American Economic Journal: Economic Policy 2021, 13(2): 100–134
https://doi.org/10.1257/pol.20180529

Heterogeneous Workers and Federal Income Taxes in a Spatial Equilibrium†

By Mark Colas and Kevin Hutchinson*

We study the geographic incidence and efficiency of an income tax by estimating a spatial equilibrium model with heterogeneous workers. The US income tax shifts households out of high-productivity cities, leading to locational inefficiency of 0.25 percent of output. Removing spatial tax distortions increases inequality because more educated households are more mobile and own larger shares of land. Flattening the tax schedule, or introducing cost-of-living adjustments or local wage adjustments leads to efficiency gains but causes substantial increases in inequality. Differences in mobility and land ownership across skill groups create an equity-efficiency trade-off that is unique to spatial settings. (*JEL* H24, H22, D31, J31, J24, R23)

Wages vary greatly across local labor markets in the United States, reflecting, at least partially, underlying heterogeneity in location-specific productivity.[1] This observation implies that a federal income tax can lead to a geographic misallocation of resources. In particular, households in highly productive locations will be taxed at a higher amount, thus providing a disincentive to live and work in these places. Some policymakers and researchers have argued for reducing the tax burden in high-income cities on precisely these grounds.[2]

However, these policy recommendations must also acknowledge the distributional aspects of such a policy. More educated households have been shown to be significantly more mobile than less educated households and therefore may be better able

* Colas: University of Oregon, 1285 University of Oregon, Eugene, OR 97403 (email:mcolas@uoregon.edu); Hutchinson: Uber Technologies Inc., 1455 Market St, San Francisco, CA 94103 (email:kphutchinson2@gmail.com). Dan Silverman was coeditor for this article. We thank three anonymous referees. We wish to thank Chris Taber, John Kennan, and Jesse Gregory for their guidance throughout. We also greatly appreciate advice and suggestions from Naoki Aizawa, David Albouy, David Evans, Chao Fu, Kyle Herkenhoff, Tom Holmes, Rasmus Lentz, Keaton Miller, Eduardo Morales, John M. Morehouse, Ronni Pavan, Nicolas Roys, Ananth Seshadri, Jim Walker, and Woan Foong Wong. Finally, we extend our thanks to countless visiting seminar speakers and graduate students at the University of Wisconsin. This work benefited from access to the University of Oregon high performance computer, Talapas. All remaining errors are our own.

† Go to https://doi.org/10.1257/pol.20180529 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] For example, Moretti (2011) documents that in the manufacturing industry total factor productivity in the most productive US counties is three times as large as in the least productive counties.

[2] In early work, Wildasin (1980) shows that income taxes lead to inefficient spatial sorting. Albouy (2009) argues that the unequal tax burden across cities leads to substantial deadweight loss and that indexing taxes to local price levels can reduce this inefficiency. Moody and Hoffman (2003, 7) argue that a "flat tax rate on all income with no standard deduction, for example, would not discriminate against taxpayers living and working in high-cost areas," while Eeckhout and Guner (2017) argue that the optimal federal tax code in a spatial setting should be less progressive than the current tax code.

to capture the gains from tax changes via migration. Further, educated households are more likely to own land and may disproportionately benefit from changes to the tax code that increase landowner profits. Finally, educated households disproportionately reside in more productive cities and therefore will benefit from reductions to taxes in those cities. Allowing for heterogeneity in mobility, land ownership and the distribution of skill groups across space means that tax changes aimed to reduce locational inefficiency can reduce deadweight loss, but at the cost of widening welfare inequality between households. The main contribution of this paper is to better understand the spatial dimension of the classic trade-off between equity and efficiency.

To accomplish this we develop and estimate a spatial equilibrium model using US census data, similar to those recently used by Diamond (2016) and Piyapromdee (2017). In the model, the decision of where to live is a static discrete choice; heterogeneous households choose the city that yields the highest utility in terms of *after-tax* wages, rents, and amenities. Households own diversified portfolios of land that can be developed for housing and rented. Wages, rents, and landowner profits are determined in equilibrium by the location and housing consumption choices of households. Changes to the tax code will alter the distribution of after-tax wages across cities, thus leading to changes in the spatial distribution of pretax and posttax wages, rents, populations, profit, and welfare in equilibrium.

In the model, changes to the tax code may differentially affect unskilled and skilled households for four main reasons.[3] First, we allow households of different skill groups to vary in the responsiveness of their location choices to after-tax wages, and thus to be differentially responsive to changes in the tax code. If skilled households are more mobile, as we estimate, then they may be better able to capture the gains from changes in income tax policy via migration. Second, households of different skill levels are imperfect substitutes in production, so that the relative supplies of heterogeneous labor in each location determine local *pretax* wages. Thus, taxes affect households directly via their effect on posttax wages, but also by changing the distribution of households across cities and therefore changing equilibrium wages and rents. Third, land ownership varies by skill level. If changes in the tax code increase profits from land ownership, this will disproportionately benefit skilled households, who own larger shares of land. Finally, we allow skill groups to vary in their preferences over cities. Tax reductions in high productivity cities will benefit skill groups who prefer to live in these cities.

One of the main benefits of the framework here is the transparency with which the key parameters are identified. The household sorting component of the model, which determines the elasticity of household location choices with respect to wages, is estimated using a two-step procedure similar to that in Berry, Levinsohn, and Pakes (2004). The first step transforms the nonlinear discrete choice problem into one that can be solved using standard Instrumental Variables. The second step uses changes in national tax policy to generate variation in after-tax wages at the local labor market level. We estimate the elasticity of location choices with respect to after-tax income using this tax instrument to account for the potential correlation of

[3] Throughout the paper, we define a household's skill level by the education level of the household head. If the household head has graduated from college, the household is classified as "skilled."

wages with unobserved amenities. The labor demand and housing supply curves are estimated using the immigrant enclave instrument proposed by Card (2009); immigrant supply shocks from sending countries create variation in the relative supplies of labor and variation in housing demand.

Our first counterfactual experiment measures the deadweight loss and incidence of the current tax code. We find that the current tax code leads to a resorting of households into lower productivity places, particularly by skilled households. This leads to a deadweight loss caused by the misallocation of households across locations of 0.25 percent of output.[4] Next, we decompose the distortion of the current tax code into that caused by state income taxes, the payroll tax, and the federal income tax. We find that removing state income tax distortions reduces deadweight loss to 0.18 percent of output, and removing payroll taxes and state income tax distortions reduces deadweight loss to 0.14 percent of output. The federal income tax is responsible for this final deadweight loss of 0.14 percent of output.

We then calculate the incidence of switching from the current tax code to non-distortionary lump-sum taxes in which tax revenue collected from each skill group is held constant. When we switch to these lump-sum taxes, deadweight loss from locational inefficiency drops to zero.[5] However, the average well-being for unskilled households decreases despite the fact that tax revenue raised from skilled and unskilled households is held constant. Lump-sum taxes shift the distribution of households toward high-wage cities with inelastic housing supply curves and therefore increase landowner profits, which disproportionately benefits skilled households. This implies that the current income tax subsidizes workers to live in areas with elastic housing supply curves, which increases unskilled relative to skilled household well-being. Further, as skilled households are relatively more mobile than unskilled households, they are better able to adjust spatially to tax changes. This introduces a version of the equity-efficiency trade-off that is unique to the spatial setting: reducing the tax burden in higher-income cities increases the size of the pie but reduces the fraction that goes to unskilled households.

Lump-sum taxes are efficient but may not be practical to implement. Therefore, we next consider several alternative tax codes that may lead to efficiency gains. We begin by evaluating a revenue-neutral flat tax. The flattening of the tax code leads to a small relocation toward higher-income cities, which leads to a reduction in deadweight loss, from 0.25 percent to 0.16 percent of baseline output.

Next, we simulate two location-specific tax adjustments aimed to reduce locational inefficiency. First, we consider cost-of-living adjustments, so that households living in cities with high housing prices receive a tax break. Second, we consider a local wage adjustment, so that households living in cities with higher average wages receive a tax deduction. These local tax adjustments have several features that we find attractive. First, they can potentially reduce the deadweight loss caused by spatial misallocation by reducing the tax burden in more productive

[4] Throughout the paper, we focus on the effects of the tax on location choice, and abstract away from other potential labor supply responses to taxes. The full deadweight loss of the current tax in a model that allowed for adjustment in hours worked or labor force participation, for example, would be larger.

[5] We prove that the competitive equilibrium with lump-sum taxes is Pareto optimal in online Appendix A1.

cities. Second, within a location, these tax adjustments maintain the same progressivity as the current tax code, which helps to reduce inequality. Finally, we believe that these adjustments would be relatively straightforward to implement, in that local housing prices, rental rates, and wages are observed.

The cost-of-living adjustments lead to large increases in population in high-rent, high-wage cities. This reduces deadweight loss from 0.25 percent to 0.09 percent. However, the adjustment also leads to increases in between group well-being inequality as unskilled households are made worse off and skilled households are made better off. The drop in utility for the average unskilled household is equivalent to a \$600 decrease in yearly income, while the increase in average skilled household utility is equivalent to \$1,420 increase in yearly income. When we index taxable income to the local wages of skilled households, the population in high-income cities increases even more dramatically, resulting in a decrease in deadweight loss from 0.25 percent to 0.04 percent. However, between-group well-being inequality also increases dramatically.

To better understand the increases in inequality associated with each counterfactual tax code, we perform a model based decomposition to understand the mechanisms behind these increases in inequality. For each of our three counterfactual taxes (flat taxes, cost-of-living adjustments, and local wage adjustments), we decompose the effects on inequality into (i) mechanical effects—increases in inequality holding location choices and prices constant, (ii) sorting effects—increases in inequality due to differential mobility across skill groups, and (iii) increases in inequality due to equilibrium responses in prices and landowner profits.

We find that the inequality increase associated with a flat tax is almost entirely due to mechanical effects. This is because flat taxes lead to large changes in taxes paid across households, but lead to a relatively small decrease in the variation in taxes across locations for a given household. However, for cost-of-living adjustments and local wage adjustments we find the opposite effects. Neither of these tax changes leads to a large mechanical increase in inequality. However, sorting effects due to differential mobility and equilibrium responses in prices and profits lead to large increases in between-group well-being inequality. We therefore conclude that differential mobility and land ownership across skill groups lead to a quantitatively important equity-efficiency trade off that is unique to the spatial setting. We believe this version of the equity-efficiency trade off has not been quantified previously.

Two papers closely related to ours are Albouy (2009) and Eeckhout and Guner (2017), who also study the efficiency losses associated with a federal income tax in a spatial equilibrium. In Albouy (2009), as in the models of Rosen (1979) and Roback (1982), households are identical and perfectly mobile. A fixed supply of land is used in the production of home goods and tradeable goods, thus creating downward-sloping labor demand curves and upward-sloping home good supply curves. Locations may vary in their amenity values; wages and rents adjust so that utility is equal in all inhabited locations. Albouy (2009) calculates the deadweight loss of income taxes and presents policies that can alleviate this deadweight loss.

Eeckhout and Guner (2017) also study the spatial effects of an income tax in a model that includes perfectly mobile homogeneous households, firms that produce a consumption good, and construction firms that combine the consumption good and land in the production of housing. Additionally, the model allows for agglomeration economies and a congestion externality—higher city populations increase productivity and reduce the amenity value of living there. The presence of these externalities implies a role for a corrective income tax to improve welfare; characterizing this optimal tax is the main goal of their paper, which turns out to still be progressive, but less so than the current US tax code. We view these results to be highly complementary to ours. In particular, our work focuses on distributional effects, which is made possible by relaxing the assumption of homogeneous households.

Our work departs from these two papers in three main ways. The first distinguishing feature of our research is the addition of heterogeneity across households. In particular, we allow the values of the amenity associated with each location to separately vary across 32 different demographic groups.[6] All of these groups are also allowed to have differential attachments to their state of birth. Households are defined as either skilled or unskilled households. These skill groups are assumed to be imperfect substitutes in production and vary in the responsiveness of their location choices to wage differentials across cities. This allows us to study the impact of tax reform on between-group well-being inequality.

Second, not only do we allow for different skill levels (skilled and unskilled) and demographic groups, but *within* each group, households are allowed to have idiosyncratic preferences over locations, so that the labor supply curves in each location are upward sloping. Furthermore, a household's utility in a location is a function of the distance between the location and the household's birthplace. These distinctions have important implications for the elasticity of location choice with respect to wages and thus tax incidence and deadweight loss. Given the importance of the elasticity of location choices in determining tax incidence, we believe our relaxation of the perfect elasticity assumption helps us to accurately estimate the incidence associated with the tax codes we consider. Finally, rather than use calibration, we estimate a structural spatial equilibrium model. This allows us to consider a much richer setup in terms of observable and unobservable heterogeneity.

This paper is also related to Fajgelbaum and Gaubert (2018) (FG), who characterize the efficient spatial transfers in a model with spillovers across heterogeneous households. In our paper, we do not consider spillover effects, but we analyze what we believe to be a more practically implementable set of tax policies, such as changes to the progressivity of the tax code or local tax deductions. In contrast, FG consider a planner that can choose demographic and location specific taxes and transfers. Further, as FG consider a planner that can lump-sum transfer wealth across heterogeneous groups, there is no trade-off between equity and efficiency. As such, while FG focus on characterizing the first-best solution in their environment, we focus on calculating the welfare effects and distortions associated with these more practically implementable, and generally not first-best, policies.

Furthermore, this paper is related to a series of papers measuring the distortionary effects of state income and corporate taxes in a spatial equilibrium setting. Moretti and Wilson (2017) analyze the effects of state taxes on the geographic mobility of

[6]These groups are defined as combinations of education, potential experience, marital status, and number of children.

star scientists. Fajgelbaum et al. (2018) measure the efficiency losses caused by dispersion in state taxes using a general equilibrium model with imperfectly mobile firms and households and costly trade between locations. Coen-Pirani (2018) analyzes the role of progressive taxation on migration and welfare in a dynamic model with homogeneous locations. Suárez Serrato and Zidar (2016) study the incidence of state corporate taxes on households, firms, and land owners in an equilibrium model that incorporates imperfectly mobile firms and households with endogenously determined wages and rent levels. Our work contributes to this literature by quantifying the incidence of a tax on heterogeneous households.

To reiterate, the central contributions of this paper are to (i) quantify the effects of various tax codes in a model with heterogeneous households, (ii) measure the incidence of these tax codes, and (iii) analyze the equity-efficiency trade-off associated with these tax codes within a spatial framework.

I. Model

We build a model of spatial equilibrium, similar to those used by Diamond (2016) and Piyapromdee (2017). Locations vary along three dimensions: wages, rents, and amenities. The choice of location is modeled as a static discrete choice; households choose the city that yields the highest utility. We assume that a single, tradeable good is produced in each location and households of different skill levels are imperfect substitutes in its production, so that the supplies of heterogeneous labor in each location determine local wages. Households own diversified portfolios of land across all cities and the size of a household's land ownership portfolio may vary by the household's skill level. The cost of developing this land for housing varies across cities, implying that changes in the distribution of households across cities can change total landowner profits. Thus, wages, rents, and population in each location are determined endogenously as equilibrium outcomes.

Our model extends the Rosen-Roback framework along four dimensions, all of which are important for answering our question. First, we allow for two imperfectly substitutable types of labor in production: skilled and unskilled. As emphasized by Stiglitz (1982), the elasticity of substitution between skilled and unskilled labor plays a key role in the welfare effects of taxation. In our setting, the substitutability of labor determines how *local* wages react to changes in the local supply of each group. For example, if tax reform induces more skilled households to choose a certain location, this will increase the wage of the unskilled households in that location.

Second, we allow for rich heterogeneity in households' preferences. In particular, we allow households' preferences to include a premium for living close to the state of birth of the household head. Bayer, Keohane, and Timmins (2009) demonstrate the importance of doing so; the authors find that ignoring imperfect mobility of this kind can result in substantially biased estimates of the other preference parameters.[7] Additionally, households' preferences may vary based on their marital status, number of children, experience level, and education.

[7] In particular, they find that the marginal willingness to pay for air quality is understated by a factor of three when perfect mobility is assumed.

Third, we allow for heterogeneity in preferences over locations, *conditional* on household type, thus relaxing the assumption of perfect mobility. More precisely, we allow for idiosyncratic location-specific preference shocks, where the variance of the shock is allowed to differ by household skill. As emphasized by Kline and Moretti (2014) and Busso, Gregory, and Kline (2013), this sort of preference heterogeneity is crucial for analyzing policies when households are mobile. In our case, these variances govern the location choice elasticities, which are essential for quantifying the impact of tax reform on the resulting spatial equilibrium.

Finally, we incorporate the federal income, state income, and payroll taxes, as well as the corresponding credits which depend on marital status and number of children. We assume the government uses tax revenue to fund a public good which benefits all households, regardless of their location.[8]

A. *Households*

Locations are indexed by j and time is indexed by t.[9] Households, indexed by i, maximize utility by (i) allocating their resources between a nationally traded consumption good, c_t, and housing, h_{jt}, and (ii) choosing the location j that yields the highest utility. Households belong to a narrowly defined demographic group, d, which is defined by marital status, education level, age, and number of children. Households of different demographic groups vary in their tax levels, productivity, and preferences over locations. Households are also characterized as either a skilled (S) or an unskilled (U) household, which is defined by the education level of household head. We index these broad skill groups by e. Households of different skill groups are imperfect substitutes in production.

We proceed by first solving the households' maximization problem, conditional on location. The price of housing is denoted by r_{jt}, and we normalize the price of the tradeable consumption good to 1. Notice that the price of the consumption good is constant across all locations, reflecting the law of one price, which applies because c is tradeable. Locations are also distinguished by their amenity value, Γ_{ijt}. Preferences over the consumption good and housing are assumed to be Cobb-Douglas and are written as

$$u_{it}(j, c_t, h_{jt}) = (1-\alpha^e)\log(c_t) + \alpha^e \log(h_{jt}) + \Gamma_{ijt}, \tag{1}$$

where α^e is a parameter that corresponds to the agent's optimal budget share of housing.[10]

We assume households of a demographic group d inelastically supply labor at their chosen location. Income received in location j by households of group d in

[8] As we do not directly model utility from the public good, we hold the amount of tax revenue fixed across counterfactuals.

[9] The model is static. The t indexes are useful in keeping track of the different cross sections that are used in estimation.

[10] See Davis and Ortalo-Magné (2011) for evidence that the budget share of housing is indeed constant across metropolitan areas and time. Furthermore, allowing these parameters to vary by skill, rather than by demographic groups, significantly reduces the number of parameters to be estimated.

time t, I_{jt}^d, is equal to the sum of their earned income and their share of profits from landownership: $I_{jt}^d = w_{jt}^d + s_t^e \mathbf{\Pi}_t$, where w_{jt}^d is labor income, $\mathbf{\Pi}_t$ is the total profits from land ownership across all locations, and s_t^e is the share of landowner profits received by a household of skill level e.[11]

Further, let $\mathcal{T}_{jt}^d\left(I_{jt}^d\right)$ denote the household's tax burden, which includes federal income, state income, and federal payroll taxes. Note that, in addition to income level, the tax also depends on demographics, location, and time. These dependencies account for differences in state income taxes, differences in income tax deductions by demographic group, and changes in the tax code over time. We can now write the households' budget constraint as

$$c_t + r_{jt} h_{jt} = I_{jt}^d - \mathcal{T}_{jt}^d\left(I_{jt}^d\right), \tag{2}$$

and solving the households' problem yields the following indirect utility associated with choosing location j:

$$v_{ijt} = \log\left(I_{jt}^d - \mathcal{T}_{jt}^d\left(I_{jt}^d\right)\right) - \alpha^e \log\left(r_{jt}\right) + \Gamma_{ijt}. \tag{3}$$

As households inelastically supply labor in all locations, the only relevant labor supply decision is where to live, which is now a static discrete choice defined by equation (3).[12] Households choose the location that maximizes indirect utility.

We now decompose the amenity term, Γ_{ijt}, into five distinct components. In particular,

$$\Gamma_{ijt} = \gamma_{hp,t}^d \mathbf{1}\left(j \in Bstate_i\right) + \gamma_{dist,t}^d \phi\left(j, Bstate_i\right) + \gamma_{dist2,t}^d \phi^2\left(j, Bstate_i\right) + \xi_{jt}^d + \sigma^e \epsilon_{ijt},$$

where $\mathbf{1}\left(j \in Bstate_i\right)$ is an indicator for location j being in the household head's birth state; $\phi\left(j, Bstate_i\right)$ and $\phi^2\left(j, Bstate_i\right)$ are the distance and squared distance, respectively, between the birth state and location j; ξ_{jt}^d is a common, unobservable component of amenities; ϵ_{ijt} is an idiosyncratic, stochastic term meant to capture the fact that some households are more or less attached to certain locations; and σ^e measures the dispersion in ϵ_{ijt}.[13] We assume that ϵ_{ijt} follows a Type 1 Extreme Value distribution.

The term measuring the dispersion of the preference shock, σ^e, plays an important role in determining the elasticity of household choices with respect to after-tax income. A large value of σ^e implies that idiosyncratic factors play a large role in determining households' location choices relative to wages and rents, and thus changes in the tax code will not have a large effect on location choices. A small value of σ^e implies that households will be closer to the margin between two cities

[11] It would be straightforward to allow the share of landowner profits to vary by demographic group, d, rather than by skill. As we are mostly concerned with the welfare of skilled relative to unskilled households, we choose this simpler approach.

[12] The assumption of inelastic labor supply is typical in this literature. See, for example, Moretti (2013); Kline and Moretti (2014); or Busso, Gregory, and Kline (2013).

[13] We do not allow for the possibility of endogenous amenities, as in Diamond (2016). As such, the no-tax equilibrium in our model is efficient.

and thus will be more responsive to changes in the tax code. We allow σ^e to vary by skill level to allow for differential responsiveness to income changes across skill groups.

B. *Firms*

Perfectly competitive firms in each labor market use the following constant elasticity of substitution (CES) production function to produce an identical tradeable good, using capital, K, skilled labor, S, and unskilled labor, U, as inputs:

$$Y_{jt} = A_{jt} K_{jt}^{1-\eta} L_{jt}^{\eta}.$$

The labor supply aggregator, L_{jt}, is given by

$$L_{jt} = \left[(1-\theta_{jt}) U_{jt}^{\rho} + \theta_{jt} S_{jt}^{\rho}\right]^{\frac{1}{\rho}}, \tag{4}$$

where U_{jt} and S_{jt} are defined as the total efficiency units of labor supplied by unskilled and skilled households, respectively. In particular, a household of demographic group d supplies $\mathcal{H}^d$ efficiency units of labor, which captures both differences in hours worked and productivity across demographic groups.[14] Notice that the factor intensities (i.e., the θ_{jt} terms) are allowed to vary across labor markets and over time, while the elasticity of substitution between skilled and unskilled labor, $\varsigma \equiv 1/(1-\rho)$, is restricted to be the same across locations and time. Total factor productivity (TFP), A_{jt}, is also allowed to vary across labor markets and over time.[15]

The production function exhibits constant returns to scale, and thus relative wages are determined by the ratio of skilled to unskilled labor.[16] The extent to which changes in the ratio of skilled to unskilled labor affect relative wages is governed by the parameter ρ. A high value of ρ implies a high elasticity of substitution, ς, and therefore that changes in the ratio of skilled to unskilled labor will lead to small changes in the wage ratio.[17]

[14] For married households, this gives the sum of the efficiency units for both spouses.

[15] Another option is to allow for capital-skill complementarity, as in Krusell et al. (2000). However, as we do not have data on capital at the local labor market level, we instead use the production function presented here, which can be estimated without data on physical capital.

[16] We have assumed away the possibility of agglomeration in the production function, as in Glaeser and Gottlieb (2009) or Baum-Snow, Freedman, and Pavan (2018). We leave an investigation of the role of income taxes in a model with agglomeration effects and heterogeneous workers for future work.

[17] Another option is to allow for imperfect substitution between multiple education groups, instead of the two we have allowed here. Ottaviano and Peri (2012) estimate a high elasticity of substitution between high school dropouts and graduates and between agents with some college and college graduates.

Labor markets are perfectly competitive, so that skilled and unskilled labor are paid their marginal revenue products, which yields the following expressions for wages:

$$w_{jt}^{S} = \frac{Y_{jt}\eta}{L_{jt}} L_{jt}^{1-\rho}\theta_{jt} S_{jt}^{\rho-1}, \tag{5}$$
$$w_{jt}^{U} = \frac{Y_{jt}\eta}{L_{jt}} L_{jt}^{1-\rho}\left(1-\theta_{jt}\right) U_{jt}^{\rho-1},$$

where, as before, the price of the tradeable output good is normalized to 1. We assume that each city is a small open economy and therefore that the price of the output good is exogenous. The wage income for a household of demographic type d is simply this wage multiplied by the number of efficiency units supplied by this demographic group, $w_{jt}^{d} = w_{jt}^{e}\mathcal{H}^{d}$, where e is the skill level associated with demographic group d. We assume that the capital supply is perfectly elastic and has a rental rate of $\bar{\mathcal{R}}_t$.[18]

Finally, taking logs of the two marginal revenue products in (5) and differencing yields the relative labor demand curve, which we will use for estimating the model:

$$\log\left(\frac{w_{jt}^{S}}{w_{jt}^{U}}\right) = \log\left(\frac{\theta_{jt}^{S}}{1-\theta_{jt}^{S}}\right) - \frac{1}{\varsigma}\log\left(\frac{S_{jt}}{U_{jt}}\right). \tag{6}$$

C. *Housing Supply and Land Ownership Profits*

Households own diversified portfolios of land, which they may choose to develop and rent for housing. Land varies in how costly it is to develop for housing. These plots of land form an upward-sloping housing supply curve: for small quantities of housing, the plots with the lowest cost are rented, implying low housing costs. As quantity increases, increasingly costly plots of land must be rented. Rental rates are taken as given and housing is chosen to maximize profits by setting the marginal cost of production equal to the rental rate.

In particular, we follow Kline and Moretti (2014) and parameterize the marginal cost curve such that the inverse housing supply curve in city j, time t, is given by

$$r_{jt} = z_{jt} H_{jt}^{k_j}, \tag{7}$$

where H_{jt} is quantity of housing, z_{jt} is a parameter, and k_j is a parameter equal to the inverse elasticity of the housing supply curve $\left(\text{i.e., } \partial \log r_{jt} / \partial \log H_{jt} = k_j\right)$.

As households' optimal fraction of after-tax income spent on housing is given by α^e, we can write total housing demand as

$$H_{jt} = \sum_{d} N_{jt}^{d} \frac{\alpha^{e}\left(I_{jt}^{d} - \mathcal{T}_{jt}^{d}\left(I_{jt}^{d}\right)\right)}{r_{jt}}, \tag{8}$$

[18] See online Appendix A1 for details on the firm's choice of capital.

where N_{jt}^{d} is the total number of households of demographic d living in city j. Plugging this equation for housing demand into the housing supply curve, taking logs, and rearranging yields the following reduced-form relationship:

$$\text{(9)} \qquad \log(r_{jt}) \;=\; \left(\nu_1 + \nu_2 \psi_j^{WRI}\right)\log\left[\sum_d N_{jt}^{d} \alpha^{e}\left(I_{jt}^{d} - \mathcal{T}_{jt}^{d}\left(I_{jt}^{d}\right)\right)\right] + \zeta_{jt},$$

where $\zeta_{jt} = \log(z_{jt})$ and the inverse elasticity housing supply with respect to rents is parameterized as $k_j/(1+k_j) = \left(\nu_1 + \nu_2 \psi_j^{WRI}\right)$. Gyourko, Saiz, and Summers (2008) use the Wharton Regulation Survey to produce municipality-level measures of the strictness of land use regulations. We aggregate their measures up to the core-based statistical area (CBSA) level to obtain our measure of land use regulations, ψ_j^{WRI}. Increasing housing supply is more costly in CBSAs with stricter land use policies, so we expect ν_2 to be positive.

Profits for a given plot of land are given by the difference between rents and the cost of developing the land. Therefore, total realized landowners profits in city j are given by the area above the housing supply curve and below the rental rate of housing:

$$\Pi_{jt} \;=\; \max_{H}\left[\int_0^{H}\left(r_{jt} - z_{jt}x^{k_j}\right)dx\right].$$

Plugging the above equation for housing demand into landowner profits, we obtain the following expression of equilibrium landowner profits in city j:

$$\text{(10)} \qquad \Pi_{jt} \;=\; \frac{k_j}{1+k_j}\sum_d N_{jt}^{d}\alpha^{e}\left(I_{jt}^{d} - \mathcal{T}_{jt}^{d}\left(I_{jt}^{d}\right)\right).$$

Total landowner profits are the sum of all landowner profits across cities: $\mathbf{\Pi}_t = \sum_{j'} \Pi_{j't}$.

To get a sense of how changes in the distribution of households across cities affect landowner profits, consider moving one household from a city j' to another city j while holding income levels constant. This will lead to a change in total landlord profit of

$$\Delta\mathbf{\Pi}_t \;=\; \alpha_e\left(\frac{k_j}{1+k_j}\left(I_{jt}^{d} - \mathcal{T}_{jt}^{d}\left(I_{jt}^{d}\right)\right) - \frac{k_{j'}}{1+k_{j'}}\left(I_{j't}^{d} - \mathcal{T}_{j't}^{d}\left(I_{j't}^{d}\right)\right)\right).$$

Total landowner profits, $\mathbf{\Pi}_t$, will increase as households move into areas with higher posttax income and more inelastic housing supply curves (i.e., higher k_j). The effect of a change in taxes on landowner profit will therefore largely depend on the extent to which the tax incentivizes households to move toward high-paying cities and toward cities with more inelastic housing supply curves.

D. *Equilibrium and Efficiency*

A formal definition of the competitive equilibrium is included in online Appendix A1. An equilibrium is a set of wages, rents, landowner profits, labor, capital, and household choices that is consistent with optimization by all agents, labor and housing markets clearing, the government budget constraint, and aggregate feasibility.[19]

Efficiency.—In online Appendix A1 we prove that the competitive equilibrium with lump-sum taxes is Pareto efficient in our setting. The proof follows from the fact that all markets are competitive and all landower profits are distributed lump-sum to households. The proof is by contradiction, we show that any allocation that Pareto dominates the competitive equilibrium with lump-sum taxes must violate aggregate feasibility.[20]

E. *Alternative Tax Codes*

The current tax code is progressive and leads to inefficiency because households pay higher taxes if they live in higher paying cities. This is not corrected by a flat tax; households in high-wage cities still pay higher tax levels under a flat tax. Furthermore, lump-sum taxes may not be practically implementable. As alternatives to lump-sum taxes, we consider two local tax deductions that can approximate lump-sum taxes in some cases.

Suppose households in location j are taxed $\mathcal{T}_{jt}^{d}\left(\bar{\kappa}_t I_{jt}^{d}/\kappa_{jt}\right)$, where κ_{jt} is a tax index for location j and $\bar{\kappa}_t$ is a parameter. As an example, suppose $\kappa_{jt} = I_{jt}^{\hat{d}}$, the income level of a given demographic group $\hat{d}$ in city j. In this case, taxes for households of this demographic $\hat{d}$ in location j are equal to $\mathcal{T}_{jt}^{\hat{d}}(\bar{\kappa}_t)$. Further, if the tax function does not differ by location j (for example if there are no state or local taxes), then households of demographic group $\hat{d}$ are taxed the same amount, $\mathcal{T}_{t}^{\hat{d}}(\bar{\kappa}_t)$, regardless of where they live.

We consider two possibilities for κ_{jt}. The first is a cost-of-living adjustment, similar to that suggested by Albouy (2009). With a cost-of-living adjustment, taxable income is indexed by the local price of a market basket, such that households are taxed on real income. Specifically, let $\kappa_{jt}^{COLA} = \bar{h}_t r_{jt} + \bar{c}_t$, where $\bar{h}_t$ and $\bar{c}_t$ are the average consumption of housing and the tradeable good across all households. Therefore, κ_{jt}^{COLA} gives the cost of a market basket given the prices of location j. As locations with high housing prices generally have higher wages, this deduction lowers taxable income in high-wage locations.

Perhaps a more direct way to reduce locational inefficiency is to index taxable income to local wage levels. In particular, we consider indexing taxable income to local skilled wages, such that $\kappa_{jt}^{WAGE} = w_{jt}^{S}$.[21] Ultimately, as neither tax scheme is first best, the efficiency and welfare consequences are a quantitative question.

[19] Bayer and Timmins (2005) provide a proof of existence in this type of model with one type of household. Piyapromdee (2017) extends this proof to the case with different types of households.

[20] Fajgelbaum and Gaubert (2018) prove a similar result in a slightly different setting as long as there are no agglomeration or endogenous amenities under the assumption that workers are perfectly mobile. Eeckhout and Guner (2017) use a two city model to show that the decentralized equilibrium in their setting is efficient when housing wealth is fully redistributed to households and there are no externalities. We make no assumption of perfect mobility in this proof, the social planner is free to chose each household's location.

[21] This could be implemented by taking average wages of college graduates across cities after controlling for experience, race, and other observables which might affect individual wages levels. We have also tried indexing to income levels of various skilled demographic groups and found similar results.

We calculate the effects of these various tax schemes using our estimated model in Section VB.

II. Data

Large samples are imperative for our analysis, given its local labor market nature. As such, we use the US Integrated Public Use Microdata Series (IPUMS) to draw data from the 5 percent samples of the 1980, 1990, and 2000 US census. We also use the 3 percent, three-year aggregated American Community Survey (ACS) for the years 2005–2007 (Ruggles et al. 2020).[22]

A. *Geography*

The two most important considerations we face in choosing a local labor market concept are that (i) locations correspond to distinct labor markets, and (ii) they can be compared over time. As such, we use CBSAs as our geographic definition. CBSAs naturally satisfy requirement (i), as they are the Office of Management and Budget's (OMB) official definition of a metropolitan area.[23] We achieve consistency and fulfill our second requirement by mapping the most disaggregated geographic units available in the IPUMS data, County Groups (CGs) in 1980 and Public Use Microdata Areas (PUMAs) in 1990 and after, into CBSAs.[24]

There is a fundamental trade-off between the size of the choice set and sample size. We choose to use the 70 largest CBSAs, as defined by population in 1980. Although a relatively small subset of the 929 CBSAs, these 70 locations make up approximately 60 percent of the entire US population. Further, we map individuals that do not live in one of these 70 areas into their corresponding census division, creating nine additional choices.[25]

B. *CBSA-Level Data*

We use four main variables at the CBSA level: wages, labor supply, housing rents, and a measure of housing supply elasticity. To maintain comparability with the broader wage inequality literature, we follow Autor, Katz, and Kearney (2008) (AKK) closely in constructing both the wage and labor supply series, though some choices are necessarily different given that we use different datasets and different geographies. To calculate rents, we run a hedonic regression of gross rent (which includes utilities) on a set of housing characteristics and a set of CBSA fixed effects, separately by year. The rent index is then generated by the predicted values from the

[22] We do not use ACS data after 2007 because hours worked are only reported in intervals. In principle, one could impute hours in order to extend the analysis. We prefer to use the nonimputed data.

[23] The OMB replaced the Metropolitan Statistical Area (MSA) concept with CBSAs in 2003.

[24] This is the same procedure used by Dorn (2009) and Autor and Dorn (2013) to map CGs/PUMAs into commuting zones.

[25] It is worth noting here that we take the set of 70 cities as a model primitive even though which cities constituted the largest 70 cities in 1980 was itself an equilibrium outcome.

hedonic regressions, holding the set of housing characteristics fixed across CBSAs and time. We describe both approaches in detail in online Appendix A3.

Finally, we use the Wharton Residential Land Use Regulation Index (WRLURI) proposed by Gyourko, Saiz, and Summers (2008).[26] The index is based on a nationwide survey of local land use regulations, with the basic idea being that regulation makes it more costly to build, making the local housing supply curve more inelastic.

C. *Household Choice Data*

We restrict the household choice sample to agents who identify themselves as the head of household and work full-time, full-year. We make this restriction to minimize concerns about labor force attachment. Further, we make use of information on a household head's state of birth, which we will use to define a home premium in our model.

As mentioned above, we allow for a rich set of observable heterogeneity. First, based on the education level of the household head, we split households into four education categories: high school equivalents, some college, college graduate, and post college. We define college graduates and beyond as "skilled." We further split the sample by marital status (single or married) and work experience (those with less than 20 years of potential experience are defined to be "less experienced," while those with more than 20 years are categorized as "more experienced"). Our final demographic characteristic is number of children. For married households, we use three categories: zero children, one child, and two or more children. The vast majority of single households in our sample do not have children, so we make the assumption that all single households have no children. This gives us 32 distinct demographic groups.

D. *Tax Calculations*

We perform our tax calculations using the NBER's TAXSIM, a tax calculator that replicates the federal income tax, payroll tax, and state income codes in a given year, accounting for differences in state income taxes, the deduction of state income taxes in calculating federal taxable income, and the differential credits afforded to various demographic groups (Feenberg and Coutts 1993). Throughout the paper, our measure of federal income taxes includes the federal income tax itself and the employee's contribution of payroll taxes. We approximate state income taxes with state-specific linear income taxes.

To get a better sense of how incomes taxes might influence household decisions, Figure 1 displays the federal income and payroll tax burden in 2017 as a function of household income, for single households without children and for married households with two children. The blue line shows the full federal income tax we consider in the model, which includes both the federal income tax and the employees contribution of the payroll tax. The dotted red line shows the federal income tax, including

[26] This dataset can be downloaded at http://real.wharton.upenn.edu/gyourko/landusesurvey.html.

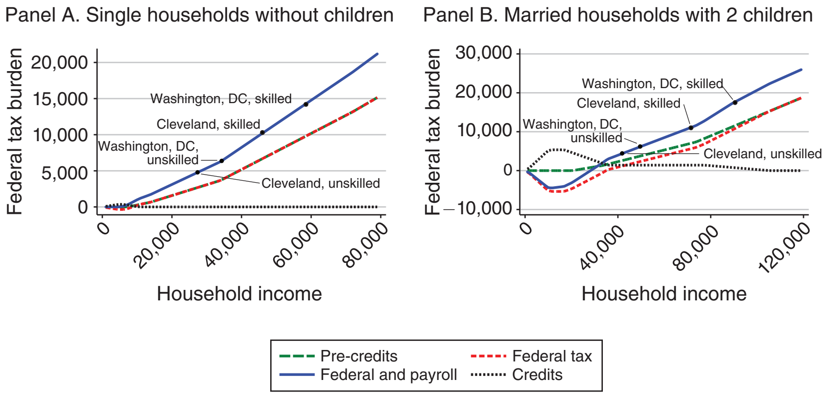

Figure 1

Notes: The panel on the left displays the 2017 federal income tax as a function of household income for single households with no children. The blue line shows the full federal income tax we consider in the model, which includes both the federal income tax and the employees contribution of the payroll tax. The dotted red line shows the federal income tax, including credits, while the green dashed line shows federal income tax before credits are applied. The four dots show the estimated earned income for Cleveland and Washington, DC, for households with less experience who are either high school equivalents or who are college graduates. The panel on the right repeats the exercise for married households with two children.

credits, while the green dashed line shows federal income tax before credits are applied. In the panel on the left, these two lines are mostly overlapping, as credits for single households without children are small. Finally, the black dotted line shows tax credits, which is equal to the difference between the red and green lines.

To this graph we add earned income, w_{jt}^{d}, in Cleveland, a low-income city, and Washington, DC, a high-income city, from 2007. Specifically, we show w_{jt}^{d} for less experienced households whose household head is either a high school equivalent or college graduate. For example, the first dot in the panel on the left indicates that, for single households with less experience, a high school equivalent household in Cleveland has earned income of $27,000 and therefore faces a federal tax burden of roughly $4,800. In contrast, a household with the same demographics living in Washington, DC, has earned average income of $32,000 and therefore faces a federal tax burden of roughly $5,900. For single households with no children who are college graduates, a household in Cleveland faces a federal tax burden of $10,300, while a similar household in Washington, DC, faces a federal tax burden roughly $14,200.[27]

The panel on the right shows the tax burden for married households with two children. For these families, high school equivalent households in Cleveland face a tax burden of $4,500, while unskilled households in Washington, DC, face a tax burden of $6,000. For college graduate households, a household of this demographic in Cleveland faces a tax burden of $11,000, while a household in Washington, DC, faces a tax burden of $17,500.[28]

[27] Note that these graphs do not show differences in state income taxes across cities.

[28] For the most part, the households in our model are not affected by the phaseout of tax credits. It is important to remember that our incomes are estimated using full-time, full-year households and also are composition adjusted

III. Estimation

The parameters to be estimated include the parameters of the production function, the parameters of the housing supply curve, the land ownership shares, and the household preference parameters. For computational simplicity, we estimate these sets of parameters separately.

A. *Households*

Our approach for estimating the labor supply component of the model closely mirrors the procedure commonly used for estimating differentiated product demand systems with microdata (i.e., Berry, Levinsohn, and Pakes 2004, which we refer to as BLP throughout). In particular, we estimate the parameters in two steps, where the first step estimates the home premiums, distance and distance squared parameters, and mean utilities using maximum likelihood; and the second step uncovers the wage and rent preference parameters, using instrumental variables to deal with the endogeneity of wages and rents.

Equation (3) is the basis for estimating the underlying preference parameters. We proceed by normalizing both the location and scale of this equation and redefining the parameters accordingly. In particular, we normalize the mean utility of location one to zero and divide through by σ^e, which yields

$$v_{ijt} = \delta_{jt}^d + \beta_{hp,t}^d \mathbf{1}(j \in Bstate_i) + \beta_{dist,t}^d \phi(j, Bstate_i) + \beta_{dist2,t}^d \phi^2(j, Bstate_i) + \epsilon_{ijt}, \tag{11}$$

where

$$\delta_{jt}^d = \beta_w^e \log\left(I_{jt}^d - \mathcal{T}_{jt}^d\left(I_{jt}^d\right)\right) + \beta_r^e \log(r_{jt}) + \xi_{jt}^d, \tag{12}$$

and $\beta_w^e \equiv 1/\sigma^e$, $\beta_r^e \equiv -\alpha^e/\sigma^e$, $\beta_{hp,t}^d \equiv \gamma_{hp,t}^d/\sigma^e$, $\beta_{dist,t}^d \equiv \gamma_{dist,t}^d/\sigma^e$, and $\beta_{dist2,t}^d \equiv \gamma_{dist2,t}^d/\sigma^e$. Abusing notation, we also have $\delta_{jt}^d \equiv \delta_{jt}^d/\sigma^e$, $v_{ijt} \equiv v_{ijt}/\sigma^e$, and $\xi_{jt}^d \equiv \xi_{jt}^d/\sigma^e$. Our goal is to estimate the vectors $\boldsymbol{\delta}_t^d$ and $\boldsymbol{\beta}_t^d \equiv \left[\beta_{hp,t}^d \beta_{dist,t}^d \beta_{dist2,t}^d\right]$, which is done in the first step, along with β_w^e and β_r^e, which is done in the second step. Assuming that ϵ_{ijt} is distributed i.i.d. according to the Type 1 Extreme Value distribution, we can estimate $\boldsymbol{\delta}_t^d$ and $\boldsymbol{\beta}^d$ using maximum likelihood.[29]

for differences in race and other demographic characteristics. Further, while we do consider 32 different demographic groups, we still aggregate over some groups. A single-parent minority household where the parent is not working full-time, for example, would likely experience significant reductions in the Earned Income Tax Credit as a result of moving from a low-income to a high-income city.

[29] The choice probabilities and likelihood function are given in online Appendix A1. Computationally, we invert the choice probabilities using the contraction mapping in Berry (1994) to obtain the unique $\boldsymbol{\delta}_t^d$ associated with every $\boldsymbol{\beta}_t^d$.

The second step uses our estimates from the first step to uncover the underlying preference parameters. In particular, we pool all the δ_{jt}^{d} terms within each skill group and estimate

$$\delta_{jt}^{d} = \beta_{w}^{e} \log\left(I_{jt}^{d} - \mathcal{T}_{jt}^{d}\left(I_{jt}^{d}\right)\right) + \beta_{r}^{e} \log\left(r_{jt}\right) + \xi_{jt}^{d} \tag{13}$$

using IV to address the endogeneity of wages and rents. Note that by instrumenting for wage change with these instruments, we can recover the structural parameters, β_{w}^{e} and β_{r}^{e}, and not just the reduced form elasticities of household mobility with respect to taxes. This differs from the strategy employed in Moretti and Wilson (2017), whose reduced form estimates of location choice capture both the labor supply and labor demand responses to taxes.

First, consider identification of the wage coefficient, β_{w}^{e}. The equilibrium nature of the model mechanically induces a correlation between after-tax income, $I_{jt}^{d} - \mathcal{T}_{jt}^{d}\left(I_{jt}^{d}\right)$, and the unobserved amenity, ξ_{jt}^{d}. If some place becomes unobservably more attractive for only skilled households, this will induce in-migration of skilled households, thus driving down their wages.

To address this issue, we take advantage of the fact that our model explicitly accounts for income taxes. Typically, sorting models of this type (Diamond 2016, Piyapromdee 2017) need variation in labor demand, but in our case we need either variation in labor demand or variation in income taxes. We use the latter by constructing instruments similar to those developed in Gruber and Saez (2002). In particular, we use changes in the tax code to generate local labor market variation in after-tax income. We implement this instrument by calculating the change in after-tax income that would have occurred from changes in the tax code, had there been no local changes in labor demand. More formally, we construct our instrument as

$$\Delta Z_{j}^{d}(t) = \frac{\mathcal{T}_{jt}^{d}\left(I_{jt-1}^{d}\right) - \mathcal{T}_{jt-1}^{d}\left(I_{j,t-1}^{d}\right)}{I_{jt-1}^{d} - \mathcal{T}_{jt-1}^{d}\left(I_{jt-1}^{d}\right)}, \tag{14}$$

where $\mathcal{T}_{jt}^{d}(\cdot)$ maps pretax income into taxes, where the subscript on $\mathcal{T}_{jt}^{d}$ denotes the tax code in year t given demographics d. In words, equation (14) holds local income, $I_{j,t-1}^{d}$ constant while using changes in federal and state tax codes.[30]

To identify β_{r}^{e}, we follow Diamond (2016) and interact $\Delta Z_{j}^{d}(t)$ with our measure of housing supply elasticity. The idea is that as households in-migrate to take advantage of higher after-tax income, cities with more inelastic housing supply curves will see rents bid up faster. This creates variation in rent levels across cities that is assumed to be exogenous to changes in unobserved amenities. Households' responsiveness to these rent changes is used to identify β_{r}^{e}.

In our preferred specification, we also use a Bartik instrument when estimating labor supply (Bartik 1991). As the approach is rather standard, we include further

[30] Online Appendix A2 provides some suggestive evidence that changes in the national tax code in the data have led to significant changes in differences in after-tax income and shows an example of the local labor market variation generated by these tax changes.

details on this instrument in online Appendix A2.[31] Our estimates of β_w^e and β_r^e when we do not use the Bartik instrument are qualitatively similar but slightly larger in magnitude.

Estimation of Profit Share.—Next, we need to estimate s_t^e, the share of landowner profits owned by households of each skill level. We estimate these shares using data on interest, dividend, and rental income from the census and ACS data. We estimate that the share of landowner profits that is owned by all skilled households has increased substantially over time, from 0.44 in 1980 to 0.71 in 2007.[32] Therefore, in later years, increases in landowner profits can have important implications for inequality as most of the profits are captured by skilled households.

B. *Labor Demand*

We estimate the labor demand parameters of our model using the relative labor demand curve defined in (6). First, we parameterize the factor intensity parameters as

$$\log\left(\frac{\theta_{jt}}{1-\theta_{jt}}\right) = \alpha_0^t + \alpha_1\left(t \times \log Pop\,80_j\right) + \alpha_2\left(t \times \log Pop\,80_j^2\right) + \mu_j + \varepsilon_{jt},$$

where $\log Pop\,80_j$ is the log population of city j, measured in 1980. The term α_0^t allows for skill-biased technical changes that affect all cities equally. The next two terms allow for the possibility that initially larger cities may experience differential changes in relative labor demand, as documented by Baum-Snow and Pavan (2013).[33] Finally, μ_j allows for city-specific differences in relative demand for skilled labor that are persistent over time.

Taking differences of the relative labor demand curve yields the following estimating equation:

$$\Delta\log\left(\frac{w_{jt}^S}{w_{jt}^U}\right) = \Delta\,\alpha_0^t + \alpha_1 \log Pop\,80_j + \alpha_2 \log Pop\,80_j^2 - \frac{1}{\varsigma}\Delta\log\left(\frac{S_{jt}}{U_{jt}}\right) + \Delta\varepsilon_{jt}. \tag{15}$$

The concern in estimating equation (15) is that unobserved changes in skill-biased labor demand $(\Delta\varepsilon_{jt})$ induce changes in the quantities of skilled labor (S_{jt}) across labor markets. This will lead to a correlation between unobserved changes in skill-biased labor demand and skilled labor, and, therefore, biased estimates of the elasticity of substitution ς.

[31] Online Appendix A2 shows alternative estimates of labor supply parameters in which we only use our tax instruments and in which we only use the Bartik instrument.

[32] That is, s_t^e for skilled households multiplied by the total number of skilled households is 71 percent of total interest, dividend, and rental income in 2007.

[33] Baum-Snow, Freedman, and Pavan (2018) argue that this increase in wage inequality in larger cities is partially driven by skill-biased agglomeration, a force that is absent from our model.

Therefore, we estimate equation (15) using the instruments proposed in Card (2009) and Moretti (2004) to instrument for changes in the labor supply ratio $\left(\Delta\log\left(S_{jt}/U_{jt}\right)\right)$. The Card "ethnic enclave" instrument interacts the lagged geographical distribution of immigrants from different countries with current, aggregate inflows from those countries. The Moretti instrument interacts the long-term trend of increasing educational attainment with the lagged age structure of labor market. For instance, labor markets that are disproportionately young or old are predicted to have larger increases in skilled labor. Formal definitions of these instruments can be found in online Appendix A1.

C. *Housing Supply*

Now we turn to the estimation of the housing supply curve in equation (7). Taking first differences of equation (9), the reduced-form equation for housing rents, we obtain our estimating equation for housing supply:

$$(16) \quad \Delta\log\left(r_{jt}\right) = \left(\nu_1 + \nu_2 \psi_j^{WRI}\right)\Delta\log\left[\sum_d N_{jt}^d \hat{\alpha}^e\left(I_{jt}^d - \mathcal{T}_{jt}^d\left(I_{jt}^d\right)\right)\right] + \Delta\tilde{\zeta}_{jt},$$

where $\hat{\alpha}^e$ is the estimate of the budget share of housing (described above), ν_1 and ν_2 are parameters to be estimated, and r_{jt}, N_{jt}^d, and $I_{jt}^d - \mathcal{T}_{jt}^d\left(I_{jt}^d\right)$ are observed in the data.

The estimation strategy is to use variation in housing demand to identify housing supply. However, as with the wage equations, the concern with estimating equation (16) via least squares is that $\Delta\tilde{\zeta}_{jt}$ will be correlated with $\Delta\log\left(N_{jt}^d\right)$ because households prefer locations with lower rents. Therefore, we again utilize the Card instrument and Bartik instrument to instrument for changes in population.

IV. Results

A. *Parameter Estimates*

Estimates for all the key parameters are displayed in Table 1. This section discusses each of the estimates related to the labor demand, housing supply, and labor supply components of the model. First, we estimate an elasticity of substitution, ς, of 3.63, which is on the higher end of estimates in the literature, but within reason. In particular, studies that use local labor market variation tend to estimate higher elasticities of substitution than those using time-series variation at the national level. Card (2009), for example, using variation in immigrant inflows across cities, estimates elasticities of substitution between 2.5 and 4. Katz and Murphy (1992) and Heckman, Lochner, and Taber (1998), using national level variation, estimate elasticities of substitution between 0.9 and 1.4.

The next panel shows the parameters governing the housing supply curves. We estimate ν_1 to be 0.33 and ν_2, the interaction between housing consumed and the Wharton Regulation Index, to be 0.13. The signs of these estimates imply that more housing consumed in a local labor market leads to higher rents in that location and this rent increase is larger in labor markets with tighter land use restrictions. The mean inverse housing supply elasticity is 0.57, similiar to Saiz (2010)

Table 1—Parameter Estimates

Panel I. Labor demand			
ς: Elasticity of substitution	3.63		
	(1.36)		
Partial F-statistic	10.30		
Panel II. Housing supply			
ν_1: Baseline	0.33	ν_2: Regulation	0.13
	(0.07)		(0.05)
Partial F-statistic	22.38		
	Unskilled	Skilled	
Panel III. Labor supply			
β_w^e: Wage	7.15	12.54	
	(1.35)	(2.20)	
β_r^e: Rent	−5.23	−5.23	
	(0.96)	(0.95)	
α^e: Share housing	0.73	0.42	
	(0.03)	(0.02)	
Cragg-Donald Wald F-statistic	12.84	9.86	

Note: Standard errors are in parentheses.

and Piyapromdee (2017), and higher than the mean housing supply elasticity in Diamond (2016). Furthermore, as we show in online Appendix A2, inverse housing supply elasticities and city-level earnings are highly correlated—higher wage cities tend to have more inelastic housing supply curves. Therefore, tax changes that shift the distribution of households to more productive cities also imply a shift toward cities with more inelastic housing supplies.

The bottom panel of Table 1 contains estimates for the labor supply component of the model. We estimate β_w^e to be 12.54 and 7.15, respectively, for skilled and unskilled households.[34] To get a better sense of what our estimates imply for mobility, we simulate the equilibrium changes in local population associated with a local increase in income. We do this by separately giving a 1 percent income subsidy for households of a given education level in each city and calculating the resulting equilibrium.[35] We find that these general equilibrium elasticities are considerably lower than the partial equilibrium elasticities, with a mean general equilibrium elasticity of 2.0 for unskilled households and 5.7 for skilled households.

Our estimates of β_w^e are larger than the parameters on pretax wages in Diamond (2016) and Piyapromdee (2017). As the tax is progressive, changes in log pretax wages translate to smaller changes in log posttax income. This should lead to larger coefficient estimates. Moretti and Wilson (2017) find a 1 percentage point increase in local income tax leads to a 4.0 percent decrease in the population of star scientists living in a state over a ten year period. This implies a long-run elasticity of location

[34] Online Appendix A2 shows alternative estimates of labor supply parameters in which we only use our tax instruments and in which we only use the Bartik instrument. When we only use the tax instrument, we estimate β_w^e to be 13.83 and 9.57 for skilled and unskilled households, respectively, and β_r^e to be −5.63 and −7.58.

[35] Specifically, we assume that wages for skill group e in city j' are given by $\hat{I}_{j't}^e = 1.01 I_{j't}^e$, where $I_{j't}^e$ is the equilibrium wage without the subsidy.

choice with respect to posttax wages of 4, slightly smaller than our estimated equilibrium elasticity for skilled households.

Next, we estimate β_r^e, the partial equilibrium elasticity of the location choice probability with respect to rents, to be -5.23 for both skilled and unskilled households. Together, our estimates of β_w^e and β_r^e imply the budget shares of housing, α^e, to be 0.42 and 0.73, respectively, for skilled and unskilled households. Estimates of α vary within the local labor market literature. Suárez Serrato and Zidar (2016), using data from the CEX, estimate a budget share of housing of 0.3. Moretti (2011) finds households in the 2000 census spend 41 percent of their income on housing. Diamond (2016) calibrates a expenditure share of local goods of 0.62.[36]

The estimates of the birthplace premium, distance, and distance squared estimates for each demographic group are available on request. Generally speaking, the birthplace premium and the distance cost are decreasing in education, reflecting the greater propensity of skilled agents to live away from their birth state. The estimates of the distance squared parameters suggest that cost-of-living far from one's birth is concave in distance.

B. *Model Fit*

In this section, we analyze how well our model can replicate the data. Recall that we estimate separate unobserved amenities for each city-demographic group combination, which implies that we will exactly match the population of each demographic group in each city. Therefore, we plot the simulated and observed average log distance between an agent's birth state and chosen city for each city. Figure 2 plots these average log distances for unskilled and skilled households for the year 2007. Each dot represents a CBSA, and the size of the dot is proportional to population. Overall, the model appears to fit this aspect of the data fairly well. The results for other years are similar and are available in online Appendix A2.

V. Counterfactuals

There are two main goals of the counterfactuals section. The first is to measure the distortion of the current tax code. In Section VA, we calculate the deadweight loss and incidence of the 2017 income tax code and compare the distribution of wages, rents, and populations to those under non-distortionary lump-sum taxes in which tax revenue from each demographic group is held constant. We then proceed to separately analyze the incidence of the federal income tax, state income taxes, payroll taxes, and tax credits.

Lump-sum taxes are efficient but may not be practical to implement. Therefore, the second goal is to evaluate several alternative tax codes which may lead to efficiency gains. In Section VB we assess the efficiency gains from switching to these

[36] As noted by Moretti (2013), the price of housing is generally correlated with the price of other local goods. Therefore, our estimates of α may also be capturing expenditure on local goods whose prices are correlated with the local price of housing. In Section VI, we examine the sensitivity of our results to alternative estimates of α. We find a deadweight loss of the current tax code equal to 0.31 percent of output (compared to 0.25 percent of output given our baseline estimates) when we set $\alpha = 0.4$ for all groups.

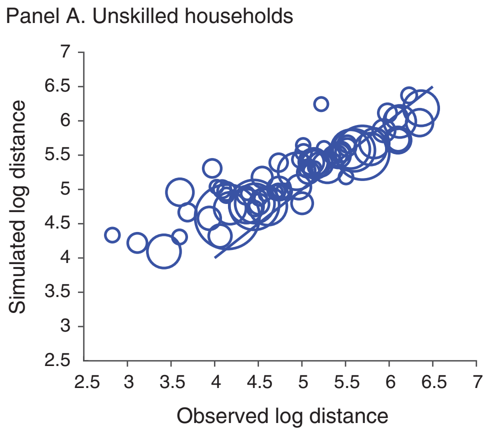

FIGURE 2

Notes: This figure shows the average log distance traveled from birth state in each city in the data and predicted by the model for unskilled households and skilled households in 2007. The horizontal access is the average log distance from birth state in the data; the vertical access is the model's prediction. The results for other years are similar and are available in online Appendix A2.

various tax programs, and calculate the effects of these new taxes on between group welfare inequality.

For all counterfactuals, we fix households' characteristics (demographics, education, birth states, etc.), the value of unobserved amenities for each demographic group, and labor demand parameters to their 2007 values. We then calculate the new equilibrium allocation under each counterfactual tax code. We calculate deadweight loss as the sum of the equivalent variation of switching from the lump-sum tax to each tax code in question.[37] Our measure of deadweight loss is the sum of total equivalent variation measured as a fraction of total output under the baseline tax code. To calculate total output, we assume a labor share of output of $\eta = 0.66$.[38]

A. *The Incidence of the Current Tax Code*

In this section, we calculate the effects of the current income tax codes. The main results are summarized in Table 2. Column A shows the results with the full set of federal and state income taxes. Each of the following columns removes state taxes, payroll tax, and tax credits and replaces them with lump-sum taxes which keep tax revenue from each demographic group constant. Column E therefore shows the case when only the non-distortionary lump-sum taxes remain. In order to visually represent the distortion caused by each tax code, Figure 3 graphs the changes in population, rents, and wages relative to the equilibrium with lump-sum taxes. Specifically, we graphically compare each counterfactual allocation to the allocation

[37] This corresponds with the definitions of deadweight loss suggested by Mohring (1971) and Kay (1980). Calculating the forgone tax revenue of not having access to a lump-sum tax holding utility constant (similar to the measure in Albouy 2009) yields nearly identical results. For a discussion of various measures of deadweight loss and excess burden, see Auerbach (1985).

[38] This is a relatively standard calibration. See Gollin (2002) for a discussion.

Table 2

	A. Full fed and state	B. A – state taxes	C. B – payroll tax	D. C – credits	E. Only lump sum
Panel I. Percent high-income cities					
Unskilled	16.3	16.6	17.1	17.1	18.4
Skilled	25.9	27.1	27.6	27.6	32.9
Panel II. Percent outside large cities					
Unskilled	44.8	44.6	43.8	43.7	41.1
Skilled	30.7	30.2	29.4	29.3	23.7
Panel III. Deadweight loss	0.25	0.18	0.14	0.14	0.00
Panel IV. Landowner profits	8.76	8.84	8.95	8.95	9.57
Panel V. Average tax incidence (\$1,000)					
Unskilled	−0.3	−0.3	−0.3	−0.3	0.0
Skilled	1.3	1.1	0.9	0.9	0.0

Notes: This table shows the main effects of the current income tax code. Column A shows the results with the full set of federal and state income taxes, including the federal income tax, state income taxes, the employee contribution of the payroll tax, and tax credits. Each of the following columns removes state taxes, payroll tax, and tax credits and replaces them with lump-sum taxes, which keep tax revenue from each demographic group constant. Column B removes state income taxes, column C removes payroll taxes, and column D removes tax credits. Column E shows the case when only the non-distortionary lump-sum taxes remain. High-income cities are the 10 cities with the highest average skilled income in 2007. Tax incidence is the average equivalent variation of moving from the lump-sum tax to the tax code under consideration, measured in thousands of dollars. Deadweight loss and landowner profits are measured as a percentage of baseline output.

with demographic specific lump-sum taxes such that tax revenue raised from each demographic group is the same as the 2017 tax code.[39]

Panel I of Table 2 shows the percent of skilled and unskilled households in the cities that had the highest skilled wages in 2007. Under the current tax code, 16.3 percent of unskilled households and 25.9 percent of skilled households live in these high productivity cities. Compared to non-distortionary lump-sum taxes (column E), this implies the population of skilled households in the 10 highest paying cities decreases by 21 percent while the population of unskilled households decreases by 11 percent. As we can see in Figure 3, this leads to a decrease in rents and in the college share in high-income cities. As a result, pretax wages for unskilled households in high-productivity cities decrease while pretax wages for skilled households in these cities increase. Next, in panel II, we examine the percent of each group that does not live in one of the 70 largest cities we consider in the model.[40] These areas outside of large cities generally have lower incomes than large cities. Compared to the non-distortionary equilibrium, the number of households outside of the 70 largest cities increases by 9 percent and 30 percent for unskilled and skilled households, respectively. Overall, these distortions lead to a deadweight loss of 0.25 percent of total output.

[39] All lump-sum taxes are efficient. Keeping revenue from each group constant allows us to focus on the distortions caused by income taxes, without transferring wealth across demographic groups.

[40] As explained earlier, outside of the 70 CBSAs we consider, we group all other locations into 9 census divisions.

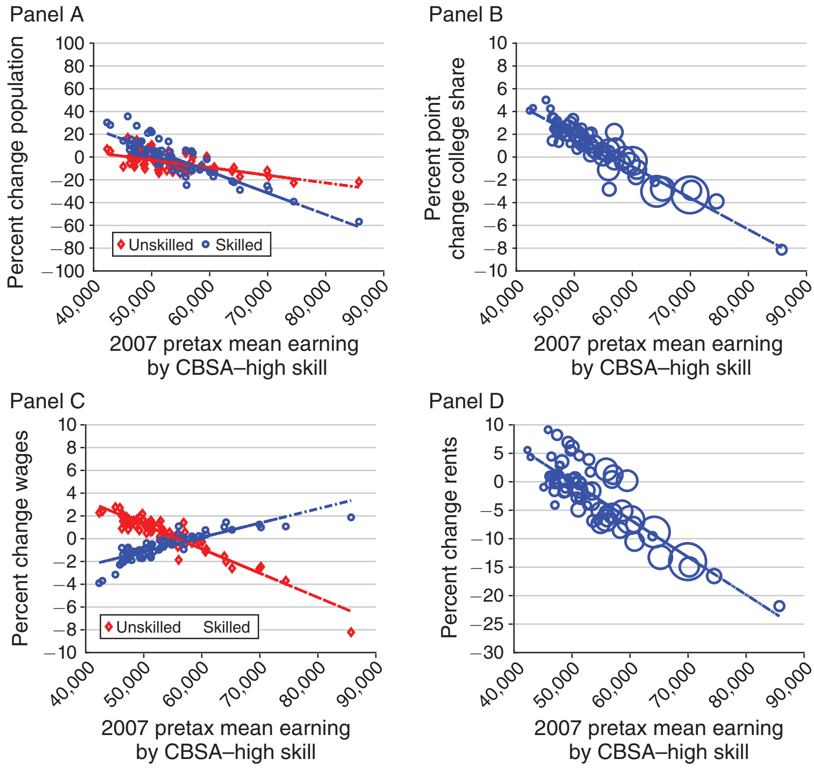

Figure 3

Notes: This figure shows the counterfactual changes in population levels (panel A), college shares (panel B), log wages (panel C), and rent (panel D) as a result of changing from an equilibrium with non-distortionary lump-sum taxes to the current tax code. Each dot represents a CBSA. The horizontal axis is the 2007 pretax mean earnings for skilled households. Unobserved amenities, agent demographics, and labor demand parameters are held fixed at their 2007 levels. Bubbles are proportional to the 1980 CBSA population in panels B and D.

Columns B through E of Table 2 sequentially remove state income taxes, payroll taxes, tax credits, and the federal income tax and replace them with demographic-specific lump-sum taxes such that tax revenue collected from each demographic group is constant. Removing distortions from state income taxes reduces the deadweight loss from 0.25 percent of baseline output to 0.18 percent of baseline output as households move towards more high-income cities.[41] Additionally removing payroll taxes further decreases deadweight loss by another 0.04 percentage points. Finally, when only lump-sum taxes remain, deadweight loss is 0.

Panel IV shows landowner profits, measured as a percentage of baseline output. When we remove tax distortions, households move toward high-income cities with more inelastic housing supply curves. As we can see, landowner profit increases as a result, from 8.76 percent of baseline output in the baseline case to 9.57 percent of baseline output with non-distortionary taxes.

[41] This 0.07 percent decrease in deadweight loss equals roughly 0.5 percent of tax revenue, nearly the same as the increase in household welfare as a result of harmonizing state taxes in Fajgelbaum et al. (2018).

As we sequentially remove tax distortions, the distribution of the tax incidence across skilled and unskilled households changes as well. Panel V shows the tax incidence of each tax code measured as average equivalent variation of moving from the lump-sum tax to the tax code in question. In general, all tax codes except for lump-sum taxes make skilled households worse off while making unskilled households better off, despite the fact that tax revenue from each demographic group is kept constant. Under the current tax code, the average equivalent variation for skilled households is equal to \$1,300 per household, compared to −\$300 per household for unskilled households. This version of the equity-efficiency trade-off is unique to the spatial setting: when we reduce the tax burden in higher income cities we increase the size of the pie but reduce the fraction that goes to unskilled households.

B. *Counterfactual Tax Codes*

We have seen that the current tax code leads to a deadweight loss from locational inefficiency equal to 0.25 percent of output. In this section we consider three counterfactual tax changes aimed at reducing this inefficiency. For each of the three counterfactual tax codes, we keep state income taxes constant and chose either the flat tax percentage or the parameter $\bar{\kappa}_t$ to keep total tax revenue from the sum of federal and state taxes equal to the baseline tax revenue.[42]

A Flat Tax.—The main effects of the flat tax are summarized in the second column of Table 3 and the left panel of Figure 4, which shows the skilled and unskilled population levels in each city relative to an equilibrium with non-distorionary lump-sum taxes.[43] Moving from the current tax code to a flat tax leads to a moderate increase in the number of households living in high-income cities. The percent of unskilled households in the top 10 highest paying cities increases from 16.3 percent of all unskilled households to 16.9 percent and the percent of total skilled households increases from 25.9 percent to 27.0 percent. However, from the first panel of Figure 4, we can see that the number of households living in high paying cities is still substantially lower than the case with non-distortionary taxes. Deadweight loss with a flat tax is 0.16 percent of baseline output.

Next, we turn to the effects of changing from the current tax code to a flat tax on between-group inequality. Changing the tax code can lead to an increase in inequality for a number of reasons. First, there is a mechanical effect; holding prices and household locations constant, changing the tax code may differentially affect skilled and unskilled households. Second, there is a sorting effect: skilled households are more mobile than unskilled households and may be better able to respond to changes in the tax code. Finally, there are price and landowner profit effects: changes in the sorting of households lead to changes in prices and landowner profits that may differentially affect skilled relative to unskilled households.

[42] We consider the case with harmonized state taxes in online Appendix A2. The results with harmonized state taxes are qualitatively similar but with slightly larger efficiency gains.

[43] For the flat tax, we also remove tax credits and payroll tax. We choose the flat tax so that revenue from the flat tax is equal to the total tax revenue from the current tax code, including credits, and the payroll tax.

TABLE 3

	Baseline	Flat	COLA	Wage adjustment
Panel I. Percent high-income cities				
Unskilled workers	16.3	16.9	18.6	18.1
Skilled workers	25.9	27.0	33.9	31.8
Panel II. Percent outside large cities				
Unskilled workers	44.8	43.7	40.7	41.1
Skilled workers	30.7	29.3	23.1	23.9
Panel III. Deadweight loss	0.25	0.16	0.09	0.04
Panel IV. Landowner profits	8.76	8.83	9.63	9.47

Notes: This table shows the main effects of counterfactual tax codes. High-income cities are the 10 cities with the highest average skilled income in 2007. Deadweight loss and landowner profits are measured as a percentage of baseline output. See the text for details.

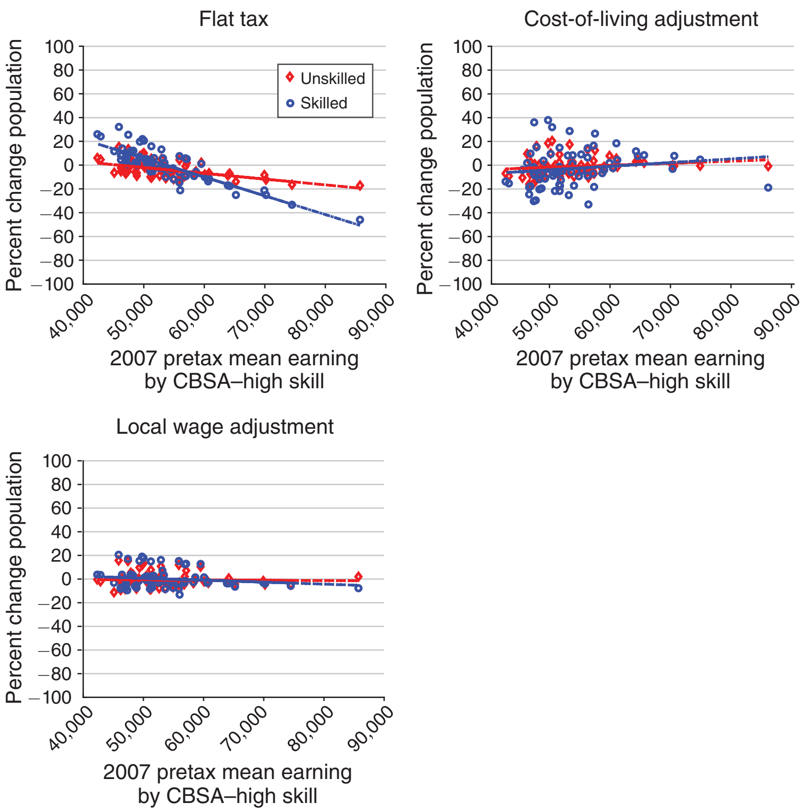

FIGURE 4

Notes: This figure shows the counterfactual population relative to lump-sum taxes for a flat tax, cost-of-living adjustments, and local wage adjustments. Each dot represents a CBSA. The horizontal axis is the 2007 pretax mean earnings for skilled households. Unobserved amenities, agent demographics, and labor demand parameters are held fixed at their 2007 levels.

We attempt to understand the contribution of these three mechanisms on between-group well-being inequality in the first column of Table 4. Panel I measures the mechanical effect of the tax change by calculating the change in tax burden

Table 4

	Flat	COLA	Wage adjustment
Panel I. Mechanical			
Unskilled	1.97	0.62	0.39
Skilled	−3.74	0.43	−0.42
Difference	5.71	0.19	0.81
Panel II. Mechanical + Sorting			
Unskilled	1.96	0.08	0.20
Skilled	−3.76	−1.49	−1.55
Difference	5.72	1.57	1.75
Panel III. Full effects			
Unskilled	2.01	0.60	0.43
Skilled	−3.73	−1.42	−1.30
Difference	5.74	2.02	1.74

Notes: This table shows the average equivalent variation of moving from the current tax code to each of the counterfactual tax codes, measured in thousands of dollars. Panel I shows the effect of changing from the current tax to each counterfactual tax code, holding all location choices, prices, and profits constant at their baseline levels (the "mechanical effect"). In panel II households are allowed to optimally chose their location in response to the tax change, but prices and landowner profits are held at their baseline levels. Panel III incorporates equilibrium effects on price and landowner profits. Therefore, each entry captures the full effect of switching from the current tax code to each of the counterfactual tax codes.

for each skill group holding all prices and location choices constant. Switching from a progressive tax to a flat tax raises the tax rate for lower income households and decreases the tax rate for high-income households, mechanically increasing between-group well-being inequality. Specifically, the average unskilled household experiences a drop in utility equal to a \$1,970 decrease in yearly income, while the average skilled household experiences an increase in utility equal to a \$3,740 increase in yearly income.

In panel II, we consider the tax incidence when households are allowed to sort in response to tax changes, but equilibrium wages, rents, and profits are held at their baseline levels. Equivalent variation for both groups decreases slightly, implying small utility gains when households are allowed to reoptimize their location choices in response to the tax change. These small gains are consistent with the result we have seen from Table 3 that the distribution of households across cities with a flat tax is not drastically different than the baseline case.

Finally, panel III incorporates equilibrium effects on wages, rents, and profits, and therefore measures the full effect of changing to flat taxes on inequality. Relative to panel II, incorporating these effects deceases well-being slightly for both skill levels, but more for unskilled households. This occurs because the tax change leads to an increase in landowner profits by shifting the distribution of households toward cities with more inelastic housing supply curves and higher wages. As landowner profits are disproportionately distributed toward skilled households, this leads to an increase in well-being inequality. Taken together, these results imply a difference in average equivalent variation of \$5,740, most of which is due to the mechanical effect of the tax change.

Cost-of-Living Adjustments.—Next we consider the two location-specific tax deductions we introduced in Section IE: a cost-of-living adjustment, and a local wage adjustment. As described in greater detail in Section IE, under these two location-specific tax deductions, individuals in location j are taxed $\mathcal{T}^d_{jt}\left(\bar{\kappa}_t I^d_{jt}/\kappa_{jt}\right)$, where κ_{jt} is a tax index for location j and $\bar{\kappa}_t$ is a parameter. With a cost-of-living adjustment, taxable income is indexed by the local price of a market basket, such that individuals are taxed on real income.

Cost-of-living adjustments lower the tax burden of households living in high cost-of-living cities. From the upper right panel of Figure 4 and the second column of Table 3, we can see that the introduction of this tax leads to a large relocation towards high-income cities. Compared to the current tax code, the percent of unskilled households in the top 10 highest income cities increases from 16.3 to 18.6 percent, while the percent of skilled households in these cities increases from 25.9 to 33.9 percent. Similarly, the percent of unskilled households outside of large cities decreases from 44.8 to 40.7 percent while the percent of skilled households outside big cities decreases from 30.7 to 23.1 percent. These relocations toward more productive cities reduce the deadweight loss to 0.09 percent of baseline output. Further, landowner profits increase from 8.76 percent of baseline output under the current tax code to 9.63 percent.

In the second column of Table 4, we examine the effects of cost-of-living adjustments on between-group well-being inequality. Mechanically, adding cost-of-living adjustments does not lead to large changes in welfare inequality. However, once we allow households to sort in response to the tax change, the difference in average equivalent variation increases to \$1,570, as skilled households take advantage of the tax deduction by moving to higher rent cities. Finally, when we allow prices and profits to change, the difference in average equivalent variation increases again, to \$2,020, as skilled households benefit disproportionately from the increased landowner profits. Ultimately, introducing cost-of-living adjustments to the current tax code decreases average unskilled household utility by an amount equivalent to \$600 and increases average skilled household utility by an amount equivalent to \$1,420. This increase in between-group well-being inequality is driven almost entirely by the differential mobility and land ownership of skilled versus unskilled households.

Local Wage Adjustments.—Finally, we turn to the effects of local wage adjustments, which decrease the tax burden in high wage cities.[44] When the tax adjustment is introduced, the percent of unskilled households in the top 10 highest income cities increases from 16.3 to 18.1 percent and the percent of skilled households in these cities increases from 25.9 to 31.8 percent, while the percent of unskilled households outside of large cities decreases from 44.8 to 41.1 percent and the percent of skilled households outside big cities decreases from 30.7 to 23.9 percent. From Figure 4, we can see that the distribution of households across cities is almost the same as the distribution with non-distortionary taxes. Deadweight loss is therefore very low, at only 0.04 percent of baseline output.

[44] We have experimented with indexing taxable wages to skilled wages w^S_{jt} and income levels of various skilled demographic groups. The results are qualitatively similar.

Under the current tax code, skilled households are more likely to live in high-wage cities than unskilled households. Therefore, introducing local wage adjustments mechanically benefits skilled households and mechanically harms unskilled households, leading to a difference in average equivalent variation of \$810. When households can sort in response to the tax deduction, the difference in average equivalent variation increases to \$1,750, as skilled households move to cities with larger tax deductions. Allowing prices and profits to change implies a difference in average equivalent variation equal to \$1,740. All together, introducing local wage adjustments to the current tax code leads to a decrease in unskilled household utility equal to a \$430 decrease in income and in increase in skilled household utility equal to \$1,300 increase in income. The increase in between-group well-being inequality is driven largely by the differential mobility of skilled versus unskilled households. The size of the pie has increased, but the proportion going to unskilled households has decreased.

Finally, to better understand the distributional effects of each counterfactual tax code, Figure 5 displays the distribution of equivalent variation for skilled and unskilled households of moving from the current tax code to each of the counterfactual tax codes. While the average skilled household benefits from each tax change and the average unskilled household is harmed, we can see that there are significant distributional effects within each skill group. This is especially true for the cost-of-living adjustment and wage adjustment. Many unskilled households live in high cost-of-living and high-wage cities and therefore benefit from the adjustments while many skilled households live in low cost-of-living and low wage cities and are harmed benefit by the adjustments.

VI. Sensitivity to Parameter Estimates

The incidence of the tax depends crucially on the elasticity of household location choice with respect to earnings. In this section, we first examine the sensitivity of our deadweight loss and welfare calculations to this elasticity. Specifically, we recalculate deadweight loss of the current tax code for a range of values for the scale of the preference shock, σ^e, for skilled and unskilled households. Lower values of σ^e imply that households are more responsive to differences in earnings across locations, and therefore an income tax will lead to larger distortions.[45]

The deadweight loss and tax incidence for unskilled and skilled households for a range of values of σ^e for skilled and unskilled households are shown in Figure 6. In the upper left panel, we display the deadweight loss on the z- (vertical) axis, $\sigma^e/(\sigma^e+1)$ for unskilled households on the x-axis, and $\sigma^e/(\sigma^e+1)$ for skilled households on the y-axis. Note that $\sigma^e/(\sigma^e+1) = 0$ implies households are perfectly mobile as $\sigma^e = 0$, and $\sigma^e/(\sigma^e+1) = 1$ implies no mobility as $\sigma^e = \infty$. We can see that deadweight loss is strongly decreasing in σ^e for skilled households, but the effects of changes in this parameter for unskilled households are small.

[45] For each counterfactual in which we change σ^e, we recalculate the amenity values ξ_{jt}^d such as to keep the mean utility δ_{jt}^d equal to it's baseline level. Therefore, the distribution of households of each demographic group given the observed tax function will be equal to the baseline distribution with the original values of β_r^e and β_w^e.

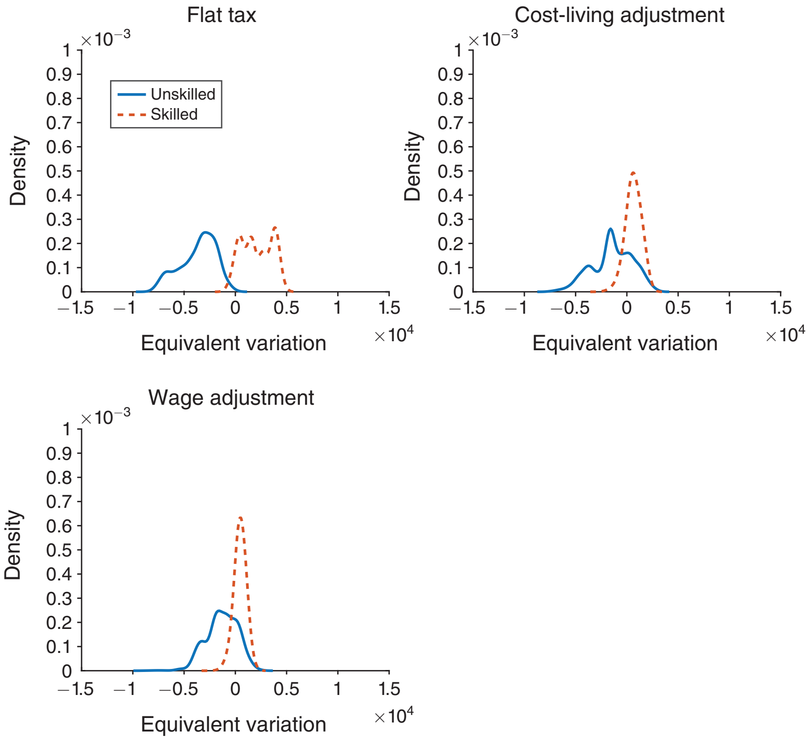

FIGURE 5

Notes: This figure shows the distribution of equivalent variation of moving from the current tax code to each of the counterfactual tax codes, measured in thousands of dollars. The solid line shows the distribution for skilled workers and the dotted line shows the distribution for unskilled workers.

As $\sigma^e/(\sigma^e + 1)$ approaches 1 for both groups, deadweight loss approaches and eventually reaches 0, as location choices are not distorted by the tax change. Lower values of σ^e imply large distortions in location choice as a result of the tax. As σ^e for skilled households approaches 0, deadweight loss exceeds 0.6 percent of output, compared to 0.25 percent of output given the estimated values of σ^e.

The next two graphs show the tax incidence for skilled and unskilled households as a function of $\sigma^e/(\sigma^e + 1)$ for skilled and unskilled households. Tax incidence is the average equivalent variation of moving from lump-sum tax to the tax code under consideration, measured in thousands of dollars. First, consider the second graph, which shows the tax incidence of skilled households. When households become more mobile (smaller σ^e), essentially two things happen. First, as we have seen in the previous figure, there is an efficiency cost. Second, the tax moves more households away from high-income cities and therefore reduces landowner profits. As skilled households own a larger share of landowner profits, this implies an increase in the share of the tax burden of skilled households. Essentially, as household mobility increases, the size of the pie decreases *and* the fraction of the pie going to skilled households decreases. Therefore, increasing mobility unambiguously increases skilled households' tax incidence.

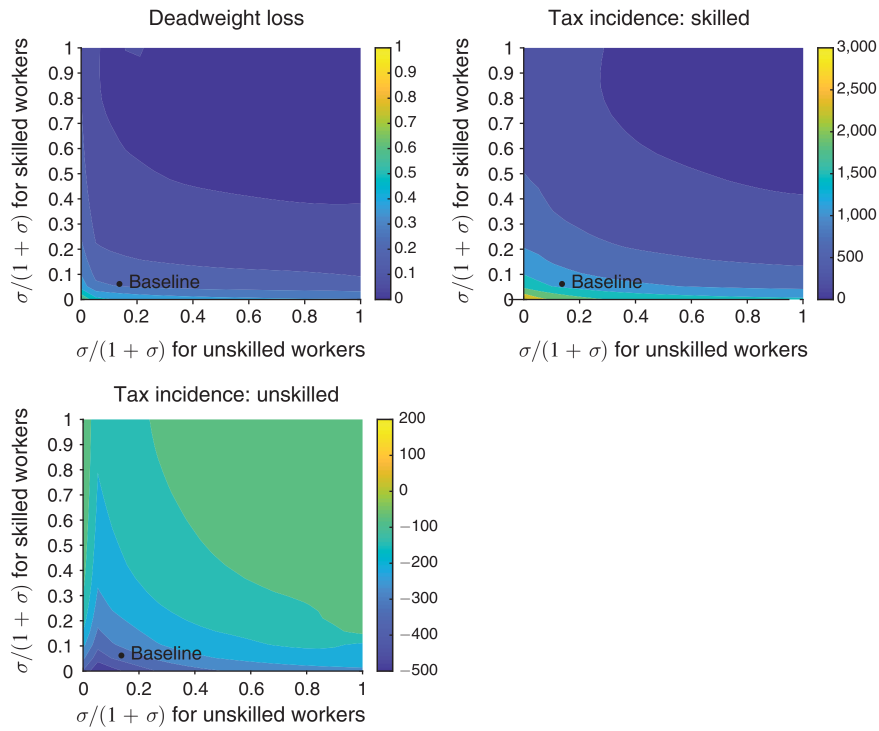

Figure 6

Notes: This figure shows the deadweight loss and tax burden for unskilled and skilled households for a range of values of $\sigma^e/(\sigma^e+1)$ for skilled and unskilled households. In each panel, we display $\sigma^e/(\sigma^e+1)$ for unskilled households on the x-axis and $\sigma^e/(\sigma^e+1)$ for skilled households on the y-axis. For each simulation, we recalculate the amenity values ξ^d_{jt} such as to keep the mean utility δ^d_{jt} equal to its baseline level.

The effects of increased mobility on unskilled households are quite a bit more complicated. Again, when household mobility increases, deadweight loss increases. However, higher mobility implies unskilled households bear a smaller fraction of the total tax burden, as landowner profits decrease. Essentially, when mobility increases, the size of the pie decreases, but the fraction of the pie going to unskilled households increases. As a result, the average tax burden of unskilled households is nonmonotonic in mobility. Starting from the no-mobility case ($\sigma^e/(\sigma^e+1) = 1$ for both groups), increasing mobility for either group decreases the tax burden of unskilled households as fraction of the tax burden borne by unskilled households decreases. However, for very low values of σ^e, the efficiency costs dominate and average unskilled household welfare decreases slightly.

Next, we examine the role of the elasticity of substitution between skilled and unskilled households. We estimated $\rho = 0.725$, which implies an elasticity of substitution of 3.63. In Table 5, we simulate the model under alternative values of the elasticity of substitution, ranging from perfect substitutes to perfect complements.[46]

[46] In the case of perfectly complementary skilled and unskilled labor, the firm's optimality condition is no longer characterized by the marginal products of labor. Instead, firms hire labor to keep factor ratios constant.

TABLE 5

	Elasticity of Substitution	ρ	Deadweight Loss	Skilled EV (\$1,000)	Unskilled EV (\$1,000)
I. Perfect substitutes	∞	1	0.33	1.1	−0.0
II. Baseline	3.63	0.725	0.25	1.3	−0.3
III. Elasticity of substitution = 2	2	0.5	0.24	1.6	−0.5
IV. Elasticity of substitution = 1.5	1.5	0.33	0.23	1.9	−0.7
V. Cobb Douglas	1	0	0.22	2.3	−0.9
VI. Perfect complements	0	$-\infty$	0.21	13.7	−7.4

Notes: This table shows the deadweight loss, and tax burden for unskilled and skilled households for a range of values of the elasticity of substitution between skilled and unskilled labor. For each simulation, we recalculate the factor intensity parameters, θ_{jt} and the TFP parameters, A_{jt}, such that wages are equal to baseline wages given the baseline distribution of labor across cities.

For each simulation, we recalculate the factor intensity parameters, θ_{jt} and the TFP parameters, A_{jt}, such that wages are equal to baseline wages given the baseline distribution of labor across cities. As we decrease the elasticity of substitution, deadweight loss drops slightly. This is because increasing the degree of complementarity between households decreases the general equilibrium elasticity of location choices with respect to wages. For example, as skilled labor exits a city, local skilled wages rise, which disincentivizes further skilled households from leaving the city. When households are perfect substitutes, the factor ratio has no effect on wages, and deadweight loss is 0.33 percent of output. However, when production is Leontief, deadweight loss drops to 0.21 percent of output.

In the following three columns, we show the implications of changing the elasticity of substitution on the tax incidence of skilled and unskilled households. Decreasing the elasticity of substitution shifts the tax burden from unskilled toward skilled households. The tax shifts skilled households out of high-paying cities, and therefore increase wages in low-paying cities where unskilled households are likely to live. The smaller the elasticity of substitution, the larger the wage increase for unskilled labor in these cities.

Next, we examine the sensitivity of our results to alternative estimates of α, the parameter that determines the budget share of housing. We rerun our counterfactuals but set $\alpha = 0.4$, based on the estimates of Moretti (2011). We find a deadweight loss of the current tax code equal to 0.31 percent of output when we set $\alpha = 0.4$ for all groups. With $\alpha = 0.4$, the average skilled household has an equivalent variation equal to \$1,200 from the current tax code and the average unskilled household has an equivalent variation of −\$200. The results are quite similar to the baseline case, with slightly larger deadweight loss and slightly smaller effects on between-group inequality.

VII. Conclusion

This paper uses a spatial equilibrium model to measure the deadweight loss and incidence of the income tax code. Our quantitative model builds upon previous work that evaluates optimal spatial taxes in two key respects. First, we relax the assumption that households are perfectly mobile, which is crucial for measuring the distortionary effects of the tax; if households' location choices do not respond strongly to taxes, then the adverse effects of the tax are mitigated. Second, we relax the assumption that households are completely homogeneous, which effectively removes any redistribution incentive from the optimal tax problem. Put simply, our contribution is to measure the incidence of income tax across heterogeneous households and to incorporate the classic equity-efficiency trade-off into the literature studying the spatial distortions of the federal income tax.

We find that the current tax code leads to a moderate deadweight loss due to locational inefficiency. However, by removing spatial tax distortions, we encourage households to live in places with higher income and more inelastic housing supply curves and thereby increase landowner profits. As land is mostly owned by skilled households, this essentially redistributes wealth to skilled households. Further, as skilled households are relatively more mobile than unskilled households, they are better able to adjust spatially to these tax changes. Therefore, if the government is not able to redistribute efficiency gains from the winners to the losers, removing spatial distortions to the tax code implies an equity-efficiency trade-off. Changes to the tax code aimed at reducing locational inefficiencies should take this into account.

REFERENCES

Albouy, David. 2009. "The Unequal Geographic Burden of Federal Taxation." *Journal of Political Economy* 117 (4): 635–67.

Auerbach, Alan J. 1985. "The Theory of Excess Burden and Optimal Taxation." In *Handbook of Public Economics*, Vol. 1, edited by Alan J. Auerbach and Martin Feldstein, 61–127. Amsterdam: Elsevier.

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

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2008. "Trends in US Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics* 90 (2): 300–23.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: Upjohn Press.

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

Baum-Snow, Nathaniel, and Ronni Pavan. 2013. "Inequality and City Size." *Review of Economics and Statistics* 95 (5): 1535–48.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and Hedonic Valuation: The Case of Air Quality." *Journal of Environmental Economics and Management* 58 (1): 1–14.

Bayer, Patrick, and Christopher Timmins. 2005. "On the Equilibrium Properties of Locational Sorting Models." *Journal of Urban Economics* 57 (3): 462–77.

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics* 25 (2): 242–62.

Berry, Steven, James Levinsohn, and Ariel Pakes. 2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy* 112 (1): 68–105.

Busso, Matias, Jesse Gregory, and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review* 103 (2): 897–947.

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

Colas, Mark, and Kevin Hutchinson. 2021. "Replication data for: Heterogeneous Workers and Federal Income Taxes in a Spatial Equilibrium." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E117661V1.

Coen-Pirani, Daniele. 2018. "Geographic Mobility and Redistribution." Unpublished. http://www.pitt.edu/~coen/research/mobredis.pdf.

Davis, Morris A., and François Ortalo-Magné. 2011. "Household Expenditures, Wages, Rents." *Review of Economic Dynamics* 14 (2): 248–61.

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

Dorn, David. 2009. "Essays on Inequality, Spatial Interaction, and the Demand for Skills." PhD diss. University of St. Gallen.

Eeckhout, Jan, and Nezih Guner. 2017. "Optimal Spatial Taxation: Are Big Cities Too Small?" Center for Monetary and Financial Studies Working Paper 1705.

Fajgelbaum, Pablo D., and Cecile Gaubert. 2018. "Optimal Spatial Policies, Geography, and Sorting." Unpublished. https://economicdynamics.org/meetpapers/2018/paper_1319.pdf.

Fajgelbaum, Pablo D., Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar. 2018. "State Taxes and Spatial Misallocation." *Review of Economic Studies* 86 (1): 333–76.

Feenberg, Daniel, and Elisabeth Coutts. 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management* 12 (1): 189–94.

Glaeser, Edward L., and Joshua D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47 (4): 983–1028.

Gollin, Douglas. 2002. "Getting Income Shares Right." *Journal of Political Economy* 110 (2): 458–74.

Gruber, Jon, and Emmanuel Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications." *Journal of Public Economics* 84 (1): 1–32.

Gyourko, Joseph, Albert Saiz, and Anita Summers. 2008. "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." *Urban Studies* 45 (3): 693–729.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics* 1 (1): 1–58.

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

Kay, John A. 1980. "The Deadweight Loss from a Tax System." *Journal of Public Economics* 13 (1): 111–9.

Kline, Patrick, and Enrico Moretti. 2014. "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs." *Annual Review of Economics* 6: 629–62.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull, and Giovanni L. Violante. 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* 68 (5): 1029–53.

Mohring, Herbert. 1971. "Alternative Welfare Gain and Loss Measures." *Economic Inquiry* 9 (4): 349–68.

Moody, Scott, and David K. Hoffman. 2003. *Federal Income Taxes and the Cost of Living: States with Lower Cost of Living Enjoy Greater Federal Tax Advantage.* Tax Foundation. https://files.taxfoundation.org/legacy/docs/8c4b12f729020a3fb81d7721c43fb7ec.pdf.

Moretti, Enrico. 2004. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data." *Journal of Econometrics* 121 (1): 175–212.

Moretti, Enrico. 2011. "Local Labor Markets." In *Handbook of Labor Economics*, Vol. 4B, edited by Orley Ashenfelter and David Card, 1237–313. Amsterdam: Elsevier.

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

Moretti, Enrico, and Daniel J. Wilson. 2017. "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists." *American Economic Review* 107 (7): 1858–903.

Ottaviano, Gianmarco I.P., and Giovanni Peri. 2012. "Rethinking the Effect of Immigration on Wages." *Journal of the European Economic Association* 10 (1): 152–97.

Piyapromdee, Suphanit. 2017. "The Impact of Immigration on Wages, Internal Migration, and Welfare." Puey Ungphakorn Institute for Economic Research Discussion Paper 69.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–78.

Rosen, Sherwin. 1979. "Wage-Based Indexes of Urban Quality of Life." In *Current Issues in Urban Economics*, edited by Peter Mieszkowski and Mahlon Straszheim, 74–104. Baltimore, MD: John Hopkins University Press.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. 2020. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V10.0.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics* 125 (3): 1253–96.
Stiglitz, Joseph E. 1982. "Self-Selection and Pareto Efficient Taxation." *Journal of Public Economics* 17 (2): 213–40.
Suárez Serrato, Juan Carlos, and Owen Zidar. 2016. "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms." *American Economic Review* 106 (9): 2582–624.
Wildasin, David E. 1980. "Locational Efficiency in a Federal System." *Regional Science and Urban Economics* 10 (4): 453–71.