Place-Based Redistribution*

Cecile Gaubert

Patrick Kline

Damian Vergara

Danny Yagan

February 2025

Abstract

We study optimal income taxation in a spatial equilibrium model with heterogeneous locational preferences, labor supply decisions, and competitive housing and labor markets. Expressions characterizing the optimal tax schedule in each community are provided that capture the fiscal externalities associated with migration and the effects of redistribution between households and landlords. Correlation between skill and locational preferences yields optimal transfers to poor areas, while sorting based on comparative advantage can motivate transfers in either direction. A calibration to areas targeted by the U.S. Empowerment Zone program yields sizable optimal spatial transfers that are sensitive to assumed levels of migration responsiveness.

Keywords: Place-based policies, equity-efficiency tradeoff, taxes.

*Gaubert, Kline, Yagan: UC-Berkeley and NBER. Vergara: University of Michigan. We thank David Atkin, Alan Auerbach, Roger Gordon, Nathan Hendren, Louis Kaplow, Emmanuel Saez, Ben Sprung-Keyser, Nancy Stokey, Juan Carlos Suarez-Serrato, Dmitry Taubinsky, and Matthew Weinzierl for helpful discussions and Akcan Balkir, Jimmy Chin, Katie Donnelly Moran, Clancy Green, and Dustin Swonder for outstanding research assistance. We thank Patrick Kennedy for a crosswalk from Census tracts to Empowerment Zones. Thanks also go to seminar participants at Yale, Stanford, MIT, Columbia, TSE, Dartmouth, Brown, Wharton, the Harris School, the Paris Trade Seminar, Oslo, the Women in Macro Conference, the Paris-London Public Economics Conference, the NBER Cities, Labor Markets, and the Global Economy Conference, the 2020 UEA meetings, and the NBER Summer Institute. This research was funded by the Berkeley Opportunity Lab and the Smith Richardson Foundation. Cecile Gaubert acknowledges support from NSF CAREER grant #1941917.

1 Introduction

Place-based policies tie economic benefits to geographic locations and are prevalent throughout the world (Glaeser and Gottlieb, 2008; Kline and Moretti, 2014b; Ehrlich and Overman, 2020). The espoused rationale for such programs is often redistributive: because poor households are concentrated in certain places, targeting these areas helps the disadvantaged. However, national governments already redistribute to such households through progressive income taxes. Should poor residents of poor places receive an extra transfer based on their location?

Glaeser (2008) articulates the traditional answer of economists that have studied these programs:

> "Help poor people, not poor places"...is something of a mantra for many urban and regional economists... [Place-based] aid is inefficient because it increases economic activity in less productive places and decreases economic activity in more productive places.

In line with this efficiency-based view, most academic research on place-based policies has focused either on their efficiency costs (e.g., Glaeser and Gottlieb, 2008; Albouy, 2009; Austin et al., 2018; Fajgelbaum et al., 2018; Gaubert, 2018) or the potential for such programs to correct market failures by internalizing productivity spillovers or other local externalities (e.g., Kline, 2010; Kline and Moretti, 2014a; Austin et al., 2018; Fu and Gregory, 2019; Rossi-Hansberg et al., 2019; Fajgelbaum and Gaubert, 2020). To date, however, little effort has been devoted to formalizing the redistributive goals that often motivate these policies in the first place.

In this paper, we study conditions under which place-based redistribution schemes are able to improve on the equity-efficiency tradeoffs posed by "place-blind" transfers implemented through income taxation. Our approach applies mathematical tools from the public economics literature on optimal labor income taxation to a spatial equilibrium model in the modern urban economics tradition. Residential sorting is modeled via a standard discrete choice formulation of household location decisions (McFadden, 1978; Bayer et al., 2007; Busso et al., 2013; Kline and Moretti, 2014b; Ahlfeldt et al., 2015) augmented to accommodate income effects. There are two locations: *Distressed* and *Elsewhere*, the latter of which may have better amenities, greater labor productivity (i.e., higher wages), and a higher cost of living. Households differ in their skill levels and in their relative tastes for the Distressed location. Each location has a competitive labor market that maps skill levels to wages. These mappings may differ from each other, generating spatial comparative advantage. Each household chooses where to live and how much to earn given the tax system, local costs of living, and local amenities. Housing is supplied by landlords and the cost of living in each community adjusts to clear housing markets.

A utilitarian planner chooses income tax schedules in the two communities to maximize a weighted average of household utilities and landlord profits. As in classic optimal tax problems (e.g., Mirrlees, 1971), we assume no market failures are present but household types are private information, with the planner observing only household earnings levels and location choices. Our analysis therefore differs fundamentally from studies considering optimal spatial transfers when household types are directly observed by the planner (e.g., Albouy, 2012; Fajgelbaum and Gaubert, 2020).

To develop intuition, we begin by studying the optimal choice of a lump-sum transfer to residents of Distressed financed by a head tax on residents of Elsewhere under a fixed (potentially sub-optimal) income tax system. At an optimum, the equity gains of the transfer to residents of Distressed equal their corresponding efficiency costs. The optimal transfer grows large when less-skilled households are concentrated in Distressed, when few households are indifferent between the two locations, or when productivity and rent differences across areas are small. The formula highlights the earnings effects associated with migration responses as a "sufficient statistic" in the sense described by Chetty (2009) and Kleven (2021) for the efficiency costs of place-based policies, providing guidance for future empirical research on place-based transfers.

An important question left unanswered by this formula is whether spatial subsidies can improve on an optimal income tax system. The answer to this question turns out to depend crucially on the forces generating sorting of less skilled households to Distressed. When sorting is driven by a propensity for higher skilled households to exhibit stronger tastes for residence in Elsewhere ("skill-taste correlation") then place-based transfers will tend to be welfare improving even when taxes are set optimally. Spatial transfers can also be welfare improving when sorting is driven by locational productivity differences generating spatial comparative advantage. However, when locational preferences are homogeneous and sorting is driven entirely by income effects, the planner may find it optimal to abstain from spatial transfers, relying only on income taxes for redistribution.

This last finding mirrors classic results in public economics establishing conditions under which redistribution via differential taxation of commodities can improve welfare over and above redistribution via optimal nonlinear income taxes and transfers (Atkinson and Stiglitz, 1976; Saez, 2002; Ferey et al., 2024). The prototypical result is that differential commodity taxation will tend to be superfluous whenever heterogeneity in consumption bundles across earnings groups is entirely attributable to income effects. As our modeling framework formalizes, place-based taxation differs from traditional commodity taxation problems in several important respects: locational choices are discrete (non-differentiable) decisions, they can directly affect the wage faced by households due to spatial productivity differences, and there are good reasons to expect poor households to locate in poor places for reasons besides that they are poor.

Relaxing the assumption that place-based transfers must be lump-sum, we develop general results characterizing optimal place-specific income tax schedules with unrestricted marginal tax rate (MTR) schedules. In effect, place-based redistribution can now vary by income. The differential equations characterizing these optimal MTRs resemble classic results in the optimal tax literature describing standalone economies (Mirrlees, 1971; Diamond, 1998; Saez, 2001). However, new terms emerge capturing considerations novel to our framework. One is that raising MTRs in a location yields mobility responses that incur a fiscal externality (i.e., a change in income tax revenue) due both to cost of living and productivity differences across locations. Another is that MTR changes generate equilibrium rent adjustments in both communities, which yield redistribution between households and landlords.

To study the quantitative implications of our planning framework, we solve calibrated versions of the model numerically, yielding optimal nonlinear tax schedules in each location. Complementing our theoretical analysis of a place-based head tax, we find that when sorting is generated entirely by skill-taste correlation, Distressed receives not only a larger demogrant (i.e., transfer to zero earners) but also substantially lower MTRs than Elsewhere. This finding accords with standard intuition from the literature on commodity taxation, where it has long been understood that taxing goods that high-ability households differentially prefer can improve on income taxation (Mirrlees, 1976). While that literature constrains the planner to consider linear commodity taxes, the optimal spatial transfers turn out to be decidedly nonlinear in our setting.

In simulations where sorting is driven by comparative advantage, the sign of the optimal transfer to Distressed residents is found to hinge on the magnitude of migration elasticities as governed by the dispersion in idiosyncratic locational preferences. When household migration elasticities are low, the optimal tax system subsidizes Distressed residents and yields lower MTRs in Distressed than Elsewhere at all but the highest income levels. When migration elasticities are large, the optimal tax system yields transfers to Elsewhere because fiscal externalities dominate. For intermediate household migration elasticities, it is optimal to redistribute towards Distressed at low earnings levels and towards Elsewhere at high earnings levels, highlighting the limitations of lump-sum place-based transfers. These findings complement recent work on optimal sectoral taxation that features analogous productivity differences attributable to comparative advantage but lacks cost of living differences (Scheuer, 2014; Gomes et al., 2018), variation in pre-tax incomes within location (Ales and Sleet, 2022), or subnational variation in the tax schedule (Rothschild and Scheuer, 2013).

Finally, in simulations where sorting is driven by income effects, small demogrants to Elsewhere are optimal accompanied by elevated MTRs in Distressed. However, as the variance of idiosyncratic locational preferences diminishes, the optimal spatial transfer approaches zero. This finding accords with theoretical results establishing that idiosyncratic preference heterogeneity violates the necessary conditions for commodity taxation to be superfluous (Kaplow, 2008).

We conclude our analysis with a detailed quantitative calibration, investigating the potential magnitude and direction of optimal place-based transfers in two scenarios, both anchored by Census data on the residents of the distressed areas targeted by the U.S. Empowerment Zone (EZ) program. A first calibration attempts to capture the tradeoffs involved in the design of urban EZs, which target particular neighborhoods in large cities. A second calibration mimics the design tradeoffs of rural EZs, which target collections of rural counties. The urban scenario assumes high migration and no comparative advantage, while the rural scenario assumes low migration and substantial comparative advantage. In the urban EZ calibration, modest demogrants to targeted areas are found to be optimal, along with reductions in MTRs at lower earning levels. The net transfer turns out to be close to the magnitude of actual transfers provided by the EZ program's wage tax credits. For rural EZs, the optimal tax system involves smaller demogrants but larger reductions in MTRs, leading to sizable tax advantages for the typical zone resident.

Sensitivity analysis reveals that, in both cases, raising migration elasticities lowers the size of optimal transfers and can even reverse their sign. When sorting into rural EZs is driven by comparative advantage rather than income effects, optimal transfers grow much larger. Both these findings highlight the potential value of additional empirical research into the forces driving sorting behavior. Interestingly, placing less weight on landlord profits yields less generous subsidies to urban EZs but more generous subsidies to rural EZs. This asymmetry is attributable to the larger housing supply elasticities estimated to be present in rural EZs, illustrating the complex interplay between housing and labor markets in our analysis.

While calibration exercises are inevitably speculative, our numerical results strongly corroborate the theoretical message of our paper that place-based redistribution can serve as a useful complement to place-blind taxation when income groups are geographically segregated. That is, sorting ultimately eases the efficiency costs of redistributing across household earnings levels. The urban economist's mantra appears to warrant revision: there is good reason to consider helping poor people *and* poor places.

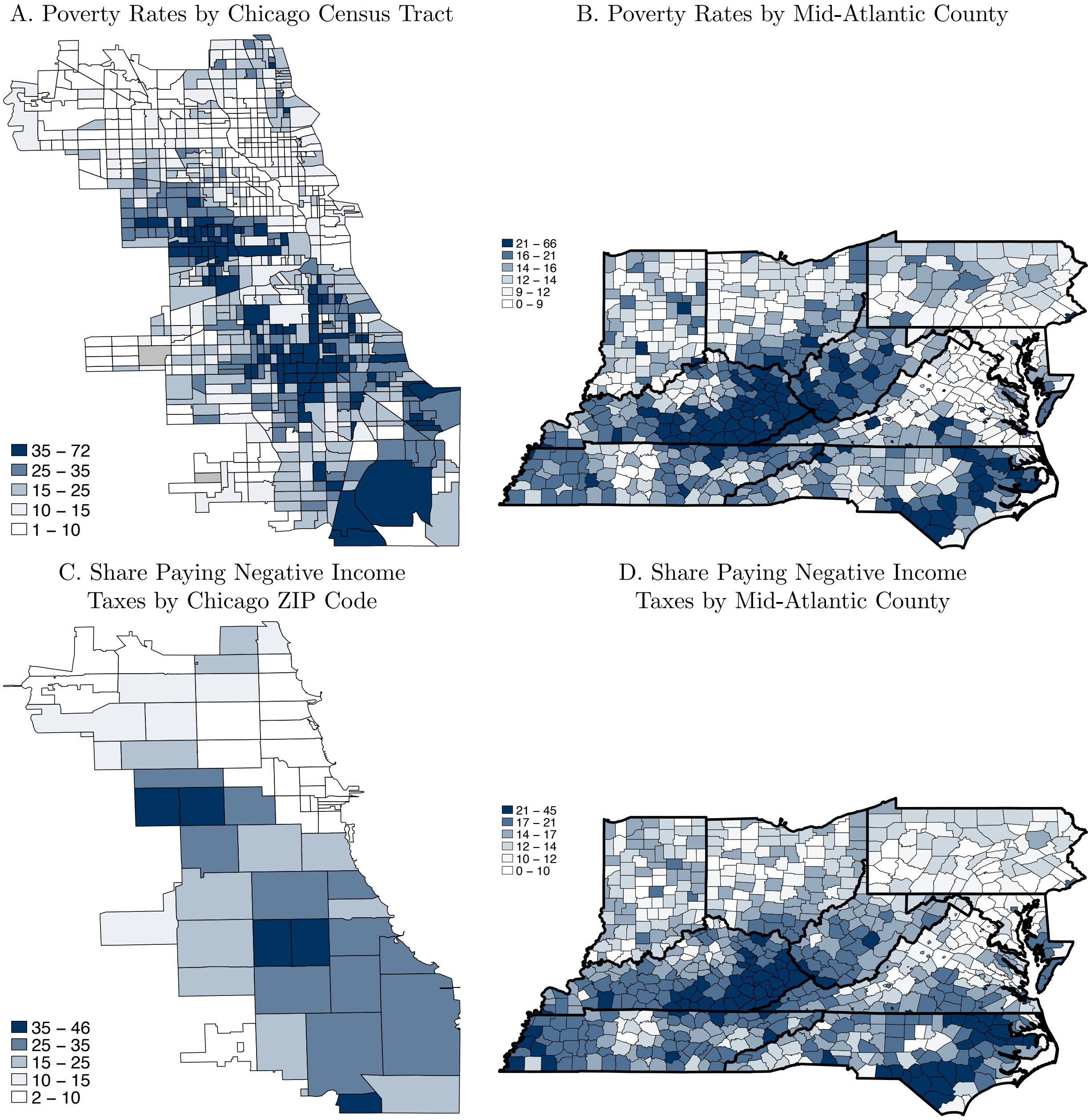
2 Motivating facts

We begin by briefly highlighting the uneven spatial distribution of economic outcomes in the United States. The key motivation underlying place-based redistribution is that disadvantage is spatially concentrated across states, counties, and Census tracts (Jargowsky, 1997; Reardon and Bischoff, 2011; Reardon et al., 2018; Gaubert et al., 2021). Poverty is especially concentrated across Census tracts, which are spatially contiguous land areas with typically between 1,200 and 8,000 people. A number of prominent place-based policies have been defined in terms of Census tracts including urban Empowerment Zones (EZs), and more recently, Opportunity Zones. We use EZs to anchor our quantitative exercises below. In pooled estimates from the 2016-2020 waves of the American Community Survey, Census tracts in the top centile of poverty rates have an average poverty rate of 64%, despite a national poverty rate over this period of only 13%.

Panel A of Figure 1 illustrates the spatial concentration of poverty in Chicago, Illinois – America's third largest city. Darker areas indicate higher poverty tracts in the 2016-2020 ACS. Tracts on the West Side and on the South Side have poverty rates exceeding 50%, while tracts in and around the Gold Coast neighborhood in the northeast of the city have near-zero poverty rates. Place-based redistribution from the Gold Coast to the West and South Sides has the potential to yield equity gains. In fact, Chicago's EZ comprises a contiguous section of the West Side and a separate contiguous section of the South Side.

Similar patterns exist at the regional level. For example, rural EZs are primarily collections of counties. Panel B of Figure 1 shows county-level poverty rates in Mid-Atlantic states. Poverty is heavily clustered in the rural Appalachian Mountain area of Eastern Kentucky, with rates comparable to those found in Chicago's West and South Sides. Place-based redistribution to residents of these heavily impoverished Eastern Kentucky counties has the potential to yield equity gains. Three of these counties compose the Kentucky Highlands EZ. Notably, while Chicago's EZ is very disproportionately Black or Hispanic, the Kentucky Highlands EZ is over 95% white (Census, 2024).

Figure 1: Poverty and Income-Based Transfers Are Spatially Concentrated



Notes: Panels A and B use the 2016-2020 American Community Survey to plot the share of households below the poverty line, for each Census tract in the city of Chicago and for each county in the Mid-Atlantic region. Panels C and D use 2020 Internal Revenue Service Statistics of Income ZIP-level aggregates to plot the share of tax filers receiving a net transfer from the federal income tax due to the refundable Earned Income Tax Credit, for each Chicago ZIP code and for Mid-Atlantic county.

Of course, place-blind transfers based on household income also redistribute to residents of poor areas. Panel C of Figure 1 uses a ZIP-code map of Chicago to plot the share of tax filing units paying negative federal income taxes due to the Earned Income Tax Credit. Half of tax filers in parts of the West and South Sides have negative income tax bills. Similar rates of negative income tax filing are present in the

Appalachian region, as shown in Panel D of Figure 1.

The U.S. tax system also makes transfers within income levels based on differences in family structure. Online Appendix A investigates the spatial impact of conditioning taxes and transfers on marital and parental status. We find that these two factors essentially cancel out, implying that tax incentives for particular family arrangements do not yield substantial *de facto* place-based transfers within income levels.

Is it desirable to redistribute to residents of poor communities using place-based transfers, or should governments rely solely on place-blind transfers? The analysis below formulates rigorous answers to this question. The analysis reveals that the desirability of optimal place-based transfers depends not just on the extent to which poverty is spatially concentrated but also on the nature of the economic forces driving concentration.

3 Model

In this section, we lay the foundation of a spatial model in which heterogeneous households optimize over both location choice and labor supply decisions. To connect with traditional results on income taxation, place-based redistribution schemes are modeled as transferring income directly to households. In practice, place-based policies often channel spatial transfers through capital or wage subsidies to businesses (Slattery and Zidar, 2020) or through federal transfers to local governments (Oates, 1999). To the extent that the incidence of such policies falls on households, the guiding principles derived here will hold.

An additional goal of business tax incentives can be to correct market imperfections, for example by exploiting agglomeration economies or endogenous amenities (Kline and Moretti, 2014a; Diamond, 2016; Bartik, 2020; Fajgelbaum and Gaubert, 2020); likewise, grants to local governments can be designed to correct fiscal externalities (Flatters et al., 1974; Albouy, 2012) or to compensate for under-investment in children (Chyn and Daruich, 2023). We leave to future work the task of integrating such Pigouvian corrections into a common theoretical framework with the redistributive motives studied here. Similarly, we consider only a central government and leave interactions with and incidence on subnational governments and local public services (e.g., Oates 1972; Gordon 1983; Kleven et al. 2020; Gordon 2023), such as via local property tax revenues, to future work.

3.1 Preliminaries

A unit mass of heterogeneous households chooses to live in one of two communities: 1 ("Distressed") or 0 ("Elsewhere"). Households are characterized by a two-dimensional type $\Theta = (\theta, \phi)$, where θ indexes the household's skill, while ϕ is an idiosyncratic preference for (or negative cost to) living in Distressed over Elsewhere. These types are distributed according to a continuous two-dimensional cumulative distribution function (CDF) $H : [\underline{\theta}, \bar{\theta}] \times [\underline{\phi}, \bar{\phi}] \to [0,1]$, which we factor into a marginal CDF F of θ and a conditional CDF G_θ of ϕ given θ, with corresponding densities f and g_θ. To economize on integral notation, it will be convenient to make use of the expectations operator $\mathbb{E}_\Theta [x] : x(\Theta) \mapsto \int x(\Theta)\, dH(\Theta)$ when describing average household behavior. We use $\mathbb{E}_\theta [x] : x(\Theta) \mapsto \int x(\Theta)\, dF(\theta)$ to denote averages integrated over only skills.

The two locations—a word we will use interchangeably with communities—may differ in four ways. First, since we are interested in studying optimal taxes indexed by earnings *and* place, each location $j \in \{0, 1\}$ may have its own income tax schedule $T_j : \mathbb{R}_{\geq 0} \to \mathbb{R}$. Communities can also differ in their level of amenities a_j, their rental cost of housing r_j, and their productivity, as reflected in local wage rates. Specifically, each location has a wage schedule $w_j : \mathbb{R} \to \mathbb{R}_+$ that is an increasing function of household skill. A household of skill type θ that resides in location j must supply $\ell = \frac{z}{w_j(\theta)}$ units of labor to generate pre-tax earnings z. This formulation allows for some skill types to possess a productive comparative advantage in a given location.

3.2 Household preferences

Households exhibit common preferences over the consumption of a homogeneous traded good c, the amenity level a of their community of residence, and their labor supply ℓ, represented by the following subutility function

$$(c, a, \ell) \mapsto U(c, a) - \psi(\ell).$$

Preferences over consumption and amenities are captured by the function $U(c, a)$, with derivatives $U_c > 0$, $U_a > 0$, and $U_{cc} \leq 0$, while the disutility of hours worked is captured by a convex function $\psi(\ell)$.

The price of the traded good is common in both communities and taken as the numeraire. Additionally, households inelastically demand a single unit of housing. Given these constraints, the indirect subutility for skill θ in location j can be written as

$$v_j(\theta) = \max_z U(z - T_j(z) - r_j, a_j) - \psi\left(\frac{z}{w_j(\theta)}\right).$$

We assume the relative preference ϕ for living in Distressed over Elsewhere impacts utility additively. The total indirect utility of a household of skill type θ residing in location j can therefore be written as

$$V_j(\Theta) = v_j(\theta) + j\phi. \tag{1}$$

Households choose whichever community offers higher indirect utility. The choice of location made by a household of type Θ will be denoted by

$$j^*(\Theta) = \arg\max\{V_0(\Theta), V_1(\Theta)\}. \tag{2}$$

Likewise, for every variable $x \in \{c, z, v, V\}$ we use $x_j^*(\Theta)$ to denote the value of x that would be chosen (or attained) by a household of type Θ if forced to reside in location j, while $x^*(\Theta) = j^*(\Theta)\, x_1^*(\Theta) + [1 - j^*(\Theta)]\, x_0^*(\Theta)$ gives the value of x actually realized by such a household.

3.3 Landlords

Housing is provided competitively by atomistic landlords. Each location has a housing supply elasticity ϱ_j yielding a supply equation $H_j(r_j) = \underline{H}_j r_j^{\varrho_j}$, where $\underline{H}_j \in \mathbb{R}_{\geq 0}$ is a community specific intercept. Define $L_j = \mathbb{E}_\Theta[j \cdot j^* + (1-j)\cdot(1-j^*)]$ as the share of households choosing to locate in community j. Each household demands a single unit of housing. The housing market clears, ensuring $L_j = H_j(r_j)$, and generates landlord profits $\Pi_j(r_j) = \frac{1}{1+\varrho_j} r_j H_j$. As shown in Appendix B.1.1, these expressions for $H_j(r_j)$ and $\Pi_j(r_j)$ constitute standard solutions to the landlords' problem under isoelastic costs of providing housing.

4 Spatial inequality

Two forces can generate locational income differences in the model. First, locations can exert a causal impact on earnings, in which case incomes can vary conditional on skill. Second, the skill composition of Distressed may differ from Elsewhere because of spatial sorting.

4.1 Location effects on within-skill earnings

Whenever locational choice exerts a causal effect on household earnings, there exists a fiscal cost to place-based redistribution. Conditional on locating in community j, the worker's optimal choice of earnings $z_j^*(\Theta)$ solves the first-order condition

$$\left[1 - T_j'\left(z_j^*(\theta)\right)\right] w_j(\theta) U_c\left(z_j^*(\theta) - T_j\left(z_j^*(\theta)\right) - r_j, a_j\right) = \psi'\left(\frac{z_j^*(\theta)}{w_j(\theta)}\right),$$

where we have used the fact that $z_j(\Theta) = z_j(\theta)$ as a result of the additive separability of ϕ in household utility. Optimal earnings equate the marginal disutility of earning an additional pre-tax dollar to its after-tax consumption value.

A type-θ household's optimal earnings in Distressed $z_1^*(\theta)$ can differ from its optimal earnings level in Elsewhere $z_0^*(\theta)$ for four reasons. First, the household's wage $w_j(\theta)$ might vary across locations. Second, the household's marginal utility of leisure is potentially shaped by the quality of local amenities (a_j), which can differ across locations. Pleasant amenities may encourage more leisure time; however, locations with poor amenities might also depress hours of work (e.g., if they lead to poor health). Third, if labor supply choices exhibit income effects, higher cost-of-living locations will induce higher earnings, all else equal. Finally, differences in location-specific tax systems $T_j(z)$ can alter incentives to work.

The empirical literature has established that earnings do tend to adjust when workers move between cities (Glaeser and Mare, 2001; Baum-Snow and Pavan, 2011; Dauth et al., 2022; Card et al., Forthcoming). A typical finding is that earnings fall when moving to smaller, less dense, metropolitan areas, which suggests employing assumptions that generate $z_1^*(\theta) \leq z_0^*(\theta)$. Less evidence is available, however, on whether earnings adjust when households move between neighborhoods in the same city. Since commuting behavior allows access to a common set of jobs from several different addresses within a city, it is reasonable to expect moves between residential neighborhoods to yield smaller earnings adjustments than moves between regions.

4.2 Spatial sorting

Another key stylized fact established by the empirical literature is that skill groups sort spatially (Diamond and Gaubert, 2022). The model encompasses three potential drivers of spatial sorting highlighted in the Economic Geography literature: 1) households with different skills may have different preferences for locations (skill-taste correlation), 2) their skills may be differentially productive in different locations (comparative advantage), and 3) the cost of living may drive poorer, lower skill households out of some locations (income-based sorting). It is useful to review the mechanics of these drivers of sorting, as they will have different policy implications.

Equations (1) and (2) imply that locational choice is governed by a simple threshold rule. Let $\phi_\theta = v_0^*(\theta) - v_1^*(\theta)$ denote the value of ϕ that would make a type θ household indifferent between the two locations. Households with idiosyncratic preference $\phi \leq \phi_\theta$ will locate in Elsewhere, while the rest will locate in Distressed. Hence, a share $G_\theta(\phi_\theta)$ of type θ households will locate in Elsewhere. When this share varies with θ, spatial sorting ensues. We introduce below three polar cases where such sorting arises from a single motive. These examples temporarily assume away taxes. We return to these polar cases when analyzing optimal place-based taxes to illustrate that not only the existence of sorting, but the *motive* for sorting matters for optimal taxation.

Example 1 (Skill-Taste Correlation)**.** Suppose that wages and housing costs are identical in the two locations (i.e., $w_0(\theta) = w_1(\theta)$ for all θ and $r_0 = r_1$) and that amenities enter utility $U(c, a)$ additively. Optimal earnings for any household are then equal between communities: $z_0^*(\theta) = z_1^*(\theta)$. Consequently, the preference cutoff for choosing to reside in Distressed is simply $\phi_\theta = a_0 - a_1$, which does not depend on θ. Let the distribution G_θ be weakly increasing in θ, capturing that tastes ϕ are negatively dependent on skills. It follows that $G_\theta(a_0 - a_1)$ increases with θ; therefore high skilled households will sort into Elsewhere.

Example 2 (Comparative Advantage)**.** Suppose that cities are identical except for their wage schedule and that preferences do not vary with skills (i.e., $G_\theta = G$). Elsewhere is a more productive location for all skill levels but especially so for the highly skilled, who have a comparative advantage in production there. Formally, we assume $w_0(\theta) \geq w_1(\theta)$ and $\gamma_0(\theta) > \gamma_1(\theta)$ for all θ, where $\gamma_j(\theta) = \frac{d \log w_j(\theta)}{d \log \theta} > 0$ is the elasticity of wage rates with respect to skill in location j. We show in Appendix B.2.1 that, under these assumptions, the preference cutoff ϕ_θ for choosing Distressed increases with θ so long as U is not too concave in consumption.[1] Consequently, the highly skilled sort into Elsewhere.

Example 3 (Income-based Sorting)**.** Suppose that Elsewhere is more expensive than Distressed ($r_0 > r_1$) and U is concave in consumption. Furthermore, assume that locational tastes are independent of skills

[1] When U is too concave, income effects may induce non-monotone sorting as the labor supply of high-θ types became less sensitive to differences in wages if leisure is a normal good. We characterize the concavity threshold as a function of γ_0/γ_1 in the appendix. A stricter sufficient condition is given by $x \mapsto xU_c(x)$ being non-decreasing, which is met by the functional forms assumed in the simulation exercises below.

($G_\theta = G$) and that $w_0(\theta) = w_1(\theta)$ for all skill levels θ. We show in Appendix B.2.2 that the higher cost of living in Elsewhere leads ϕ_θ to increase with θ because higher-skilled households are less impacted by the cost of living. Consequently, the highly skilled sort into Elsewhere.

Following the discussion above, we will typically think of Distressed as a location with weakly lower productivity, less skill complementarity, worse amenities, and lower housing prices than Elsewhere. This state of affairs can be characterized by the inequalities $w_0 \geq w_1$, $\gamma_0 \geq \gamma_1$, $a_0 \geq a_1$, $r_0 \geq r_1$, with G_θ weakly increasing in θ. However, the formulas we derive apply more generally to any configuration of locational fundamentals.

5 Planning problem

A planner evaluates allocations subject to a weighted utilitarian welfare criterion where households have exogenous Pareto weights $\omega_H(\Theta)$ while landlords have weight ω_L. Formally, the social welfare function is

$$SWF = \mathbb{E}_\Theta\left[\omega_H V^*\right] + \omega_L\left[\Pi_0(r_0) + \Pi_1(r_1)\right]. \tag{3}$$

Preferences for redistribution may arise either from concavity of utility in consumption or from heterogeneity in household weights ω_H. In what follows, we will be interested in cases where the households' weights depend only on skill type θ, as dependence on tastes can be difficult to motivate.[2] Weights obeying $\frac{d\omega_H(\theta)}{d\theta} < 0$ motivate redistribution from high-θ to low-θ households.

The planner evaluates allocations subject to the following budget constraint

$$\mathbb{E}_\Theta\left[j^* T_1(z_1^*) + (1 - j^*) T_0(z_0^*)\right] \geq R, \tag{4}$$

where R is an exogenous revenue requirement. Under standard regularity conditions, the budget constraint holds with equality. The marginal value of public funds is captured by the multiplier of the budget constraint, which we denote by Γ.

The social marginal welfare weight of type θ households who reside in community j is defined as $\lambda_j(\theta) = \omega_H(\theta)U_c(c_j, a_j)/\Gamma$. The assumption that ϕ enters household utility in an additively separable fashion implies that the welfare weights do not depend directly on locational preferences. Rather, the weights depend

[2] Though we do not pursue this idea, higher Pareto weights on households with relative tastes for Distressed could be justified as an attempt to capture a planner's desire to right perceived historical wrongs. For example, Gulf Opportunity Zones were instituted in 2005 for areas devastated by Hurricane Katrina.

indirectly on ϕ through the location choice of the household. The quantity $\lambda_L = \omega_L/\Gamma$ gives the social marginal welfare weight of landlords. These weights summarize the redistributive tastes of the planner. At the optimum, the planner is indifferent between giving one more dollar to a household θ living in j or having $\lambda_j(\theta)$ more dollars of public funds.

If household types could be observed, the planner would equalize social marginal welfare weights across households via type-specific lump-sum taxes and transfers. Because household types are not observable, the planner instead implements an incentive-compatible second-best tax system that trades off equity and efficiency. The next two sections explicitly characterize these equity-efficiency tradeoffs.

6 Optimal place-based transfers (lump-sum case)

We first study the optimal design of a simplified place-based redistribution scheme where, in addition to a place-blind income tax schedule $T : \mathbb{R}_{\geq 0} \to \mathbb{R}$, the planner can introduce a lump-sum place-based transfer indexed by $t \in \mathbb{R}$ from residents of Elsewhere to residents of Distressed. Despite its simplicity, this case is rich enough to highlight the key tradeoffs involved in place-based redistribution. Such a restricted place-based instrument may be practically relevant under political and administrative constraints. For example, Empowerment Zones provide a capped wage subsidy that yields a constant per-worker transfer when the covered worker is employed full-time.

To ease exposition, we make two simplifying assumptions in this section. First, we impose that optimal earnings exhibit no income effects by assuming quasi-linear consumption utility. In this case, redistributive motives derive exclusively from non-constant Pareto weights $\omega_H(\theta)$. Second, we temporarily treat rents as fixed, therefore abstracting in particular from the incidence of taxes on landlords. Our general derivations of Section 7 include these forces.

6.1 A simple place-based redistribution scheme

Consider a place-based redistribution scheme in which each of the households in Elsewhere face a head tax t/L_0, while each of the households in Distressed receive t/L_1. Hence, the place-based tax faced by a household of type Θ is

$$t \cdot \frac{L_1 - j^* (\Theta, t)}{L_0 L_1}, \tag{5}$$

where we have indexed the function j^* by t in order to highlight the potential influence of the place-based transfer scheme on location choices. The tax schedule is therefore given by $T_j(z) = T(z) + t \cdot \frac{L_1 - j}{L_0 L_1}$. In this restricted policy scheme, the planner's budget constraint reduces to

$$\mathbb{E}_\Theta \left[T\left(z^*\right) + t \cdot \frac{L_1 - j^*}{L_0 L_1} \right] = R.$$

The amount $\mathbb{E}_\Theta \left[T\left(z^*\right)\right]$ is the net fiscal revenue of the place-blind income tax, while $\mathbb{E}_\Theta \left[t \cdot \frac{L_1 - j^*}{L_0 L_1} \right]$ is the net fiscal revenue generated by the place-based transfer scheme. By construction, the place-based policy is ex-ante budget balanced since $\mathbb{E}_\Theta \left[t \cdot \frac{L_1 - j^*}{L_0 L_1} \right] = 0$.

6.2 Optimality conditions

Consider now the welfare effect of an infinitesimal place-dependent tax reform dt. There are no mechanical effects of this reform on fiscal revenue because the reform is ex-ante budget neutral. Therefore, the total welfare effect $dSWF/dt$ of the reform starting from $t = 0$ is the sum of only two components: a direct impact on welfare, denoted dW/dt, and a fiscal cost of the reform attributable to behavioral responses, denoted dB/dt.

Implementing the place-based transfer (5) generates a net transfer of utility from residents of Elsewhere to those of Distressed equal to

$$\frac{dW}{dt} = \mathbb{E}_\Theta \left[\lambda_1 \frac{j^*}{L_1} - \lambda_0 \frac{1 - j^*}{L_0} \right] = \bar{\lambda}_1 - \bar{\lambda}_0, \tag{6}$$

where $\bar{\lambda}_j$ is the average social marginal welfare weight of households located in community j. The expression in (6) is the *equity gain* from the transfer, which will be positive so long as the average social marginal welfare weight of Distressed inhabitants is higher than that of residents of Elsewhere. Households that migrate in response to the reform do not contribute to dW/dt because marginal movers are initially indifferent between locations. When the weights depend only on θ, a sufficient condition for $\bar{\lambda}_1 > \bar{\lambda}_0$ is that there is spatial sorting such that the skill distribution in Elsewhere first-order stochastically dominates that of Distressed (Atkinson, 1970).

Two forces help this condition arise. First, Distressed residents may earn sufficiently less than Elsewhere residents that consumption is lower at every consumption rank in Distressed than in Elsewhere. Second, whenever consumption and amenities are q-substitutes (i.e., $\partial^2 U/\partial c \partial a < 0$), the lower amenity level in

Distressed raises the marginal utility of consumption at any fixed level of consumption, thereby raising λ_1. If preferences are sufficiently strong q-substitutes, first-order stochastic dominance will ensue. In contrast, the Cobb-Douglas preferences $U = ac$ typically used in the economic geography literature impose q-complementarity (i.e., $\partial^2 U/\partial c \partial a > 0$), which will serve to reduce the equity gain of redistributing to Distressed.

The equity gain of a place-based transfer must be weighed against its corresponding efficiency loss. Although the tax reform is *ex-ante* budget neutral, some households may change their location as a response to the policy. As discussed above, these moves do not mechanically generate a first order fiscal externality when starting from a place-blind tax system, nor do they generate a first order welfare effect. However, workers who move may reduce their earnings, which will generate a first order fiscal cost.

To compute this fiscal externality, we define $m_\theta(t) = d(1 - G_\theta(\phi_\theta(t)))/dt = -g_\theta(\phi_\theta(t))\, d\phi_\theta(t)/dt > 0$ as the number of households of skill level θ induced to move from Elsewhere to Distressed by a marginal increase in the transfer. The overall fiscal cost of movers starting from a place blind system is

$$\frac{dB}{dt} = \mathbb{E}_\theta \left\{ m(0) \cdot \left[T(z_1^*) - T(z_0^*) \right] \right\}. \tag{7}$$

Note that movers to Distressed change earnings in a way that depends only on their skill θ, not on ϕ – that is, $z_j^*(\Theta) = z_j^*(\theta)$ – because idiosyncratic preferences ϕ are additively separable.

Equity-Efficiency Tradeoff The following result summarizes the equity-efficiency tradeoff of a small place-based transfer to Distressed starting from a place-blind tax system:

Lemma 1. *The first order effect on social welfare of a small per-capita transfer from Elsewhere to Distressed starting from a place-blind system is* $dSWF/dt = \bar{\lambda}_1 - \bar{\lambda}_0 + \mathbb{E}_\theta \left\{ m(0) \cdot \left[T(z_1^*) - T(z_0^*) \right] \right\}$.

The expression makes clear that, for a given degree of spatial inequality $\bar{\lambda}_1 - \bar{\lambda}_0$, the place-based transfer to Distressed is more likely to be desirable when mobility responses are small or are dominated by households for whom earnings differences across space are small. Large mobility responses among households for whom earnings differences across space are large can yield a desirable place-based transfer to Elsewhere.

Lemma 1 can additionally be used to characterize the magnitude of the optimal place-based transfer. Starting from an optimal t^*, a small place-based reform will have no first-order effect on welfare but it will have a first-order fiscal externality when starting at a $t^* \neq 0$. Movers from Elsewhere to Distressed generate a fiscal loss per capita of $t^*\left(\frac{1}{L_1}+\frac{1}{L_0}\right) = \frac{t^*}{L_0 L_1}$ as they go from being net contributors to becoming net beneficiaries of the place-based redistribution scheme. Equating $dSWF(t)/dt$ to zero yields the optimal place-based transfer t^*.

Proposition 1. *The optimal place-based transfer t^* under any place-blind tax schedule $T(\cdot)$ obeys*

$$t^* = \frac{\bar{\lambda}_1(t^*) - \bar{\lambda}_0(t^*) + \mathbb{E}_\theta\left\{m(t^*)\cdot\left[T(z_1^*) - T(z_0^*)\right]\right\}}{\mathbb{E}_\theta\left[m(t^*)\right]/\left[L_0(t^*)\,L_1(t^*)\right]}.$$

Proof. See Online Appendix B.3.1 □

Note that Proposition 1 holds regardless of whether the place-blind tax schedule $T(\cdot)$ is optimal. The size of the optimal transfer is increasing in the degree of inequality between Elsewhere and Distressed, which is captured by the difference $\bar{\lambda}_1(t^*) - \bar{\lambda}_0(t^*)$ in average social marginal welfare weights evaluated at the optimal value of the transfer. All else equal, the optimal transfer is larger if mobility is low or if the earnings responses to migration are small.

Appendix B.3.2 extends the analysis here to settings featuring non-additive locational preferences. A difficulty that arises in this extension is that welfare weights then depend on idiosyncratic tastes, leading to pitfalls that the appendix also discusses.

6.3 Examples revisited

We now revisit the polar examples of sorting discussed in Section 4.2. Lemma 1 is used to study whether place-based redistribution can be unambiguously desirable in these three polar cases even when place-blind income taxes are present and set optimally. The first two examples maintain that household preferences are quasi-linear and rely on the planner's Pareto weights $\omega_H(\theta)$ to generate an equity motive for place-based redistribution. The third example brings back curvature in the utility function to analyze income effects in location choice.

Example 4 (Skill-taste correlation, continued)**.** Returning to Example 1, assume that the Pareto weights $\omega_H(\theta)$ strictly decline with skill. Under skill-taste correlation, the high-skilled are more likely to locate in Elsewhere and hence the equity gain $\bar{\lambda}_1 - \bar{\lambda}_0$ to a small increase in t is positive. Moreover, there is no fiscal cost to a small increase in t starting from $t = 0$ because movers do not experience wage changes and therefore do not change their income tax payments. Consequently, the right hand side of the condition depicted in

Lemma 1 must be positive, implying that a small positive place-based head tax is unambiguously desirable, even in the presence of an optimal place-blind income tax.

Example 5 (Comparative advantage, continued)**.** Suppose that the high-skilled enjoy a discrete wage advantage in Elsewhere: $w_1(\theta) = \theta$ and $w_0(\theta) = \theta + b \cdot 1\{\theta \geq \theta^*\}$, for some $\theta^* > \underline{\theta}$ and $b > 0$. Households with skill levels higher than θ^* enjoy a comparative advantage in Elsewhere. Consider the simplified case where utility is quasi-linear, labor supply is isoelastic, Pareto weights $\omega_H(\theta)$ are decreasing in θ, and all households have the same value of ϕ. As detailed in Appendix B.3.4, these assumptions imply there exist values of ϕ such that all highly skilled ($\theta \geq \theta^*$) households strictly prefer locating in Elsewhere, while all less skilled ($\theta < \theta^*$) households strictly prefer Distressed. In this environment, a small $t > 0$ generates equity gains by redistributing from the highly skilled to the less skilled but incurs no fiscal cost because the transfer avoids triggering any migration. Thus, introducing a small t is welfare-improving, even if the place-blind tax system is optimal. As shown in the appendix, this conclusion also holds in the presence of taste heterogeneity if the higher skill types remain inframarginal but the low-skilled households move in response to the transfer. These moves have no fiscal consequences because individuals with $\theta < \theta^*$ do not change their earnings upon moving.

In the examples above, taxing location choices is unambiguously desirable. A related literature beginning with Atkinson and Stiglitz (1976) provides conditions under which commodity taxation is superfluous in the presence of an optimal nonlinear income tax. The key sufficient condition for commodity taxes to be superfluous is that if high earners were hypothetically forced to earn less, they would purchase the same consumption bundle as their lower earning peers (Saez, 2002; Ferey et al., 2024).[3] Example 4 violates this condition: skill-taste correlation directly implies the high-skilled would be more likely to reside in Elsewhere even if they earned less. Example 5 involves wage changes omitted in commodity tax models that also induce the high-skilled to locate in Elsewhere even if forced to earn less.

In our final example, we reintroduce income effects in order to study optimal place-based redistribution in a model with sorting driven solely by income effects in location choice. Atkinson and Stiglitz (1976) showed that commodity taxes are superfluous when consumption differences are driven entirely by income effects. This final example demonstrates that place-based transfers can also be superfluous when locational sorting

[3]While the results in these papers yield similar intuition, the theoretical results do not directly apply to our setting: both Saez (2002) and Ferey et al. (2024) consider optimal taxation of continuous consumption decisions, while the present discussion concerns taxation of a discrete choice of location.

is driven solely by income effects.

Example 6 (Income-based sorting, continued)**.** Returning to the setting in Example 3, suppose that locational preferences are homogenous (i.e., $\underline{\phi} = \bar{\phi} = 0$) and that household utility takes the form $U(z - T(z) - r_j, a_j) - \psi\left(\frac{z}{\theta}\right)$ with $U_{cc} < 0$ and $U_{ca} > 0$. Suppose also that Elsewhere has better amenities but higher rents than Distressed ($a_0 > a_1$ and $r_0 > r_1$). In this case, all households with skill above some threshold level will sort into Elsewhere to benefit from higher amenities, while all those will lower skills, who are more sensitive to the cost of living, will sort into Distressed. These preferences obey the key condition for superfluous commodity taxes: if high earners were hypothetically forced to work less such that they earn and consume as little as lower earners, they would purchase the same consumption bundle as lower earners. We show in Appendix B.3.5 that, indeed, a tax system with $t > 0$ cannot improve welfare relative to a properly designed place-blind tax system. Our proof relies on the absence of idiosyncratic locational preferences; we explore numerically below how locational preferences can generate non-zero optimal place-based transfers when income effects drive skill sorting.

7 Optimal place-based transfers (general case)

This section analyzes the general model of Section 3, acknowledging the potential for income effects in labor supply, incorporating equilibrium adjustments in housing rents, and lifting the restriction that place-based transfers take the form of lump-sum transfers.

7.1 Optimality conditions

The planner chooses location-specific tax schedules $\{T_0(\cdot), T_1(\cdot)\}$ to maximize the social welfare function (3) subject to the budget constraint (4). As is standard in the optimal taxation literature, we characterize the solution to this problem using optimal control methods. Denote the location-specific marginal utility of consumption by $U^*_{cj}(\theta) = U_c(z^*_j(\theta) - T_j(z^*_j(\theta)) - r_j, a_j)$. In what follows, we assume that the marginal rate of substitution between labor earnings and consumption $-\psi'(z^*_j(\theta)/w_j(\theta))/(w_j(\theta) \cdot U^*_{cj}(\theta))$ is decreasing in θ in each location, which ensures that the usual single-crossing property holds.[4] The nonlinear schedule of optimal taxes in the two locations is characterized by the marginal tax rate schedules $T'_0(\cdot)$ and $T'_1(\cdot)$ and

[4]This assumption allows us to rely on local incentive compatibility constraints to derive the optimal tax schedule as in, e.g., Mirrlees (1971) and Rothschild and Scheuer (2013). Note that multidimensional heterogeneity does not undermine standard screening arguments here because the tax schedules are location-specific and ϕ does not affect optimal earnings conditional on location.

the transfers to zero earners ("demogrants") $-T_0(0)$ and $-T_1(0)$.

Proposition 2. *The optimal place-specific marginal tax rates are characterized by the differential equations*

$$\frac{T_0'(z_0^*(\theta))}{1-T_0'(z_0^*(\theta))} = \frac{1+\eta_0^U(\theta)}{\eta_0^C(\theta)}\frac{\gamma_0(\theta)}{\theta f(\theta)G_\theta(\phi_\theta)} \times U_{c0}^*(\theta)\left\{\int_\theta^{\bar\theta}\frac{1-\lambda_0(s)}{U_{c0}^*(s)}G_s(\phi_s)\,dF(s)+\Delta\tau^+(\theta)+(\lambda_L-1)\,\Delta r^+(\theta)\right\}, \quad (8)$$

$$\frac{T_1'(z_1^*(\theta))}{1-T_1'(z_1^*(\theta))} = \frac{1+\eta_1^U(\theta)}{\eta_1^C(\theta)}\frac{\gamma_1(\theta)}{\theta f(\theta)(1-G_\theta(\phi_\theta))} \times U_{c1}^*(\theta)\left\{\int_\theta^{\bar\theta}\frac{1-\lambda_1(s)}{U_{c1}^*(s)}(1-G_s(\phi_s))\,dF(s)-\Delta\tau^+(\theta)-(\lambda_L-1)\,\Delta r^+(\theta)\right\}, \quad (9)$$

where

$$\Delta\tau^+(\theta) \equiv \int_\theta^{\bar\theta}\left[T_1(z_1^*(s))-T_0(z_0^*(s))\right]g_s(\phi_s)\,dF(s),$$

$$\Delta r^+(\theta) \equiv \left(\frac{r_1}{\varrho_1}-\frac{r_0}{\varrho_0}\right)\cdot\int_\theta^{\bar\theta}g_s(\phi_s)\,dF(s).$$

The difference in demogrants in the two locations is characterized by the restriction

$$L_0L_1\int_{\underline{\theta}}^{\bar\theta}\left(\frac{1}{L_1}(\lambda_1(\theta)-1)\frac{1}{U_{c1}^*(\theta)}(1-G_\theta(\phi_\theta))-\frac{1}{L_0}(\lambda_0(\theta)-1)\frac{1}{U_{c0}^*(\theta)}G_\theta(\phi_\theta)\right)dF(\theta) + (\lambda_L-1)\,\Delta r^+(\underline{\theta})+\Delta\tau^+(\underline{\theta}) = 0. \quad (10)$$

Proof. See Online Appendix B.4.1 □

The function $\eta_j^U(\theta)=\frac{d\log z_j^*(\theta)}{d\log 1-\tau_j}$ denotes the uncompensated elasticity of earnings with respect to the net of tax rate in community j for households of skill type θ, while $\eta_j^C(\theta)=\frac{\partial\log z_j^*(\theta)}{\partial\log 1-\tau_j}>0$ gives the corresponding compensated (i.e., Frisch) elasticity, where $\tau_j=T_j'\left(z_j^*(\theta)\right)$ is the place-specific MTR at the optimum. The term $\Delta\tau^+(\theta)$ measures the fiscal externality induced by movers from Elsewhere to Distressed for all skill levels $s\geq\theta$. At each skill level s, there are $g_s(\phi_s)f(s)$ households indifferent between the two communities who, if induced to move by a tax reform, will change their tax revenue by an amount $T_1(z_1^*(s))-T_0(z_0^*(s))$. Similarly, $\Delta r^+(\theta)$ measures the impact of movers from Elsewhere to Distressed with skill level higher than θ on net rent payments to landlords. Evaluating these expressions at $\theta=\underline{\theta}$ yields the total fiscal externality of movers and their total impact on net rent payments to landlords, respectively.

7.2 Demogrants

Equation (10) pins down the difference in demogrants conditional on the schedule of MTRs. At the optimal tax schedule, a small decrease in $T_1(0)$ financed by a small increase in $T_0(0)$ generates welfare effects captured by the first two terms of equation (10): the first term captures welfare changes between workers residing in different locations, while the second term captures welfare effects derived from redistribution between workers and landlords as rents adjust. These welfare effects generate fiscal cost $\Delta\tau^{+}(\underline{\theta})$. At the optimal place-specific tax schedules, the total equity gain of this small perturbation to the tax system must be offset by the total fiscal externality. Notice that when housing supply is infinitely elastic and income effects are absent from labor supply, equation (10) reduces to $\overline{\lambda}_1 - \overline{\lambda}_0 + \frac{\Delta\tau^{+}(\underline{\theta})}{L_0 L_1} = 0$, which is equivalent to the expression obtained in Proposition 1 after restricting MTRs to be place-blind.

Perturbing the tax system affects equilibrium rental rates, as movers change the price of housing for all infra-marginal households. Subsidizing Distressed leads households to move from Elsewhere to Distressed, which raises rents in Distressed and lowers them in Elsewhere. On net, this generates a transfer $\Delta r^{+}(\theta)$ from renters to landlords. The sign of $\Delta r^{+}(\theta)$ depends on the ratio of rents to housing supply elasticities $\left(\frac{r_1}{\varrho_1} - \frac{r_0}{\varrho_0}\right)$. Baum-Snow and Han (2024) find that rural high-poverty U.S. areas exhibit lower rents and higher housing supply elasticities than the rest of the country on average, which suggests $\Delta r^{+}(\theta) < 0$ for place-based transfers to distressed rural areas. Such transfers lead households to move out of congested areas, which causes rents to react downwards strongly in those areas while rents reacts mildly upwards in destination distressed rural areas. Assuming that landlords tend to have lower welfare weight than the average household ($\lambda_L < 1$), the rent term generates a motive for place-redistribution to poor rural areas.

In contrast, Baum-Snow and Han (2024) find that urban high-poverty U.S. neighborhoods exhibit lower rents but also lower housing supply elasticities than the rest of the country on average. In this case, the rent effect can turn negative because place-based transfers to distressed urban neighborhoods induce moves from the rest of the country to those relatively congested neighborhoods. The rent term can therefore generate a motive for place-based transfers away from distressed urban neighborhoods when $\lambda_L < 1$ and the rent differences between distressed urban areas and elsewhere are small enough relative to their housing supply elasticity differences for the housing supply elasticity effect to dominate.

7.3 Marginal tax rates

The optimal tax policy involves place dependent MTRs that exhibit a familiar structure. Note that if we drop the terms $\Delta\tau^{+}(\theta)$ and $(\lambda_L - 1)\,\Delta r^{+}(\theta)$ from equations (8) and (9) these expressions reduce to standard results providing the optimal MTR schedules for standalone economies (e.g., equation 25 of Saez, 2001).

Putting the new terms aside for a moment, a first reason why MTRs may differ across places is sorting: if skilled households are over-represented in Elsewhere then MTRs will be higher there, a phenomenon captured by the inverse hazard ratio terms in equations (8) and (9) familiar from standard optimal tax results (Diamond, 1998). A given labor supply distortion at middle incomes raises less revenue in Distressed where there are proportionally fewer high earners.[5] Second, the skill composition within location affects how the average welfare weights in a location compare to the average welfare weight in the economy. Third, there may be differences in welfare weights across locations conditional on type. When households of a given skill level have higher social marginal welfare weights in Distressed ($\lambda_1(\theta) > \lambda_0(\theta)$) then optimal MTRs in Elsewhere will be elevated.

The new term $\Delta\tau^{+}(\theta)$ reflects the fiscal externality generated by movers of skill θ and above who respond to changes in taxes across locations. If taxes in Elsewhere increase, households migrate to Distressed. When within-skill earnings are lower in Distressed, this term is negative. This force limits tax differences between locations, particularly when locational effects on earnings and migration responses are large. An analogous term features in Scheuer (2014)'s analysis of occupation-specific taxes, where occupational switches govern fiscal externalities when workers and entrepreneurs have different tax liabilities. Unlike in that paper, however, the fiscal externalities here can be mediated both by income effects stemming from cost of living differences across locations and by comparative advantage.

Finally, the term $(\lambda_L - 1)\,\Delta r^{+}(\theta)$ captures the net transfers from renters to landlords as a result of rent changes induced by taxation. When $\lambda_L < 1$, the planner may use place-specific taxes to redistribute from landlords to renters. This force tends to lower relative MTRs in Elsewhere whenever moving to Distressed generates transfers from renters to landlords on net ($\Delta r^{+}(\theta) > 0$).

[5] Kremer (2003) makes an analogous hazard-rate-based argument for lower MTRs on young earners relative to middle-age earners. The true inverse hazard rate that matters is the inverse hazard rate of the earnings distribution as emphasized by Saez (2001), which is why the formulas contain the additional ratio $\gamma_j(\theta)/\theta$.

8 Sorting motives and optimal taxes: numerical results

This section explores quantitatively the extent to which different sorting motives generate different optimal place-based redistribution schemes. To isolate the role of each motive on the structure of optimal taxes, we solve calibrated versions of the model numerically where a single sorting motive – skill-taste correlation, comparative advantage, or income-based sorting – drives income differences between communities, while varying migration responsiveness. Additional details on these simulations can be found in Appendix C. The subsequent section allows sorting to be driven by combinations of these forces.

In the numerical exercises that follow, household utility takes a form commonly found in the urban economics literature: a Cobb-Douglas aggregate of amenities, consumption (including unit housing), and an idiosyncratic location preference. Without loss of generality, common valuations of the unobserved amenities are folded into the location taste term ϕ. To allow for labor supply decisions, this canonical specification is augmented with an isoelastic disutility of labor supply. Formally, the household utility function can be written:

$$\ln\left(z - T_j\left(z\right) - r_j\right) - \frac{\eta}{1+\eta}\left(\frac{z}{W\theta^{\gamma_j}}\right)^{\frac{1+\eta}{\eta}} + j\phi. \tag{11}$$

Note that, with unit housing demand, low-income households are particularly sensitive to rent differentials, which generates income-based sorting.

Location preferences ϕ are assumed to be logistically distributed with mean $\mu - \beta F(\theta)$ and standard deviation κ. The parameter $\mu = a_1 - a_0$ captures the difference in amenities between the two locations. To accommodate the possibility of skill-taste correlation, the conditional distribution $G_\theta(\phi)$ of location preferences given skill type is assumed to take the form $G_\theta(\phi) = \left[1 + \exp\left(-\left(\phi - \mu + \beta F(\theta)\right)/\kappa\right)\right]^{-1}$. The parameter β governs the strength of skill-taste correlation, with $\beta = 0$ yielding independence. The elasticity of migration with respect to the mean relative utility of Distressed is $d\ln G_\theta / d\ln\mu = 1/\kappa \cdot [1 - G_\theta(\phi)]$. Thus, a higher κ generates lower migration elasticities.

The distribution of θ is assumed to be log-normal with mean and variance parameters (ξ, σ). As in Mankiw et al. (2009), we additionally allow a mass point at $\theta = 0$ of "disabled" households and a Pareto distributed right tail with parameter $p > 0$. Note that equation (11) assumes $w_j(\theta) = W\theta^{\gamma_j}$. We normalize $\gamma_0 = 1$ throughout our analysis. Hence, whenever $\gamma_1 < 1$, higher-skilled workers have a comparative advantage at working in Elsewhere and households with $\theta > 1$ also enjoy an absolute advantage in Elsewhere. Labor supply is governed by a Frisch elasticity $\eta > 0$. Finally, concavity in the consumption aggregate generates income-based sorting when rental rates differ ($r_0 \neq r_1$).

Table 1 details the values chosen for (η, ξ, σ, p) along with other key calibration choices. Throughout this section, Distressed is taken to be a small area with an over-representation of low earners that mimics the urban Empowerment Zones in the 2016-2020 American Community Survey, while Elsewhere is calibrated to resemble the rest of the United States under the existing place-independent tax system. The calibration chooses three parameters – μ, W, and a sorting parameter (either β, γ_1, or $\underline{H}_1$) capturing one of the three sorting motives – to match three empirical moments: (1) Distressed covers 1.7% of the population, (2) 39% of households nationwide earn under \$50,000, and (3) 56% of Distressed residents earn under \$50,000. Although parameters are chosen jointly to match the three moments, μ can be thought of as governing the size of Distressed (moment 1) and W the scale of average earnings given a skill distribution (moment 2). The sorting parameter can be thought of as rationalizing moment (3) conditional on the other parameters.

Table 1: Parameters That Are Fixed across Simulations

Parameter	Value	Source
Panel A. Features of the current tax system		
Current place-blind lump-sum transfer	\$11,214	Piketty et al. (2018)
Current tax brackets	{\$0K, \$20K, \$500K}	Piketty et al. (2018)
Current marginal tax rates	{44.6%, 28.1%, 49.4%}	Piketty et al. (2018)
Exogenous revenue requirement, R	\$14,746	Implied under symmetric benchmark
Panel B. Preference and skill distribution parameters		
Labor supply elasticity, η	0.5	Chetty et al. (2011)
Mean of log-normal skill distribution, ξ	2.757	Mankiw et al. (2009)
Std. dev. of log-normal skill distribution, σ	0.5611	Mankiw et al. (2009)
Pareto parameter of Pareto skill distribution, p	2	Mankiw et al. (2009)
Grid size for skill distribution, N	578	-
Comparative advantage in Elsewhere, γ_0	1.00	Normalization
Household Pareto weights, $\omega_H(\theta)$	1.00	-
Panel C. Features of Elsewhere ($j = 0$) and Distressed ($j = 1$)		
Distressed population share, L_1 (under current taxes)	1.7%	2016-2020 American Community Survey
Elsewhere population share, L_0 (under current taxes)	98.3%	2016-2020 American Community Survey
Elsewhere housing supply elasticity, ϱ_0	0.34	Baum-Snow and Han (2023)
Elsewhere housing supply shifter, $\underline{H}_0$	0.50	Implied
Elsewhere rent, r_0 (under current taxes)	\$7,284	2016-2020 American Community Survey

Notes: This table enumerates parameters and empirical quantities used in our numerical simulations that remain fixed across all scenarios. Panel A shows the tax system under which we calibrate free parameters to match targeted moments in each simulation scenario. Panel B shows primitives governing household and social preferences and the parameters governing the shape of the skill distribution. Finally, Panel C shows features of Elsewhere and Distressed. The quantity we use for the cost of housing in Elsewhere r_0 is the 20th percentile rent, according to the American Community Survey (ACS) microfiles. Dollar amounts are in 2020 dollars.

The Elsewhere housing supply elasticity ϱ_0 is set to 0.34 based on estimates from Baum-Snow and Han (2024). The Elsewhere housing supply shifter $\underline{H}_0$ is chosen to yield ACS-based Elsewhere rent $r_0 = \$7,284$. The Distressed housing supply elasticity ϱ_1 is also set in this section to 0.34. In the income-based sorting calibrations we choose $\underline{H}_1$ to match the sorting moment. Otherwise, $\underline{H}_1$ is set to ensure $r_1 = r_0 = \$7,284$ under the existing tax system. Household Pareto weights $\omega_H(\theta)$ are set to one for all θ to ensure that redistribution across households is exclusively driven by decreasing marginal consumption utility.

The landlord weight ω_L is chosen to ensure that the social marginal welfare weight of landlords λ_L equals one in the symmetric benchmark below, which in practice leads to values of λ_L at place-based optima very close to one. This choice ensures the planner has no motive to redistribute away from (or towards) landlords on the margin.

Finally, to probe the dependence of our findings to the degree of migration responsiveness assumed, we vary κ across a range of values. For large values of κ, the sorting moments were unable to be rationalized by income-based sorting alone (i.e., matching the moments would have required setting $\underline{H}_1 < 0$). Income-based sorting results are therefore depicted for a restricted range of κ values.

8.1 No-sorting benchmark

The no-sorting symmetric benchmark has no skill-taste correlation ($\beta = 0$), no comparative advantage ($\gamma_1 = \gamma_0 = 1$), and no income-based sorting ($\underline{H}_1$ is set such that $r_1 = r_0 = \$7,284$). As detailed in Appendix C, the optimal tax system is found by numerically solving the system of equations listed in Proposition 2.

As expected, the optimal tax systems are identical for the two communities in this symmetric benchmark. Panel A(i) of Figure 2 displays the optimal marginal tax rate schedule, which exhibits the classic U-shape when top skills are Pareto distributed (Diamond, 1998; Saez, 2001).[6] Panel B(i) of Figure 2 shows the same marginal tax rate patterns when indexing households by skill quantile $F(\theta)$ rather than earnings. Panels A(ii) and B(ii) of Figure 2 display average tax rates, which rise monotonically with earnings and skill. Low-earnings households receive sizable transfers leading to negative average tax rates. The demogrant transfer to zero-earners $-T_0(0) = -T_1(0)$ is \$27,243.[7] This demogrant is taxed away as earnings increase, leading to positive average tax rates for households in the top sixty percent of the earnings distribution.

Optimal taxes will differ across communities as we move away from this symmetric benchmark. We use two metrics, reported in Table 2 below, to summarize differences in taxes between communities. Let $\Delta_0 = T_0(0) - T_1(0)$ denote the amount by which the demogrant in Distressed exceeds the demogrant in

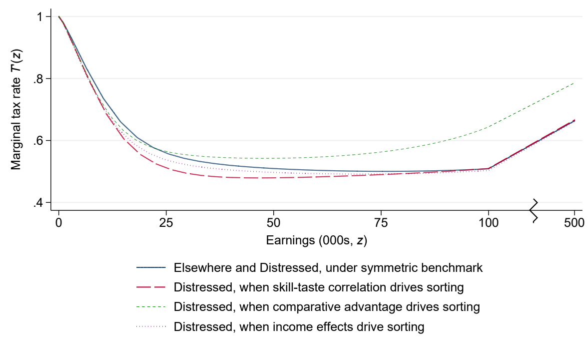
[6] Since the bottom type in our simulations does not generate labor earnings ($\underline{\theta} = 0$), the optimal bottom MTR is 1 (Piketty and Saez, 2013).

[7] Optimal tax simulations based on log consumption utility routinely find larger demogrants transfers and top marginal tax rates than prevail in the United States (Saez, 2001; Mankiw and Weinzierl, 2010).

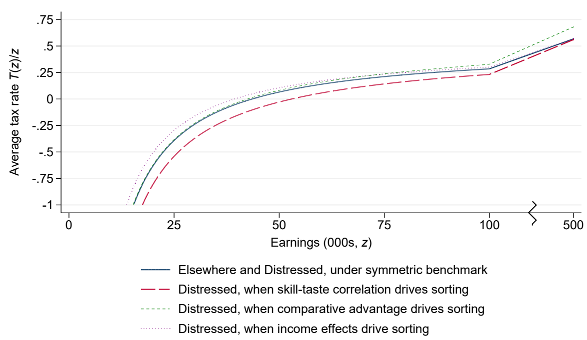
Figure 2: Optimal Tax Schedules by Sorting Motive

A. By Earnings z

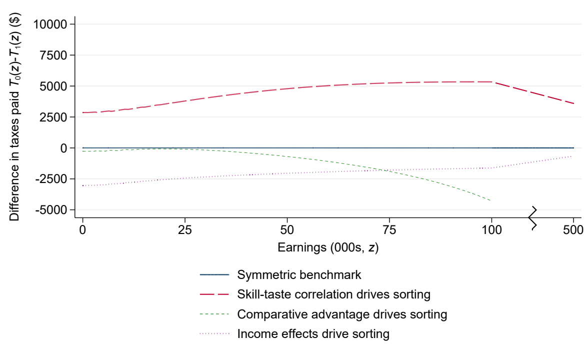
(i) Marginal Tax Rates



(ii) Average Tax Rates

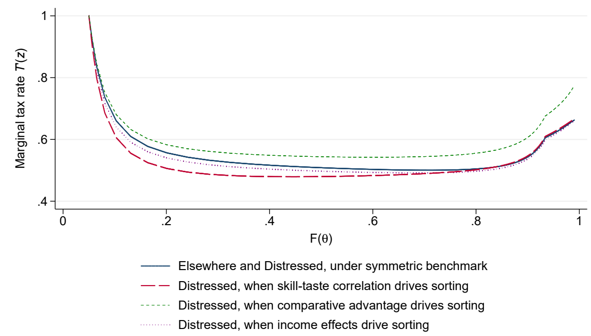


(iii) Difference in Taxes Paid

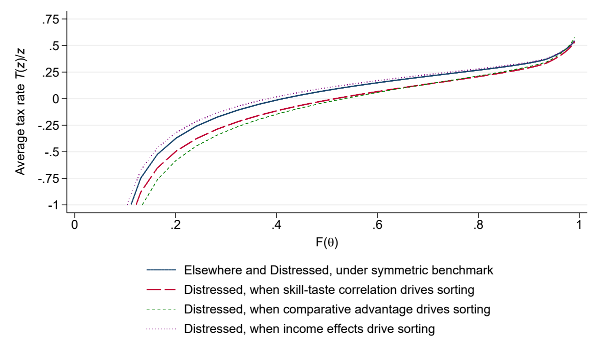


B. By Skill Quantile $F(\theta)$

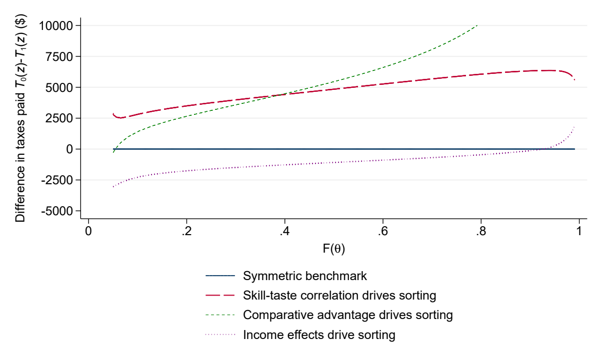
(i) Marginal Tax Rates



(ii) Average Tax Rates



(iii) Difference in Taxes Paid



Notes: This figure shows how optimal place-based redistribution varies with the driver of sorting between Distressed and Elsewhere. Panels A and B show optimal tax schedules faced by households at different earnings and skill levels, respectively. Panels A(i) and B(i) show marginal tax rates (MTRs). Panels A(ii) and B(ii) show average tax rates (ATRs). Appendix Figure 1 shows that MTRs and ATRs are similar in Elsewhere across sorting scenarios. Panels A(iii) and B(iii) show the difference in annual tax amounts, equal to taxes in Elsewhere minus taxes in Distressed. Each panel depicts series for four theoretical scenarios: one in which there is no driver of sorting, one in which skill-taste correlation is the only driver of sorting, one in which comparative advantage is the only driver of sorting, and one in which income effects (via rent differences across communities) are the only driver of sorting. The standard deviation of idiosyncratic preferences for living in Distressed κ is set to 0.5. The substantial contrast between the comparative advantage series in Panel A(iii) and Panel B(iii) reflects productivity differences: a given household is more productive in Elsewhere and therefore earns more and would pay more tax even under a place-blind tax system. Our preferred measure of the average place-based transfer Δ_z reflects only tax differences.

Elsewhere. This quantity measures the optimal place-based transfer to Distressed for disabled households, who earn zero.

To summarize tax differences across all households, we use a second measure based on the within-earnings difference in taxes paid $T_0(z) - T_1(z)$, rather than the within-skill difference in taxes paid $T_0(z_0^*(\theta)) - T_1(z_1^*(\theta))$, as the latter measure would reflect not only tax differences but also productivity and rent impacts on earnings. Specifically, our preferred measure of the average place-based transfer

$$\Delta_z = 1/2 \cdot \mathbb{E}_\theta \left\{ \left[T_0(z_0^*(\theta)) - T_1(z_0^*(\theta)) \right] + \left[T_0(z_1^*(\theta)) - T_1(z_1^*(\theta)) \right] \right\},$$

is an equally weighted average of two indices of tax differences across the two communities. The first index is the average difference in taxes paid that would result if every household was taxed based on their optimal Elsewhere earnings $z_0^*(\theta)$. The second index is the average tax difference that would emerge if every household was taxed based on their optimal Distressed earnings $z_1^*(\theta)$.

Finally, note that the discrepancy $\Delta_z - \Delta_0$ between our two transfer measures reflects only differences in MTR schedules across the two communities. We return to this point frequently below. Section 8.6 decomposes the forces generating differences in optimal MTR schedules across locations.

Table 2: Optimal Place-Based Redistribution by Sorting Motive

	Income-constant average tax differences, Δ_z	Difference in demogrants, Δ_0	Skill-taste correlation, β	Comparative advantage, γ_1	Calibrated rent ratio, r_1/r_0
	(1)	(2)	(3)	(4)	(5)
A. High migration: Std. dev. of idiosyncratic preferences for living in Distressed $\kappa = 0.5$					
Skill-taste correlation	4,805	2,862	0.85	1.00	1.00
Comparative advantage	-2,763	-268	0.00	0.924	1.00
Income-based sorting	-2,225	-3,042	0.00	1.00	0.07
B. Low migration: Std. dev. of idiosyncratic preferences for living in Distressed $\kappa = 4$					
Skill-taste correlation	10,918	6,608	6.83	1.00	1.00
Comparative advantage	7,091	3,740	0.00	0.906	1.00

Notes: This table displays optimal tax results under three sorting scenarios: one in which skill-taste correlation is the only driver of sorting, one in which comparative advantage is the only driver of sorting, and one in which income effects (via rent differences across communities) are the only driver of sorting. Panel A reports results for scenarios with κ set to 0.5, implying stronger migration responses. Panel B reports results for scenarios with κ set to 4, implying weaker migration responses. Panel B omits the income-based sorting row because the moments cannot be rationalized with non-negative Distressed rent when κ is greater than approximately 0.55. Columns 1-2 show our measures of place-based redistribution: the summary measure Δ_z of Elsewhere-minus-Distressed tax differences across all households, and the Elsewhere-minus-Distressed difference Δ_0 in taxes on zero-earners only. Columns 3-5 report the sorting parameters (r_1/r_0 is implied by the calibrated income-based-sorting parameter $\underline{H}_1$ and other assumed values). See Appendix Table 2 for all simulation parameters underlying these scenarios.

8.2 Skill-taste correlation

When sorting is generated entirely by skill-taste correlation, the optimal place-based transfer to Distressed is implemented via both a substantial per capita transfer and lower marginal income tax rates. There are fewer households with high incomes to tax in Distressed, therefore high marginal tax rates yield less revenue per dollar of distortion in Distressed than in Elsewhere. Consequently, as depicted in Panels A(i) and B(i) of Figure 2, the entire nonlinear tax schedule is shifted down in Elsewhere.

Panel A of Table 2 reports that when sorting is solely driven by skill-taste correlation and $\kappa = 0.5$, a strong level of migration responsiveness approximately corresponding to our urban calibrations described in the next section, the optimal tax system entails residents of Distressed receiving an average place-based transfer of $4,805. The difference in demogrants is smaller ($2,862), implying that about 40% of the average place-based transfer arises through lower MTRs in Distressed. The lower MTRs in Distressed yield larger transfers to Distressed at middle incomes than at low incomes, as displayed in Panel A(iii) of Figure 2.

Panel B of Table 2 repeats the analysis for $\kappa = 4$, entailing relatively weak migration responsiveness that approximately corresponds to our rural calibrations discussed in the next section. This eight-fold increase in κ requires an eight-fold increase in β to rationalize the sorting moments. The weaker migration response reduces the fiscal externality to place-based redistribution and more than doubles the optimal place-based transfer to Distressed. These results highlight the quantitative importance of migration responsiveness for the levels of optimal place-based redistribution.

8.3 Comparative advantage

Panel A of Table 2 reveals that when sorting is solely driven by comparative advantage and $\kappa = 0.5$, the optimal tax system entails a *negative* average place-based transfer to Distressed. Equivalently, the optimal tax system entails a positive average transfer to Elsewhere in the amount of $2,763. In contrast, demogrants are essentially equal in the two locations, implying that the optimal tax system is nearly place-neutral for the least skilled. Evidently, the transfers to Elsewhere are facilitated via higher MTRs in Distressed, a pattern apparent from Panels A(i) and B(i) of Figure 2.

Panel B of Table 2 shows that reducing migration responsiveness reverses the direction of transfers. Setting $\kappa = 4$ and repeating the calibration, the resulting optimal tax system yields an average place-based transfer to Distressed of $7,091, with an optimal demogrant that is $3,740 larger in Distressed than in

Elsewhere. The finding that the average tax difference Δ_z exceeds the difference in demogrants Δ_0 implies that MTRs are lower in Distressed on average. As illustrated in Appendix Figure 2, MTRs turn out to be lower in Distressed at all but the highest income levels. This finding echoes the intuition developed in Section 6: the greater the migration responses, the higher the efficiency cost of the policy compared to its equity gains. With sufficiently high migration responsiveness, the direction of optimal transfers is reversed.

8.4 Income-based sorting

Panel A of Table 2 reveals that a rent ratio of 0.07 is needed to rationalize sorting entirely by income effects when $\kappa = 0.5$. This extreme rent difference yields large income effects that amplify the fiscal externalities associated with moves to Distressed. It also generates strong motives for redistribution towards Elsewhere since Elsewhere households are burdened by much higher rents. Evidently, the redistributive motive generated by higher rents dominates the redistributive motive generated by low-skill sorting in this parameterization because of the strong non-homotheticity in housing consumption.

When sorting is driven purely by income effects, the optimal tax system involves an average place-based transfer away from Distressed and toward Elsewhere of $2,225. This transfer partially compensates low-skilled types for the higher rent in Elsewhere, an effect that grows with the dispersion of locational tastes. Figure 3 shows that as the standard deviation of idiosyncratic location preferences κ shrinks, optimal place-based transfers tend toward zero, which mirrors the analytical result in Example 6 and accords with recent results in the optimal commodity taxation literature (Kaplow, 2008; Allcott et al., 2019).

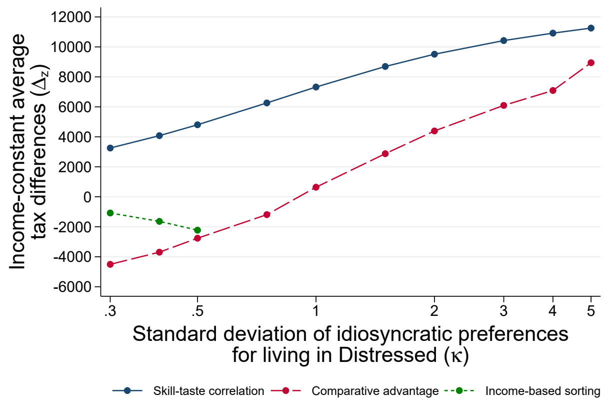
8.5 Sensitivity to κ

Figure 3 illustrates the sensitivity of optimal place-based transfers to migration responsiveness under each sorting motive by repeating the Table 3 analysis at various values of $\kappa \in [0.3, 5]$. When comparative advantage drives sorting, the average place-based transfer is negative for values of κ less than approximately one and positive for higher values. By contrast, in the skill-taste correlation case, κ influences the size of the transfer but not its sign.

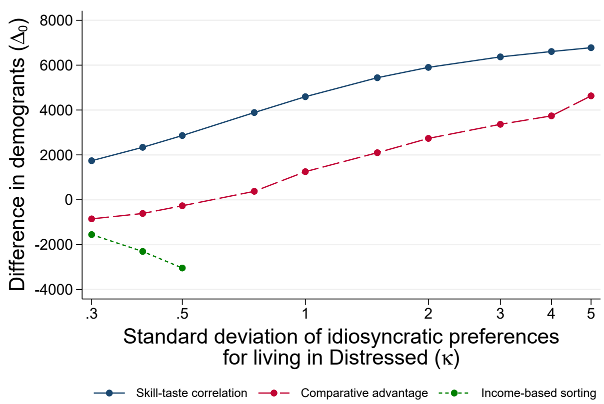
Note that Δ_0 and Δ_z have opposite signs when comparative advantage drives sorting and κ is near 0.75. At such levels of intermediate migration responsiveness, the planner optimally provides higher transfers to the poor in Distressed, while imposing higher taxes on the rich in Distressed. For example, at $\kappa = 1$, the poorest in Distressed receive a place-based transfer of $1,254 (Figure 3B), while households in Distressed

Figure 3: Optimal Place-Based Redistribution by Dispersion of Preferences for Distressed

A. Income-constant Average Tax Difference (Δ_z)



B. Difference in demogrants (Δ_0)



Notes: This figure shows how summary measures of place-based redistribution depend on the standard deviation of idiosyncratic preferences for living in Distressed κ displayed on a log scale. Panel A shows the summary measure Δ_z of Elsewhere-minus-Distressed tax differences across all households. Panel B shows the Elsewhere-minus-Distressed difference Δ_0 in taxes on zero-earners only. Each panel depicts series for three theoretical scenarios: one in which skill-taste correlation is the only driver of sorting, one in which comparative advantage is the only driver of sorting, and one in which income effects (via rent differences across communities) are the only driver of sorting. Income-based sorting was only capable of rationalizing the target moments for a restricted set of κ values that imply relatively high mobility.

at the 99th percentile skill level pay a place-based tax of \$12,398.[8] Hence, the optimal transfer is highly nonlinear.

Section 6 discussed how when the planner is only able to implement flat transfers, the optimal policy becomes a horserace between the fiscal costs of migration and the equity benefits of redistribution that can yield net transfers in either direction. When nonlinear instruments are available, the place-based transfer can be targeted to dampen the migration responses of particular income groups. As migration elasticities fall, the advantage of using nonlinear taxes to stem migration of high-income households fades. This phenomenon is apparent from the contrast between Figure 2 and Appendix Figure 2: with high migration responses, MTRs at top incomes are larger in Distressed, a pattern that is reversed when migration responses are low.

8.6 Decomposing MTRs

Section 7.3 observed that sorting tends to generate lower MTRs in Distressed. However, Section 8.3 found that comparative advantage generated higher MTRs in Distressed when $\kappa = 0.5$. To illuminate the economic forces driving spatial MTR differences, we use equations (8) and (9) to decompose optimal MTR schedules into three components reflecting different planning objectives. A "within-community" component captures tradeoffs between equity and efficiency within communities (Mirrlees, 1971; Saez, 2001), while a second "between-community" component captures the fiscal externalities associated with migration. A third "landlord" component captures the welfare effects of redistributing between households and landlords.

For Distressed, the three MTR components are proportional to the following expressions:

$$\frac{1+\eta_1^U(\theta)}{\eta_1^C(\theta)} \frac{\gamma_1(\theta)}{\theta f(\theta)(1-G_\theta(\phi_\theta))} U_{c1}^*(\theta) \int_\theta^{\overline{\theta}} \frac{1-\lambda_1(s)}{U_{c1}^*(s)} (1-G_s(\phi_s))\, dF(s) \quad \text{(Within)}$$

$$-\frac{1+\eta_1^U(\theta)}{\eta_1^C(\theta)} \frac{\gamma_1(\theta)}{\theta f(\theta)(1-G_\theta(\phi_\theta))} U_{c1}^*(\theta)\, \Delta\tau^+(\theta) \quad \text{(Between)}$$

$$\frac{1+\eta_1^U(\theta)}{\eta_1^C(\theta)} \frac{\gamma_1(\theta)}{\theta f(\theta)(1-G_\theta(\phi_\theta))} U_{c1}^*(\theta)(1-\lambda_L)\, \Delta r^+(\theta) \quad \text{(Landlord)}$$

When scaled by one plus their sum, totaling these three expressions yields the optimal MTR schedule in Distressed. The corresponding expressions for Elsewhere are symmetric.

Table 3 reports the optimal MTR, the scaled within-community component, and the scaled between-community component for the median skill type (i.e., for $\theta = F^{-1}(0.5)$). The within and between components sum approximately to the average MTR because the omitted landlords component is nearly zero when $\lambda_L \approx 1$.

[8] That is, $1/2 \cdot \left[T_0\left(z_0^*\left(F^{-1}(0.99)\right)\right) - T_1\left(z_0^*\left(F^{-1}(0.99)\right)\right)\right] + 1/2 \cdot \left[T_0\left(z_1^*\left(F^{-1}(0.99)\right)\right) - T_1\left(z_1^*\left(F^{-1}(0.99)\right)\right)\right] = -12,398$.

Decompositions are reported for each sorting motive and for two choices of κ.

Table 3: Optimal Marginal Tax Rates at the Median Skill Type

	Distressed			Elsewhere		
	Marginal tax rate	Within-community component	Between-community component	Marginal tax rate	Within-community component	Between-community component
	(1)	(2)	(3)	(4)	(5)	(6)
A. High migration: Std. dev. of idiosyncratic preferences for living in Distressed $\kappa = 0.5$						
Skill-taste correlation	0.479	0.246	0.234	0.510	0.514	-0.004
Comparative advantage	0.544	-0.111	0.655	0.509	0.521	-0.012
Income-based sorting	0.499	0.523	-0.027	0.510	0.509	0.000
B. Low migration: Std. dev. of idiosyncratic preferences for living in Distressed $\kappa = 4$						
Skill-taste correlation	0.442	0.374	0.068	0.510	0.511	-0.001
Comparative advantage	0.446	0.232	0.214	0.510	0.514	-0.004

Notes: This table reports a decomposition of the optimal marginal tax rate for the median skill type in three theoretical scenarios: one in which skill-taste correlation is the only driver of sorting, one in which comparative advantage is the only driver of sorting, and one in which income effects (via rent differences across communities) are the only driver of sorting. Panel A reports simulation results for scenarios with κ set to 0.5, implying a higher level of migration. Panel B reports results for scenarios with κ set to 4, implying a lower level of migration. Panel B omits the income-based sorting row because the moments cannot be rationalized with non-negative Distressed rent when κ is greater than approximately 0.55 Columns 1-3 report results for Distressed, and columns 4-6 report results for Elsewhere. Columns 1 and 4 reports the marginal tax rate. Columns 2 and 4 report the component of the marginal tax rate attributable to an equity-efficiency tradeoff within communities. Columns 3 and 6 report the component of the marginal tax rate attributable to fiscal externalities arising from migration between communities. The marginal tax rate equals the within-community component plus the between-community component.

Column 2 row 2 of Table 3A reveals that the standard within-community equity-efficiency tradeoff depresses the MTR of the median skill type in Distressed when $\kappa = 0.5$ and comparative advantage drives sorting. This finding arises because comparative advantage and high Distressed MTRs (Figure 2) at the optimum cumulate into high consumption differences at higher skill levels, which leads to few higher-skill types in Distressed. However, column 3 indicates that the between-community component pushes toward very high MTRs, consistent with each marginal mover to Distressed yielding a large negative fiscal externality. The two components combine to yield moderately higher MTRs in Distressed (column 1) than in Elsewhere (column 4).

In contrast, Panel B of Table 3 reports a much smaller Distressed between-community component when $\kappa = 4$, reflecting the smaller fiscal externalities that arise at low levels of migration responsiveness. Lower Distressed MTRs at the optimum cumulate into lower average tax rate differences at higher skill levels, which leads to more high-skill types in Distressed and thus more revenue to be raised by high MTRs at middle skill levels. With less sorting at the optimum, the Distressed within-community component becomes positive and substantial. The two components sum to a Distressed MTR that is moderately lower than the Elsewhere MTR.

9 How Large Might Optimal Place-Based Transfers Be?

The previous section investigated the structure of optimal place-based transfers when a single sorting motive was at play. This section studies the size and shape of optimal place-based transfers when sorting is generated by a plausible mix of forces. The analysis proceeds under two alternate sets of assumptions corresponding to different sorts of place-based policies. We use the mean characteristics of the thirty urban and the ten rural Empowerment Zones (EZs), respectively, to anchor these calibrations.

The first scenario assumes that sorting is driven by income effects and skill-taste correlation. This "urban" calibration is meant to correspond to a setting where subsidies impact residential location choice within a city, with each skill type having access to identical wage opportunities regardless of their chosen neighborhood. Next, we consider a "rural" scenario where sorting is driven by income effects and comparative advantage. To the extent that skill-taste correlation is also an important contributor to sorting into rural EZs, neglecting this force will lead us to overstate the contribution of comparative advantage. As illustrated in Figure 3, misattributing sorting to comparative advantage rather than to skill-taste correlation should lead to an underestimate of the true optimal level of redistribution to rural areas.

9.1 Urban baseline parameterization

Urban EZs cover 1.7% of the U.S. population in the 2016-2020 ACS. We use recent Census-tract-level estimates by Baum-Snow and Han (2024) to set Elsewhere's housing supply elasticity ϱ_0 to 0.34 (the non-urban-EZ mean) and Distressed's elasticity ϱ_1 to 0.24 (the urban EZ mean).

Elsewhere's housing supply intercept $\underline{H}_0$ is chosen to ensure that when Elsewhere covers 98.3% of the population Elsewhere's rent r_0 is $\$7,284$, the 20th percentile U.S. rent, which roughly corresponds to the median of the 39% of households earning under \$50,000 nationwide in the 2016-2020 ACS. Distressed's housing supply intercept $\underline{H}_1$ is chosen to ensure that when Distressed covers 1.7% of the population r_1/r_0 equals 0.86: the ratio of the median rent in urban EZs to median rent in the rest of the country according to 2016-2020 ACS tract-level aggregates.[9]

The parameters (μ, W, β) are chosen to match the three empirical moments introduced in the previous section. The standard deviation of location preferences κ is set to yield an elasticity of Distressed population with respect to Distressed wages equal to the 0.82, which is the value estimated empirically for urban EZs by

[9] We use the ratio of median rents rather than 20th percentile rents because quantiles other than the median are not available in the tract level tabulations needed to aggregate to urban EZs.

Busso et al. (2013).[10] In this exercise, the skill-taste-correlation parameter β governs the over-representation of low incomes in Distressed beyond that which is explained by income effects.

The calibration is performed under an approximation of the current (place-blind) tax system, which we empirically estimate using Piketty et al. (2018)'s microfiles aggregated across years 2016-2020, each of which is inflated to 2020 dollars using the CPI-U. Finally, we continue to set $\omega_H(\theta) = 1$ for all θ and to choose ω_L to ensure λ_L equals one at the symmetric benchmark described in the previous section. Additional details on the calibration are described in Appendix C.

Table 4: How Large Might Optimal Place-Based Transfers Be?

	Income-constant average tax differences, Δ_z	Difference in demogrants, Δ_0	Std. dev. of idiosyncratic preferences for living in Distressed, κ	Skill-taste correlation, β	Comparative advantage, γ_1
	(1)	(2)	(3)	(4)	(5)
A. Urban scenarios					
Urban baseline	3,143	1,462	0.44	0.61	1.000
1/2x migration	7,655	4,420	1.52	2.48	1.000
2x migration	-383	-284	0.09	-0.02	1.000
No rent differences	3,265	1,779	0.30	0.51	1.000
1.5x rent differences	2,917	1,139	0.49	0.64	1.000
Swap housing elasticities	4,256	2,137	0.62	0.93	1.000
75% weight on landlords	1,846	236	0.44	0.61	1.000
B. Rural scenarios					
Rural baseline	4,329	532	4.06	0.00	0.900
1/2x migration	6,906	1,496	8.23	0.00	0.885
2x migration	573	-1,755	1.99	0.00	0.900
No rent differences	8,841	4,311	4.32	0.00	0.874
1.5x rent differences	2,238	-1,481	3.95	0.00	0.904
Swap housing elasticities	3,721	-68	3.97	0.00	0.900
75% weight on landlords	5,127	1,154	4.06	0.00	0.900

Notes: This table reports optimal place-based transfers across scenarios. Columns 1-2 show our measures of place-based redistribution: the summary measure Δ_z of Elsewhere-minus-Distressed tax differences across all households, and the Elsewhere-minus-Distressed difference Δ_0 in taxes on zero-earners only. Columns 3-5 report key calibrated parameters. Panel A reports results from an "urban" baseline scenario that assumes sorting is driven by income effects and skill-taste correlation, assumes Distressed housing is supplied with elasticity $\varrho_1 = 0.24$, and targets a migration elasticity of 0.82, a Distressed-to-Elsewhere rent ratio of 0.86, and a share of Distressed households earning under \$50,000 of 56%. Panel B reports results from a "rural" baseline scenario that assumes sorting is driven by income effects and comparative advantage, assumes Distressed housing is supplied with elasticity $\varrho_1 = 0.60$, and targets a migration elasticity of 0.20, a Distressed-to-Elsehwere rent ratio of 0.54, and a share of Distressed households earning under \$50,000 of 60%. The migration scenarios target half or twice the assumed baseline migration elasticity. The rent scenarios set the Distressed housing supply shifter to yield no rent difference or 50% greater rent difference with Elsewhere. The Swap housing elasticities scenario sets the Elsewhere housing supply elasticity ϱ_0 equal to the baseline Distressed elasticity and set the Distressed elasticity $\varrho_1 = 0.34$. The 75% weight on landlords scenario reduces the value of landlord weight ω_L by 25% so that the social planner at the symmetric benchmark in Figure 2 values a marginal dollar to landlords three-quarters as much as a marginal dollar to households. See Appendix Table 2 for all simulation parameters underlying these scenarios. See Appendix Table 3 for additional scenarios that vary economic characteristics of Distressed.

Panel A of Table 4 displays the baseline results. The calibrated value of κ is 0.44, close to the 0.5 value

[10] EZs offer a 20% wage subsidy to zone jobs for zone residents, implying a reduction in the net of tax rate of approximately $\ln(1.2) = 0.182$. Busso et al. (2013, Table 10) report that EZs increased local jobs held by zone residents by 15 log points, implying a behavioral elasticity of $0.15/0.182 = 0.82$. Note however that the EZ wage credits only affected the subset of zone residents who also worked in the zone and that we have abstracted from the subsidy's \$3,000 cap and from the program's local block grant. These omissions likely bias the elasticity in opposite directions.

used in the illustrative simulations of Panel A of Table 2. At the optimum, residents of Distressed enjoy an average place-based transfer Δ_z of \$3,143, which is approximately two-thirds of the optimal value reported in Panel A of Table 2 under the slightly higher κ and when skill-taste correlation drives all sorting. Distressed residents enjoy a larger demogrant Δ_0 of \$1,462, implying that just over half of the average place-based transfer arises through lower marginal tax rates.

By comparison, actual EZs provide a slightly larger transfer of \$3,000 for all eligible full-time-employed residents, which suggests that the magnitude (but not the capped nature) of the urban EZ transfers may be close to optimal. However, actual EZs involved additional parameters whose optimality we cannot assess using our theoretical framework.

9.2 Robustness: alternative economic assumptions

The remaining rows of Panel A of Table 4 repeat the baseline specification under alternative scenarios. In each row, we recalibrate the model by changing either the value of some parameter or the value of some targeted moment. The first two rows vary the migration target to either half or twice the baseline urban EZ migration elasticity, yielding different calibrated values for κ. The next two rows set Distressed rent equal to either 72% of Elsewhere rent or 100%. The greater the rent differential, the more sorting is driven by income effects and the less is driven by skill-taste correlation. The next two rows halve or double the Distressed housing supply elasticity.

Optimal place-based transfers to Distressed rise to \$7,655 when when the target migration elasticity is doubled and become slightly negative when it is halved. Varying the rent has offsetting effects. For example, while eliminating rent differences increases the skill-taste correlation needed to explain sorting and thus pushes toward greater place-based transfers, it also raises migration responsiveness (lowers κ) which tempers place-based transfers. Making Distressed housing supply relatively elastic rather than inelastic increases place-based transfers by about \$1,000. As detailed in Appendix Table 3, optimal place-based transfers are relatively sensitive to the low-income share of households in Distressed but not to the size of Distressed.

9.3 Robustness: alternative weight on landlords

Thus far, the landlord weight ω_L has been chosen to ensure the planner is indifferent about redistribution between landlords and households. Nevertheless, landlord incidence has been an important indirect feature of all our results, as the previous subsection's housing elasticity result showed.

The final row of Panel A of Table 4 reduces the value of ω_L by 25% so that $\lambda_L = 0.75$ at the symmetric benchmark in Section 8.1. The value 0.75 approximately equals the ratio of the marginal utility of consumption of a household at the 82nd percentile of the skill distribution to that of the average household at the place-based optimum. We choose the 82nd percentile because it is the U.S. real-estate-ownership-weighted-mean income percentile in the 2016-2020 Piketty et al. (2018) microfiles.

Placing a 75% weight on landlords relative to households reduces optimal place-based transfers to Distressed by about \$1,000. In the urban baseline, Distressed housing is less elastically supplied than Elsewhere housing. As a result, $\frac{r_1}{\varrho_1} - \frac{r_0}{\varrho_0} > 0$ in equations (8)-(10), implying the planner has a motive to move households to Elsewhere in order to reduce aggregate rent.

9.4 Rural parameterization

To conclude, we examine optimal place-based transfers to or from areas where local productivity levels are lower. In this calibration, wages are lower in Distressed, and the over-representation of high-skilled households in Elsewhere stems, in part, from a spatial comparative advantage in production. We refer to these specifications as "rural" EZ parameterizations because productivity differences between distressed rural counties and the rest of the country are especially plausible.

Distressed's housing supply intercept $\underline{H}_1$ is chosen to achieve a rent ratio r_1/r_0 of 0.54, the ratio of the median rent in rural EZs to median rent in the rest of the country according to 2016-2020 ACS tract-level aggregates. Distressed's housing supply elasticity ϱ_1 is set to 0.60, the rural EZ mean in Census-tract-level estimates by Baum-Snow and Han (2024), while Elsewhere's elasticity ϱ_0 is preserved at 0.34.

Skill-taste correlation is set to zero and the parameters (μ, W, γ_1) are chosen to match the following three empirical moments drawn from the 2016-2020 ACS: (1) Distressed covers 1.7% of the population, as in the urban calibrations; (2) 39% of households nationwide earn under \$50,000, as in the urban calibrations; and (3) 60% of Distressed residents earn under \$50,000, as rural EZs are lower income than Urban EZs. The migration elasticity moment is drawn from Sprung-Keyser et al. (2022), who report an elasticity of Commuting Zone population with respect to wages of 0.20.[11] The smaller migration elasticity accords with rural EZs being larger than urban EZs (Kleven et al., 2020).

The first row of Panel B of Table 4 reports the results for the rural baseline. The calibrated value of

[11] Sprung-Keyser et al. (2022) study Commuting Zones which, much like Rural EZs, are collections of counties or large parts of counties. They estimate that an approximately 3% increase in local wages causes an average increase in local population of approximately 0.6%.

κ is 4.06, similar to the $\kappa = 4$ simulations of Panel B of Table 2. At the optimum, Distressed residents receive an average place-based transfer of \$4,329, most of which arises from lower MTRs in Distressed. The average place-based transfer is lower than that found in Panel B of Table 2 when comparative advantage exclusively drove sorting under a similar value of κ, rather than the rural baseline's combination of comparative advantage and income-based sorting. Lower MTRs in Distressed were also observed and discussed in the comparative advantage case of Panel B of Table 2.

The remaining rows of Panel B of Table 4 report the alternative scenarios analogous to those discussed above for the urban exercise. Doubling the migration elasticity yields near-zero optimal place-based transfers on average. Contrary to the urban exercise, placing 75% weight on landlords yields approximately \$1,000 greater optimal place-based transfers to Distressed. Unlike in the urban scenario, Distressed in the rural scenario enjoys greater housing supply elasticity than Elsewhere. As a result, $\frac{r_1}{\varrho_1} - \frac{r_0}{\varrho_0} < 0$ in equations (8)-(10). Thus, the planner redistributes away from landlords by using place-based transfers to shift more households to Distressed.

Table 5: Crosswalk from Urban Baseline to Rural Baseline

Scenario	Income-constant average tax differences, Δ_z (1)	Difference in demogrants, Δ_0 (2)	Std. dev. of idiosyncratic preferences for living in Distressed, κ (3)	Skill-taste correlation, β (4)	Comparative advantage, γ_1 (5)
Urban baseline	3,143	1,462	0.44	0.61	1.000
+ migration moment from rural baseline	9,870	5,612	3.85	6.46	1.000
+ calibrate γ_1 instead of β	6,402	3,227	3.81	0.00	0.909
+ rent moment from rural baseline	2,780	-450	3.89	0.00	0.916
+ sorting moment from rural baseline	4,140	457	3.82	0.00	0.900
+ housing supply elasticities from rural baseline	4,329	532	4.06	0.00	0.900

Notes: This table begins with the Urban baseline results of Table 4A and progressively alters the targeted moments or the assumed Distressed housing supply elasticity in order to yield the rural baseline results of Table 4B. The second row replaces the migration elasticity target of 0.82 with 0.20. The third row calibrates comparative advantage γ_1 instead of skill-taste correlation β to match the moments. The fourth row calibrates $\underline{H}_1$ to achieve a Distressed-to-Elsewhere rent ratio of 0.54 instead of 0.86. The fifth row targets a Distressed share of households earning under \$50,000 of 60% instead of 56%. The sixth row replaces the Distressed housing supply elasticity of 0.24 with 0.60. The columns are the same as those in Table 4; see the notes to that table for details.

Table 5 provides a crosswalk between the urban and rural baselines by altering key moments and parameters in succession. Starting from the urban baseline, matching the lower migration response from the rural baseline triples the optimal place-based transfer to Distressed. Subsequently replacing skill-taste correlation with comparative advantage reduces optimal place-based transfers half-way back to the urban baseline. Additionally altering the Distressed housing supply shifter to match the rural baseline's lower Distressed rent has two effects: it increases income-based sorting and it reduces the marginal utility of consumption in

Distressed especially at the bottom. This combination yields attenuated place-based transfers to Distressed on average and a small place-based transfer to Elsewhere at the bottom. Finally, matching the rural baseline's higher sorting moment increases comparative advantage and nearly yields the rural baseline's average place-based transfer, with the remaining difference accounted for by the Distressed housing supply elasticity from the rural baseline.

10 Conclusion

This paper has found that place-blind taxation is generally suboptimal in a spatial equilibrium model with idiosyncratic locational preferences. The optimal location-specific tax systems are strongly influenced by the economic motives generating this sorting. When sorting is generated by skill-taste correlation, location serves as a valuable proxy for skill that motivates spatially targeted transfers to Distressed communities. When sorting is driven by productive comparative advantage, spatial transfers will tend to be optimal but their direction depends on the nature and strength of migration responses to taxation. While migration elasticities are already a standard estimand in the empirical literature on place-based policies, little work exists exploring heterogeneity in migration responses by income level, which is a key determinant of the magnitude of fiscal externalities. More work exploring the characteristics, motives, and earnings changes of marginal (as opposed to average) movers would be valuable for future quantitative policy assessments.

Finally, the scope for spatial targeting to improve on place-blind taxation hinges critically on the degree of sorting present in an economy. Mean income has been diverging across U.S. states and counties for decades. In recent years, much of this pattern is driven by rising concentration of high income households (Gaubert et al., 2021). An interesting subject for future research is understanding the forces generating these trends. Quantitative evidence on this question would help to inform whether the potential gains from spatial targeting of top tax rates have grown relative to the gains from spatial targeting of anti-poverty programs.

References

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf, "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Albouy, David, "The Unequal Geographic Burden of Federal Taxation," *Journal of Political Economy*, 2009, *117* (4), 635–667.

— , "Evaluating the efficiency and equity of federal fiscal equalization," *Journal of Public Economics*, 2012, *96* (9-10), 824–839.

Ales, Laurence and Christopher Sleet, "Optimal Taxation of Income-Generating Choice," *Econometrica*, 2022, *90* (5), 2397–2436.

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky, "A theory of regressive sin taxes, with an application to the optimal soda tax," *Quarterly Journal of Economics*, 2019, *134* (3), 1557–1626.

Atkinson, Anthony B, "On the measurement of inequality," *Journal of Economic Theory*, 1970, *2* (3), 244–263.

Atkinson, Anthony Barnes and Joseph E Stiglitz, "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 1976, *6* (1-2), 55–75.

Austin, Benjamin, Edward Glaeser, and Lawrence Summers, "Jobs for the heartland: Place-based policies in 21st-century America," *Brookings Papers on Economic Activity*, 2018, pp. 151–232.

Bartik, Timothy J, "Using Place-Based Jobs Policies to Help Distressed Communities," *Journal of Economic Perspectives*, 2020, *34* (3), 99–127.

Baum-Snow, Nathaniel and Lu Han, "The microgeography of housing supply," *Journal of Political Economy*, 2024, *136* (6), 1897–1946.

— **and Ronni Pavan**, "Understanding the city size wage gap," *The Review of Economic Studies*, 2011, *79* (1), 88–127.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A unified framework for measuring preferences for schools and neighborhoods," *Journal of Political Economy*, 2007, *115* (4), 588–638.

Busso, Matias, Jesse Gregory, and Patrick Kline, "Assessing the incidence and efficiency of a prominent place based policy," *American Economic Review*, 2013, *103* (2), 897–947.

Card, David, Jesse Rothstein, and Moises Yi, "Location, Location, Location," *American Economic Journal: Applied Economics*, Forthcoming.

Census, "QuickFacts Clinton County, Kentucky; Wayne County, Kentucky; Jackson County, Kentucky; Accessed August 21, 2024," 2024.

Chetty, Raj, "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annual Review of Economics*, 2009, *1* (1), 451–488.

— **, John N Friedman, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter**, "The opportunity atlas: Mapping the childhood roots of social mobility," *Working Paper*, 2020.

Chyn, Eric and Diego Daruich, "An equilibrium analysis of the effects of neighborhood-based interventions on children," *Working Paper*, 2023.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum, "Matching in cities," *Journal of the European Economic Association*, 2022, *20* (4), 1478–1521.

Davis, Morris A and Jesse Gregory, "Place-based redistribution in location-choice models," *Working Paper*, 2022.

Diamond, Peter A, "Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates," *American Economic Review*, 1998, *88* (1), 83–95.

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

— **and Cecile Gaubert**, "Spatial Sorting and Inequality," *Annual Review of Economics*, 2022, *14* (1), 795–819.

Ehrlich, Maximilian and Henry G Overman, "Place-based policies and spatial disparities across European cities," *Journal of Economic Perspectives*, 2020, *34* (3), 128.

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

Fajgelbaum, Pablo D, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar, "State taxes and spatial misallocation," *The Review of Economic Studies*, 2018, *86* (1), 333–376.

Ferey, Antoine, Benjamin B Lockwood, and Dmitry Taubinsky, "Sufficient statistics for nonlinear tax systems with general across-income heterogeneity," *American Economic Review*, 2024, *114* (10), 3206–3249.

Flatters, Frank, Vernon Henderson, and Peter Mieszkowski, "Public goods, efficiency, and regional fiscal equalization," *Journal of Public Economics*, 1974, *3* (2), 99–112.

Fu, Chao and Jesse Gregory, "Estimation of an Equilibrium Model With Externalities: Post-Disaster Neighborhood Rebuilding," *Econometrica*, 2019, *87* (2), 387–421.

Gaubert, Cecile, "Firm sorting and agglomeration," *American Economic Review*, 2018, *108* (11), 3117–53.

_ , Patrick Kline, Damián Vergara, and Danny Yagan, "Trends in US spatial inequality: Concentrating affluence and a democratization of poverty," *AEA Papers and Proceedings*, 2021, *111*, 520–525.

Glaeser, Edward and Joshua Gottlieb, "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, 2008, *39.*

Glaeser, Edward L, *Cities, agglomeration, and spatial equilibrium*, Oxford University Press, 2008.

_ and David C Mare, "Cities and skills," *Journal of Labor Economics*, 2001, *19* (2), 316–342.

Gomes, Renato, Jean-Marie Lozachmeur, and Alessandro Pavan, "Differential taxation and occupational choice," *The Review of Economic Studies*, 2018, *85* (1), 511–557.

Gordon, Roger H, "An optimal taxation approach to fiscal federalism," *Quarterly Journal of Economics*, 1983, *98* (4), 567–586.

_ , "Fiscal Federalism and the Role of the Income Tax," *Working Paper*, 2023.

Jargowsky, Paul A, *Poverty and place: Ghettos, barrios, and the American city*, Russell Sage Foundation, 1997.

Kaplow, Louis, "On the undesirability of commodity taxation even when income taxation is not optimal," *Journal of Public Economics*, 2006, *90* (6-7), 1235–1250.

_ , "Optimal policy with heterogeneous preferences," *The BE Journal of Economic Analysis & Policy*, 2008, *8* (1).

Kleven, Henrik, Camille Landais, Mathilde Muñoz, and Stefanie Stantcheva, "Taxation and migration: Evidence and policy implications," *Journal of Economic Perspectives*, 2020, *34* (2), 119–142.

Kleven, Henrik J, "Sufficient statistics revisited," *Annual Review of Economics*, 2021, *13* (1), 515–538.

Kline, Patrick, "Place based policies, heterogeneity, and agglomeration," *American Economic Review*, 2010, *100* (2), 383–87.

_ and Enrico Moretti, "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 2014, *129* (1), 275–331.

_ and _ , "People, places, and public policy: Some simple welfare economics of local economic development programs," *Annual Review of Economics*, 2014, *6* (1).

Kremer, Michael, "Should taxes be independent of age?," *Working Paper*, 2003.

Mankiw, N Gregory and Matthew Weinzierl, "The optimal taxation of height: A case study of utilitarian income redistribution," *American Economic Journal: Economic Policy*, 2010, *2* (1), 155–76.

_ , _ , **and Danny Yagan**, "Optimal taxation in theory and practice," *Journal of Economic Perspectives*, 2009, *23* (4), 147–74.

McFadden, Daniel, "Modeling the choice of residential location," *Transportation Research Record*, 1978, (673).

Mirrlees, James A, "An exploration in the theory of optimum income taxation," *The Review of Economic Studies*, 1971, *38* (2), 175–208.

_ , "Optimal tax theory: A synthesis," *Journal of Public Economics*, 1976, *6* (4), 327–358.

Oates, Wallace E, *Fiscal federalism*, Harcourt Brace Jovanovich Inc, 1972.

_ , "An essay on fiscal federalism," *Journal of Economic Literature*, 1999, *37* (3), 1120–1149.

Piketty, Thomas and Emmanuel Saez, "Optimal labor income taxation," in "Handbook of Public Economics," Vol. 5, Elsevier, 2013, pp. 391–474.

_ , _ , **and Gabriel Zucman**, "Distributional national accounts: Methods and estimates for the United States," *Quarterly Journal of Economics*, 2018, *133* (2), 553–609.

Reardon, Sean F and Kendra Bischoff, "Income inequality and income segregation," *American Journal of Sociology*, 2011, *116* (4), 1092–1153.

_ , _ , **Ann Owens, and Joseph B Townsend**, "Has income segregation really increased? Bias and bias correction in sample-based segregation estimates," *Demography*, 2018, *55* (6), 2129–2160.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Felipe Schwartzman, "Cognitive Hubs and Spatial Redistribution," *Working Paper*, 2019.

Rothschild, Casey and Florian Scheuer, "Redistributive taxation in the Roy model," *Quarterly Journal of Economics*, 2013, *128* (2), 623–668.

Saez, Emmanuel, "Using elasticities to derive optimal income tax rates," *The Review of Economic Studies*, 2001, *68* (1), 205–229.

_ , "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," *Journal of Public Economics*, 2002, *83* (2), 217–230.

Scheuer, Florian, "Entrepreneurial taxation with endogenous entry," *American Economic Journal: Economic Policy*, 2014, *6* (2), 126–163.

Slattery, Cailin and Owen Zidar, "Evaluating state and local business incentives," *Journal of Economic Perspectives*, 2020, *34* (2), 90–118.

Sprung-Keyser, Ben, Nathaniel Hendren, Sonya Porter et al., "The radius of economic opportunity: Evidence from migration and local labor markets," *Working Paper*, 2022.

ONLINE APPENDIX

A Net Within-Earnings Transfers to Distressed Locations Due to Marriage and Children

The United States conditions taxes and transfers on marital status and parental status in addition to earnings. The residents of distressed areas are less likely to be married and more likely to have children. To the extent that marital status and parental status serve as proxies to direct transfers to the residents of distressed areas within earnings levels, it is valuable to know if such net indirect transfers are substantial.

No publicly available micro dataset contains both fine geocoding as well as comprehensive measures of taxes and transfers. However, the 2016-2020 American Community Survey (ACS) provides tract-level counts of earnings×marriage×children cells, and the 2016-2020 Piketty et al. (2018) distributional national accounts (DINA) data provide estimates of total taxes and transfers by these same cells. We aggregate the tract-level ACS data to produce counts of families by these cells for "Distressed" (all forty urban and rural Empowerment Zones pooled) and "Elsewhere" (the rest of the country). We use those counts to produce mean within-earnings net taxes in Distressed and Elsewhere, equal to the community-specific marriage×children weighted average of DINA net taxes within each earnings cell, integrated over the *nationwide* earnings distribution. The difference between those two weighted averages equals our estimate of implicit place-based transfers.

Mathematically, we estimate the implicit place-based transfer to Distressed as

$$\mathbb{E}_z\left[\mathbb{E}_{mc|0z}\left[T\left(z,m,c\right)\right]\right]-\mathbb{E}_z\left[\mathbb{E}_{mc|1z}\left[T\left(z,m,c\right)\right]\right] \tag{12}$$

where z denotes earnings, $m \in \{0,1\}$ is an indicator for being married, $c \in \{0,1\}$ is an indicator for having children, $T\left(z,m,k\right)$ denotes net taxes within each earnings×marriage×children cell, $\mathbb{E}_z[\cdot]$ denotes the expectation over the nationwide distribution of earnings z, $\mathbb{E}_{mc|jz}\left[\cdot\right]$ denotes the expectation over community j's probability mass function of married-kids at earnings level z, $j=0$ denotes Elsewhere, and where $j=1$ denotes Distressed.

The ACS data possess sixty-four earnings×married×children bins on resident families: {sixteen earnings groups} × {married or unmarried} × {has kids or does not have kids}. We aggregate the DINA data to the tax unit level and compute DINA taxes as all taxes gross of refundable tax credits minus social insurance contributions, which equals DINA's main raw tax variable. DINA transfers are computed as all individualized benefits (e.g., refundable tax credits, medicaid, SNAP) minus social insurance benefits, which equals DINA's main raw transfers variable minus collective consumption expenditure (e.g., national defense). DINA net taxes equals DINA taxes minus DINA transfers. DINA earnings equals fiscal income excluding capital gains (the `fninc` variable) inflated to 2020 dollars, which is similar to ACS's "total income during the past 12 months" earnings concept for most of the income distribution.

As a first pass, we take the DINA data as given and report results in Panel A.i of Appendix Table 1. Column 1 reports that the average family in Elsewhere pays \$842 less in net taxes annually than the average family in Distressed within earnings levels. Holding earnings constant, the demographic makeup of Distressed yields \$842 more in net taxes paid per family than in Elsewhere. That value is the sum of paying \$623 less in taxes and receiving \$220 more in transfers.

Column 4 row 3 repeats the analysis while ignoring marital status and reports $\mathbb{E}_{zm}\left[\mathbb{E}_{c|0z}\left[T\left(z,m,c\right)\right]\right]-\mathbb{E}_{zm}\left[\mathbb{E}_{c|1z}\left[T\left(z,m,c\right)\right]\right]$. Column 7 row 3 repeats the analysis while ignoring parental status and reports

$\mathbb{E}_{zc}\left[\mathbb{E}_{m|0z}\left[T\left(z,m,c\right)\right]\right]-\mathbb{E}_{zc}\left[\mathbb{E}_{m|1z}\left[T\left(z,m,c\right)\right]\right]$. These results indicate that within-earnings differences in taxes and transfers by marital status drive most of the column 1 result.

A drawback to the preceding analysis is that the DINA data do not use children to impute all transfers. DINA federal taxes are drawn directly from public-use IRS tax returns, and state and local tax data can be imputed fairly reliably based on federal tax return information. However, transfers are often functions of variables not present in federal tax returns and are not always taken up. Consequently, the DINA data impute transfers using the Current Population Survey (CPS). Some imputations like Medicaid are performed within earnings-octile×married×number-of-children bins. However, due to thinner annual data, other DINA imputations like SNAP do not condition on children despite children being an important input to the SNAP benefits formula.

We therefore replace major DINA transfers available in the CPS: SNAP, TANF, Social Security Disability, Supplemental Security Income, veterans benefits, "other benefits" like utility bill reductions (which DINA imputes as a fraction of SNAP benefits), and Medicaid (though Medicaid benefits barely change as expected because DINA Medicaid benefits already condition on children). As in DINA, for each benefit, we blow up each cell's amount by a constant factor across cells in order to maintain aggregates.

Panel A.ii reports results using the CPS-augmented DINA data. In accordance with transfers being greater for families with children, column 4 reports that considering only differences in parental status yield the average Elsewhere family paying $130 more in net taxes than the average Distressed family within earnings levels. However, column 1 reports that when considering both marital status and parental status, the average Elsewhere family pays $713 less in taxes. Hence, the CPS adjustment attenuates the DINA-only-based estimate toward zero while preserving its sign.

Overall, the CPS-adjusted estimates accord with intuition. Within earnings levels, taxes are on average lower for married families filing jointly than for single families. Transfers are also larger for married families within earnings levels, which is substantially driven by the fact that two adults rather than only one qualify for Medicaid and veterans health benefits. Transfers like SNAP are greater for families with children than those without. On net, however, the marital status differences dominate.

Panel A considered all families. Panel B repeats the analysis only for families earning $30,000 or less, for whom taxes are much smaller and transfers are much larger. Panel B.i finds a trivial difference in net taxes between the two communities within these low earnings levels, as low-earners in Distressed enjoy a greater transfer advantage due to parental status and a smaller tax penalty due to marital status. Panel B.ii. similarly finds that the average low-earning Elsewhere family pays $274 more in net taxes than the average low-earning Distressed family within earnings levels.

In sum, all of our column 1 net taxes estimates are below $1,000 in magnitude and neither of our preferred CPS-augmented estimates indicate large *de facto* within-earnings transfers to Distressed locations due to marital and parental status differences. While our results are not entirely dispositive, these findings strongly suggest that redistribution between marital and parental demographic groups does not induce large transfers towards distressed areas.[12]

[12] This exercise is not entirely dispositive because taxes and transfers are imputed to families in distressed areas based on the nationwide taxes and transfers within earnings, marriage, and children cells rather than measuring taxes and transfers directly. For example, if Distressed families with children have more children on average than Elsewhere families with children, our exercise could be biased against finding implicit place-based transfers to Distressed, as benefits like Medicaid and SNAP depend positively on the number of children.

B Proofs and derivations

B.1 Section 3

B.1.1 Landlords' problem

Omit subscripts j for simplicity. The landlords' problem is given by $\max_H Hr - \varphi(H)$, where $\varphi(.)$ is the cost of supplying housing. The FOC is given by $r = \varphi'(H)$. Assume $\varphi(.)$ is isoelastic with elasticity ϱ

$$\varphi(H) = \frac{\underline{H}}{\frac{1}{\varrho}+1}\left(\frac{H}{\underline{H}}\right)^{\frac{1}{\varrho}+1}.$$

Then, the FOC implies that $H = \underline{H}r^{\varrho}$. Moreover

$$\Pi(r) = \max_H Hr - \frac{\underline{H}}{\frac{1}{\varrho}+1}\left(\frac{H}{\underline{H}}\right)^{\frac{1}{\varrho}+1} = \underline{H}r^{\varrho}r - \frac{\underline{H}}{\frac{1}{\varrho}+1}\left(\frac{\underline{H}r^{\varrho}}{\underline{H}}\right)^{\frac{1}{\varrho}+1} = \underline{H}r^{\varrho+1}\left(1 - \frac{1}{\frac{1}{\rho}+1}\right) = \frac{\underline{H}r^{\varrho+1}}{1+\varrho} = \frac{Hr}{1+\varrho}.$$

B.2 Section 4

B.2.1 Example 2 (sorting under comparative advantage)

If rents and amenities are identical in the two cities, we can normalize the indirect utility functions to

$$V_j(\Theta) = U\left(z_j^*(\theta)\right) - \psi\left(\frac{z_j^*(\theta)}{w_j(\theta)}\right) + j\phi.$$

The cutoff for sorting is

$$\phi_\theta = U\left(z_0^*(\theta)\right) - \psi\left(\frac{z_0^*(\theta)}{w_0(\theta)}\right) - U\left(z_1^*(\theta)\right) + \psi\left(\frac{z_1^*(\theta)}{w_1(\theta)}\right), \tag{B.1}$$

where income is chosen optimally given wages according to the household first-order condition

$$U_c\left(z_j^*(\theta)\right) - \frac{1}{w_j(\theta)}\psi'\left(\frac{z_j^*(\theta)}{w_j(\theta)}\right) = 0. \tag{B.2}$$

We aim to show that $\frac{d\phi_\theta}{d\theta} > 0$. Differentiating equation (B.1), plugging in equation (B.2), and using the envelope theorem yields

$$\frac{d\phi_\theta}{d\theta} = \frac{1}{\theta}\left(\gamma_0 z_0^*(\theta)U_c\left(z_0^*(\theta)\right) - \gamma_1 z_1^*(\theta)U_c\left(z_1^*(\theta)\right)\right). \tag{B.3}$$

To examine the sign of equation (B.3), note that the implicit function theorem applied to equation (B.2) (using $U_{cc} < 0$ and $\psi'' > 0$) yields

$$\frac{dz_j^*(\theta)}{dw_j(\theta)} > 0. \tag{B.4}$$

We have that $\gamma_0 > \gamma_1$ and $z_0^*(\theta) \geq z_1^*(\theta)$. This implies that $\frac{d\phi_\theta}{d\theta} > 0$ so long as U is not too concave in c given ψ, that is, if

$$\frac{\gamma_0 z_0^*(\theta)}{\gamma_1 z_1^*(\theta)} > \frac{U_c\left(z_1^*(\theta)\right)}{U_c\left(z_0^*(\theta)\right)},$$

which, at the optimum, simultaneously means that ψ is not too convex in l given U. Note that a sufficient condition is that $xU(x)$ is non-decreasing in x, a condition is met in the functional form considered in the numerical simulations, $U(x) = \log(x)$.

B.2.2 Example 3 (sorting under income-based sorting)

Given that wages and amenities are identical in the two locations, we can normalize wages to θ and write the indirect utility function as

$$V_j(\Theta) = U\left(z_j^*(\theta) - r_j\right) - \psi\left(\frac{z_j^*(\theta)}{\theta}\right) + j\phi.$$

The cutoff for sorting is

$$\phi_\theta = U\left(z_0^*(\theta) - r_0\right) - \psi\left(\frac{z_0^*(\theta)}{\theta}\right) - U\left(z_1^*(\theta) - r_1\right) + \psi\left(\frac{z_1^*(\theta)}{\theta}\right),$$

where income is chosen optimally given rents according to the household first-order condition

$$U_c\left(z_j^*(\theta) - r_j\right) = \frac{1}{\theta}\psi'\left(\frac{z_j^*(\theta)}{\theta}\right) \tag{B.5}$$

We aim to show that $\frac{d\phi_\theta}{d\theta} > 0$, so that higher skill households sort into the expensive location. Using the envelope theorem applied to income choice, we have

$$\frac{d\phi_\theta}{d\theta} = \frac{1}{\theta}\left(z_0^*(\theta)U_c\left(z_0^*(\theta) - r_0\right) - z_1^*(\theta)U_c\left(z_1^*(\theta) - r_1\right)\right). \tag{B.6}$$

To examine the sign of equation (B.6), note that

$$\frac{dU_c\left(z_j^*\left(r_j\right) - r_j\right)}{dr_j} = U_{cc}\left(z_j^*(\theta) - r_j\right)\left(\frac{dz_j^*(\theta)}{dr_j} - 1\right). \tag{B.7}$$

The implicit function theorem applied to equation (B.5) yields

$$\frac{dz_j^*(\theta)}{dr_j} = \frac{U_{cc}\left(z_j^*(\theta) - r_j\right)}{U_{cc}\left(z_j^*(\theta) - r_j\right) - \frac{1}{\theta^2}\psi''\left(\frac{z_j^*(\theta)}{\theta}\right)}.$$

Then, $U_{cc} < 0$ and $\psi'' > 0$ implies that $\frac{dz_j^*(\theta)}{dr_j} \in (0,1)$. Plugging this into equation (B.7) yields $\frac{dU_c\left(z_j^*(r_j) - r_j\right)}{dr_j} > 0$. Finally, $\frac{dz_j^*(\theta)}{dr_j} > 0$ implies that $z_0^*(\theta)U_c\left(z_0^*(\theta) - r_0\right) > z_1^*(\theta)U_c\left(z_1^*(\theta) - r_1\right)$ whenever $r_0 > r_1$, hence $\frac{d\phi_\theta}{d\theta} > 0$.

B.3 Section 6

B.3.1 Proposition 1

At the optimal t^*, $\frac{dSWF}{dt} = 0$, where $\frac{dSWF}{dt} = \frac{dW}{dt} + \frac{dB}{dt}$. Equations (6) and (7) account for for $\frac{dW}{dt}$ and $\frac{dB}{dt}$ when starting from $t = 0$. When $t \neq 0$, movers from Elsewhere to Distressed stop paying $\frac{t^*}{L_0}$ and start receiving $\frac{t^*}{L_1}$, so the $\frac{dB}{dt}$ term is augmented by t^* per mover, generating an additional fiscal externality of

$\frac{t^* \mathbb{E}_\theta[m(t^*)]}{L_0 L_1}$. The result then proceeds from isolating t^* from $\frac{dW}{dt} + \frac{dB}{dt} = 0$ after correcting the term $\frac{dB}{dt}$.

B.3.2 Section 6 results under alternative specifications of locational preferences

Rather than assuming that idiosyncratic choices of locations are driven by additive preference heterogeneity as in the main text, we consider here a more general formulation where households have idiosyncratic productivity in both locations, as well as idiosyncratic preferences for location. They are therefore characterized by $\Theta = \{w_0, w_1, \phi\}$ distributed according to the CDF $F(.)$. Furthermore, we do not restrict the preference heterogeneity to enter additively in the utility function. That is, we assume that

$$u_j(\Theta) = U\left(c, h, a_j, \frac{z}{w_j}, \phi\right).$$

We first discuss how the results carry through to these more general cases. We then discuss the pitfalls of non-additive idiosyncratic preferences in the context of normative questions.

The logic of the derivations in Section 6 is unchanged, but notations need to be adjusted. In particular, we define the share of households who live in Distressed when the transfer is of size t as

$$L_1(t) = \int_{\Theta \in \mathbb{R}^4} j^*(\Theta, t)\, dF(\Theta).$$

We have

$$\frac{dL_1}{dt} = \lim_{t \to 0} \int_{\Theta \in \mathbb{R}^4} \left[\frac{j^*(\Theta, t) - j^*(\Theta, 0)}{t}\right] dF(\Theta)$$

The fiscal cost of movers still corresponds to the earnings losses of movers, which now writes more generally

$$\begin{aligned}
\frac{dB}{dt} &= \lim_{t \to 0} \int_{\Theta \in \mathbb{R}^4} \left[\frac{[j^*(\Theta, t) - j^*(\Theta, 0)]\,[T(z_1^*(\Theta, t)) - T(z_0^*(\Theta, 0))]}{t}\right] dF(\Theta) \\
&= \lim_{t \to 0} \int_{\Theta \in \mathbb{R}^4} \left[\frac{j^*(\Theta, t) - j^*(\Theta, 0)}{t}\,[T(z_1^*(\Theta, 0)) - T(z_0^*(\Theta, 0))]\right] dF(\Theta)
\end{aligned}$$

where the last line follows because $T(z_1^*(\Theta, t)) = T(z_1^*(\Theta, 0))$: absent an income effect on labor supply, stayers do not adjust their earnings following a lump-sum tax/subsidy. We write this expression with a more convenient notational shortcut

$$\begin{aligned}
\frac{dB}{dt} &= \mathbb{E}[T(z_1^*(\Theta, 0)) - T(z_0^*(\Theta, 0))\,|move]P(move) + \mathbb{E}[T(z_1^*(\Theta, 0)) - T(z_0^*(\Theta, 0))\,|stay]P(stay) \\
&= \mathbb{E}[T(z_1^*(\Theta, 0)) - T(z_0^*(\Theta, 0))\,|move]\frac{dL_1}{dt}
\end{aligned}$$

Lemma 1 writes as follows (nothing is changed in the equity computations, only in the efficiency cost computations): The first order effect on welfare of a small PBR reform starting from a place-blind system is

$$\frac{dSWF}{dt} = \bar{\lambda}_1 - \bar{\lambda}_0 + \frac{dL_1}{dt}\mathbb{E}_\Theta\left\{[T(z_1^*(.,0)) - T(z_0^*(.,0))]\,|move\right\} \tag{B.8}$$

Similar notational adjustments hold for Proposition 1. Technical results are therefore similar to what is in the main text with an additive formulation of idiosyncratic preferences, except that they call for a more cumbersome notation. We now discuss the advantage of choosing additively separable idiosyncratic preferences for location when it comes to normative questions.

Pitfalls of non-additive idiosyncratic tastes With additively separable idiosyncratic preferences for location, the social welfare weights $\lambda(\Theta)$ are not direct functions of the idiosyncratic preference ϕ - they are only indirectly impacted by idiosyncratic preferences through their effect on choice of city j. The reason why this is an advantage is that welfare weights – hence welfare implications of policies – do not depend on the specification and values of the unobserved taste ϕ, they only depend on the observed location choice j. In contrast, when the λ's directly depend on the value of ϕ, the definition of ϕ obviously matters for welfare. Unfortunately, as we show in the example below, one can easily build examples where two alternative models of ϕ lead to observationally equivalent equilibria, hence they cannot be disentangled using data, but have opposite welfare implications. It makes it undesirable to rest a normative argument on such a model.[13] Finally, it is easy to see that a similar argument applies to the case where idiosyncratic preferences for location are additively separable but the planner has concave preferences over levels of indirect utility, i.e. $SWF = \int G\left(V^{\Theta}\right) d\Theta$.

Example Assume that households are indexed by $(\theta, \varepsilon_0, \varepsilon_1)$ where the ε's are idiosyncratic preferences for 0 and 1. They have the simple utility function

$$U\left(c + \varepsilon_j\right)$$

with $U(.)$ concave. Households supply labor inelastically. Type θ gets income $z_j\left(\theta\right)$ in city j. Households choose city $j = 1$ iff

$$U\left(z_1\left(\theta\right) + \varepsilon_1\right) > U\left(z_0\left(\theta\right) + \varepsilon_0\right),$$

i.e., iff

$$\varepsilon_1 - \varepsilon_0 > z_0\left(\theta\right) - z_1\left(\theta\right). \tag{B.9}$$

Note that the values of ε_1 and ε_0 separately play no role in any of the observable choices of households, so that ε_1 and ε_0 are not separately identified. We then consider two alternative models: in model (a), $\varepsilon_0 = 0$ while $\varepsilon_1 = \phi$ is an iid random variable with some positive variance. In model (b), $\varepsilon_1 = 0$ while $\varepsilon_0 = -\phi$. Both models can rationalize the exact same same sorting equilibrium. The two models are therefore observationally equivalent. Interestingly though, they have opposite welfare (place-based related) implications. To make the point very stark, assume that $z_0\left(\theta\right) = z_1\left(\theta\right) \left(\equiv z\left(\theta\right)\right)$. We compute $\lambda(\Theta)$ the social welfare weight of type $\Theta = \{\theta, \phi\}$ in city j

$$\lambda(\Theta) = U'\left(z\left(\theta\right) + \varepsilon_j\right)\frac{1}{\Gamma}$$

We now compare λ_0 in location 0 to λ_1 in location 1, which correspond to the same income, to determine the direction of desirability of redistribution across locations implied by the two models. In both models, households are in 1 iff $\phi > 0$.

[13] Davis and Gregory (2022) discuss a related point: multiplicative preference heterogeneity, commonly used in Economic Geography model, is typically not identified but influences the marginal utility of consumption, and therefore social preferences. They propose to adjust the planner's problem to neutralize the influence of these preferences on marginal utility of consumption.

Therefore, given that U is concave, and $\phi > 0$ when 1 is chosen, we have in model (a)

$$\lambda_1 = U'(z+\phi)\frac{1}{\Gamma} < U'(z)\frac{1}{\Gamma} = \lambda_0$$

and redistribution from 1 to 0 is desirable.

In model (b), $\phi > 0$ too when 1 is chosen, so $\phi \leq 0$ when 0 is chosen. Therefore, in contrast to model (a)

$$\lambda_1 = U'(z)\frac{1}{\Gamma} > U'(z-\phi)\frac{1}{\Gamma} = \lambda_0$$

and redistribution from 0 to 1 is desirable. We conclude that two observationally-equivalent model predict opposite directions of redistribution. This example illustrate the pitfalls of allowing welfare weights to directly depend on idiosyncratic tastes, rather than indirectly based on city choice only.

B.3.3 Example 4 (stylized case of place-based redistribution desirability under skill-taste correlation)

The fact that $\overline{\lambda}_1 - \overline{\lambda}_0 > 0$ follows directly from Example 1. Although Example 1 assumes away taxes, the argument holds for any place-blind tax system because it only requires optimal earnings to be the same across locations. Conditional on j, the FOC for earnings is given by

$$[1 - T'(z_j^*(\theta))]w(\theta) = \psi'\left(\frac{z_j^*(\theta)}{w(\theta)}\right).$$

Because of the lack of income effects, t does not affect labor supply. No other primitive is location-specific, so $z_0^*(\theta) = z_1^*(\theta)$, for all θ. Then, the fiscal externality departing from $t = 0$ is zero, yielding the result.

B.3.4 Example 5 (stylized case of place-based redistribution desirability under comparative advantage)

Assume $\theta \in [\underline{\theta}, \overline{\theta}]$, with $\underline{\theta} \geq 0$ and $\overline{\theta} < \infty$, and that $\omega(\theta)$ is decreasing in θ. We consider the utility function

$$z - T(z) - r_j + a_j - \frac{\eta}{1+\eta}\left(\frac{z}{w_j(\theta)}\right)^{\frac{1+\eta}{\eta}} + j\phi$$

and a wage function $w_1(\theta) = \theta$ for all θ, $w_0(\theta) = \theta$ if $\theta < \theta^*$, and $w_0(\theta) = \theta + b$ if $\theta \geq \theta^*$, with $b > 0$ and $\theta^* \in (\underline{\theta}, \overline{\theta})$. WLOG, we assume that $a_0 = a_1 = r_0 = r_1 = 0$.

The condition that guarantees that a household (θ, ϕ) prefers Elsewhere is $\phi \leq \phi_\theta$, with $\phi_\theta = v_0(\theta) - v_1(\theta)$, where $v_j(\theta)$ is the indirect sub-utility when residing in j. Since the wage is the only primitive that varies across place, and, therefore, the only place-specific primitive that affects labor supply, we have that $z_j^*(\theta) = z^*(w_j(\theta))$. Therefore, we can rewrite $v_j(\theta) \equiv \mathcal{V}(w_j(\theta))$, where $\mathcal{V}(w_j(\theta)) = z^*(w_j(\theta)) - T(z^*(w_j(\theta))) - \frac{\eta}{1+\eta}\left(\frac{z^*(w_j(\theta))}{w_j(\theta)}\right)^{\frac{1+\eta}{\eta}}$, so $\phi_\theta = \mathcal{V}(w_0(\theta)) - \mathcal{V}(w_1(\theta))$.

When $\theta < \theta^*$, $w_0(\theta) = w_1(\theta)$, so $\phi_\theta = 0$. When $\theta \geq \theta^*$, define $z_D(\theta)$ as the optimal earnings when facing the wage schedule of Distressed, that is, $z_D(\theta) = \arg\max_z \left(z - T(z) - \frac{\eta}{1+\eta}\left(\frac{z}{w_1(\theta)}\right)^{\frac{1+\eta}{\eta}}\right)$. Assume

$z_D(\theta) > 0$ for all $\theta \geq \theta^*$.[14] Then, for $\theta \geq \theta^*$

$$\begin{aligned}\mathcal{V}(w_1(\theta)) &= z_D(\theta) - T(z_D(\theta)) - \frac{\eta}{1+\eta}\left(\frac{z_D(\theta)}{w_1(\theta)}\right)^{\frac{1+\eta}{\eta}} \\ &< z_D(\theta) - T(z_D(\theta)) - \frac{\eta}{1+\eta}\left(\frac{z_D(\theta)}{w_0(\theta)}\right)^{\frac{1+\eta}{\eta}} \\ &\leq \mathcal{V}(w_0(\theta)),\end{aligned}$$

where we used that $w_0(\theta) > w_1(\theta)$, for all $\theta \geq \theta^*$. It follows that $\phi_\theta > 0$ for all $\theta \geq \theta^*$ and, therefore, that $\min_{\theta\in[\theta^*,\overline{\theta}]} \phi_\theta \equiv \phi_\theta^{\min} > 0$.[15]

Consider the introduction of a small place-based transfer t such that $(\phi_\theta^{\min} - \overline{\phi}) > t > 0$. We now show that this transfer can be welfare improving in two separate cases.

Case with no taste heterogeneity Assume ϕ equals $\overline{\phi}$ for all individuals, with $0 < \overline{\phi} < \phi_\theta^{\min}$. It follows that $G(0) = 0$ and $G(\phi_\theta^{\min}) = 1$. Thus, all individuals with $\theta < \theta^*$ inframarginally locate in Distressed, and all individuals with $\theta \geq \theta^*$ inframarginally locate in Elsewhere. It follows that the transfer t is unambiguously welfare-improving starting from a place-blind tax system. For $\theta \geq \theta^*$, $\phi_\theta \geq \phi_\theta^{\min} - t$, which implies $\phi_\theta > \overline{\phi}$. Hence, no type with $\theta \geq \theta^*$ moves in response to the tax reform if $t < \phi_\theta^{\min} - \overline{\phi}$. Likewise, individuals with $\theta < \theta^*$ remain inframarginal given that $t > 0$. Thus, the reform does not generate fiscal externalities. Since $\omega(\theta)$ decreases with θ, the reform generates equity gains. Therefore, the reform is welfare-improving.

Case with taste heterogeneity Consider a taste distribution $G_\theta(\phi)$ with support $[\underline{\phi}, \overline{\phi}]$, where $\underline{\phi} < 0$ and $0 < \overline{\phi} < \phi_\theta^{\min}$. Individuals with $\theta \geq \theta^*$ remain inframarginally located in Elsewhere and do not migrate after the introduction of t. Since 0 is in the support of G_θ, individuals with $\theta < \theta^*$ split between the two communities. Suppose $G_\theta(0)$ does not depend on θ. Then all types $\theta < \theta^*$ have the same split between communities. Hence, after the introduction of t, a share of individuals with $\theta < \theta^*$ may migrate to Distressed. However, any such moves will not have fiscal effects because these individuals do not change their earnings after migrating. Therefore, the introduction of t is welfare-improving.

B.3.5 Example 6 (place-based redistribution is superfluous under income-based sorting and homogeneous preferences)

Consider an equilibrium with place-blind tax system $T(\cdot)$ and place-based transfers t_j, where t_j is a lump-sum tax on j. Utility is given by $U(z - T(z) - t_j - r_j, a_j) - \psi\left(\frac{z}{\theta}\right)$, with $U_c > 0$, $U_{cc} < 0$, $\psi'' > 0$. An individual with income z is located in $j(z) = \arg\max_j \{U(z - T(z) - t_j - r_j, a_j)\}$ and gets an (indirect) sub-utility of consumption denoted $V(z, T(\cdot), t_0, t_1)$. We show that, in this context, place-based transfers are superfluous compared to using the income tax alone.

The proof closely follows Kaplow (2006). Consider a new tax system $\widetilde{T}(z)$ with no place-based transfers, which is chosen such that $\forall z$, $V(z, T(\cdot), t_0, t_1) = V(z, \tilde{T}(\cdot), 0, 0)$. Such a tax schedule exists due to the continuity of the sub-utility $U(z, a)$ in income. Since $V(z, T(\cdot), t_0, t_1) - \psi\left(\frac{z}{\theta}\right) = V(z, \tilde{T}(\cdot), 0, 0) - \psi\left(\frac{z}{\theta}\right)$ by construction of $\tilde{T}(\cdot)$ for all z, individuals make the same optimal choice of labor z in the new tax system

[14] Any incentive compatible tax system should incentivize individuals with $\theta > 0$ to have positive labor supply. Then, this assumption is innocuous if we assume the tax system is optimal.

[15] To see why, consider any convergent sequence $\theta^{(n)}$ with $\theta^{(n)} \in [\theta^*, \overline{\theta}]$ such that $\phi_\theta^{(n)} \geq \phi_\theta^{(n+1)}$. Denote the limit as $(\theta^{\lim}, \phi_\theta^{\lim})$. Because $[\theta^*, \overline{\theta}]$ is closed, then $\theta^{\lim} \in [\theta^*, \overline{\theta}]$, so $\phi_\theta^{(n)} \geq \phi_\theta^{\lim} > 0$.

(this result comes from the assumption of separability of labor disutility). Therefore, individuals have the same utility as with the initial tax system.

We now show that under this new tax system, government revenues increase. First, for individuals whose choice of j did not change, taxes paid are unchanged. Second, consider individuals who changed their location choice from i to $j \neq i$ because they strictly prefer j to i. Since i was their optimal choice initially, $V(z, T(\cdot), t_0, t_1) = U(z - \tilde{T}(z) - r_i, a_i)$. Now assume (by contradiction) that under the new tax system where j is strictly preferred to i, the individual could still afford the initial city i. Since they can afford it but don't choose it, it must be that $V(z, \tilde{T}(\cdot), 0, 0) = U(z - \tilde{T}(z) - r_j, a_j) > U(z - \tilde{T}(z) - r_i, a_i)$, where the first equality is the definition of the indirect utility in the new equilibrium, and the inequality comes from the assumption that one can still afford i but does not choose it. In turn, $V(z, T(\cdot), t_0, t_1) > U(z - \tilde{T}(z) - r_i, a_i)$, using the the sub-utility of consumption did not change by construction of the new tax schedule. This yields a contradiction. It follows that these individuals cannot afford i in the new equilibrium, i.e. $z - \tilde{T}(z) - r_i < 0$, while initially $z - T(z) - t_i - r_i = 0$. It follows that $\tilde{T}(z) > T(z) + t_i$, so that these individuals pay more taxes and government revenues increase.

We have built an income tax reform that gets rid of place-based transfers, keeps the utility of all individuals constant and increases government revenues. Hence, the place-based transfers were superfluous. This argument holds for an arbitrary T function; in particular, for the optimal one.

B.4 Section 7

B.4.1 Proposition 2

With the assumption that the utility function satisfies single-crossing, we proceed by solving the "relaxed" version of the screening problem that replaces the global incentive compatibility constraints with local incentive compatibility constraints that become sufficient when allocations are monotone in skill within each location. We take the usual approach of dropping the monotonicity condition, solving the resulting version of the planning problem, and then verifying monotonicity ex-post in our numerical simulations.

Following the notation developed in the main text, let indirect sub-utilities be denoted by $v_j(\theta)$ and indirect utility be denoted by $V_j(\theta, \phi) = v_j(\theta) + j\phi$. Let $j^*(\theta, \phi)$ denote optimal location decision given primitives (θ, ϕ). Then, we define $V(\theta, \phi) = j^*(\theta, \phi)V_1(\theta, \phi) + (1 - j^*(\theta, \phi))V_0(\theta, \phi)$ as the utility of individuals of type (θ, ϕ) at the optimal location decision. We similarly define $T(\theta, \phi)$ as the taxes paid by individuals of type (θ, ϕ) at the optimal location decision. Using the the definition of ϕ_θ, housing demands are given by $H_0^D(r) = \int_{\underline{\theta}}^{\bar{\theta}} G_\theta(\phi_\theta) f(\theta) d\theta$ and $H_1^D(r) = \int_{\underline{\theta}}^{\bar{\theta}} (1 - G_\theta(\phi_\theta)) f(\theta) d\theta$, where $f(\theta) = dF(\theta)$ is the pdf of θ and housing rents, $r = (r_0, r_1)$, are implicit in ϕ_θ. To simplify notation, the derivation omits the asterisk superscript to denote optimal decisions since incentive compatibility ensures that outcome.

The planner's problem consists on maximizing the social welfare function (SWF) subject to incentive compatibility constraints, total budget constraint, and housing market clearing conditions. As discussed in the main text, we consider Pareto weights $\omega_H(\theta)$ that are only functions of skill θ. WLOG, we normalize weights on renters so that they sum to 1 across skills: $\mathbb{E}_\theta[\omega_H] = \int_{\underline{\theta}}^{\bar{\theta}} \omega_H(\theta) f(\theta) d\theta = 1$. Since the planner does not observe (θ, ϕ), she is constrained to second-best allocations with taxes as a function of earnings and location. The problem can be written as

$$\max_{c,z,T} \int_{\underline{\theta}}^{\bar{\theta}} \omega_H(\theta) \left[\int_{\underline{\phi}}^{\bar{\phi}} V(\theta,\phi)\; g_\theta(\phi)\, d\phi \right] f(\theta)\, d\theta + \omega_L \left[\Pi_0(r_0) + \Pi_1(r_1) \right],$$

$$\text{s.t.} \int_{\underline{\theta}}^{\bar{\theta}} \left[\int_{\underline{\phi}}^{\bar{\phi}} T(\theta,\phi)\; g_\theta(\phi)\, d\phi \right] f(\theta)\, d\theta \geq R,$$

$$\int_{\underline{\theta}}^{\bar{\theta}} \left(G_\theta(\phi_\theta) - \underline{H}_0 r_0^{\varrho_0} \right) f(\theta)\, d\theta = 0, \qquad \int_{\underline{\theta}}^{\bar{\theta}} \left([1 - G_\theta(\phi_\theta)] - \underline{H}_1 r_1^{\varrho_1} \right) f(\theta)\, d\theta = 0,$$

$$\text{and } \forall \theta', \phi' : \quad \hat{V}(\theta,\phi;\theta,\phi) \geq \hat{V}(\theta',\phi';\theta,\phi),$$

where $g_\theta(\phi) = dG_\theta(\phi)$ is the marginal conditional distribution of ϕ given θ and $\hat{V}(\theta',\phi';\theta,\phi)$ is the utility an individual of type (θ,ϕ) gets from mimicking an individual of type (θ',ϕ'), with $\hat{V}(\theta,\phi;\theta,\phi) = V(\theta,\phi)$. The first constraint is the budget constraint. The second constraints are the housing market equilibrium in both locations. The third set of constraints are the incentive compatibility constraints.

Recalling that $\phi_\theta = v_0(\theta) - v_1(\theta)$, we can write

$$\int_{\underline{\phi}}^{\bar{\phi}} V(\theta,\phi)\; g_\theta(\phi)\, d\phi = v_0(\theta) G_\theta(\phi_\theta) + v_1(\theta)\left(1 - G_\theta(\phi_\theta)\right) + \int_{\phi_\theta}^{\bar{\phi}} \phi g_\theta(\phi) d\phi. \tag{B.10}$$

Also, let $\Psi(.,a)$ be the inverse (reciprocal) of $U(.,a)$, which is well-defined since $U_c > 0$. Then, we can write

$$T_j(z_j(\theta)) = z_j(\theta) - \Psi\left(v_j(\theta) + \psi\left(\frac{z_j(\theta)}{w_j(\theta)} \right), a_j \right) - r_j.$$

We denote by $\Psi_j(\theta)$ the function $\Psi(.,a)$ evaluated at the optimum of types θ residing in j.

With these expressions, we can recast the mechanism design problem that precedes as an optimal control problem. We define the set $x_j = \{x_j(\theta)\}$ as the collection of $x_j(\theta)$ for all θ, for $x \in \{v, z, r\}$. Then, the planner's problem can be expressed as

$$\max_{v_0,v_1,z_0,z_1,r_0,r_1} \int_{\underline{\theta}}^{\bar{\theta}} \omega_H(\theta) \left[v_0(\theta)\; G_\theta(\phi_\theta) + v_1(\theta)\; (1 - G_\theta(\phi_\theta)) + \int_{\phi_\theta}^{\bar{\phi}} \phi g_\theta(\phi)\, d\phi \right] f(\theta)\, d\theta$$

$$+ \int_{\underline{\theta}}^{\bar{\theta}} \omega_L \left(\Pi_0(r_0) + \Pi_1(r_1) \right) f(\theta) d\theta,$$

$$\text{s.t.} \int_{\underline{\theta}}^{\bar{\theta}} \left\{ \left[z_0(\theta) - \Psi\left(v_0(\theta) + \psi\left(\frac{z_0(\theta)}{w_0(\theta)} \right), a_0 \right) - r_0 \right] G_\theta(\phi_\theta) \right.$$

$$\left. + \left[z_1(\theta) - \Psi\left(v_1(\theta) + \psi\left(\frac{z_1(\theta)}{w_1(\theta)} \right), a_1 \right) - r_1 \right] (1 - G_\theta(\phi_\theta)) f(\theta)\; d\theta \geq R, \right. \tag{Γ}$$

$$\text{and} \int_{\underline{\theta}}^{\bar{\theta}} \left(G_\theta(\phi_\theta) - \underline{H}_0 r_0^{\varrho_0} \right) f(\theta)\, d\theta = 0, \qquad \int_{\underline{\theta}}^{\bar{\theta}} \left([1 - G_\theta(\phi_\theta)] - \underline{H}_1 r_1^{\varrho_1} \right) f(\theta)\, d\theta = 0, \tag{χ_j}$$

$$\text{and } \forall \theta, j : \quad r_j'(\theta) = 0, \tag{$A_j(\theta)$}$$

$$\text{and } \forall \theta, j : \quad v_j'(\theta) = \psi'\left(\frac{z_j(\theta)}{w_j(\theta)} \right) z_j(\theta) \left(\frac{\frac{\partial}{\partial \theta} w_j(\theta)}{w_j(\theta)^2} \right). \tag{$\mu_j(\theta)$}$$

The optimal tax problem, then, is equivalent to choosing allocations of earnings, utility, and rents that maximize welfare subject to respecting the government budget constraint, the housing market equilibrium, and the incentive compatibility constraint. These allocations can then be implemented with the tax system according to the expressions above. To accommodate notation and definitions to the optimal control setup, we treat housing rents as state variables, so we introduce the law of movement of rents that imposes constant rents within location across skills ($r_j'(\theta) = 0$) and, therefore, we indistinctly use r_j and $r_j(\theta)$ below. In this problem, Γ is the budget constraint multiplier, χ_j are the multipliers of the housing market clearing conditions, $A_j(\theta)$ are the multipliers for the within-location constant rents constraint, and $\mu_j(\theta)$ is the multiplier of the incentive compatibility (IC) constraint. Note that, given the separability assumption on ϕ, the "intensive margin" IC constraints are unchanged compared to the standard optimal income tax problem. The location choice IC constraint is explicitly incorporated in the objective function and constraints by substituting for $j^*(\theta, \phi)$ using the threshold rule based on ϕ_θ.[16]

Assuming no bunching, we can derive the optimal schedule of location-specific taxes by solving the following Hamiltonian by pointwise maximization

$$\begin{aligned} H(\theta) \quad = \quad & \omega_H(\theta)\left(v_0(\theta)\ G_\theta(\phi_\theta) + v_1(\theta)\ (1 - G_\theta(\phi_\theta)) + \int_{\phi_\theta}^{\overline{\phi}} \phi g_\theta(\phi)\, d\phi\right) f(\theta) + \omega_L\left[\Pi_0(r_0) + \Pi_1(r_1)\right] f(\theta) \\ & +\Gamma\left\{\left[z_0(\theta) - \Psi\left(v_0(\theta) + \psi\left(\frac{z_0(\theta)}{w_0(\theta)}\right), a_0\right) - r_0\right] G_\theta(\phi_\theta)\right. \\ & \left.+\left[z_1(\theta) - \Psi\left(v_1(\theta) + \psi\left(\frac{z_1(\theta)}{w_1(\theta)}\right), a_1\right) - r_1\right](1 - G_\theta(\phi_\theta))\right\} f(\theta) \\ & +\chi_0\left(G_\theta(\phi_\theta) - \underline{H}_0 r_0^{\varrho_0}\right) f(\theta) + \chi_1\left([1 - G_\theta(\phi_\theta)] - \underline{H}_1 r_1^{\varrho_1}\right) f(\theta) \\ & +\sum_j \mu_j(\theta)\, \psi'\left(\frac{z_j(\theta)}{w_j(\theta)}\right) z_j(\theta)\left(\frac{\frac{\partial}{\partial\theta} w_j(\theta)}{w_j(\theta)^2}\right) + \sum_j A_j(\theta) \cdot 0, \end{aligned}$$

where the state vectors are (v_0, v_1, r_0, r_1) and the control vectors are (z_0, z_1). To simplify notation, below we denote $\frac{\partial w_j(\theta)}{\partial\theta} = w_j'(\theta)$. A set of sufficient condition to maximize $H(\theta)$ is given by

$$\frac{\partial H(\theta)}{\partial v_j(\theta)} = -\mu_j'(\theta), \quad \frac{\partial H(\theta)}{\partial r_j} = -A_j'(\theta), \quad \frac{\partial H(\theta)}{\partial z_j(\theta)} = 0,$$

and the transversality conditions $\mu_j(\underline{\theta}) = \mu_j(\bar{\theta}) = A_j(\underline{\theta}) = A_j(\bar{\theta}) = 0$, for $j \in \{0, 1\}$.

To solve for the first-order-condition (FOC) with respect to $z_j(\theta)$, note that $\Psi_j'(\theta) = \frac{1}{\frac{dU_{cj}(\theta)}{dc_j(\theta)}} = \frac{1}{\frac{dU_{cj}(\theta)}{dz_j(\theta)}}$, where $\Psi_j'(\theta)$ is the derivative of the function $\Psi(., a)$ with respect to its argument, evaluated at the optimal choices of types θ located in j. Then, $\Psi_j'(\theta) = \frac{1}{U_{cj}(\theta)}$, where $U_{cj}(\theta)$ is the marginal utility of consumption

[16] Since location is observed, the IC constraint states that $\forall \theta', \theta : \hat{v}_j(\theta; \theta) \geq \hat{v}_j(\theta'; \theta)$ within each location, where $\hat{v}_j(\theta'; \theta) = c_j(\theta') - \psi\left(\frac{z_j(\theta')}{w_j(\theta)}\right)$ and $\hat{v}_j(\theta; \theta) = v_j(\theta)$. The "revelation problem" of each household solves

$$\max_{\theta'} c_j(\theta') - \psi\left(\frac{z_j(\theta')}{w_j(\theta)}\right),$$

which in an incentive compatible allocation optimally leads to $\theta' = \theta$. The first-order-condition is given by $\frac{d}{d\theta'}\left[c_j(\theta') - \psi\left(\frac{z_j(\theta')}{w_j(\theta)}\right)\right] = 0$ evaluated at $\theta' = \theta$. Then, $\frac{d}{d\theta} v_j(\theta) = \frac{d}{d\theta'}\left[c_j(\theta') - \psi\left(\frac{z_j(\theta')}{w_j(\theta)}\right)\right]_{\theta'=\theta} - \frac{\partial}{\partial\theta}\left[\psi\left(\frac{z_j(\theta')}{w_j(\theta)}\right)\right]_{\theta'=\theta} = \psi'\left(\frac{z_j(\theta)}{w_j\theta}\right)\frac{z_j(\theta) w_j'(\theta)}{w_j(\theta)^2}$, where the last equality uses the first-order-condition. The IC constraint becomes

$$v_j'(\theta) = \psi'\left(\frac{z_j(\theta)}{w_j(\theta)}\right)\frac{z_j(\theta)\, w_j'(\theta)}{w_j(\theta)^2}.$$

In the standard problem, $w_j'(\theta) = 1$, which yields the standard expression for the IC constraint.

evaluated at the optimal allocation which, as discussed in the main text, is a key input to the marginal welfare weight computation. We also note that, conditional on $v_j(\theta)$, $\frac{d\phi_\theta}{dz_j(\theta)} = 0$. Then

$$\begin{aligned}
\Gamma f(\theta) G_\theta(\phi_\theta)\left[1-\frac{1}{U_{c0}}\frac{1}{w_0(\theta)}\psi'\left(\frac{z_0(\theta)}{w_0(\theta)}\right)\right]+\mu_0(\theta)\frac{w_0'(\theta)}{w_0(\theta)}\frac{1}{w_0(\theta)}\left[\psi''\left(\frac{z_0(\theta)}{w_0(\theta)}\right)\frac{z_0(\theta)}{w_0(\theta)}+\psi'\left(\frac{z_0(\theta)}{w_0\theta}\right)\right] &= 0,\\
\Gamma f(\theta)(1-G_\theta(\phi_\theta))\left[1-\frac{1}{U_{c1}}\frac{1}{w_1(\theta)}\psi'\left(\frac{z_1(\theta)}{w_1(\theta)}\right)\right]+\mu_1(\theta)\frac{w_1'(\theta)}{w_1(\theta)}\frac{1}{w_1(\theta)}\left[\psi''\left(\frac{z_1(\theta)}{w_1(\theta)}\right)\frac{z_1(\theta)}{w_1(\theta)}+\psi'\left(\frac{z_1(\theta)}{w_1\theta}\right)\right] &= 0.
\end{aligned}$$

The FOC of the individual households is given by $\left(1-T_j'(z_j(\theta))\right) = \frac{\psi'\left(\frac{z_j(\theta)}{w_j(\theta)}\right)\frac{1}{w_j(\theta)}}{v_{cj}'}$. Then, by differentiating U_{cj} and properly manipulating terms, we can write utility primitives as a function of elasticities (see Saez (2001) for details)

$$\eta_j^U(\theta) = \frac{\frac{\psi'\left(\frac{z_j(\theta)}{w_j(\theta)}\right)}{\ell_j(\theta)}+\frac{\psi'(\ell_j(\theta))^2}{U_{cj}^{*2}}U_{cj}'}{\psi''(\ell_j(\theta))-\frac{\psi'(\ell_j(\theta))^2}{U_{cj}^2}U_{cj}'},\quad \eta_j^C(\theta) = \frac{\frac{\psi'(\ell_j(\theta))}{\ell_j(\theta)}}{\psi''(\ell_j(\theta))-\frac{\psi'(\ell_j(\theta))^2}{U_{cj}^2}U_{cj}'},$$

where $l_j(\theta) = \frac{z_j(\theta)}{w_j(\theta)}$. Then we can write

$$\frac{1+\eta_j^U(\theta)}{\eta_j^C(\theta)} = \frac{\psi''(\ell_j(\theta))+\frac{\psi'(\ell_j(\theta))}{\ell_j(\theta)}}{\frac{\psi'(\ell_j(\theta))}{\ell_j(\theta)}}. \tag{B.11}$$

This implies that

$$\psi''\left(\frac{z_j(\theta)}{w_j(\theta)}\right)\frac{z_j(\theta)}{w_j(\theta)}+\psi'\left(\frac{z_j(\theta)}{w_j(\theta)}\right) = \psi'\left(\frac{z_j(\theta)}{w_j(\theta)}\right)\left[\frac{1+\eta_j^U(\theta)}{\eta_j^C(\theta)}\right],$$

so the planner's FOC with respect to $z_j(\theta)$ reduces to

$$\Gamma f(\theta) G_\theta(\phi_\theta)\left[1-\frac{1}{U_{c0}}\frac{1}{w_0(\theta)}\psi'\left(\frac{z_0(\theta)}{w_0(\theta)}\right)\right]+\mu_0(\theta)\frac{\gamma_0(\theta)}{\theta}\frac{1}{w_0(\theta)}\psi'\left(\frac{z_0(\theta)}{w_0(\theta)}\right)\left[\frac{1+\eta_0^U(\theta)}{\eta_0^C(\theta)}\right] = 0, \tag{B.12}$$

$$\Gamma f(\theta)(1-G_\theta(\phi_\theta))\left[1-\frac{1}{U_{c1}}\frac{1}{w_1(\theta)}\psi'\left(\frac{z_1(\theta)}{w_1(\theta)}\right)\right]+\mu_1(\theta)\frac{\gamma_1(\theta)}{\theta}\frac{1}{w_1(\theta)}\psi'\left(\frac{z_1(\theta)}{w_1(\theta)}\right)\left[\frac{1+\eta_1^U(\theta)}{\eta_1^C(\theta)}\right] = 0 \tag{B.13}$$

where we replaced $\gamma_j(\theta) = \frac{w_j'(\theta)\theta}{w_j(\theta)}$.

Noting that $\frac{\partial\phi_\theta}{\partial v_0(\theta)} = \frac{\partial(v_0(\theta)-v_1(\theta))}{\partial v_0(\theta)} = 1$, the FOC with respect to $v_0(\theta)$ yields

$$\omega_H(\theta)G_\theta(\phi_\theta)f(\theta)-\Gamma\frac{1}{U_{c0}(\theta)}G_\theta(\phi_\theta)f(\theta)-\Gamma\Delta\tau(\theta)\,g_\theta(\phi_\theta)f(\theta)+(\chi_0-\chi_1)\,g_\theta(\phi_\theta)f(\theta) = -\mu_0'(\theta),$$

where we used the notation $\Delta\tau(\theta) = T_1(z_1(\theta))-T_0(z_0(\theta))$. Also, the envelope theorem implies no first order welfare effect through the sorting decisions.[17] To solve for χ_j, consider the FOC with respect to r_0

$$\omega_L\Pi_0'(r_0)f(\theta)-\Gamma G_\theta(\phi_\theta)f(\theta)-\chi_0\varrho_0\underline{H}_0 r_0^{\varrho_0-1}f(\theta) = -A_0'(\theta).$$

[17] The total derivative of the first term within the planner's objective equals $G_\theta(\phi_\theta)+v_0(\theta)g_\theta(\phi_\theta)-v_1(\theta)g_\theta(\phi_\theta)-\phi_\theta g_\theta(\phi_\theta)$, so terms cancel out considering that $\phi_\theta = v_0(\theta)-v_1(\theta)$. Intuitively, marginal movers experience no welfare effect since they are indifferent between locations.

The transversality conditions imply $\int_{\underline{\theta}}^{\bar{\theta}} A_0'(\theta) = 0$. Also, the envelope theorem applied to the landlords' implies $\Pi_0'(r_0) = H_0$, with $H_0 = \underline{H}_0 r_0^{\varrho_0}$ total housing supply in location 0.[18] Noting that integrating $G_\theta(\phi_\theta) f(\theta)$ yields total housing demand in location 0 which, in equilibrium, is equal to H_0, integrating the previous expression over θ implies that $\chi_0 = \frac{(\omega_L - \Gamma) r_0}{\varrho_0}$. A similar logic solves for χ_1. Then, the FOC with respect to $v_0(\theta)$ can be reduced to:

$$\omega_H(\theta) G_\theta(\phi_\theta) f(\theta) - \Gamma \frac{1}{U_{c0}(\theta)} G_\theta(\phi_\theta) f(\theta) - \Gamma \Delta\tau(\theta)\, g_\theta(\phi_\theta) f(\theta) + (\omega_L - \Gamma)\left(\frac{r_0}{\varrho_0} - \frac{r_1}{\varrho_1}\right) g_\theta(\phi_\theta) f(\theta) = -\mu_0'(\theta). \tag{B.14}$$

An analog expression can be derived from the FOC with respect to $v_1(\theta)$

$$\omega_H(\theta)(1 - G_\theta(\phi_\theta)) f(\theta) - \Gamma \frac{1}{U_{c1}(\theta)} (1 - G_\theta(\phi_\theta)) f(\theta) + \Gamma \Delta\tau(\theta)\, g_\theta(\phi_\theta) f(\theta) \tag{B.15}$$

$$- (\omega_L - \Gamma)\left(\frac{r_0}{\varrho_0} - \frac{r_1}{\varrho_1}\right) g_\theta(\phi_\theta) f(\theta) = -\mu_1'(\theta). \tag{B.16}$$

Because of the transversality conditions, $-\int_{\underline{\theta}}^{\overline{\theta}} \mu_j'(\theta) d\theta = 0$. Then, adding the two FOCs with respect to $v_j(\theta)$, (B.14) and (B.16), integrating and using $\int_{\underline{\theta}}^{\bar{\theta}} \omega_H(\theta) f(\theta)\, d\theta = 1$ yields

$$\Gamma = \left(\int_{\underline{\theta}}^{\bar{\theta}} \left(\frac{1}{U_{c0}(\theta)} G_\theta(\phi_\theta) + \frac{1}{U_{c1}(\theta)} (1 - G_\theta(\phi_\theta)) \right) f(\theta)\, d\theta \right)^{-1},$$

so the marginal value of public funds coincides with the harmonic mean of the marginal utility of consumption across types and locations. Note that this expression makes clear why $\int_{\underline{\theta}}^{\bar{\theta}} \omega_H(\theta) f(\theta)\, d\theta = 1$ is WLOG, since for computing welfare weights, any set of Pareto weights that do not average to one will not play any role in the numerator (beyond its normalized version) after properly scaling the budget constraint multiplier.

In the final step, we solve out for the functions $\mu_j(\theta)$ that appear in the FOCs with respect to $z_j(\theta)$ and those with respect to $v_j(\theta)$. Specifically, we note that, since $\mu_j(\overline{\theta}) = 0$, $\mu_j(\theta) = -\int_\theta^{\overline{\theta}} \mu_j'(\theta) d\theta$. We then solve out for $\mu_0(\theta)$ in (B.12) by integrating $\mu_0'(\theta)$ over θ in (B.14) and equating the two. Plugging in the FOC of the houseold optimal income problem[19] leads to the following expression for $j = 0$

$$\frac{T_0'(z_0(\theta))}{1 - T_0'(z_0(\theta))} = \left(\frac{1 + \eta_0^U(\theta)}{\eta_0^C(\theta)}\right) \frac{\gamma_0(\theta)}{\theta f(\theta) G_\theta(\phi_\theta)} \times U_{c0}(\theta) \left\{ \int_\theta^{\overline{\theta}} \frac{1 - \lambda_0(s)}{U_{c0}(s)} G_s(\phi_s) f(s)\, ds + \Delta\tau^+(\theta) + (\lambda_L - 1)\, \Delta r^+(\theta) \right\},$$

where we used the shorthands $\Delta\tau^+(\theta) \equiv \int_\theta^{\bar{\theta}} [T_1(z_1^*(s)) - T_0(z_0^*(s))]\, g_s(\phi_s)\, dF(s)$, $\Delta r^+(\theta) \equiv \left(\frac{r_1}{\varrho_1} - \frac{r_0}{\varrho_0}\right)$.

[18] The profit maximization problem for the landlords is $\Pi_j = \max_{H_j} H_j r_j - \varphi(H_j)$ where $\varphi(H_j)$ is the cost function for housing supply, so the envelope theorem implies that $\Pi_j'(r_j) = H_j$. Optimal housing supply $H_j = \underline{H}_j r_j^{\varrho_j}$ arises then as the optimal solution when cost functions have standard isoelastic forms (see Appendix B.1).

[19] The FOC in location j is: $1 - T_j'(z) = \frac{\psi'\left(\frac{z}{w_j(\theta)}\right) \frac{1}{w_j(\theta)}}{v_j'(c)}$

$\int_{\theta}^{\bar{\theta}} g_s(\phi_s)\, dF(s)$, $\lambda_j(\theta) = \frac{\omega_H(\theta) U_{cj}(\theta)}{\Gamma}$, and $\lambda_L = \frac{\omega_L}{\Gamma}$. A similar procedure yields

$$\frac{T_1'(z_1(\theta))}{1 - T_1'(z_1(\theta))} = \left(\frac{1 + \eta_1^U(\theta)}{\eta_1^C(\theta)}\right) \frac{\gamma_1(\theta)}{\theta f(\theta)(1 - G_\theta(\phi_\theta))}$$
$$\times U_{c1}(\theta) \left\{ \int_{\theta}^{\bar{\theta}} \frac{1 - \lambda_1(s)}{U_{c1}(s)} (1 - G_s(\phi_s)) f(s)\, ds - \Delta\tau^+(\theta) - (\lambda_L - 1) \Delta r^+(\theta) \right\}.$$

Finally, to derive equation (10), we use the transversality condition $\int_{\underline{\theta}}^{\bar{\theta}} -\mu_0'(\theta)\, d\theta = 0$. Computing $\int_{\underline{\theta}}^{\bar{\theta}} -\mu_0'(\theta)\, d\theta$ from equation (B.14) yields

$$\int_{\underline{\theta}}^{\bar{\theta}} (\lambda_0(\theta) - 1) \frac{1}{U_{c0}(\theta)} G_\theta(\phi_\theta) f(\theta)\, d\theta - \Delta\tau^+(\underline{\theta}) - (\lambda_L - 1) \Delta r^+(\underline{\theta}) = 0. \tag{B.17}$$

A similar expression can be derived by using the transversality condition for $j = 1$

$$\int_{\underline{\theta}}^{\bar{\theta}} (\lambda_1(\theta) - 1) \frac{1}{U_{c1}(\theta)} (1 - G_\theta(\phi_\theta)) f(\theta)\, d\theta + \Delta\tau^+(\underline{\theta}) + (\varpi_L - 1) \Delta r^+(\underline{\theta}) = 0. \tag{B.18}$$

Multiplying (B.18) by L_0 and substracting (B.18) multiplied by L_1 yields

$$L_0 L_1 \int_{\underline{\theta}}^{\bar{\theta}} \left(\frac{1}{L_1} (\lambda_1(\theta) - 1) \frac{1}{U_{c1}(\theta)} (1 - G_\theta(\phi_\theta)) - \frac{1}{L_0} (\lambda_0(\theta) - 1) \frac{1}{U_{c0}(\theta)} G_\theta(\phi_\theta) \right) f(\theta)\, d\theta$$
$$+ \Delta\tau^+(\underline{\theta}) + (\varpi_L - 1) \Delta r^+(\underline{\theta}) = 0.$$

C Simulation Appendix

C.1 Additional relevant formulas

The following already-specified equations are used in the simulations: equations (8)-(10) (optimal taxes), (11) (parametric utility), and (B.4.1) (optimal multiplier). In addition, the following equations are used.

The conditional density of location tastes specified in Section 8 can be expressed as

$$g_\theta(\phi) = \frac{\exp\left(\frac{\phi - \mu + \beta F(\theta)}{\kappa}\right)}{\kappa \left(1 + \exp\left(\frac{\phi - \mu + \beta F(\theta)}{\kappa}\right)\right)^2} = \frac{1}{\kappa} G_\theta(\phi)(1 - G_\theta(\phi)).$$

Social marginal welfare weights on households are given by

$$\lambda_j(\theta) = \frac{1}{\Gamma \cdot (z_j(\theta) - T_j(z_j(\theta)) - r_j)},$$

where Γ is the budget constraint multiplier characterized in equation (B.4.1). The welfare weight on landlords is given by $\lambda_L = \omega_L / \Gamma$.

From equation (B.11) under our iso-elastic functional form, we have

$$\frac{1 + \eta_j^U(\theta)}{\eta_j^C(\theta)} = 1 + \frac{1}{\eta}.$$

which reduces the optimal marginal tax rate formulas to

$$\begin{aligned}\frac{T_0'(z_0(\theta))}{1-T_0'(z_0(\theta))} &= \left(1+\frac{1}{\eta}\right)\frac{\gamma_0}{\theta f(\theta)\,G_\theta(\phi_\theta)} \\ &\quad \times U_{c0}(\theta)\left\{\int_\theta^{\overline{\theta}}\frac{1-\lambda_0(s)}{U_{c0}(s)}G_s(\phi_s)\,dF(s)+\Delta\tau^+(\theta)+(\lambda_L-1)\,\Delta r^+(\theta)\right\}, \\ \frac{T_1'(z_1(\theta))}{1-T_1'(z_1(\theta))} &= \left(1+\frac{1}{\eta}\right)\frac{\gamma_1}{\theta f(\theta)\,(1-G_\theta(\phi_\theta))} \\ &\quad \times U_{c1}(\theta)\left\{\int_\theta^{\overline{\theta}}\frac{1-\lambda_1(s)}{U_{c1}(s)}(1-G_s(\phi_s))\,dF(s)-\Delta\tau^+(\theta)-(\lambda_L-1)\,\Delta r^+(\theta)\right\},\end{aligned}$$

where $U_{cj}(\theta) = (z_j(\theta) - T_j(z_j(\theta)) - r_j)^{-1}$.

C.2 Simulation procedure

Discrete skill distribution We work with a discrete approximation of the problem in θ-space. Consider a grid of θ of size N, such that $\theta \in \{\theta_1, ..., \theta_N\}$ with $\theta_1 = \underline{\theta}$, $\theta_N = \overline{\theta}$, and $\theta_i - \theta_{i-1} = \widehat{\Delta}$ for all i and with $\widehat{\Delta} = (\overline{\theta} - \underline{\theta})/(N-1)$ denoting the equally-sized bin width. The functions $T_j(z_j(\theta))$ are approximated by piece-wise linear tax systems with N different brackets for each j, where each income threshold is determined by the type-location-specific optimal earnings given the tax schedule. Concretely, let $T_j(z_j(\theta))$ be characterized by a demogrant intercept $T_j(z_j(\underline{\theta})) = -E_j$ and N marginal tax rates, $T_{ji}' = T_j'(z_j(\theta_i))$. Then

$$T_j(z_j(\theta_i)) = -E_j + \sum_{k=1}^{i} T_{ji}' \cdot (z_j(\theta_k) - z_j(\theta_{k-1})),$$

where $z_j(\theta_0)$ is normalized to 0 for $j \in \{0, 1\}$.

The discretization of θ also implies that F is approximated by a discrete CDF $\widehat{F}$ with corresponding PMF $\widehat{f}$:

$$\widehat{f}(\theta_i) = F\left(\theta_i + \frac{\widehat{\Delta}}{2}; \xi, \sigma, p\right) - F\left(\theta_i - \frac{\widehat{\Delta}}{2}; \xi, \sigma, p\right),$$

with $\widehat{F}(\theta_i) = \sum_{k=1}^{i} \widehat{f}(\theta_k)$. The PMF is normalized such that $\widehat{F}(\theta_N) = 1$.

We closely follow Mankiw et al. (2009) (MWY) for simulating the discrete PMF of θ. To ensure a mass of non-workers, we define $\underline{\theta} = 0$ and set $\widehat{f}(0) = 0.05$ as in MWY. We create an evenly-spaced grid of 50 points from zero to $\theta = 42.5$, the point in which MWY append the Pareto tail. This implies that $\widehat{\Delta} = 0.867$. For each point in this grid, we impute the PMF using the equation above, where F is the log-normal CDF evaluated at parameters $\xi = 2.757$ and $\sigma = 0.5611$, as estimated by MWY.[20]

We then build a second vector representing the Pareto tail with parameter $p = 2$ of the distribution by sequentially incrementing the θ-grid by the bin-width $\widehat{\Delta}$. Appending the Pareto tail vector to the previous vector yields a 578-entry θ-vector with a maximum value of $\overline{\theta} = 500.5$. We again recover the probabilities pertaining to each bin by applying the equation above using as F the Pareto CDF, as in MWY. Finally, we normalize the PMF to ensure the probabilities add up to one and are continuous at $\theta = 42.5$.

[20] The only exception is that we round up the lowest positive θ from 0.867 to 1, in order for comparative advantage to always yield higher wages in Elsewhere. This has no quantitative impact on our results.

Calibration In Section 8, we fix κ and find a solution of three unknowns to match the three moments as specified in the text. We proceed in Matlab by solving the non-linear system of equations using the `fsolve` command. In all calibrations, we test for multiple solutions by running local solvers on fine grids of initial conditions using `multistart`. In Section 9, we find a solution including κ as a fourth unknown and a migration elasticity (0.82 in the urban scenarios and 0.20 in the rural scenarios) as the fourth moment as specified in the text. This additional moment is jointly solved with the other three. To compute the additional simulated moment, we artificially increase the wage shifter W in Distressed by 1%, which should yield a population increase of 0.82% in the urban scenarios and 0.20% in the rural scenarios after the endogenous change in rents has been internalized. In this exercise, we exclude the mass point at $\theta = 0$ since, by construction, those households are not responsive to changes in W because their labor supply is always zero.

Optimization To solve for the optimal tax system given a set of calibrated parameters, we proceed by solving a non-linear system of $2N + 2 + 2$ equations in $2N + 2 + 2$ unknowns, again by using Matlab's `fsolve` command. The unknowns are $2N$ marginal tax rates, $\left\{T'_{ji}\right\}_{j\in\{0,1\},i\in\{1,\ldots,N\}}$, defined for each type in each location, 2 community-specific tax system intercepts, $\{E_j\}_{j\in\{0,1\}}$, and 2 rent prices, $\{r_j\}_{j\in\{0,1\}}$. The $2N + 2 + 2$ equations to solve are listed below. In what follows, it will be useful to write $T_j(z_j(\theta_i)) = -E_j + t_j(z_j(\theta_i))$, where $t_j(z_j(\theta_i)) = \sum_{k=1}^{i} T'_{ji} \cdot (z_j(\theta_k) - z_j(\theta_{k-1}))$ are gross taxes paid excluding the demogrant as defined above.

i. The routine solves for the discrete approximation of the MTR location-specific FOCs. To simplify notation, let $G_i = G_\theta(\phi_{\theta_i})$ and $g_i = g_\theta(\phi_{\theta_i})$. Then the $2N$ equations associated with the $2N$ MTR unknowns are given by:

$$\begin{aligned} \frac{T'_{0i}}{1-T'_{0i}} &= \left(1+\frac{1}{\eta}\right)\frac{\gamma_0}{\theta_i\left(\frac{\widehat{f}(\theta_i)}{\widehat{\Delta}}\right)G_i} \\ &\quad \times U_{c0}\left(\theta\right)\left\{\sum_{t=i+1}^{N}\frac{1-\lambda_0(\theta_t)}{U_{c0}(\theta_t)}G_t\widehat{f}(\theta_t)+\Delta\tau^+(\theta_i)+\left(\lambda_L-1\right)\Delta r^+(\theta_i)\right\}, \\ \frac{T'_{1i}}{1-T'_{1i}} &= \left(1+\frac{1}{\eta}\right)\frac{\gamma_1}{\theta_i\left(\frac{\widehat{f}(\theta_i)}{\widehat{\Delta}}\right)(1-G_i)} \\ &\quad \times U_{c1}\left(\theta\right)\left\{\sum_{t=i+1}^{N}\frac{1-\lambda_1(\theta_t)}{U_{c1}(\theta_t)}\left(1-G_t\right)\widehat{f}(\theta_t)-\Delta\tau^+(\theta_i)-\left(\lambda_L-1\right)\Delta r^+(\theta_i)\right\}, \end{aligned}$$

where $\Delta\tau^+(\theta_i) = \sum_{t=i+1}^{N}\left(t_1(z_1(\theta_t)) - E_1 - t_0(z_0(\theta_t)) + E_0\right) g_t\widehat{f}(\theta_t)$ and $\Delta r^+(\theta_i) = \left(\frac{r_1}{\varrho_1} - \frac{r_0}{\varrho_0}\right)\sum_{t=i+1}^{N} g_t\widehat{f}(\theta_t)$.

ii. The 2 equations associated with the 2 tax-system intercepts are the discrete approximations of the demogrant equations. First, we have the budget constraint:

$$L_0E_0 + (1-L_0)E_1 = \sum_{i=1}^{N}\left(t_0(z_0(\theta_i))G_i + t_1(z_1(\theta_i))\left(1-G_i\right)\right)\widehat{f}(\theta_i) - R,$$

with $L_0 = \sum_i^N G_i\widehat{f}(\theta_i)$ and $L_1 = 1 - L_0$. Second, we have the optimality condition on the difference in demogrants:

$$L_0 L_1 \sum_{i=1}^{N} \left(\frac{1}{L_1} (\lambda_1 (\theta_i) - 1) \frac{1}{U_{c1}(\theta_i)} (1 - G_i) - \frac{1}{L_0} (\lambda_0 (\theta_i) - 1) \frac{1}{U_{c0}(\theta_i)} G_i \right) \widehat{f} (\theta_i)$$
$$+ (\lambda_L - 1) \Delta r^+ (\underline{\theta}) + \Delta \tau^+ (\underline{\theta}) = 0.$$

iii. The 2 equations associated with the 2 rent levels are the housing market clearing conditions in each location, characterized by $L_j = \underline{H}_j r_j^{\varrho_j}$, for $j \in \{0, 1\}$.

C.3 Simulation inputs

Size and housing supply elasticities of Elsewhere and Distressed We use the 2016-2020 American Community Surveys (ACS) to establish sizes of Elsewhere and Distressed under current taxes. We do this by mapping EZs to 2010 Census tracts following a crosswalk provided to us by Patrick Kennedy, which we further crosswalk to 2020 Census tracts using data from Chetty et al. (2020) (accessed at https://opportunityinsights.org/data/). We then merge tract-level data from the 2020 ACS 5-Year estimates (covering 2016-2020) on population (table DP05), which was accessed at data.census.gov. To calculate the population share in Distressed, we sum the total population of all EZs tracts (either urban EZs or rural EZs, depending on the Distressed definition) and divide it by the total population.

The housing supply elasticities are computed using the tract-level housing supply elasticities of Baum-Snow and Han (2024). The estimates were accessed at https://sites.google.com/site/baumsnow/research in October 2022. We use the August 2022 vintage of the data. We use Baum-Snow and Han's preferred estimates, namely, the column "gamma01b_TYPE_FMM". To aggregate these tract-level elasticities up to the level of Distressed or Elsewhere, we take population-weighted averages.

Current tax system All calibrations are performed under an approximation to the current U.S. labor income tax system based on the Piketty et al. (2018) distributional national accounts (DINA) public micro files. We use the DINA public micro files for the years 2016-2020, pooled and inflating all dollar values to 2020. Attention is restricted to tax units (singles or couples) where all members are working age (i.e., are between 30 and 55 years old) and the tax unit earns nearly all of their income from labor. The latter restriction is operationalized by keeping only tax units for which wages and salaries (variable `flwag`) represent between 95% and 105% of total fiscal income including capital gains (variable `fiinc`). We define total taxes net of transfers as the sum of federal personal income taxes (`ditaf`), state personal income taxes (`ditas`), sales and excise taxes (`salestax`), and contributions for government social insurance other than pension, UI, and DI (`othercontrib`), minus social assistance benefits in cash (`dicab`) and social transfers in kind (`inkindinc`). Income is defined as personal factor labor income (`flinc`).

We regress total taxes net of transfers on a spline in income, considering three income brackets defined for annual earnings within $[\$0, \$20,000]$, $[\$20,000, \$500,000]$, and $[\$500,000, \$\infty)$. To account for the fact that taxes and transfers depend on marital status and the number of children, we control for marital status, a fourth-degree polynomial of the number of children, and the interaction of the two. We estimate empirical marginal tax rates for the three brackets (i.e., coefficients on the spline variables) of 44.6%, 28.1%, and 49.4%, respectively, with a corresponding demogrant (i.e., the coefficient on the constant which corresponds to single childless tax units) of $11,214. The U-shaped form of the estimated bracket structure is explained by the income phase-out of transfers such as food stamps or Medicaid, and the progressive structure of the income tax system.

Note that since the estimated current tax system has a bracket structure, the individual solution for calibration purposes solves the complete discrete choice problem allowing for bunching at kink points. This adjustment is not needed when solving the optimal tax scheme.

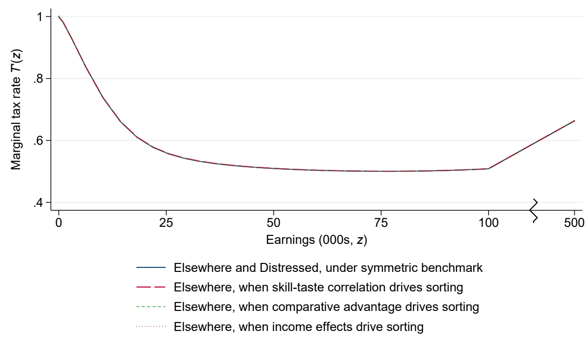
Exogenous revenue requirement and landlord weight The exogenous revenue requirement R is computed as the budget constraint residual in the symmetric benchmark (Section 8.1) calibration under current taxes, resulting in $R = \$14,746$. Given that the population is normalized to 1, this value has a per-capita interpretation. We hold this parameter fixed across all simulations.

The Pareto weight on landlords ω_L equals the equilibrium budget constraint multiplier Γ resulting from the symmetric benchmark equilibrium with optimal taxes, so the planner values landlords similar to the average household in the economy. Specifically, we solve for the optimum under the assumption that $\lambda_L = 1$, which implies that ω_L equals the value of Γ at the optimum. We obtain $\omega_L = 0.0205$. This parameter is held fixed across all simulations except the 75% weight on landlords scenarios of Table 4.

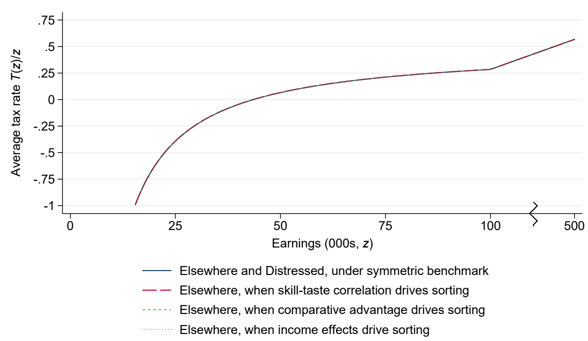
Appendix Figure 1: Optimal Taxes are Similar in Elsewhere under All Scenarios

A. By Earnings z

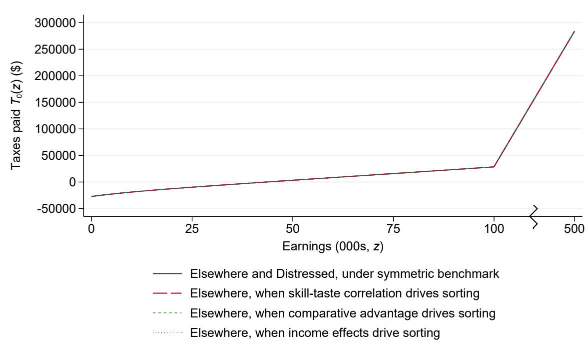
(i) Marginal Tax Rates



(ii) Average Tax Rates

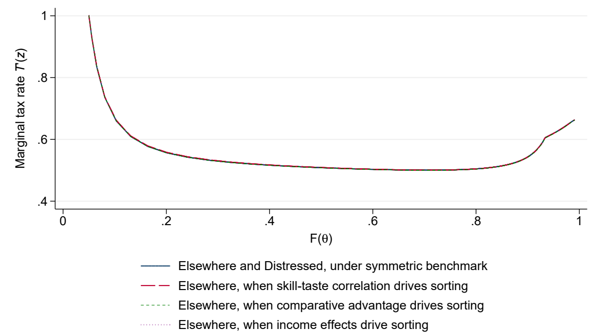


(iii) Taxes Paid

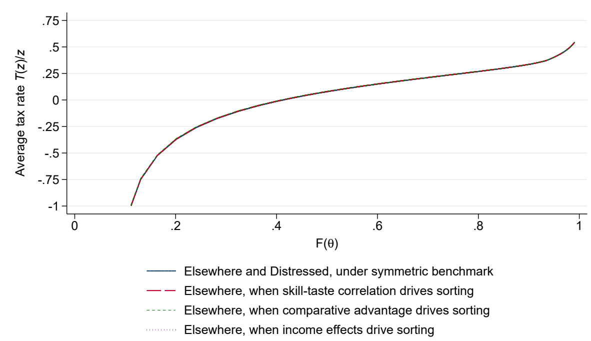


B. By Skill Quantile $F(\theta)$

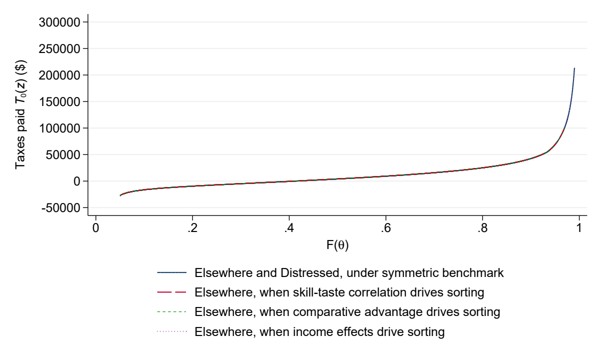
(i) Marginal Tax Rates



(ii) Average Tax Rates



(iii) Taxes Paid

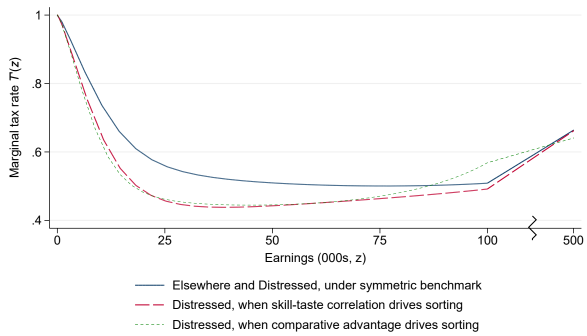


Notes: This figure replicates Figure 2 except that it plots Elsewhere tax schedules instead of Distressed tax schedules for the three sorting scenarios in the top four panels and that it shows taxes paid in Elsewhere in the bottom two panels. See the notes to that figure for additional details.

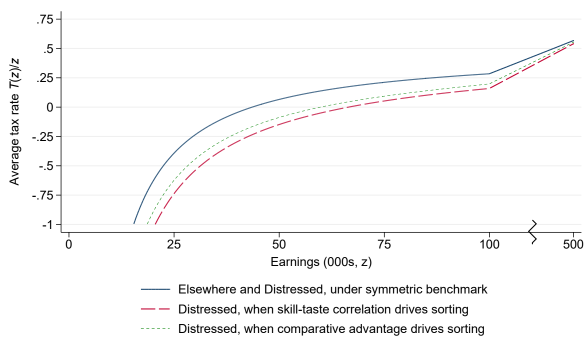
Appendix Figure 2: Optimal Tax Schedules with Less Migration

A. By Earnings z

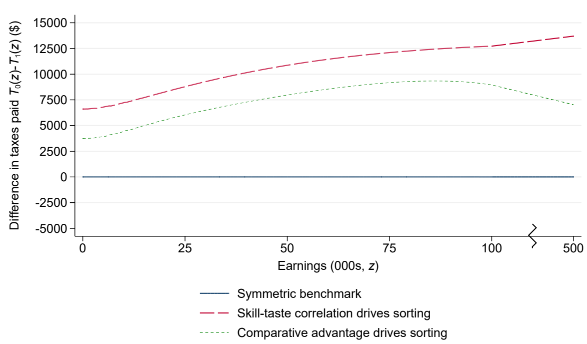
(i) Marginal Tax Rates



(ii) Average Tax Rates

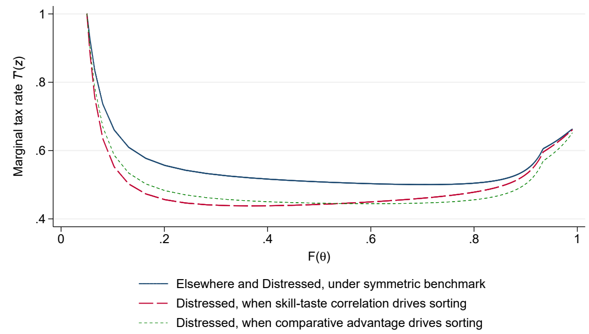


(iii) Difference in Taxes Paid

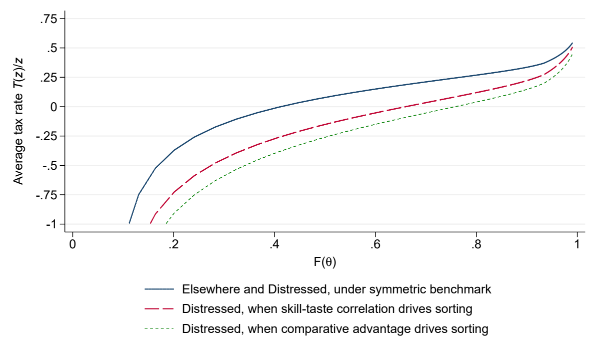


B. By Skill Quantile $F(\theta)$

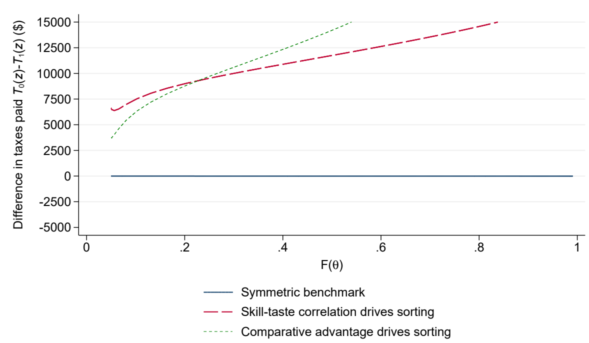
(i) Marginal Tax Rates



(ii) Average Tax Rates



(iii) Difference in Taxes Paid

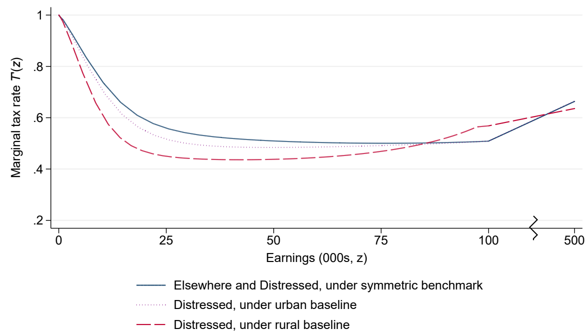


Notes: This figure replicates Figure 2 except that it sets the standard deviation of idiosyncratic preferences for living in Distressed κ is set to 4 instead of 0.5. See the notes to that figure for additional details.

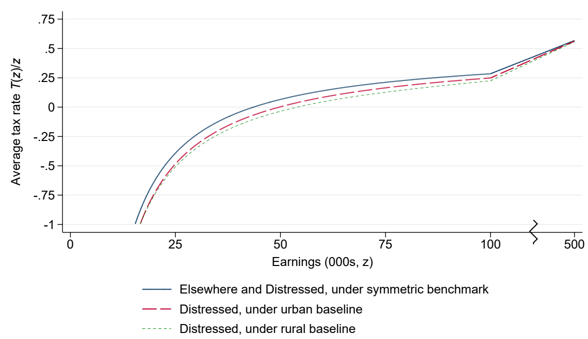
Appendix Figure 3: How Large Might Optimal Place-Based Transfers Be? MTRs and ATRs

A. By Earnings z

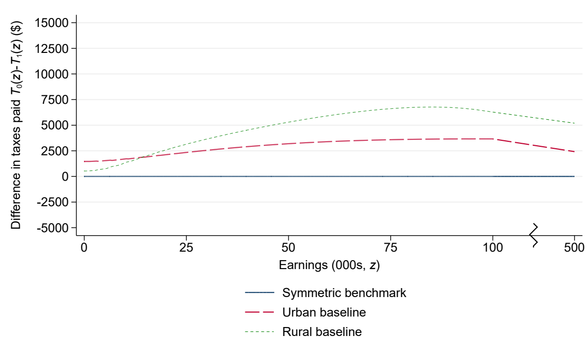
(i) Marginal Tax Rates



(ii) Average Tax Rates

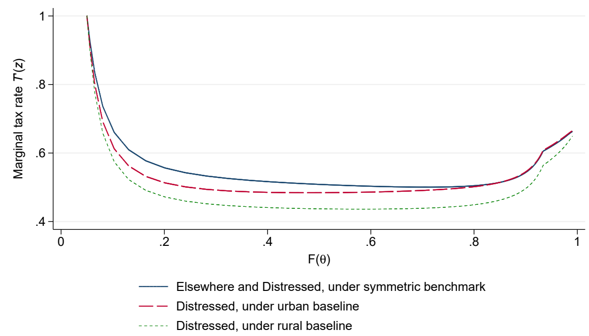


(iii) Difference in Taxes Paid

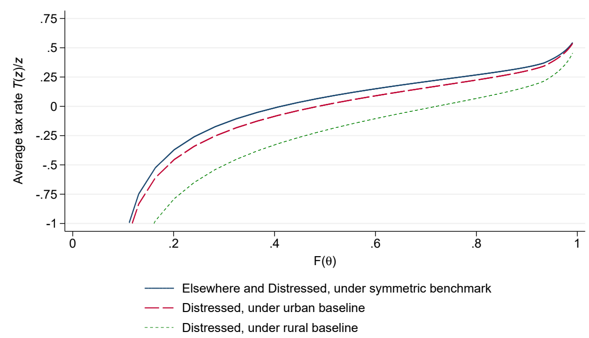


B. By Skill Quantile $F(\theta)$

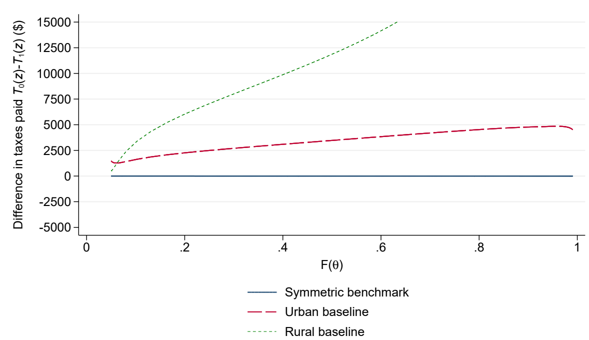
(i) Marginal Tax Rates



(ii) Average Tax Rates



(iii) Difference in Taxes Paid



Notes: This figure plots tax schedules from the symmetric benchmark (reprinted from Figure 2) and from the urban and rural baselines of Table 4. See the notes to those exhibits for details.

Appendix Table 1: Net Within-Earnings Transfers to Distressed Locations Due to Marriage and Children

	Within-earnings differences by:								
	Marriage and children			Children only			Marriage only		
	Elsewhere minus Distressed	Elsewhere	Distressed	Elsewhere minus Distressed	Elsewhere	Distressed	Elsewhere minus Distressed	Elsewhere	Distressed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. All families									
i. Distributional National Accounts (DINA) data									
Taxes	-623	34,842	35,465	-68	34,846	34,914	-513	34,844	35,357
Transfers	220	5,848	5,628	-61	5846	5,907	252	5,902	5,651
Net taxes	-842	28,994	29,836	-7	29,000	29,007	-765	28,942	29,707
ii. CPS-augmented DINA data									
Taxes	-623	34,842	35,465	-68	34,846	34,914	-513	34,844	35,357
Transfers	90	7,020	6,930	-198	7,019	7,217	264	7,060	6,796
Net taxes	-713	27,822	28,535	130	27,827	27,697	-777	27,784	28,561
B. Low-earning families only									
i. DINA data									
Taxes	-98	4,102	4,200	-45	4,103	4,148	-79	4,112	4,191
Transfers	-134	15,755	15,889	-379	15,758	16,136	397	16,079	15,682
Net taxes	36	-11,653	-11,689	334	-11,654	-11,988	-476	-11,967	-11,490
ii. CPS-augmented DINA data									
Taxes	-98	4,102	4,200	-45	4,103	4,148	-79	4,112	4,191
Transfers	-372	20,876	21,247	-447	20,885	21,333	507	21,111	20,604
Net taxes	274	-16,774	-17,048	403	-16,782	-17,185	-587	-16,999	-16,413

Notes: We aggregate the tract-level ACS data to produce counts of families by earnings-x-marriage-x-children cells for "Distressed" (all forty urban and rural EZs pooled) and "Elsewhere" (the rest of the country). We use those counts to produce mean within-earnings net taxes in Distressed and Elsewhere, equal to the community-specific marriage-x-children weighted average of Piketty et al. (2018) distributional national accounts (DINA) data's net taxes within each earnings cell, integrated over the *nationwide* earnings distribution. The difference between those two weighted averages equals our estimate of implicit place-based transfers. We aggregate the DINA data to the tax unit level. We compute DINA taxes as all taxes gross of refundable tax credits minus social insurance contributions, which equals DINA's main raw tax variable. We compute DINA transfers as all individualized benefits (e.g., refundable tax credits, medicaid, SNAP) minus social insurance benefits, which equals DINA's main raw transfers variable minus collective consumption expenditure (e.g., national defense). DINA net taxes equals DINA taxes minus DINA transfers. The CPS-augmented DINA panels use the Current Population Survey to better vary the DINA transfer data by parental status. The low-earners panels restrict attention to families with under $30,000 in annual earnings. See A for more details.

Appendix Table 2: Simulation Parameters

	Assumed Parameters									Estimated Parameters								
	Labor supply elasticity, η	Mean of log-normal skill distribution, ξ	Std.dev. of log-normal skill distribution, σ	Pareto parameter of Pareto skill distribution, p	Grid size for skill distribution, N	Comparative advantage in Elsewhere, γ_0	Housing supply elasticity in Elsewhere, ϱ_0	Housing supply elasticity in Distressed, ϱ_1	Household Pareto weights, $\omega_H(\theta)$	Elsewhere housing supply shifter, $\underline{H}_0$	Distressed housing supply shifter, $\underline{H}_1$	Wage scale, W	Preference shock mean, μ	Skill-taste correlation, β	Comparative advantage in Distressed, γ_1	Std. dev. of locational tastes, κ	Exogenous revenue requirement, R	Landlord Pareto weight, ω_L
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
A. Scenarios in Table 2																		
i. High migration: Std. dev. of idiosyncratic preferences for living in Distressed $\kappa = 0.5$																		
Skill-taste correlation	0.5	2.757	0.5611	2	578	1.00	0.34	0.34	1.00	0.50	0.009	$3,974	-1.65	0.85	1.00	0.50	$14,756	0.021
Comparative advantage	0.5	2.757	0.5611	2	578	1.00	0.34	0.34	1.00	0.50	0.009	$3,974	-1.82	0.00	0.924	0.50	$14,756	0.021
Income-based sorting	0.5	2.757	0.5611	2	578	1.00	0.34	0.34	1.00	0.50	0.021	$3,974	-2.29	0.00	1.00	0.50	$14,756	0.021
ii. Low migration: Std. dev. of idiosyncratic preferences for living in Distressed $\kappa = 4$																		
Skill-taste correlation	0.5	2.757	0.5611	2	578	1.00	0.34	0.34	1.00	0.50	0.009	$3,974	-13.2	6.83	1.00	4.00	$14,756	0.021
Comparative advantage	0.5	2.757	0.5611	2	578	1.00	0.34	0.34	1.00	0.50	0.009	$3,974	-16.0	0.00	0.906	4.00	$14,756	0.021
B. Scenarios in Table 4																		
i. Urban scenarios																		
Urban baseline	0.5	2.757	0.5611	2	578	1.00	0.34	0.24	1.00	0.50	0.011	$3,974	-1.54	0.61	1.00	0.44	$14,756	0.021
1/2x migration	0.5	2.757	0.5611	2	578	1.00	0.34	0.24	1.00	0.50	0.011	$3,974	-5.12	2.48	1.00	1.52	$14,756	0.021
2x migration	0.5	2.757	0.5611	2	578	1.00	0.34	0.24	1.00	0.50	0.011	$3,974	-0.42	-0.02	1.00	0.09	$14,756	0.021
No rent differences	0.5	2.757	0.5611	2	578	1.00	0.34	0.24	1.00	0.50	0.011	$3,974	-0.98	0.51	1.00	0.30	$14,756	0.021
1.5x rent differences	0.5	2.757	0.5611	2	578	1.00	0.34	0.24	1.00	0.50	0.011	$3,974	-1.77	0.64	1.00	0.49	$14,756	0.021
Swap housing elasticities	0.5	2.757	0.5611	2	578	1.00	0.34	0.24	1.00	0.61	0.010	$3,974	-2.14	0.93	1.00	0.62	$14,756	0.021
75% weight on landlords	0.5	2.757	0.5611	2	578	1.00	0.34	0.24	1.00	0.50	0.011	$3,974	-1.54	0.61	1.00	0.44	$14,756	0.021
ii. Rural scenarios																		
Rural baseline	0.5	2.757	0.5611	2	578	1.00	0.34	0.60	1.00	0.50	0.008	$3,974	-16.1	0.00	0.900	4.06	$14,756	0.021
1/2x migration	0.5	2.757	0.5611	2	578	1.00	0.34	0.60	1.00	0.50	0.008	$3,974	-33.2	0.00	0.885	8.23	$14,756	0.021
2x migration	0.5	2.757	0.5611	2	578	1.00	0.34	0.60	1.00	0.50	0.008	$3,974	-8.44	0.00	0.900	1.99	$14,756	0.021
No rent differences	0.5	2.757	0.5611	2	578	1.00	0.34	0.60	1.00	0.50	0.005	$3,974	-17.3	0.00	0.874	4.32	$14,756	0.021
1.5x rent differences	0.5	2.757	0.5611	2	578	1.00	0.34	0.60	1.00	0.50	0.010	$3,974	-16.0	0.00	0.904	3.95	$14,756	0.021
Swap housing elasticities	0.5	2.757	0.5611	2	578	1.00	0.34	0.60	1.00	0.30	0.011	$3,974	-16.1	0.00	0.900	3.97	$14,756	0.021
75% weight on landlords	0.5	2.757	0.5611	2	578	1.00	0.34	0.60	1.00	0.50	0.008	$3,974	-16.1	0.00	0.900	4.06	$14,756	0.021

Notes: This table reports parameters underlying the simulations whose results are reported in Tables 2 and 4. See the notes to those tables for additional details. The Table 2 scenarios fix κ and only allow one of β, γ_1, or $\underline{H}_1$ to vary. The urban scenarios fix $\gamma_1 = 1$ and the rural scenarios fix $\beta = 0$. See the notes to Tables 2 and 4 for additional details.

Appendix Table 3: How Large Might Optimal Place-Based Transfers Be? Manipulating Characteristics of Distressed

	Income-constant average tax differences, Δ_z	Difference in demogrants, Δ_0	Std. dev. of idiosyncratic preferences for living in Distressed, κ	Skill-taste correlation, β	Comparative advantage, γ_1
	(1)	(2)	(3)	(4)	(5)
A. Urban scenarios					
Urban baseline	3,143	1,462	0.44	0.61	1.000
Distressed twice as large	3,128	1,452	0.42	0.60	1.000
Distressed half as large	3,150	1,467	0.44	0.62	1.000
Distressed twice as poor	5,787	2,649	0.36	1.07	1.000
Distressed half as poor	1,520	557	0.48	0.34	1.000
B. Rural scenarios					
Rural baseline	4,329	532	4.06	0.00	0.900
Distressed twice as large	3,943	142	3.99	0.00	0.900
Distressed half as large	3,533	-347	4.13	0.00	0.900
Distressed twice as poor	11,272	5,007	3.69	0.00	0.796
Distressed half as poor	532	-1,471	4.16	0.00	0.954

Notes: This table reports results from additional scenarios that modify the urban and rural baselines of Table 4. The Distressed twice as large and Distressed half as large scenarios target a Distressed population share of 3.4% and 0.85%, respectively, instead of 1.7%. The Distressed twice as poor and Distressed half as poor scenarios target a share of Distressed households earning under $50,000 that is either double or half of the difference between the Distressed share earning under $50,000 and the nationwide share of 39%. In the urban scenarios, Distressed twice as poor targets a 73% share of Distressed earning under $50,000 relative to 56% in the urban baseline, while Distressed half as poor targets 48%. In the rural scenarios, Distressed twice as poor targets an 80% share of Distressed earning under $50,000 relative to 60% in the rural baseline, while Distressed half as poor targets 49%. See the notes to Table 4 for additional details.