing markets is non-zero, market forces whereby the housing price in $j = 1, 2$ rises if excess demand Z_j is positive and falls otherwise, lead prices to return to equilibrium. Further, this equilibrium is unique as condition (39) implies that the two $r_2(r_1)$ curves intersect at most once.

B Derivations of Section 4 for Specific Functional Forms

The proof of Proposition 6 is by construction, and is provided in Appendix s B.1 to B.6. Appendix B.7 provides the proof of Proposition 7.

We maintain Assumption 2 throughout in this Appendix , so that we can write $V_n^t = (A_n v_n^t)^\kappa$, where $v_n^t = v(r_n, w_n^t)$ and $\kappa > 1$ governs the migration elasticity.

B.1 Stone Geary

Recall equation (22) from Section 4 in the main text and use Roy's identity to write:

$$v_n^t = \frac{w_n^t - \underline{h} r_n}{p^{1-\psi}(r_n)^\psi} = \frac{w_n^t}{p^{1-\psi}(r_n)^\psi}\frac{1-s_n^t}{1-\psi}, \qquad s_n^t = \psi + (1-\psi)\underline{h}\frac{r_n}{w_n^t}, \tag{40}$$

where $\psi(0,1)$, and $\underline{h}$ can be interpreted as the subsistence level of housing if it is positive (observe that $s_n^t \in (\psi, 1)$ whenever $\underline{h} \geq 0$). The final expression for v_n^t in equation (40) is the product of a homothetic component, $\frac{w_n^t}{p^{1-\psi}(r_n)^\psi}$, and of a non-homothetic one, $\frac{1-s_n^t}{1-\psi}$.

We show in the main text that the change in welfare inequality between types s and t in equation (20) is equal to:

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \frac{\widehat{w_n^t}}{\widehat{w_n^s}}\frac{\widehat{1-s_n^t}}{\widehat{1-s_n^s}}\left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s}\right)^{-\frac{1}{\kappa}}. \tag{41}$$

This expression is a special case of equation (25) where $\alpha = b = 1$ and $a = \beta = 0$.

Under which conditions is Assumption 3 satisfied when preferences take the form in this section? In this case, we may rewrite the inequality in equation (62) as

$$0 \leq \frac{\underline{h} r_n}{w_n^t - \underline{h} r_n},$$

which is automatically satisfied whenever the wage covers the subsistence level of housing. Further,

$$\mathcal{E}_n^t(v, w) = \frac{1-\psi}{1-s_n^t}, \tag{42}$$

which is larger than one by $s_n^t \geq \psi$ and increasing in s_n^t by inspection. Thus, $\mathcal{E}_n^t(v, w)$ is decreasing in w (and hence in t). We note also that the Arrow-Pratt measure of risk-aversion, $-\mathcal{E}_n^t(v_w, w)$, is equal to zero in this case: Engel curves are linear.

B.2 PIGL

Boppart (2014) and, in a spatial context, Eckert and Peters (2022) use the following *Price-*

Independent Generalized Linear (PIGL) form for indirect utility (Muellbauer, 1975):

$$v_n^t = \left[\left(\frac{w_n^t}{p^{1-\psi}(r_n)^{\psi}}\right)^{\iota} - \frac{\nu}{\gamma}\left(\frac{r_n}{p}\right)^{\gamma\iota}\right]^{1/\iota}, \qquad \iota, \psi \in (0,1), \quad |\gamma| \geq 1, \quad \nu \in \{0,1\}, \tag{43}$$

where ι governs income effects, $\gamma\iota$ governs relative price effects, and v is an indicator parameter; preferences are non-homothetic if $\nu = 1$ and collapse to the homothetic Cobb-Douglas case if $\nu = 0$. Note that Boppart (2014) uses $\psi = 0$ and Eckert and Peters (2022) use $\gamma \leq -1$. Henceforth we set $\nu = 1$, which yields the following expenditure share of housing:

$$s_n^t = \psi + \left(\frac{w_n^t}{p^{1-\psi}(r_n)^{\psi}}\right)^{-\iota}\left(\frac{r_n}{p}\right)^{\gamma\iota}. \tag{44}$$

Combining equations (43) and (44) leads to:

$$v_n^t = \frac{w_n^t}{p^{1-\psi}(r_n)^{\psi}}\left(\frac{\gamma}{\gamma+\psi}\frac{1}{S_n^t}\right)^{1/\iota}, \qquad S_n^t = \frac{\gamma}{\gamma+\psi-s_n^t}. \tag{45}$$

Note that this expression for v_n^t is the product of a homothetic homothetic component, $\frac{w_n^t}{p^{1-\psi}(r_n)^{\psi}}$, and of a non-homothetic one, $(S_n^t)^{-\iota}$. Therefore the change in welfare inequality between types s and t in equation (20) is equal to:

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \frac{\widehat{w_n^t}}{\widehat{w_n^s}}\left(\frac{\widehat{S_n^t}}{\widehat{S_n^s}}\right)^{-1/\iota}\left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s}\right)^{-\frac{1}{\kappa}}. \tag{46}$$

That is, changes in housing expenditure shares and composition of any arbitrary municipality n, as well as an estimate of parameters κ, γ, ι, and ψ, are sufficient to compute changes in utility inequality between any pair of types. This expression is a special case of equation (25) where $\alpha = 1$, $\beta = \delta = 1/\iota$, $a = \psi$ and $b = 1/\gamma$. Note that the Stone-Geary and PIGL cases are isomorphic, and even identical in the limiting case $\iota \to 1$ and $\gamma + \psi = 1$.

Under which conditions is Assumption 3 satisfied when preferences take the form in this section? In this case,

$$\begin{aligned} \mathcal{E}_n^t(v,w) &= \left[1 - \frac{1}{\gamma}\left(\frac{r_n}{p}\right)^{\gamma\iota}\left(\frac{w_n^t}{p^{1-\psi}(r_n)^{\psi}}\right)^{-\iota}\right]^{-1} \\ &= \frac{\gamma}{\gamma+\psi-s_n} = S_n^t, \end{aligned} \tag{47}$$

where the second equality uses equation (44), and the last one follows from the definition of S_n^t in equation (45). By inspection, $\mathcal{E}_n^t(v,w)$ is larger than one, and it is increasing in s_n^t, and hence decreasing in t by $\partial s_n^t/\partial t < 0$, since housing is a necessity; hence, the inequality in equation (62) holds. The Arrow-Pratt measure of risk-aversion $-\mathcal{E}_n^t(v_w, w)$ is equal to $1 - \iota \in (0,1)$ in this case.

B.3 PIGLOG

Consider next the following *Price Invariant Generalized Logarithmic* (PIGLOG) indirect utility, which is a simplified version of Deaton and Muellbauer's *Almost Ideal Demand System* (Deaton and Muellbauer, 1980):

$$v_n^t = \left(\frac{r_n}{p}\right)^{\nu} \ln\left(\frac{w_n^t}{p^{1-\psi}\left(r_n\right)^{\psi}}\right), \qquad \nu > 0, \quad \psi \in (0,1)\,. \tag{48}$$

Under Assumption (2), shocks are Frechet distributed, which requires a cardinality for utility such that $v_n^t > 0$; in turn, this requires choosing units such that

$$\forall n \in \mathcal{C},\ \ \forall t \in \mathcal{T}; \quad w_n^t > p^{1-\psi}\left(r_n\right)^{\psi}. \tag{49}$$

Using Roy's identity, it is easy to verify that the expenditure share of housing is equal to

$$s_n^t = \psi - \nu \ln\left(\frac{w_n^t}{p^{1-\psi}\left(r_n\right)^{\psi}}\right). \tag{50}$$

Thus $s_n^t < \psi$ by equation (49). In turn, $s_n^t \in (0,\psi)$ if and only if

$$\ln\left(\frac{w_n^t}{p^{1-\psi}\left(r_n\right)^{\psi}}\right) \in \left(0, \frac{\psi}{\nu}\right).$$

Combining equations (48) and (44) leads to:

$$v_n^t = \left(\frac{r_n}{p}\right)^{\iota} \frac{\psi - s_n^t}{\nu}.$$

Therefore the change in welfare inequality between types s and t in equation (20) is equal to:

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \frac{\widehat{\psi - s_n^t}}{\widehat{\psi - s_n^s}} \left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s}\right)^{-\frac{1}{\kappa}}. \tag{51}$$

That is, changes in housing expenditure shares and composition of any arbitrary municipality n, as well as an estimate of parameters κ and ψ, are sufficient to compute changes in utility inequality between any pair of types. This expression is a special case of equation (25) where $\alpha = 0$, $\beta = 0$, $\delta = 1$, $a = 0$ and $b = 1/\psi$.

Under which conditions is Assumption 3 satisfied when preferences take the form in this section? Under Assumption 2, the inequality in equation (17) holds if and only if

$$0 \leq \ln\left(\frac{w_n^t}{p^{1-\psi}\left(r_n\right)^{\psi}}\right) = \frac{\psi - s_n^t}{\nu},$$

where the second equality follows from equation (50). Then the condition above is satisfied since $s_n^t \leq \psi$ by inspection of equation (50). We note also that the Arrow-Pratt measure of risk-aversion $-\mathcal{E}_n^t\left(v_w, w\right)$ is equal to $1-\left(\frac{r_n}{p}\right)^{\nu}$ in this case.

B.4 CRIE

Next, consider the case of preferences displaying *Constant Relative Income Elasticity* (CRIE) indirect utility (Caron, Fally, and Markusen, 2014; Eeckhout, Pinheiro, and Schmidheiny, 2014; Fieler, 2011; Hanoch, 1975):

$$v_n^t=\left[\left(\frac{w_n^t}{r_n}\right)^{-\psi}+\left(\frac{w_n^t}{p}\right)^{-\mu}\right]^{-\iota}, \tag{52}$$

where the parameters $\{\iota,\mu,\psi\}$ all have the same sign (either positive or negative). Using Roy's identity, it is easy to verify that the ratio of expenditure shares is equal to

$$\frac{1-s_n^t}{s_n^t}=\frac{\mu}{\psi}\left(\frac{w_n^t}{r_n}\right)^{\psi}\left(\frac{w_n^t}{p}\right)^{-\mu}. \tag{53}$$

Housing demand is income-inelastic if and only if $\mu<\psi$, and price-inelastic if only if $\mu,\psi<0$. This configuration is the empirically relevant one, so we focus on the case $0<\mu<\psi$, $0<\iota$. Combining equations (52) and (53) leads to:

$$v_n^t=\left[\frac{(\psi-\mu)^2}{\mu\psi}\right]^{-\iota}\left(\frac{w_n^t}{p}\right)^{\mu\iota}\left[\frac{s_n^t-\frac{\psi}{\psi-\mu}}{1+\frac{\psi-\mu}{\mu}\left(s_n^t-\frac{\psi}{\psi-\mu}\right)}\right]^{-\iota}.$$

Note that this expression for v_n^t is the product of a combination of parameters, of a homothetic homothetic component, $\left(\frac{w_n^t}{p}\right)^{\mu\iota}$, and of a non-homothetic one (the term involving s_n inside brackets). Only three among parameters κ,ι,μ,ψ can be separately identified. We set $\iota=1/\mu$ so that v_n^t is proportional to w_n^t given s_n^t. Let also define

$$S_n^t=\frac{s_n-a}{1-b\left(s_n-a\right)},\qquad a=\frac{\psi}{\psi-\mu},\quad b=-\frac{\psi-\mu}{\mu}.$$

By inspection, $0<a$ and $b<0$ hold in the combination of structural parameters of interest. Furthermore, S_n^t is positive if $s_n^t\in(a,1)$, and S_n^t is increasing in s_n^t. This normalization and the expressions above together yield the change in welfare inequality:

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s}=\frac{\widehat{w_n^t}}{\widehat{w_n^s}}\left(\frac{\widehat{S_n^t}}{\widehat{S_n^s}}\right)^{-1/\mu}\left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s}\right)^{-\frac{1}{\kappa}}. \tag{54}$$

This expression for relative welfare changes is a special case of equation (25), where $\alpha=1$, $\beta=\delta=\frac{1}{\mu}$, $a=\frac{\psi}{\psi-\mu}$ and $b=-\frac{\psi-\mu}{\mu}$.

Under which conditions is Assumption 3 satisfied when preferences take the form in this section? In this case,

$$\mathcal{E}_n^t(v,w)=\psi\frac{1}{s_n^t\mu+(1-s_n^t)\psi}. \tag{55}$$

Observe that $\mathcal{E}_n^t(v,w)$ is larger than one and increasing in s_n^t by $\mu<\psi$. Thus, $\mathcal{E}_n^t(v,w)$ is decreasing in w (and hence in t). We note also that the Arrow-Pratt measure of risk-aversion is equal to

$$\mathcal{E}_n^t(v_w,w)=1-\frac{\psi}{\psi(1-s_n^t)+\mu s_n^t}\left[\psi-(\psi-\mu)^2 s_n^t\left(1-s_n^t\right)\right],$$

which, under the parameter configuration $\psi>\mu>0$, can be shown to be bounded above by $1-\psi$.

B.5 NH-CES

Consider next the *Non-Homothetic Constant Elasticity of Substitution* (NH-CES) indirect utility (Albouy, Ehrlich, and Liu, 2016; Comin, Lashkari, and Mestieri, 2021; Hanoch, 1975; Matsuyama, 2019; Sato, 1977), defined implicitly in :

$$\left(w_n^t\right)^{1-\sigma}=\left(v_n^t\right)^{1-\sigma+\chi}\left[\Omega\left(v_n^t\right)^{\mu}(r_n)^{1-\sigma}+\left(v_n^t\right)^{\psi}p^{1-\sigma}\right],\qquad 0<\sigma,\Omega,\quad -1+\sigma-\chi<\mu<\psi.$$

We set $\chi=-\psi$ so that the elasticity of v_n^t with respect to w_n^t in equation (58) below is one, given s_n^t.[21] We can thus rewrite the expressions above as:

$$\left(w_n^t\right)^{1-\sigma}=\left(v_n^t\right)^{1-\sigma}\left[\Omega\left(v_n^t\right)^{\mu-\psi}(r_n)^{1-\sigma}+p^{1-\sigma}\right],\qquad 0<\sigma,\Omega,\quad -1+\sigma<\mu-\psi<0. \tag{56}$$

The second parameter restriction above is guided by economic logic (indirect utility is decreasing in housing and non-housing prices) and empirical evidence (housing demand is price- and income-inelastic).[22] The main text sets $\Omega=1$. This restriction is innocuous because, as we show below, this parameter enters neither the equilibrium expression for changes in welfare inequality in equation (59), nor the expression for the marginal utility of income in equation (60). Using Roy's identity, the ratio of the expenditure shares is equal to:

$$\frac{1-s_n^t}{s_n^t}=\frac{1}{\Omega}\left(v_n^t\right)^{\psi-\mu}\left(\frac{p}{r_n}\right)^{1-\sigma}. \tag{57}$$

[21] Finlay and Williams (2022) instead set $\chi=0$, which yields

$$v_n^t=\left(\frac{w_n^t}{p}\right)^{\frac{1-\sigma}{1-\sigma+\psi}}\left(1-s_n^t\right)^{\frac{1}{1-\sigma+\psi}}$$

and

$$\mathcal{E}_n^t(v,w)=\frac{1-\sigma}{1-\sigma+\mu s_n^t+\psi_{NHCES}\left(1-s_n^t\right)}$$

instead of the corresponding expressions in equations (58) and (60) below.

[22] From equation (57) below, the housing expenditure share s_n^t is increasing in its relative price $\frac{r_n}{p}$ if and only if $\sigma<1$, and decreasing in utility v_n^t if and only if $\mu<\psi$.

We can combine equations (56) and (57) to write v_n^t as a function of s_n^t and two prices among $\{p, r_n, w_n^t\}$. There are three such combinations. Solving equation (57) for $(v_n^t)^{\mu}(r_n)^{1-\sigma}$ and plugging the result into equation (56) lead to one such combination:[23]

$$v_n^t = \frac{w_n^t}{p}\left(1 - s_n^t\right)^{1/(1-\sigma)}. \tag{58}$$

This expression is the product of a homothetic component, $\frac{w_n^t}{p}$, and of a non-homothetic one (the term involving s_n^t).

The corresponding expression for the change in welfare inequality becomes:[24]

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \frac{\widehat{w_n^t}}{\widehat{w_n^s}}\left(\frac{\widehat{1 - s_n^t}}{\widehat{1 - s_n^s}}\right)^{1/(1-\sigma)}\left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s}\right)^{-\frac{1}{\kappa}}. \tag{59}$$

This expression is isomorphic to its equivalents in the Stone-Geary and PIGL cases – in equations (41) and (46), respectively. The corresponding expression for relative welfare changes is a special case of (25) where $\alpha = 1$, $\beta = 0$, $\delta = \frac{1}{1-\sigma}$ $a = 0$ and $b = 1$.

Under which conditions is Assumption 3 satisfied when preferences take the form in this section? In this case:

$$\mathcal{E}_n^t(v, w) = \frac{1-\sigma}{1 - \sigma - (\psi - \mu)\, s_n^t}. \tag{60}$$

This expression is larger than one and increasing in s_n^t by $\psi > \mu$, and hence it is decreasing in t by $\partial s_n^t/\partial t < 0$ since housing is a necessity; hence, the inequality in equation (62) holds. The equilibrium expression for the Arrow-Pratt measure of risk-aversion, $-\mathcal{E}_n^t(v_w, w)$, is too unwieldy in this case to be revealing.

---

[23] The other two are

$$v_n^t = \left[\Omega\left(\frac{r_n}{p}\right)^{1-\sigma}\frac{1 - s_n^t}{s_n^t}\right]^{1/(\psi-\mu)}$$

and

$$v_n^t = \left[\left(\frac{w_n^t}{r_n}\right)^{1-\sigma}\frac{s_n^t}{\Omega}\right]^{1/(1-\sigma+\mu-\psi)}.$$

[24] Those corresponding to the alternatives in footnote 23 are:

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \left(\frac{\widehat{1 - s_n^t}}{s_n^t}\right)^{1/(\psi-\mu)}\left(\frac{\widehat{1 - s_n^s}}{s_n^s}\right)^{-1/(\psi-\mu)}\left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s}\right)^{-\frac{1}{\kappa}}$$

and

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \left(\frac{\widehat{w_n^t}}{\widehat{w_n^s}}\right)^{(1-\sigma)/(1-\sigma+\mu-\psi)}\left(\frac{\widehat{s_n^t}}{\widehat{s_n^s}}\right)^{1/(1-\sigma+\mu-\psi)}\left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s}\right)^{-\frac{1}{\kappa}},$$

respectively.

B.6 HCD

Finally, consider the case of *Heterothetic Cobb-Douglas preferences* (HCD) analyzed by Bohr, Mestieri, and Robert-Nicoud (2024), which is a limiting case of Non-Homothetic CES preferences for $\sigma \to 1$. Solving equation (57) for v_n^t and taking the limit $\sigma \to 1$ yield:

$$\begin{aligned} v_n^t &= \lim_{\sigma \to 1} \left[\Omega \left(\frac{r_n}{p} \right)^{1-\sigma} \frac{1 - s_n^t}{s_n^t} \right]^{1/(\psi-\mu)} \\ &= \left(\Omega \frac{1 - s_n^t}{s_n^t} \right)^{1/(\psi-\mu)} . \end{aligned}$$

In this case, the expression for changes in welfare inequality becomes:

$$\frac{\hat{\mathbb{V}}^t}{\hat{\mathbb{V}}^s} = \left(\frac{\widehat{S_n^t}}{\widehat{S_n^s}} \right)^{-1/(\psi-\mu)} \left(\frac{\hat{\lambda}_n^t}{\hat{\lambda}_n^s} \right)^{-\frac{1}{\kappa}}, \qquad S_n^t = \frac{s_n^t}{1 - s_n^t}. \tag{61}$$

B.7 Proof of Proposition 7

Recall $V = \tilde{V}^\kappa$ by equations (1) and (1). Here we explicitly work with $\tilde{V}$ to study welfare changes in a transparent way. Here we allow for amenities to be type-specific and to enter utility in a flexible way (that is, we do not impose Assumption 2); in the text, we work with $\tilde{V}_n^t = A_n^t v\left(w_n^t, r_n\right)$. By definition of $\mathbb{V}^t \equiv \left[\sum_{m\in\mathcal{C}} \tilde{V}\left(w_m^t, r_m, A_m^t\right)^\kappa\right]^{1/\kappa}$ and $\mathbb{V}'^t \equiv \left[\sum_{m\in\mathcal{C}} \tilde{V}\left(w_m'^t, r_m', A_m'^t\right)^\kappa\right]^{1/\kappa}$, and making use of the hat notation, we have:

$$\begin{aligned} \ln \hat{\mathbb{V}}^t &\equiv \frac{1}{\kappa} \ln \frac{\sum_{m\in\mathcal{C}} \tilde{V}\left(w_m'^t, r_m', A_m'^t\right)^\kappa}{\sum_{m\in\mathcal{C}} \tilde{V}\left(w_m^t, r_m, A_m^t\right)^\kappa} \\ &= \frac{1}{\kappa} \ln \left\{ \sum_{m\in\mathcal{C}} \lambda_m^t \left[\frac{\tilde{V}\left(w_m^t \mathrm{e}^{\ln \hat{y}^t}, r_m, A_m^t\right)}{\tilde{V}\left(w_m^t, r_m, A_m^t\right)} \right]^\kappa \right\} \\ &\approx \frac{1}{\kappa} \ln(1) + \frac{1}{\kappa} \left\{ \sum_{m\in\mathcal{C}} \lambda_m^t \left[\frac{\tilde{V}\left(w_m^t \times 1, r_m, A_m^t\right)}{\tilde{V}\left(w_m^t, r_m, A_m^t\right)} \right]^\kappa \right\}^{-1} \\ &\times \sum_{m\in C} \lambda_m^t \kappa \left[\frac{\tilde{V}\left(w_m^t \times 1, r_m, A_m^t\right)}{\tilde{V}\left(w_m^t, r_m, A_m^t\right)} \right]^{-1+\kappa} \frac{w_m^t}{\tilde{V}\left(w_m^t, r_m, A_m^t\right)} \times \left. \frac{\partial \tilde{V}\left(w_m^t \mathrm{e}^{\ln \hat{y}^t}, r_m, A_m^t\right)}{\partial\left(w_m^t \mathrm{e}^{\ln \hat{y}^t}\right)} \right|_{\ln \hat{y}^t = 0} \times \ln \hat{y}^t \\ &= \ln \hat{y}^t \sum_{m\in\mathcal{C}} \lambda_m^t \mathcal{E}_m^t \left(\tilde{V}, w\right) \\ &= \bar{\varepsilon}^t \ln \hat{y}^t, \end{aligned}$$

where the second line uses equation (27) to substitute $\tilde{V}\left(w_m^t \hat{y}^t, r_m, A_m^t\right)^\kappa$ for $\tilde{V}\left(w_m'^t, r_m', A_m'^t\right)^\kappa$ and uses equation (1), the following one is a Taylor expansion around $\ln \hat{y}^t = 0$ (i.e., $\hat{y}^t = 1$). We then use $\hat{y}^t = 1$ to simplify terms and the elasticity notation $\mathcal{E}$.

B.8 Implications of Assumption 3

Recall Assumption 3:

$$\frac{\partial \mathcal{E}_n^t(v,w)}{\partial w} \leq 0.$$

Computing the derivative explicitly, we get: $\frac{\partial \mathcal{E}_n^t(v,w)}{\partial w} = \frac{v_w}{v}\left(1 - \frac{wv_w}{v} + \frac{wv_{ww}}{v_w}\right)$ where we use the short notation $f_x \equiv \frac{\partial f}{\partial x}$. Observe that the term outside the parenthesis in the final RHS is positive, thus

$$\frac{\partial \mathcal{E}_n^t(v,w)}{\partial w} \propto 1 - \frac{wv_w}{v} + \frac{wv_{ww}}{v_w},$$

where "$x \propto y$" is used to mean "x is of the same sign as y." Thus, Assumption 3 holds if and only if

$$1 \leq \frac{wV_w}{V} - \frac{wV_{ww}}{V_w}. \tag{62}$$

In turn, $\frac{wv_{ww}}{v_w}$ is negative whenever v is concave, which we assume. Thus a sufficient condition for Assumption 3 is:

$$-\frac{wv_{ww}}{v_w} \geq 1,$$

where $-\frac{wv_{ww}}{v_w}$ is the Arrow-Pratt measure of relative risk aversion associated with v.

C Additional Results for Section 5

Table OA.1: Summary Statistics for Expenditure Housing Shares

	Non-College						College					
	1980	1990	2000	2010	2020	Δ	1980	1990	2000	2010	2020	Δ
Median	0.18	0.20	0.19	0.24	0.24	0.05	0.17	0.18	0.17	0.19	0.20	0.02
Min	0.13	0.15	0.13	0.18	0.16	0.03	0.13	0.12	0.09	0.12	0.11	-0.02
Max	0.24	0.25	0.25	0.28	0.31	0.06	0.23	0.23	0.22	0.26	0.24	0.01
p75-p25	0.03	0.03	0.03	0.03	0.03	0.01	0.02	0.02	0.02	0.03	0.03	0.01

Note: Housing shares represent the proportion of expenditure allocated to housing. Minimum and maximum values are taken across Metropolitan Statistical Areas (MSAs). Specifically, we first calculate the median housing share within each MSA, then report the lowest (Min) and highest (Max) of these MSA-level medians. The p75-p25 range represents the interquartile range of these MSA-level medians.