Optimal Spatial Taxation: Are Big Cities too Small?*

Jan Eeckhout† and Nezih Guner‡

August 8, 2025

Abstract

We analyze the role of optimal income taxation across different local labor markets. Should labor in large cities be taxed differently than in small cities? We find that a planner, who needs to raise a given level of revenue and is constrained by free mobility of labor across cities and endogenous housing prices, does not choose equal taxes for cities of different sizes. The optimal tax schedule is place-based and tax differences between large and small cities depend on the level of government spending and the strength of agglomeration economies. Our estimates for the US imply higher optimal tax rates in bigger cities. Under the current Federal Income tax code with redistributive taxes, marginal rates are already higher in big cities which have higher wages, but the optimal difference we estimate is lower than what is currently observed. Simulating the US economy under the optimal spatial tax schedule, there are large effects on population mobility: the fraction of the population in the 5 largest cities grows by 7.65% with 3.31% of the country-wide population moving to bigger cities. The welfare gains are smaller due to the fact that much of the output gains are spent on the increased costs of housing in bigger cities. Aggregate goods consumption goes up by 0.95% while aggregate housing consumption goes down by 1.90%.

Keywords. Misallocation. Taxation. Population Mobility. City Size. General equilibrium.
JEL. H21, J61, R12, R13.

*We are grateful to seminar audiences and numerous colleagues, and in particular to Morris Davis, John Kennan, Kjetil Storesletten, Aleh Tsyvinski, and Tony Venables for detailed discussion and insightful comments. Claudio Luccioletti, Andrii Parkhomenko, and Guillem Tobías Larrauri provided excellent research assistance. Eeckhout gratefully acknowledges support by the ERC, Grant 339186. Guner gratefully acknowledges support by the ERC, Grant 263600 and the Spanish Ministry of Economy and Competitiveness Grant ECO2014-54401-P.

†UPF Barcelona, GSE-ICREA-CREI, `jan.eeckhout@upf.edu`.

‡CEMFI and Banco de España `nezih.guner@cemfi.es`.

1 Introduction

How do income taxes influence where people choose to live? Identical workers tend to earn significantly more in larger cities—a phenomenon known as the urban wage premium. For example, in the U.S., workers with the same skills earn, on average, 50% more in a large metropolitan area like New York City (with around 9 million workers) than in a smaller city such as Asheville, NC (with a workforce of about 130,000). Due to the progressive structure of federal income taxes, this wage gap translates into a nearly 5 percentage point difference in the average tax rate between the two cities, with New York workers paying more.

Setting aside other sources of heterogeneity—such as differences in skills, race, gender, or age—and abstracting from redistributive motives, we study how to design optimal spatial taxation. We do so in a general equilibrium framework where housing prices adjust endogenously, and individuals choose their location to equalize utility across cities.

Our main finding is that the optimal spatial taxation of labor income for identical workers should vary depending on where they live. We begin by computing the equilibrium allocation of workers across cities under the existing U.S. tax system. We then derive the tax schedule that would maximize overall welfare while maintaining the same level of tax revenue. In doing so, we treat each city's population as a representative agent, and assume workers are identical apart from their location. When taxes change, individuals respond by relocating, which in turn affects equilibrium housing prices. In both the current and optimal spatial tax regimes, utility is equalized across cities in equilibrium. These general equilibrium effects are central to determining the optimal spatial tax schedule and its quantitative consequences.

In this setting—where the planner is constrained by the free mobility of workers—we find that optimal income tax rates should differ across local labor markets. On one hand, the planner has an incentive to lower taxes in large, high-productivity cities to attract more workers, thereby raising total revenue. On the other hand, a larger population in these cities drives up housing prices, reducing their attractiveness to move to high productivity cities. The optimal policy balances these opposing forces. Hence, even in an economy with identical workers where the planner does not want to redistribute among identical workers, the optimal tax regime is location-specific.

We show that this trade-off between attracting workers to productive cities and the resulting increase in housing costs crucially depends on the level of government spending and its financing. As government spending rises, the optimal policy is to reduce taxes in large cities relative to small ones. There is a fiscal externality associated with revenue generation: when public spending finances economy-wide goods and services, the benefits are equally shared across locations, but the costs depend on where workers live and how much they earn. When no revenue needs to be raised, the allocation of workers across cities is efficient, and optimal tax rates are zero. But when the government requires positive revenue, it faces a trade-off. It can either (i) attract more workers to the productive city by lowering marginal tax rates—

thereby increasing the tax base and collecting more revenue despite taxing each worker less—or (ii) maintain a smaller workforce in the productive city and rely on higher marginal tax rates per worker. Because workers in large cities earn more, a broader tax base with lower rates becomes more efficient. Thus, as government spending increases, the optimal policy is to reduce marginal tax rates in large cities to attract more high-earning workers and increase total revenue.

Quantifying these forces for the US economy we find that taxes in big cities should be higher than those in smaller cities. Due to existing progressive federal income tax schedules in the US taxes in big cities are already higher than the ones in smaller cities. While there are many reasons why a redistributive tax schedule might be desirable, our spatial general equilibrium model with homogeneous workers features optimal spatial tax differences between big and small cities that should be lower than what we observe in the US tax code. Ideally, the tax schedule should have one tax rate for each city. To make this computationally feasible, in the quantitative analysis we parametrize the relation between after and before-tax wages, $\tilde{w}$ and w, as $\tilde{w} = \lambda w^{1-\tau}$, where $1 - \lambda$ is the level of taxation at mean wages and τ determines the extent of redistribution. The average tax rate is given by $1-\lambda w^{-\tau}$. Taxes are more progressive (regressive) when $\tau > 0$ ($\tau < 0$). For the benchmark economy, $\lambda = 0.856$ (i.e. the average tax rate at $w = 1$ is 14.4%) and $\tau = 0.12$. We find that the optimal value of τ is smaller, $\tau^* = 0.0123$.

For US economy, the impact of the optimal spatial tax policy is far-reaching. Implementing the optimal spatial tax schedule implies that after-tax wages increase in large cities. As a result, there is a first order stochastic dominance shift in the city size distribution. When we move from the current to optimal spatial taxes, the population in the five largest cities grows by 7.65%. About 3.31% of the workforce move from smaller to bigger cities countrywide. The aggregate output increases by 1.00%. The gains in terms of utility are, however, much smaller, only a 0.05% increase in utilitarian welfare. The small utility gain is due to the fact that most of the output gain in the more productive cities is eaten up by higher housing prices, which go up by 4.85% on average. As a result, while aggregate consumption goes up by 0.95%, aggregate housing consumption declines by 1.90%. Those moving to the big cities take advantage of the higher after tax incomes, but they end up paying higher housing prices.

The model that we use to quantify optimal spatial taxation has many features to capture the trade-offs faced by a Ramsey planner. First, the production of housing is endogenous to account for the fact that the value share of land is much higher in big cities than in small cities.[1] And it takes into account that the amount of land available for construction differs across locations. Some coastal cities are constrained by the mountains and the sea, whereas others in the interior have unconstrained capacity for expansion. Second, the model allows for congestion externalities that are increasing in city size. Third, housing is modeled in such a way that the rental price of land is retained in the economy as a transfer, while the construction cost eats up consumption goods. Fourth, we allow for amenities across different locations as the residual of the utility differences. Finally, while government

[1] See Davis and Palumbo (2008), Davis and Heathcote (2007), and Albouy and Ehrlich (2012).

expenditure is distortionary, a share of tax revenues is redistributed to the citizens, and those in less productive locations get larger per capita transfers. Even if we do not explicitly model expenditure on public goods, this accounts for the fact that tax revenues also generate benefits.[2]

We investigate in detail which features matter for the optimal level of tax differences across cities. We find that the optimal spatial tax schedule depends on the size of the government, distribution of land rents across the population, the strength of agglomeration economies, and the distribution of available land across cities. First, higher government spending reduces taxes in more productive cities relative to less productive ones, and can even imply that taxes in more productive cities are lower than less productive ones. In our benchmark economy, with a 14.4% average tax rate, the optimal level of τ^* is 0.0123. When the average tax rate rises to 17%, τ^* becomes negative (–0.0218), as it becomes optimal to raise revenue by attracting more workers to high-productivity cities through lower taxes.

Second, the distribution of land rents matters. In the benchmark, land is shared equally among all households. If land is owned by absentee landlords, the planner cannot redistribute through land rents, and the optimal spatial tax schedule implies larger tax differences between large and small cities ($\tau^* = 0.0603$), to offset the unequal benefits from rising housing prices.

Third, introducing agglomeration economies — where productivity *endogenously* increases with city size — makes it optimal to concentrate workers in large cities. This leads to a tax schedule with lower taxes in more productive cities to encourage migration ($\tau^* = -0.0323$). Indeed, welfare gains from adopting an optimal place-based tax schedule is almost three times higher when there are agglomeration economies, and implies a much larger movement of population, with top 5 cities growing by 14%.

Finally, ignoring differences in the supply of land across cities reduces taxes in more productive cities even further, making the optimal spatial tax schedule almost linear across cities. In our benchmark setup, more productive cities have less land, which constrains population growth and pushes up housing prices. If land were equally available, taxes in productive cities could be lowered further without sharp increases in housing prices, reinforcing the case for regressive taxation.

This paper is related to the work on urban accounting by Desmet and Rossi-Hansberg (2013) who analyze the effects on output from the relocation of productive resources.[3] Instead of analyzing the effect of technological change, we take the technology as exogenous and ask what the role is of federal income taxation. Our paper is also related to the literature that studies inter-state migration in the US using spatial equilibrium models, e.g. Coen-Pirani (2010), Karahan and Rhee (2014), Kaplan and Schulhofer-Wohl (2017), Davis, Fisher, and Veracierto (2021), and Albert and Monras (2022) for the US, and Ahlfeldt, Bald, Roth, and Seidel (2024) for Germany. While we study a model with homogeneous

[2] We exclusively focus on the spatial distortion at the collection side. There could also be a distortion at the benefit side, for example where big cities are more or less generous in federal benefits for the unemployed and the disabled (see for example Glaeser (1998) and Notowidigdo (2020)). Note that local benefits, just like local taxes, have no effect on the location decisions as they are financed locally. In our model, we abstract from this important channel altogether and focus on the role of active, full time workers.

[3] See also Topa, Sahin, and Violante (2014) for the role of unemployment frictions on spatial mismatch.

workers, our paper is also related to the literature that studies geographical allocation of workers with different skills, e.g. Diamond (2016). Our results on reallocation of labor across cities echoes Klein and Ventura (2009) and Kennan (2013) who analyze mobility of workers and find larger output gains. In the light of the misallocation debate in macroeconomics — the effect on aggregate output differences due to the misallocation of inputs, most notably capital, e.g. Guner, Ventura, and Yi (2008), Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) —, we add a different insight. Due to existing income taxation schemes, also labor is substantially misallocated across cities within countries. Hsieh and Moretti (2015), Herkenhoff, Ohanian, and Prescott (2018), and Parkhomenko (2023) also study spatial misallocation of labor across cities. They focus, however, on misallocation caused by restrictive housing policies. More closely related to our paper, Fajgelbaum, Morales, Suárez Serrato, and Zidar (2018) study state taxes as a potential source of spatial misallocation in the United States, and find that tax dispersion across states leads to aggregate losses.

The idea that taxation affects the equilibrium allocation is of course not new. Tiebout (1956) analyzes the impact of tax competition by local authorities on the optimal allocation of citizens across communities. Wildasin (1980), Helpman and Pines (1980) and Hochman and Pines (1993), among others, explicitly consider federal taxation and argue that it creates distortions. A common result in this literature is that a tax on the immobile factor, land, is necessary to achieve the efficient allocation. This literature, however, often studies highly stylized models that are not amenable to quantitative work. In the legal literature, Kaplow (1995) and Knoll and Griffith (2003) argue for the indexation of taxes to local wages. Albouy (2009) analyzes the question quantitatively. Starting from the Rosen-Roback trade-off with equalizing differences across locations, he calibrates the model and concludes that any tax other than a lump sum tax is distortionary. He does not, however, attempt to characterize the optimal spatial tax structure. Albouy, Behrens, Robert-Nicoud, and Seegert (2019) study optimal city size in a model where both the city size and the number of cities are allowed to vary and reach a similar conclusion to ours, i.e. that big cities are too small. What sets our work apart from the existing literature is a comprehensive quantitative framework that fully takes into account the *general equilibrium* effects, the *endogeneity* of housing prices and consumption, which in turn allows us to focus on the *optimality* of taxation. Finally, recent work builds our analysis: Colas and Hutchinson (2021) extends our model to an environment with heterogenous agents where the planner has incentives to redistribute across agents, and Huggett and Luo (2023) characterize Mirrleesean taxes in a spatial model like ours.

2 The Model

Population. The economy is populated by a continuum of identical workers. There are J locations (cities), $j \in \mathcal{J} = \{1, ..., J\}$. The amount of land in a city is fixed and denoted by T_j. Workers are identical and the total workforce in city j is denoted by l_j. The country-wide labor force is given by

$\mathcal{L} = \sum_j l_j$.

Preferences, Amenities and Congestion. All citizens have Cobb-Douglas preferences over consumption c, and housing h, with a housing expenditure share $\alpha \in [0, 1]$. This choice is motivated by Davis and Ortalo-Magné (2011), who find that expenditure shares on housing are constant across U.S. metropolitan statistical areas. The consumption good is a tradable numeraire with price normalized to one. Workers are perfectly mobile and can relocate instantaneously and at no cost.[4] Thus, in equilibrium, identical workers obtain the same utility level wherever they choose to locate. Therefore for any two cities, it must be the case that corresponding consumption bundles for a worker satisfy

$$u_j(c_j, h_j) = u_{j'}(c_{j'}, h_{j'}) \text{ for all } j \text{ and } j'.$$

Cities inherently differ in their attractiveness that is not captured in productivity (wages), but rather is valued directly by its citizens. This can be due to geographical features such as bodies of water (rivers, lakes and seas), mountains and temperature, but also due to man-made features such as cultural attractions (opera house, sports teams, etc.). We denote the city-specific amenity by a_j. We will interpret the amenities as unobserved heterogeneity that will account for the non-systematic variation between the observed outcomes and the model predictions.[5] In addition to city-specific amenities, to capture the cost of commuting, we allow for a congestion externality. The congestion depends on the city size and is given by l_j^δ, where $\delta < 0$.[6]

The utility in city j from consuming the bundle (c, h) is therefore written as

$$u_j(c, h) = a_j l_j^\delta c^{1-\alpha} h^\alpha.$$

Our assumption of a perfectly mobile workforce where workers respond instantaneously to changes in prices and taxes is a direct consequence of the Rosen-Roback framework. In reality there are of course frictions and our frictionless benchmark is only meant as a description of the economy in the long run. Frictions and the resulting transition will surely affect the welfare calculations. That said, this only pertains to net migration (the difference between in and out-migration), and gross migration is roughly three times net migration (Davis, Fisher, and Veracierto (2021)).

Technology. Cities differ in their total factor productivity (TFP), denoted by A_j. TFP is exogenously given. In each city, there is a linear technology operated by a representative firm that has access to a city-specific TFP, given by

$$F(l_j) = A_j l_j.$$

[4] The model could be extended to allow for mobility costs and location-specific preference heterogeneity, as in Fajgelbaum, Morales, Suárez Serrato, and Zidar (2018).

[5] In Diamond (2016) amenities depend endogenously on the characteristics, such as income, of individuals living in a city.

[6] See Eeckhout (2004).

Housing Supply. The existing housing stock in each city j is denoted by H_j and the production of new houses are denoted by N_j. New houses are produced by builders using capital K_j (forgone consumption) and the exogenously given land area in city j, T_j, with a CES production technology:

$$N_j = B\left[(1-\theta)K_j^\sigma + \theta T_j^\sigma\right]^{1/\sigma}, \tag{1}$$

where $\theta \in [0,1]$ indicates the relative importance of capital and land in housing production, and B the total factor productivity of the construction sector. The elasticity of substitution between K and T is given by $\frac{1}{1-\sigma}$. We assume that housing capital is paid for with consumption goods, and hence the marginal rate of substitution between consumption and housing is equal to one and the rental price of capital is equal to the numeraire. The rental price of land is denoted by r_j. Given this constant returns technology, we assume a continuum of competitive construction firms with free entry.

As in Kaplan, Mitman, and Violante (2020), builders sell new housing units to competitive firms called renters. Renters own housing units and rent them out to households. Let $\widetilde{p}_j$ be the price that developers charge to renters, and p_j be the rental price that renters charge to households. The housing stock depreciates at rate δ_h each period so in the steady state $\delta_h H_j = N_j$. In the model, households solve a static maximization problem. Renters live forever and each period, they face a new generation of households that solve a static problem that is identical to that solved by the previous generation.

Land Rents. While the housing capital to build structures is foregone consumption, the land rents stay in the economy. We assume that land is owned in equal shares by each worker in the economy in the form of a bond that is a diversified portfolio of the country's land supply. As a result, each worker receives rents R, given by

$$R = \frac{\sum_j r_j T_j}{\sum_j l_j}. \tag{2}$$

In an extension, we also consider the polar opposite case where all rents are received by absentee landlords.

Taxation. The federal government imposes an economy-wide taxation schedule. Its objective is to raise an exogenously given level of revenue G to finance government expenditure. Denote the pre-tax income by w and the post-tax income by $\tilde{w}$. Denote by t_j the specific tax rate that applies to workers in city j. Then $\tilde{w}_j = (1-t_j)w_j$. We assume that the tax schedule can be represented by a two-parameter family that relates after-tax income $\tilde{w}$ to pre-tax income w as:

$$\tilde{w}_j = \lambda w_j^{1-\tau},$$

where λ is the level of taxation and τ indicates the degree of redistribution ($\tau > 0$). As a result, in the quantitative analysis, instead of looking for optimal t_j for each city, we find the optimal τ that, given city-specific wages, characterizes city-specific taxes.

This is the tax schedule proposed by Bénabou (2002). Recent papers, e.g. Heathcote, Storesletten, and Violante (2017), Guner, Lopez-Daneri, and Ventura (2016) and Kindermann and Krueger (2022), use the same function to study optimal redistribution of income taxation in the U.S. The average tax rate is given by $1 - \lambda w_j^{-\tau}$ and the marginal tax rate is $1 - \lambda(1-\tau)w_j^{-\tau}$. Taxes are the same in each city when $\tau = 0$, with the average rate and the marginal rates given by $1 - \lambda$. If $\tau \neq 0$, average income taxes are city specific. In particular, $\tau > 0$, average tax rate in more productive cities with higher wages are higher than in less productive cities with lower wages.

The total tax collection is given by $G = \sum_j t_j w_j l_j$. A share ϕ of total tax revenue is transferred back to households as transfers. To capture redistribution across geographic areas by the federal government, we assume that transfers are city-specific, denoted by TR_j, with $\sum_j TR_j l_j = \phi G$.

Four caveats are in order about the way we model taxes and transfers: First, the analysis abstracts from local taxes. In the model, each location is populated by a representative household, hence the planner has no incentive to treat households within a city differently. Furthermore, taxes in one location only affect the location decisions of workers. Allowing a local tax that is entirely spent locally does not affect the location decision.[7]

Second, we focus on taxation of labor income. The federal tax schedule in the US also determines taxes paid by S-corporations (pass-through entities).[8] Hence, changes in spatial taxes can affect not only the location decisions of workers but also businesses. In the current analysis, we assume that some cities are more productive and offer higher wages to workers. If these cities are also more productive for businesses, this can create an additional motive for the planner to locate more economic activity in more productive locations. function in labor, hence there are no profits.

Third, government spending does not affect production, as part of G is rebated to households. Part of this rebate might be used for local infrastructure that can affect aggregate productivity, A_j, or amenities a_j. In particular, if local spending is more productive in higher wage cities that are densely populated, this can affect the planner's optimal choice for city-specific taxes and transfers.[9]

Finally, we assume a linear production function in labor; hence, there are no profits. Profits would raise two issues that affect the planner's problem. On the one hand, as we mentioned above, income taxes can affect how these profits are taxed. Second, our quantitative results show that land ownership matters for optimal spatial taxation. When the land is owned equally across the whole population, the planner is less concerned about higher housing prices, as it benefits everyone via higher rent transfers. With profits, a similar concern would arise about the ownership of firms. For example, if, by lowering

[7] That is not the case if some local taxes are used for expenditures in other locations. However, the coexistence of a local tax authority together with a federal tax authority would imply modeling multiple planners who interact in a fiscal union and engage in tax competition, which is very interesting but beyond the main focus of the analysis here.

[8] On the increase in S-corporation in the US in recent decades, see Dyrda and Pugsley (2024).

[9] On the effect of infrastructure on city growth, see, among others, Haughwout (2002) and Duranton and Turner (2012).

taxes in more productive cities, planners could attract more workers and firms to more productive cities, it would matter who benefits from higher profits in more productive locations. If these profits go to a large share of the population, the planner will be more willing to concentrate economic activity in high-productivity locations.

The linear technology in the model only has labor as an input. Although technology is linear, the utility in city j depends on the number of workers there through congestion, adding curvature in the utility function. Adding intermediate goods, produced in each city, as inputs in the production function will add curvature to output in labor. Such an extension can also affect tax incentives as allocating more workers in a city can have additional benefits by the production of intermediate goods. On the other hand, the current production structure allows a more tractable and transparent interpretation of our results.

It is also important to stress we abstract from mobility frictions and dynamics, which potentially bias our policy experiments. Taxes can affect allocation of people in a dynamic framework as there can be additional dynamics benefits from moving workers to high productive cities, such as opportunities to accumulate more human capital (as in Roca and Puga (2016)). When there are mobility frictions, the planner needs to take into account the monetary and utility costs of moving people across space, which can lower incentives to concentrate people in high-wage cities. Furthermore, any transaction costs associated with buying and selling houses is another source of mobility costs. Extending the model to allow for mobility frictions and dynamic decisions by the households, as in Oswald (2019), Luccioletti (2024), or Giannone, Paixão, Pang, and Li (2023), would not be trivial, and beyond the main point made in this paper.

Equilibrium. In a competitive equilibrium, workers and firms take wages w_j, housing prices $\widetilde{p}_j$ and p_j, and the rental price of land r_j as given. The price of consumption is normalized to one. Because housing capital is perfectly substitutable with consumption, the rental price of housing capital is also equal to one. All prices satisfy market clearing. The country-wide market for labor clears, $\sum_{j=1}^{J} l_j = \mathcal{L}$, and for housing, there is market clearing within each city $h_j l_j = H_j$ and $\delta_h H_j = N_j$. Under this market clearing specification, only those who work have housing. We interpret the inactive as dependents who live with those who have jobs. Workers optimally choose consumption and housing as well as their location j to satisfy utility equalization. Firms in production and construction maximize profits, which are driven to zero by free entry.

Welfare Criterion. When we evaluate social welfare, we use the utilitarian social welfare function, i.e. the sum of individual utilities. We have a representative agent economy, where all agents are ex ante identical, yet ex post households are heterogeneous in their consumption bundles and incomes depending on their location, which differ by wages and housing prices. Nonetheless, even ex post, utility is equalized across locations due to free mobility. In large cities, wages are high but also housing prices are high; the opposite in small cities. As long as there is free mobility, utility equalizes. Therefore, any location-neutral welfare function will generate the same outcome. For example, a concave social welfare function (concave in utilities, not in wages) will not affect the result since the concave social welfare function evaluates the outcome at one point only (at the representative agent's utility).

3 The Equilibrium Allocations

A representative worker in city j solves

$$\max_{\{c_j,h_j\}} u_j(c_j,h_j) = a_j l_j^{\delta} c_j^{1-\alpha} h_j^{\alpha} \tag{3}$$

subject to

$$c_j + p_j h_j \leq \tilde{w}_j + R + TR_j.$$

Taking first order conditions, the equilibrium allocations are $c_j = (1-\alpha)(\tilde{w}_j + R + TR_j)$ and $h_j = \alpha\frac{(\tilde{w}_j+R+TR_j)}{p_j}$.[10] The indirect utility for a worker in city j is given by

$$u_j = a_j l_j^{\delta} \alpha^{\alpha}(1-\alpha)^{1-\alpha}\frac{(\tilde{w}_j + R + TR_j)}{p_j^{\alpha}}. \tag{4}$$

Optimality in the location choice of any worker-city pair requires that $u_j = u_{j'}$ for all $j' \neq j$. The optimal production of goods in a competitive market with free entry implies that wages are equal to marginal product: $w_j = A_j$.

Optimality in the production of new housing in each city j requires that construction companies solve the following maximization problem:

$$\max_{K_j,\, T_j} \widetilde{p}_j B[(1-\theta)K_j^{\sigma} + \theta T_j^{\sigma}]^{1/\sigma} - r_j T_j - K_j.$$

The solution is characterized by $K_j^{\star} = \left(\frac{1-\theta}{\theta} r_j\right)^{\frac{1}{1-\sigma}} T_j$. This, together with the zero profit condition allows us to calculate the new housing supply in each city, which in turn determines a relation between the rental price of land r_j and the new house prices $\widetilde{p}_j$.

The problem of the renters is given by

$$V(H_j) = \max_{H_j'}[p_j H_j' - \widetilde{p}_j(H_j' - (1-\delta_h)H_j) + \frac{1}{1+r}V(H_j')],$$

where

$$N_j = H_j' - (1-\delta_h)H_j,$$

[10] The construction firms buy capital from households. Since the price of capital is one, however, this transaction does not affect the household budget constraint.

is the new housing demanded by the renters, and ρ is the interest rate used by renters to discount future. The FOC and Envelope conditions for this problem imply that in a steady state,

$$\widetilde{p}_j = p_j \frac{1+\rho}{\rho+\delta_h} = p_j + p_j \frac{1-\delta_h}{1+\rho} + p_j \left(\frac{1-\delta_h}{1+\rho}\right)^2 + ...$$

which equates what the rental firms pay for a new unit ($\widetilde{p}_j$) to income stream generated from renting these units. Note that if $\delta_h = 1$, i.e. with full depreciation of the housing stock, $\widetilde{p}_j = p_j$. In a stationary equilibrium, $N_j = \delta_h H_j$, or $H_j = \frac{N_j}{\delta_h}$.

In equilibrium, given amenities a_j, wages w_j, prices $\widetilde{p}_j$, p_j and r_j, taxes t_j, and transfers TR_j: i) households choose c_j and h_j optimally; ii) given wages w_j, production firms choose l_j optimally; iii) given $\widetilde{p}_j$ and r_j, construction firms choose K_j and T_j optimally; iv) given $\widetilde{p}_j$, the renters choose p_j optimally, v) markets clear to pin down prices, w_j, $\widetilde{p}_j$, p_j and r_j; vi) government budget constraint holds, $\sum_j TR_j l_j = \phi \sum_j t_j l_j w_j$; and vi) utility equalization across locations pins down l_j. Further details are provided in the Appendix.

4 The Planner's Problem

We study a Ramsey optimal spatial taxation problem where the planner chooses *tax instruments* to affect the equilibrium allocations. The planner assumes agents operate in a decentralized economy with equilibrium prices and free choice of consumption and location decisions, albeit affected by a city-specific tax t_j, where $\tilde{w}_j = (1-t_j)w_j$. Consider now a utilitarian planner who chooses the tax schedule $\{t_j\}$ to maximize the sum of utilities subject to: 1. the revenue neutrality constraint, i.e. she has to raise the same amount of tax revenue; 2. individually optimal choice of goods and housing consumption in a competitive market; and 3. free mobility – utility across local markets is equalized.

As in the case of the equilibrium allocation, the utility given optimal consumption (c, h) in a local labor market is given by (4). Then we can write the Ramsey planner's problem as:

$$\max_{\{t_j\}} \sum_j u_j l_j,$$

subject to $\sum_j t_j A_j l_j = G,\ u_j = u_{j'},\ \forall j' \neq j$, and $\sum_j l_j = \mathcal{L}$.

The solution to this problem involves solving a system of $J + J + 2$ equations (J FOCs and $J + 2$ Lagrangian constraints) in the same number of variables. We cannot derive an analytical solution, so we will characterize the optimal tax schedule from simulating the US economy in the next section. In order to provide intuition for our simulations, we start, however, by showing that the first welfare theorem holds when there is no housing production, no exogenous government expenditure ($G = 0$), and no externalities ($\delta = 0$).[11] We then analyze the Ramsey problem for a simple two-city example.

[11]We are grateful to John Kennan for pointing us to this equivalence.

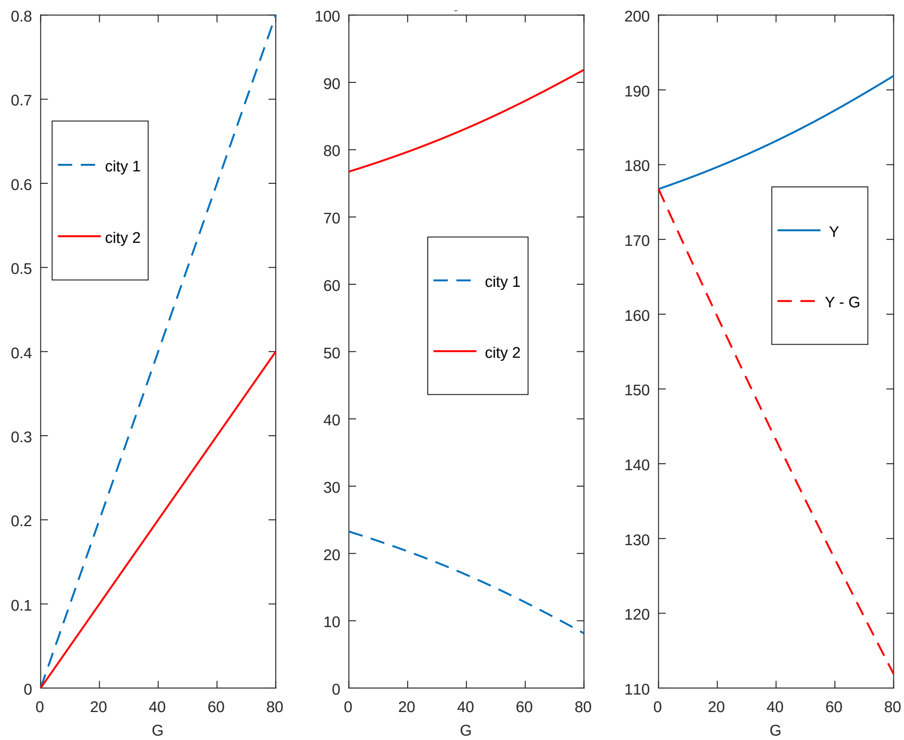

Figure 1: Optimal Ramsey taxes in a two city example as a function of government expenditure G in the benchmark setup with $A_1 = 1, A_2 = 2, \mathcal{L} = 100, \alpha = 0.28, \phi = 0.82$; A. optimal spatial tax rates t_1, t_2; B. populations l_1, l_2; C. Output Y and output net of government expenditure $Y - G$.

Proposition 1 *Let there be a two city economy with $\theta = 1, \delta = 0, \delta_h = 1$, $a_j = 1$ and preferences $u(c, h) = c^\alpha h^{1-\alpha}$. If there is no government expenditure $G = 0$, then the decentralized equilibrium allocation and the Ramsey planner's optimal allocation coincide. In the Ramsey problem, when $G > 0$, taxes are positive and increasing faster in G in the city 1 than in city 2, and population share in city 1 declines in G.*

Proof. In Appendix. ■

This result for the two-city economy illustrates how the equilibrium allocation changes with government expenditure G. Figure 1 illustrates the result in Proposition 1 for a simulation of the optimal solution to the Ramsey problem for a two-city example.

Relative taxes in big cities decrease as G increases. When $G = 0$, taxes in both cities are zero. As government expenditure G increases, the planner faces a trade-off in setting different taxes in more productive, big cities relative to less productive small cities: higher taxes in more productive cities generate bigger revenue per person, but attract fewer workers, and hence leads to a smaller tax base. We find that it is optimal to increase the base (number of people) in more productive cites: as G increases, the planner taxes those in highly productive city less to make sure that there are enough of them to pay for G. This implies a divergence of the population distribution as the large city becomes larger (Figure

1.B): higher government spending goes together with bigger population differences between small and large difference. That of course implies that output increases in government expenditure since more people live in more productive city, but the output net of government expenditure is decreasing (Figure 1.C).

In this simple model with two cities, $j = 1, 2$,given equation (4), the utility equalization implies

$$\alpha^{\alpha}(1-\alpha)^{1-\alpha}\frac{(w_1(1-t_1)+R+TR)}{p_1^{\alpha}} = \alpha^{\alpha}(1-\alpha)^{1-\alpha}\frac{(w_2(1-t_2)+R+TR)}{p_2^{\alpha}}.$$

Using $h_j = \alpha\frac{(w_j(1-t_j)+R+TR)}{p_j}$ and $h_j = \frac{H_J}{l_j}$ where H_j is the housing supply in city j, we obtain

$$\frac{l_2}{l_1} = \frac{H_2}{H_1}\left(\frac{w_2(1-t_2)+R+TR}{w_1(1-t_1)+R+TR}\right)^{\frac{1-\alpha}{\alpha}}.$$

Hence, all else equal: i) more people live in cities with higher housing supply, as housing cheaper; ii) more people live in more productive cities; iii) more people live in cities with lower taxes; iv) the elasticity of relative size of city 2 to city 1 with respect to relative net earnings is $\frac{1-\alpha}{\alpha}$, which is higher the lower the share of housing in the utility function (as people care relatively more about their net income).

5 Quantifying the Optimal Spatial Tax

We now quantify the magnitude of the impact spatial taxation. We proceed as follows: First, given the U.S. data on the distribution of labor force across cities (l_j) and wages in each city (w_j), we back out the productivity parameters A_j. Second, given (l_j, w_j), a parametric representation of current US taxes on labor income, $(\lambda^{US}, \tau^{US})$, and land area of each city (T_j), we compute amenity values a_j under the assumption that the current allocation of the labor force across cities is an equilibrium, i.e. utility of agents are equalized across cities. Third, given a_j, for any $\tau \neq \tau^{US}$, we compute the counterfactual distribution of labor across cities. In these counterfactuals, we assume revenue neutrality, and for any τ, find the level of λ such that the government collects the same amount of revenue as it does in the benchmark economy. Finally, we find the level of τ that maximizes welfare.

5.1 Labor Force and Wages

The data on the distribution of labor across cities (l_j) and wages in each city (w_j) are calculated from 2015 American Community Survey (ACS). For 254 Metropolitan Statistical Areas (MSA), we compute l_j as the population above age 16 who are in the labor force. While the model economy is populated by identical workers, average wages in each MSA in the data reflect several permanent worker characteristics, such as education. If more educated workers sort themselves into more productive cities, then higher average wages in more productive cities would partly be due to higher average human capital of workers in these cities.[12]

In order to mitigate this problem, we calculate w_j as weekly residual average wages for each MSA that controls for workers' education, age, gender and race, by estimating

$$\log(w_{ij}) = \kappa + \mu_1 \text{Education}_i + \mu_2 \text{Hispanic}_i + \mu_3 \text{White}_i + \mu_4 \text{Age} + \mu_5 \text{Age}^2 + \gamma_j + \varepsilon_{ij}, \quad (5)$$

where w_{ij} is weekly wage of worker i in MSA j, calculated as the total annual earnings divided by total number of weeks worked. Education is a dummy variable for educational attainment (with high-school dropouts, high school graduates, college graduate categories), Hispanic and White are dummies for race, and γ_j is an MSA fixed-effect. The residual wage for each worker is then calculated as $\kappa + \gamma_j + \varepsilon_{ij}$.[13]

Figure 13.A and B in the Appendix show the distribution of population and wages across MSAs. The average labor force is 485,301, with a maximum (New York, NY-Northeastern NJ) of about 9 million and a minimum (Asheville, NC) of 43,619. The population distribution is highly skewed, close to log-normal, where the top 5 MSAs account for 21.4% of total labor force.[14] Average weekly wages are $918. The highest weekly residual wage is 50% higher than the mean (Stamford, CT) and the lowest one is 20% of the mean level (Muncie, IN). Figure 2 shows the positive relation between population size and wages, the well-known urban wage premium in the data. We take both population and wage data as inputs to simulate the benchmark economy. The elasticity of wages with respect to population size is about 0.043. In Figure 2, as well as in all other figures below, we indicate the ten most populated MSAs together with five MSAs with the highest and the lowest wages.

5.2 Taxes

The relation between after and before tax wages is given by $\tilde{w} = \lambda w^{1-\tau}$, where λ is the level of taxation and τ indicates the degree of redistribution ($\tau > 0$). In order to estimate λ and τ for the US economy, we use the OECD tax-benefit calculator that gives the gross and net (after taxes and benefits) labor income at every percentage of average labor income on a range between 50% and 200% of average labor income, by year and family type.[15] The calculation takes into account different types of taxes (central government, local and state, social security contributions made by the employee, and so on), as well as many types of deductions and cash benefits (dependent exemptions, deductions for taxes paid, social assistance, housing assistance, in-work benefits, etc.).[16] Non-wage income taxes (e.g., dividend income,

[12]The evidence suggest that even though the average skill levels are constant across cities, the variance is increasing in city size, see Eeckhout, Pinheiro, and Schmidheiny (2014).

[13]We remove wages that are larger than 5 times the 99th percentile threshold and less than half of the 1st percentile threshold.

[14]The five largest MSAs are New York, NY-Northeastern NJ; Los Angeles-Long Beach, CA; Chicago, IL; Dallas-Fort Worth, TX; and Washington, DC/MD/VA.

[15] http://www.oecd.org/social/soc/benefitsandwagestax-benefitcalculator.htm, accessed on March 15, 2013.

[16]In this exercise we add local and federal taxes and related it to household income. Local taxes in the US can be progressive or regressive; see Fleck, Heathcote, Storesletten, and Violante (2025). They find that for the US as a whole

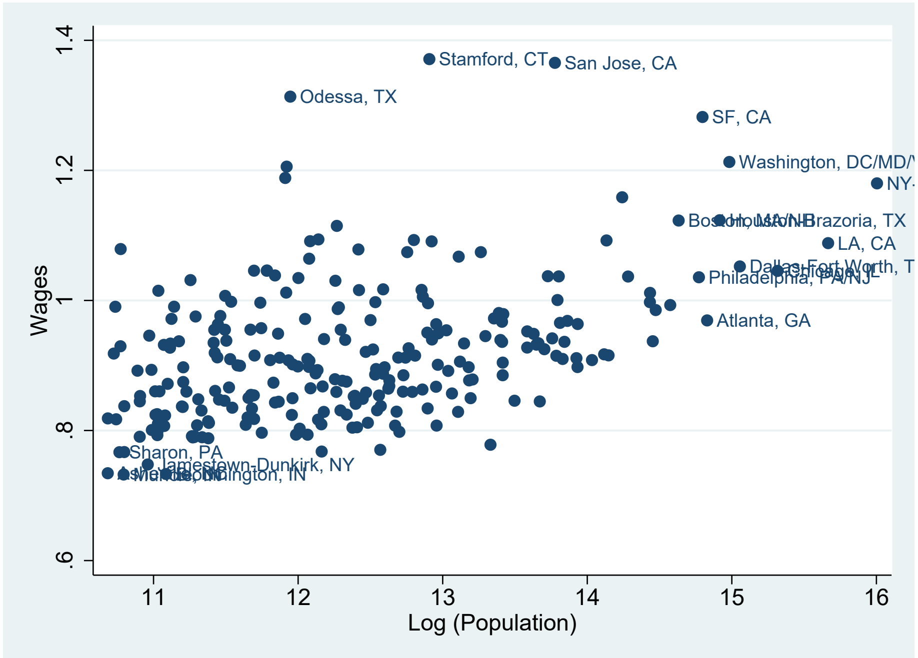

Figure 2: The Urban Wage premium.

property income, capital gains, interest earnings) and non-cash benefits (free school meals or free health care) are not included in this calculation.

We simulate values for after and before taxes for increments of 25% of average labor income. As the OECD tax-benefit calculator only allows us to calculate wages up to 200% of average labor income, we use the procedure proposed by Guvenen, Burhan, and Ozkan (2014) detailed in the Appendix, to calculate wages up to 800% of average labor income. As a benchmark specification, we calculate taxes for a single person with no dependents. Given simulated values for wages, we estimate an OLS regression

$$\ln(\tilde{w}) = \ln(\lambda) + (1 - \tau)\ln(w).$$

The estimated value of τ^{US} is 0.127. Estimating the same tax function with the U.S. micro data on tax returns from the Internal Revenue Services (IRS), Guner, Kaygusuz, and Ventura (2014) estimate lower values for τ, around 0.05. Their estimates, however, are for total income while the estimates here are for labor income. One advantage of the OECD tax-benefit calculator, compared to the micro data is that it includes social security taxes, which is not possible with the IRS data. Taking in account transfers, Heathcote, Storesletten, and Violante (2017) estimate a larger value of τ, around 0.18. Our estimates are closer to the ones provided by Guvenen, Burhan, and Ozkan (2014) who also use the OECD tax-benefit calculator to estimate tax rates using a more flexible functional form.

Below we report results with estimates for τ from Guner, Kaygusuz, and Ventura (2014) and Heath-

the local taxes and transfers are proportional, i.e., the estimated τ would be close to zero. The differences in estimated τ across states is small, between -0.02 and 0.02.

cote, Storesletten, and Violante (2017) as a robustness check. The parameter λ determines the average level of taxes. We obtain $\lambda^{US} = 0.856$, i.e., at mean wages ($w = 1$) taxes are about 14.4% of GDP in the benchmark economy. This is the average value for sum of personal taxes and contributions to government social insurance program as a percentage of GDP for 1990-2015 period.[17] As a result, the size of the government in the model is limited by the extent of tax collection from personal income taxes. Tax rates at $w = 0.5$ and $w = 2$ are 6.5% and 21.6%, respectively.

Figure 3 shows what our representation of the effective Federal Taxes in the US implies for how tax rates differ across cities. The average tax rate in San Jose, CA, for example, is about 7% points higher than it is in Bloomington, IN. Clearly, the Federal Tax Schedules do not differ across cities in the US. However, since taxes are progressive, higher average wages in a city map into higher taxes. The parametric representation captures differences across cities in average tax rates in a parsimonious way. Furthermore, it allows us to search for a single parameter, τ, to maximize welfare.

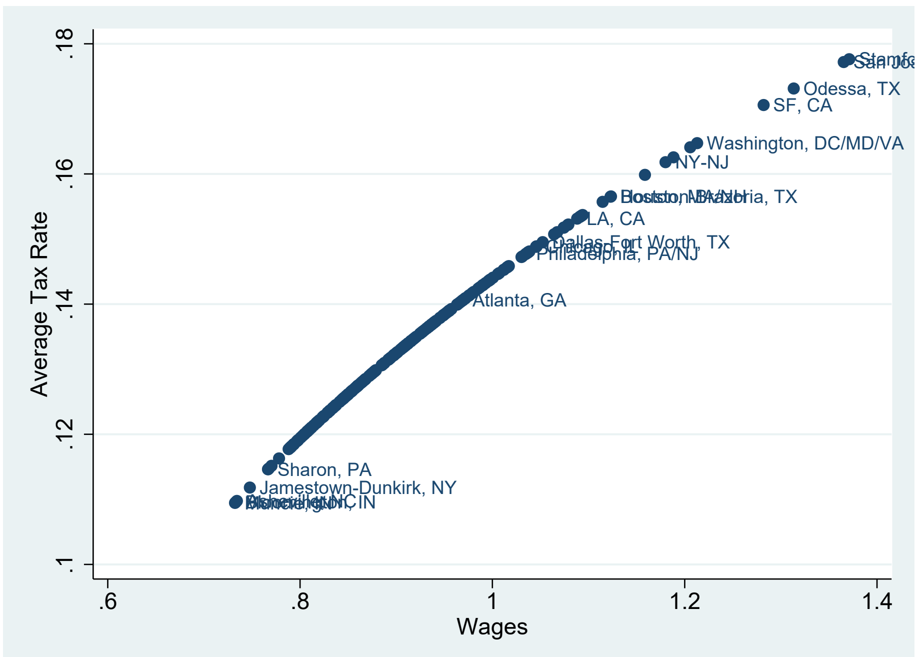

Figure 3: Taxes across cities

5.3 Transfers

To set ϕ, the share of tax revenue that is rebated to households, we assume that government expenditure (excluding defense) provides a direct income to households. The share of defense expenditure in the Federal Government's budget was about 18% in the US for the 1990-2015 period.[18] Therefore, we

[17] National Income and Product Accounts, Bureau of Economic Analysis, Table 3.2. Federal Government Current Receipts and Expenditures, http://www.bea.gov/iTable/index_nipa.cfm

[18] National Income and Product Accounts, Bureau of Economic Analysis, Table 3.16. Government Current Expenditures by Function, http://www.bea.gov/iTable/index_nipa.cfm

assume that the rest, 82% of taxes, is rebated back to households, i.e. $\phi = 0.82$. We assume that city-specific transfers, TR_j, are a declining function of city-specific wages, given by

$$TR_j = \eta_1 + \eta_2 w_j, \text{ with } \eta_1 > 0, \text{ and } \eta_2 < 0.$$

Since, total transfers must be equal to total government spending,

$$\sum_j TR_j L_j = \eta_1 \sum L_j + \eta_2 \sum w_j L_j = \phi G.$$

To calibrate parameters η_1 and η_2, we use the Bureau of Economic Analysis regional economic accounts, and choose η_1 and η_2 so that the model reproduces the relation between per-capita transfers (net of social security and disability payments) and per-worker wage and salary earnings across MSAs. Figure 4 shows the data. On average the poorest MSAs get a per capita transfer of around 6,000$, which declines below 4,000$ in richer MSAs. Further details on the data and the calibration of η_1 and η_2 are provided in the Appendix.

Figure 4: Local Transfers

These local transfers already capture part of the place-based policy. Of course, also local taxes are an important part of the fiscal tools which can be place-based. We have taken the view in the paper that most local taxes are also spent locally, and there is no redistribution across cities. Since we do not model local public goods, our model has nothing to say about this. The extent to which local taxes are also used for redistribution (maybe state taxes are used to redistribute across cities within the state) would alter the interpretation.

5.4 Housing Production

The CES housing supply technology basically stipulates that the cost of construction of housing is increasing in the size of the house, but at a (weakly) decreasing rate. If housing capital and land are complements (the elasticity of substitution is less than one), then the housing cost is decreasing in the size of the house. For example, small apartments still need a bathroom and a kitchen, so the unit cost per square meter is higher, or, it is more expensive per unit of housing to build a high-rise than a stand alone home. The implication of this is that the share of land in the value of housing is increasing in the population density, as transpires from the data.

The data on land areas of cities (MSAs), T_j, is taken from Lutz and Sand (2019), who update land availability measure constructed by Saiz (2010). For each MSA, the available land is measured as the fraction of land that can be developed in a 50 km radius from an MSA's center. Land that can't be developed are either mountains (or steep slopes) or water (e.g., oceans, lakes, etc.). The maximum available land for an MSA is the area of a circle with a 50 km radius, 7875 km^2. The fraction of available land varies from about 0.12, i.e. 921 km^2, for Honolulu (HI) to 1, i.e. 7875 km^2, Anchorage (AK). The average fraction is 0.72, i.e. 5703 km^2. Appendix Figure 14 shows the distribution of available land across MSAs. Figure 15 depicts the relation between wages and available land across MSAs; more productive cities tend to have less available land (the correlation between available land and wages is -0.23).

Davis, Larson, Oliner, and Shui (2021) show that the land share in housing is around one-third across MSAs, while Davis, Larson, Oliner, and Shui (2021) document that it is about 26.6% across US counties. We set $\theta = 0.2750$ to match an average land share of one-third in the benchmark economy. Davis, Larson, Oliner, and Shui (2021) also show that the land share ranges from 15.7% (10th percentile) to 41.5% (90th percentile). The parameter $\sigma = -0.1412$ is chosen to a maximum land shares of 41.5%.[19] Then, we set $B = 0.00738$ such that on average housing consumption is about 200m^2.[20] Following Davis and Heathcote (2007) and Kaplan, Mitman, and Violante (2020), we set $\delta_h = 0.015$. Finally, we assume that the rental companies discount future at 4%.

5.5 Preferences and Productivity

Given $w_j = A_j$, and $\tilde{w}_j = \lambda w_j^{1-\tau}$, we calculate amenities a_j from utility equalization condition across cities. Given the indirect utility function in equation (4), for any two locations j and j', the following

[19] We use 90th percentile of the land share distribution from Davis, Larson, Oliner, and Shui (2021) (Table 1, Annual Panel) as the maximum since model abstract from features that might generate a long tail of land shares. The 99th percentile is about 62% in the data.

[20] The average size of new single-family houses sold in the US between 1990 and 2015 was 2134.4 square feet (198.3 square meter) - of Commerce (2015), page 745.

equality must hold:

$$\begin{aligned} u_j &= a_j[(1-\alpha)^{1-\alpha}](\tilde{w}_j + R + TR_j)^{1-\alpha} l_j^{\delta-\alpha} H_j^{\alpha} \\ &= a_{j'}[(1-\alpha)^{1-\alpha}](\tilde{w}_{j'} + R + TR_j)^{1-\alpha} l_{j'}^{\delta-\alpha} H_{j'}^{\alpha} \\ &= u_{j'} \end{aligned}$$

Let $a_1 = 1$. Then, for each city j,

$$\begin{aligned} a_j &= \frac{(\widetilde{w}_1 + R + TR_j)^{1-\alpha} l_j^{\alpha-\delta} H_1^{\alpha}}{(\widetilde{w}_j + R + TR_j)^{1-\alpha} l_1^{\alpha-\delta} H_j^{\alpha}} \\ &= \frac{(\widetilde{w}_1 + R + TR_j)^{1-\alpha} l_j^{\alpha-\delta} \left[(1-\theta)\left(\frac{1-\theta}{\theta} r_1\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{\alpha/\sigma} T_1^{\alpha}}{(\widetilde{w}_j + R + TR_j)^{1-\alpha} l_1^{\alpha-\delta} \left[(1-\theta)\left(\frac{1-\theta}{\theta} r_j\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{\alpha/\sigma} T_j^{\alpha}} \end{aligned} \tag{6}$$

Calculations for a_j obviously depend, among other parameters, on the values we assume for α and δ. We set $\alpha = 0.2804$. Davis and Ortalo-Magné (2011) estimate that households on average spend about 24% of their before-tax income on housing. This would translate to a spending share of $\alpha = \frac{0.24}{\lambda} = \frac{0.24}{0.856} = 0.2804$ from after-tax income at mean income ($w = 1$).

We interpret the congestion term $l^{-\delta}$ in the utility as commuting costs and calibrate it using the available evidence on the relationship between city size and commuting costs. The elasticity of commuting time with respect to city size is estimated to be 0.13 by Gordon and Lee (2011). Average commuting time in the US is about 50 minutes (McKenzie and Rapino (2011)). Assuming a 20$ hourly wage, this 50 minutes costs about 17$ for households, which is about 11% of their daily income (17/160). Commuting also has a monetary cost. Roberto (2008) reports that households on average spend about 5% of their income on transportation expenditures, while Lipman et al. (2006) find these costs to be higher, close to 20%. If we take 10% as an intermediate value, then the total, time and money, cost of travel for households is about 20% of their income, which is simply the elasticity of the total commuting costs with respect to the commuting time. As a result, the elasticity of total commuting costs with respect to city size, which is the elasticity of the total commuting costs with respect to the commuting time times the elasticity of commuting time with respect to the city size is $(0.13)(0.2) = 0.026$.[21] Table 1 shows the parameter values for the benchmark economy.

[21] In this paper, we assume each city has a different, exogenously given, land area and there is congestion. An alternative strategy would be to endogenize land area by capturing the cost of commuting, for example as in Combes, Duranton, and Gobillon (2018), in the presence of a central business district. However, in our model there is no within city heterogeneity, and commuting costs are captured by the congestion externalities in utility, rather than in housing production. As we show in section 6, incorporating the exact land area in the model is an important ingredient to fit the data.

Parameters	Values	Comments
Taxes and Transfers		
λ	0.856	Benchmark tax rate for $w = 1$
τ	0.127	Benchmark place-based tax
ϕ	0.820	Fraction of taxes rebated
η_1	0.2249	Transfer function, $TR_j = \eta_1 + \eta_2 w_j$
η_2	-0.1063	Transfer function, $TR_j = \eta_1 + \eta_2 w_j$
Parameters set a priori		
α	0.280	Davis and Ortalo-Magné (2011)
δ	-0.026	See text
δ_h	0.015	Davis and Heathcote (2005), Kaplan et al (2020)
ρ	0.040	Standard
Parameters calibrated		
B	0.00738	Average household size, 200 m^2
θ	0.2750	Average land share in housing, 1/3
σ	-0.1412	Max land share in housing, 0.415

Table 1: Parameter Values of the Benchmark Economy

5.6 Benchmark Economy

In Figure 5.A we report the computed values of a_j, adjusted for congestion, i.e. $al^{-\delta}$, across metropolitan statistical areas. We set $a_1 = 1$ for New York-Northeastern NJ MSA. The mean value of a_j across MSAs is also about 0.72. The highest levels of a_j, about 1.04, is calculated for Los Angeles-Long Beach (CA), followed by New York-Northeastern NJ MSA (1), Miami-Hialeah, FL (0.996), Chicago, IL (0.938), and Atlanta, GA (0.920). The calibration assigns a high value of a for places like NY, LA, Chicago and Atlanta, relatively high-wage places, to account for their large size. On the other hand, a relatively low-productivity MSA like Miami-Hialeah (FL), with averages wages that are about 90% of the national average, also requires a high a to justify its current size, which possibly reflects better weather conditions. On the other hand, the lowest values of a are 0.57 for Waterbury, CT, 0.55 for Anchorage, AK and 0.53 for Odessa, TX. These are MSAs with relatively high wages but with small populations and low values of a are assigned to justify why more people are not living there, which might again reflect weather. Overall, the correlation between amenities and population size is about 0.6 in the benchmark economy.

Figure 5.B shows the relation between population size and the share of land values in housing prices, which we use as a target to calibrate housing production technology. Finally, the benchmark economy generates a distribution of equilibrium housing prices across MSAs. Estimated housing prices are highest in NY, followed by LA, while the lowest housing prices are computed for Sharon, PA and Muncie, IN. The ratio of the highest to lowest prices is around 8. While the average housing consumption is calibrated to be around 200m^2 across MSAs, those in NY live in houses that are about 83m^2 and about 10 times smaller than houses in Sharon, PA. Figure 5.C shows the relation between population

size and housing prices across MSAs in the benchmark economy. The figure implies an elasticity of housing prices with respect to population size that is about 0.35.

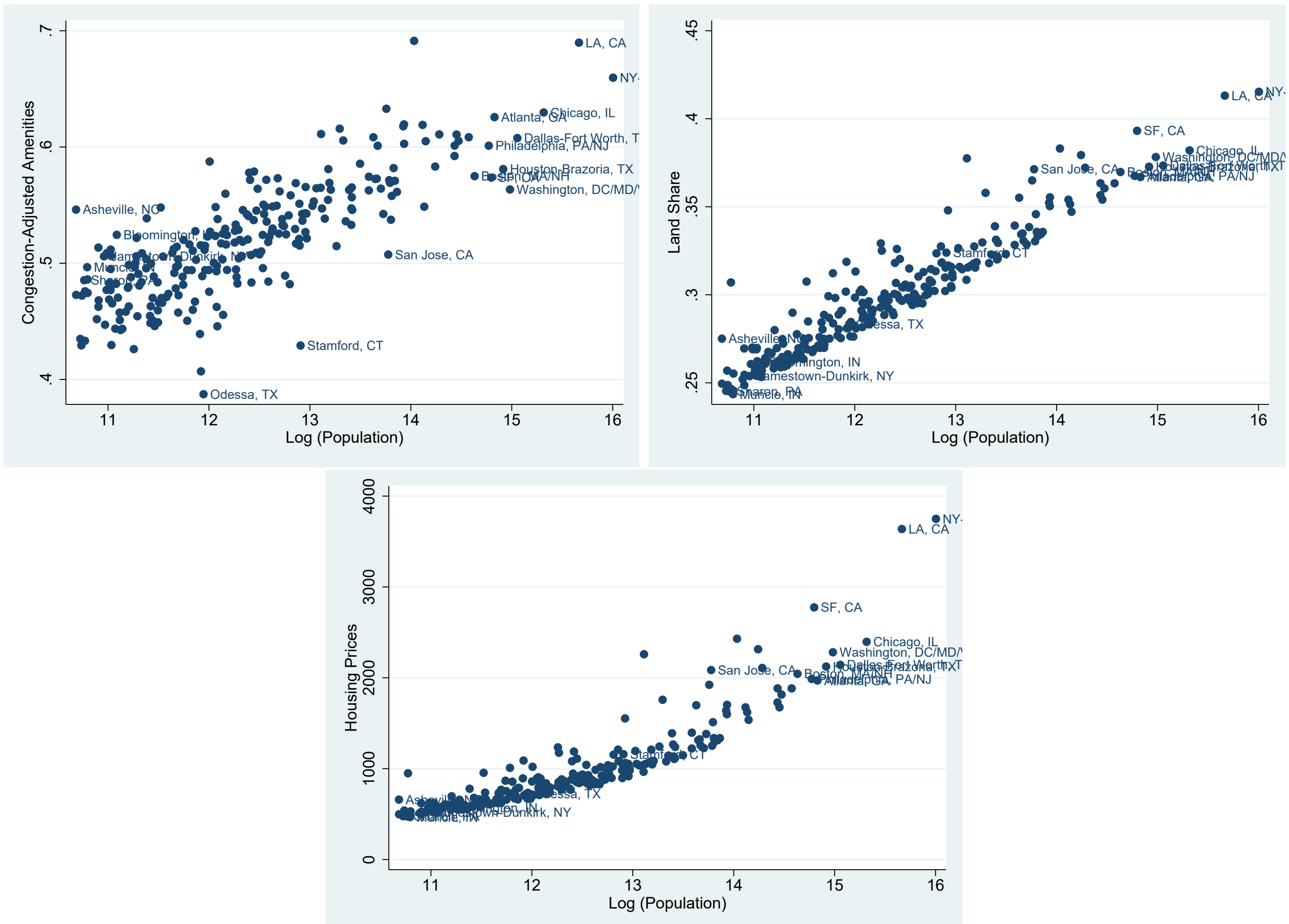

Figure 5: Benchmark Economy: A. Amenities and Population; B. Land Share in the Value of Housing and Population; C. Housing Prices and Population.

Next we compare the model outcomes, which were not directly targeted, with the data. First, we compare housing prices from the benchmark economy with housing prices in the data. In the model economy, housing is a homogeneous good with a location-specific per unit rental price p_j. In the data, on the other hand, housing units differ in many observable dimensions, and as a result, observed housing prices reflect both the location and the physical characteristics of the unit. We estimate the city-specific price level as a location-specific fixed effect in a simple hedonic regression of log rental prices on the physical characteristics, such as age, number of rooms, age of the unit, and the unit structure (one family detached unit vs. one family attached unit etc.).[22] For both the model and the data, we report prices in each city as a fraction of average prices across all cities. The model does an excellent job capturing the variation in housing prices in the data (Figure 6.A). The correlation between the model-

[22]We use 2015 American Community Survey (ACS) data on housing rentals and housing characteristics. See also Eeckhout, Pinheiro, and Schmidheiny (2014).

implied and actual prices is 0.58. The variance of housing prices in the model economy is higher than it is in the data.

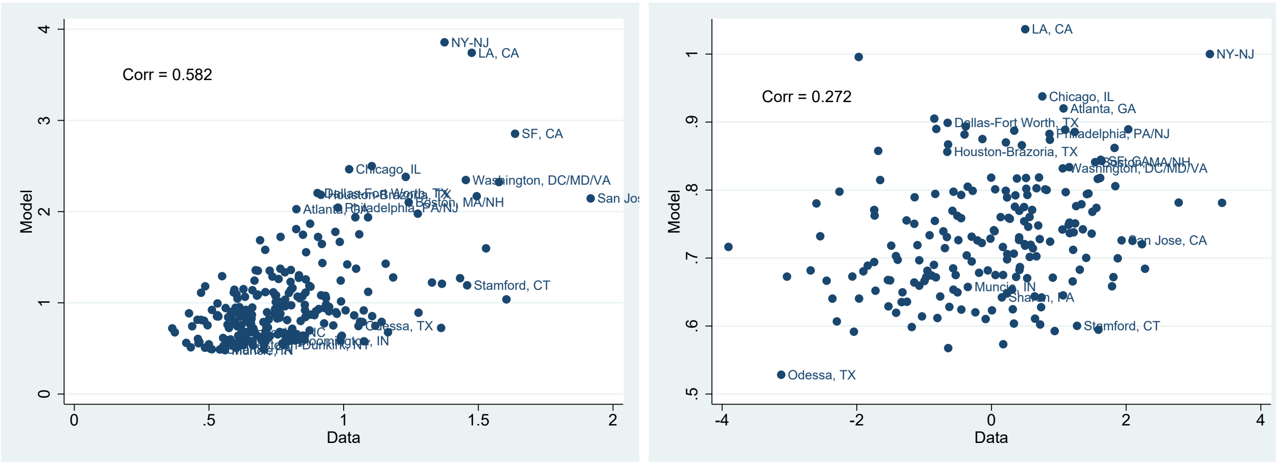

Figure 6: Housing prices (A) and Amenities (B): Data versus Model.

Second, we compare amenities from the model, a_j, with amenities in the data. We use data from Diamond (2016), who collects information on city amenities related to retail, transportation, crime, the environment, schooling and job quality. Using a principal component analysis, she summarizes the amenity into a single index for each metropolitan area and decade (years 1980, 1990, and 2000). In Figure 6.B, we use her most recent data for 2000. The correlation between amenities in the model and the data is around 0.27. Both in the data and the model, large MSAs, e.g. NY-NJ, have a high level of amenity, while high-wage, small MSAs, e.g. Odessa TX, are not attractive places to live.

5.7 Optimal Spatial Taxes

Given values for A_j and a_j, the next step is to find counterfactual allocations for any level of $\tau \neq \tau^{US}$. This is done simply by first writing equation as

$$a_j = \frac{(\lambda w_1^{1-\tau} + R + TR_1)^{1-\alpha} l_j^{\alpha-\delta} \left[(1-\theta)\left(\frac{1-\theta}{\theta} r_1\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{\alpha/\sigma} T_1^{\alpha}}{(\lambda w_j^{1-\tau} + R + TR_j)^{1-\alpha} l_1^{\alpha-\delta} \left[(1-\theta)\left(\frac{1-\theta}{\theta} r_j\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{\alpha/\sigma} T_j^{\alpha}}, \tag{7}$$

which can be used to calculate new allocations for any τ

$$l_j(\tau) = l_1(\tau) \left[a_j^{\frac{1}{\alpha-\delta}} \left(\frac{\lambda w_j^{1-\tau} + R + TR_j}{\lambda w_1^{1-\tau} + R + TR_1} \right)^{\frac{1-\alpha}{\alpha-\delta}} \left(\frac{(1-\theta)\left(\frac{1-\theta}{\theta} r_j\right)^{\frac{\sigma}{1-\sigma}} + \theta}{(1-\theta)\left(\frac{1-\theta}{\theta} r_1\right)^{\frac{\sigma}{1-\sigma}} + \theta} \right)^{\frac{\alpha}{\sigma}\frac{1}{\alpha-\delta}} \left(\frac{T_j}{T_1} \right)^{\frac{\alpha}{\alpha-\delta}} \right].$$

where $l_j(\tau)$ is the counterfactual allocation for tax schedule τ.

We want the counterfactual to be revenue neutral, so for each τ we find a value of λ such that the

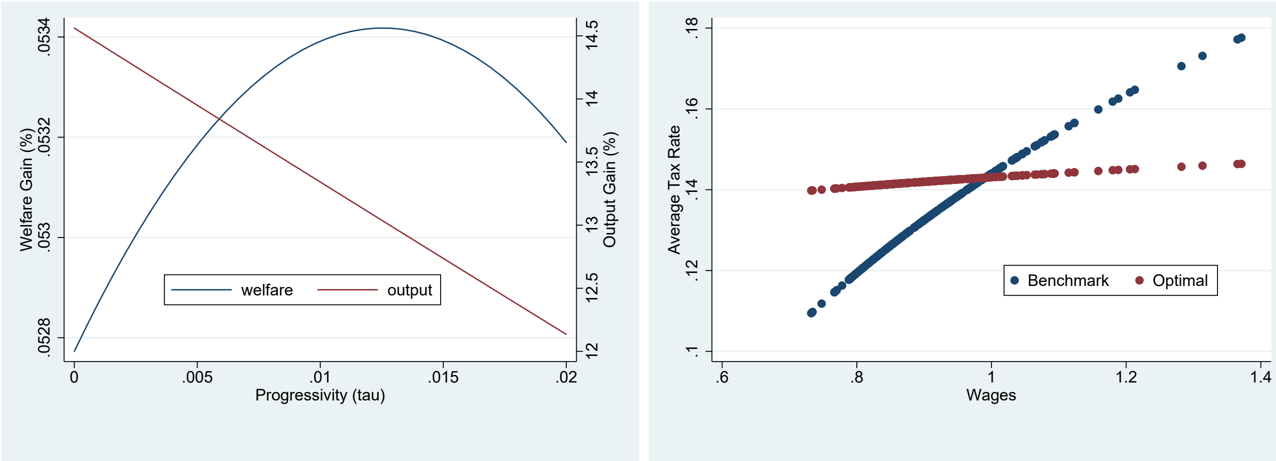

Figure 7: A. Welfare gain for different values of τ; B. The optimal spatial tax schedule $\tau^{\star}$ compared to that in the benchmark economy τ^{US}.

government collects the same tax revenue as it does in the benchmark economy, i.e.

$$\sum_j l_j(\tau)w_j(\tau)(1-\lambda w_j^{-\tau}) = \sum_j l_j w_j(1-\lambda^{US} w_j^{-\tau^{US}}). \tag{8}$$

Figure 7.A shows the percentage change in utility and output from the benchmark economy for different values of τ. The planner problem is given by

$$\max_{\tau} u(\tau),$$

subject to equation (8) and utility equalization across cities.

$$u(\tau) = u_i(\tau) = u_j(\tau), \text{ for } \forall i, j$$

An alternative objective function for the planner, following Bénabou (2002), could take into account inequality across cities, i.e.,

$$\max_{\tau} u(\tau) - \theta\Delta_J(w),$$

where $\Delta_J(w)$ is the inequality across cities in wages, and θ is inequality aversion of the planner. We could, for example, choose θ so that the observed level of τ in the US would correspond to the solution of this alternative planner problem. However, there is no reason to focus on inequality in wages and not on utilities across cities, which also takes into account housing prices. But with utility equalization, $\Delta_J(u) = 0$, this formulation coincides with ours.

The optimal value $\tau^{\star}$, is 0.0123.The optimal $\tau^{\star}$ is less than τ^{US}, i.e. taxes in big cities should be lower than those implied by the degree of redistribution of observed income taxes. However, the optimal τ is not zero. As shown in the Figure, lower values of τ results in larger movements of population to more productive cities and generates larger output gains. But it does not necessarily maximize consumer's utility as consumers are hurt by higher housing prices in larger cities. Figure 7.B shows the implied tax schedule under $(\lambda^{US}, \tau^{US})$ and $(\lambda^{\star}, \tau^{\star})$. The tax function is more flat with $(\lambda^{\star}, \tau^{\star})$. As a result, for $w = 0.5, w = 2$ and $w = 5$, the tax rates are 13.6%, 15.0% and 16.0%, respectively under the optimal τ^{*}, in contrast to 6.5%, 21.6% and 30.2% under τ^{US}.

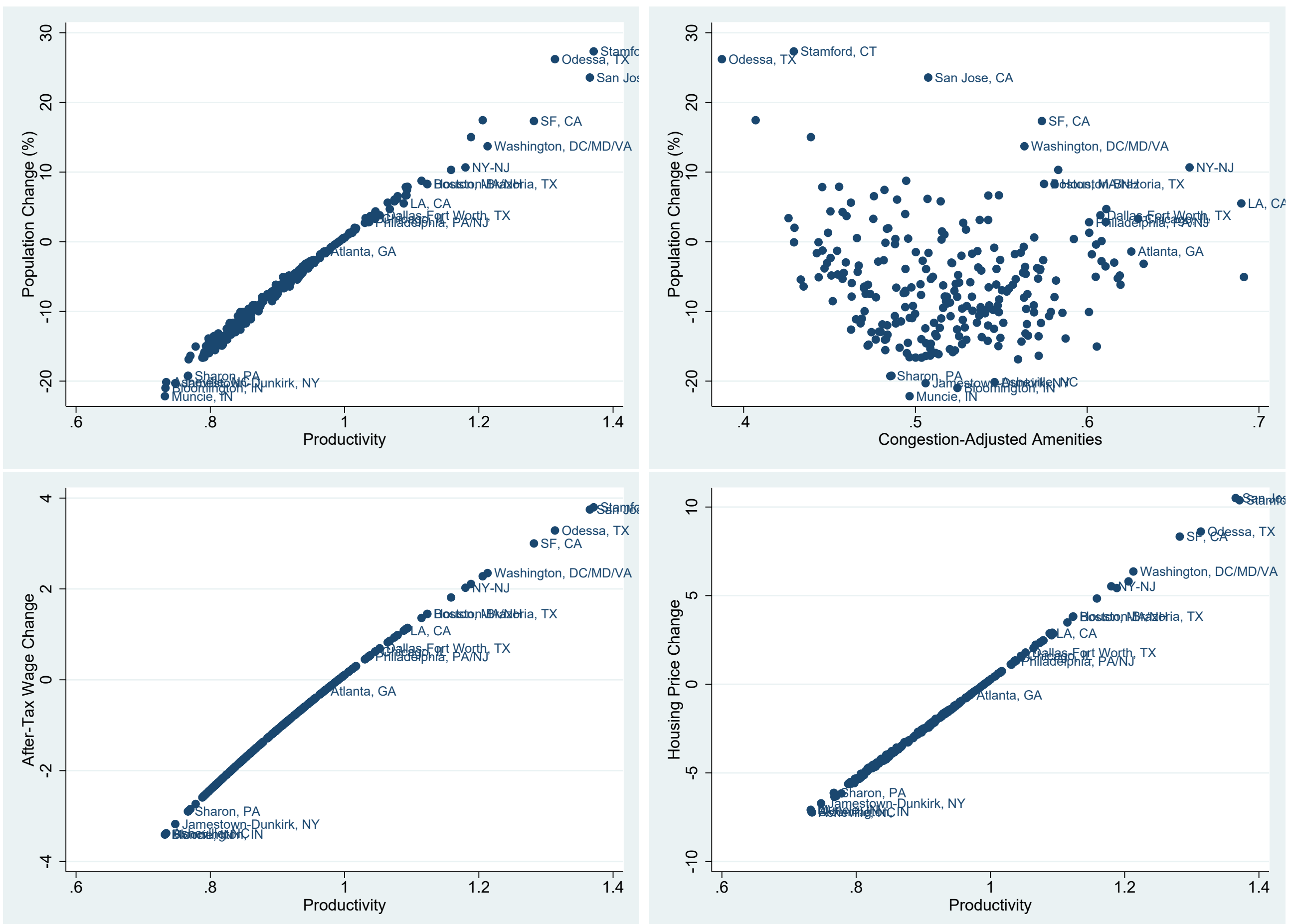

Figure 8: Implied changes of implementing the optimal policy $\tau^{\star}$. A. Change in population by TFP; B. Change in population by a; C. Change in $\tilde{w}$ by TFP; D. Change in housing prices p by TFP.

How can a change in τ from 0.127 to 0.0123 be implemented in practice? Clearly any change in the federal income tax schedule that makes it less progressive will translate into lower taxes in more productive cities in our framework. The tax reforms in the early 1980s, for example, resulted in significantly less progressive taxes in the US (as documented by Guner, Kaygusuz, and Ventura (2014) and Borella, De Nardi, Pak, Russo, and Yang (2023)). Incidentally, after the 1980s, there was also an increase in the concentration of population in larger cities.

The mortgage interest deductions in the US also lower τ, as most of these deduction accrue to higher-income households. Hence, all else equal, higher deductions might make income effective taxes lower in high-income cities. Yet, in equilibrium, lowering τ though mortgage interest payments deduction might not be an effective way of attracting workers to high-wage cities, as it will also increase housing demand, making high-wage cities less attractive. This will also be the case for other policies, such as deductions for property taxes, which is again likely to benefit high-income households. Indeed, equilibrium models of mortgage interest payment reductions with heterogenous agents (but without a spatial aspect), such as Floetotto, Kirker, and Stroebel (2016) and Sommer and Sullivan (2018), find that eliminating these deductions would be welfare improving as they would lower house prices. This also holds for other policies, such as tax deductions for property taxes, which are likely to benefit high-income households, and increase housing demand.[23]

5.8 Optimal Allocation

Now we can evaluate the implications of a tax change in the tax schedule from τ^{US} to $\tau^{\star}$, both for individual cities and in the aggregate. Consider first the impact on individual cities, which is summarized in Figure 8 and Table 2. The table gives the numerical values for those cities with extreme values either for TFP A or for amenities a.

MSA		A	a	$\%\Delta l$	$\%\Delta p$	$\%\Delta c$	$\%\Delta h$
Highest A							
	Stamford, CT	1.37	0.60	27.32	10.38	3.51	-6.22
	San Jose, CA	1.37	0.73	23.57	10.50	3.80	-6.37
	Odessa, TX	1.31	0.53	26.20	8.62	3.56	-5.15
Lowest A							
	Asheville, NC	0.73	0.72	-20.15	-7.23	-2.60	4.99
	Bloomington, IN	0.73	0.70	-20.99	-7.18	-2.61	4.92
	Muncie, IN	0.73	0.66	-22.17	-7.08	-2.62	4.80
Highest a							
	LA-Long Beach, CA	1.09	1.04	5.50	2.87	0.99	-1.83
	NY, NY-Northeastern NJ	1.18	1.00	10.67	5.53	1.84	-3.49
	Miami-Hialeah, FL	0.91	1.00	-5.06	-2.44	-0.77	1.71
Lowest a							
	Waterbury, CT	0.99	0.57	-0.08	-0.01	0.05	0.06
	Anchorage, AK	1.21	0.55	17.45	5.80	2.07	-3.52
	Odessa, TX	1.31	0.53	26.20	8.62	3.56	-5.15

Table 2: Benchmark Economy, move from τ^{USA} to $\tau^{\star}$. Outcomes for Selected Cities.

Since the optimal degree of tax difference $\tau^{\star}$ is below existing τ^{US}, the optimal policy lowers tax payments in high productivity cities. Figure 8.A. shows that the high A cities grow in size while the low productivity A cities lose population. The largest population growth rate, for Stamford (CT), is around 27% whereas Muncie (IN) loses 22% of its population. As is apparent in Figure 8.B., in contrast

[23] 2017 Tax Cut and Jobs Act (TCJA) reduced the debt limits for mortgage interest rate deductions from $1 million to $750,000. The reduction become permanent with the 2025 One Big Beautiful Bill Act.

with productivity, there is no systematic relation between amenities and population change.

The economic mechanism that drives the population mobility is the following. Due to lower marginal taxes, more productive cities pay higher after tax wages (Figure 8.C). This in turn attracts more workers relative to the benchmark equilibrium with τ^{US}. The new equilibrium is attained when utility across locations equalizes. The main countervailing force that stops further population mobility against the attractiveness of higher after tax wages is housing prices. Figure 8.D shows the change in housing prices. High productivity cities are up to 10% more expensive while low productivity cities face housing price drops of up to 7%.

Figure 9 shows the distribution of output and price changes across MSAs. Output in some MSAs grows as much as 27% while in others it declines by 20%. Output declines in the majority of MSAs, as many small MSAs lose population. Few productive, and large, MSAs on the other hand gain population. The distribution of changes in prices reflects the same forces. Prices decline in many small MSAs, and increase in few large ones. Figure 10 shows the same information in a map. The increase in population (Panel A) and prices (Panel B) are concentrated on a few locations, around NY, San Francisco, Chicago, and a few locations in Texas, while many cities experience small changes.

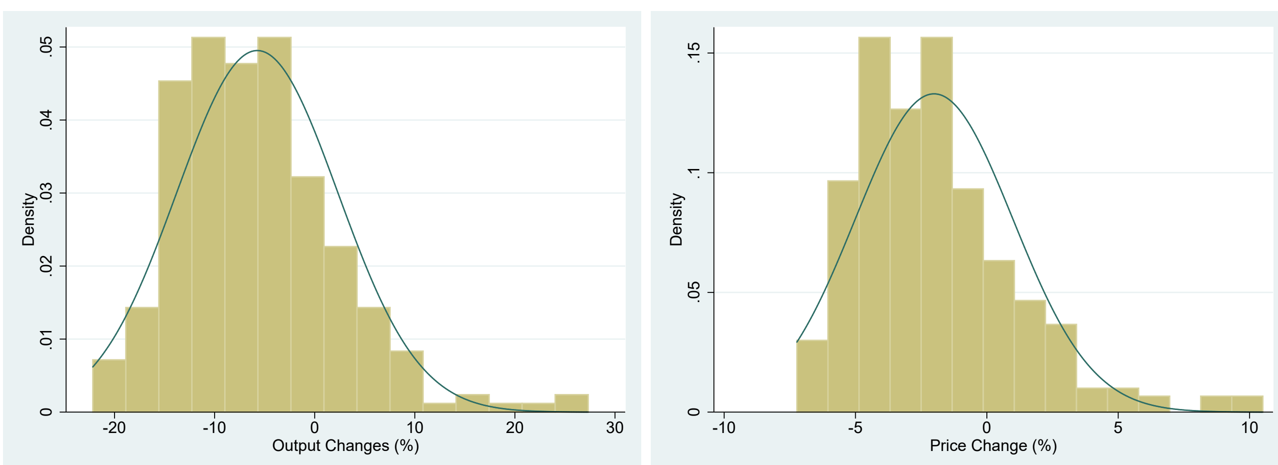

Figure 9: Distribution of changes: A. Output; B. Housing Prices.

Of course with higher housing prices goes substitution of housing for consumption (see Figure 11.A). In the high productivity cities, workers live in even smaller housing while increasing goods consumption. Housing consumption decreases by more than 6% in the high productivity cities in substitution for nearly 4% higher goods consumption. In the less productive cities housing consumption increases by up to 5% at the cost of decreased goods consumption by 2.5%. Given homothetic preferences, the marginal rate of substitution is constant.

Table 3 shows the aggregate outcomes from moving the benchmark allocation to the optimal. On average output and consumption go up by about 1.00% and 0.95%, respectively. This is driven by the population moving to the more productive cities. The population in the 5 largest cities grows by

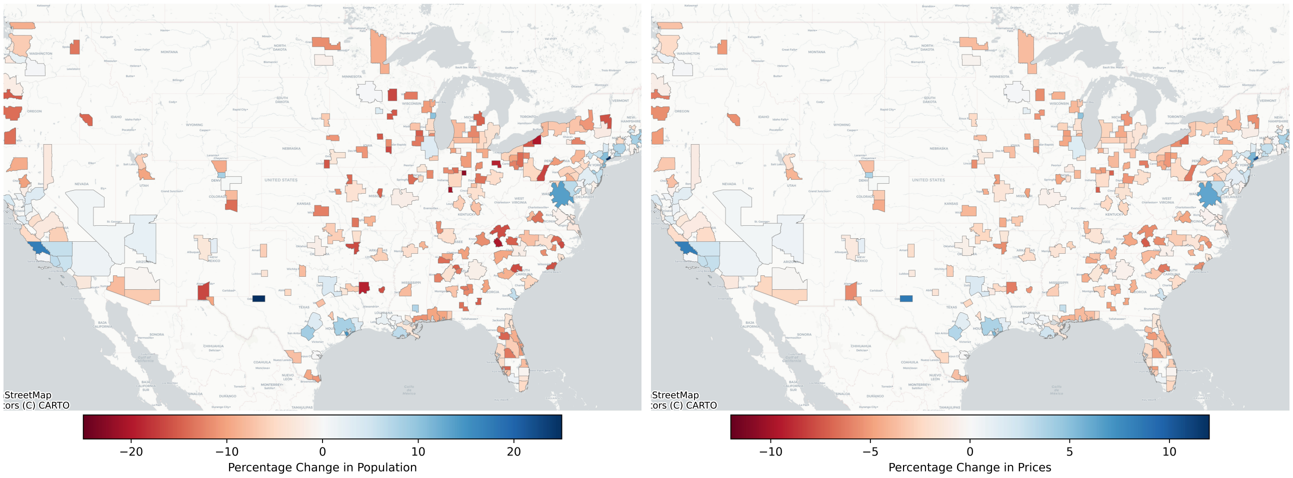

Figure 10: Distribution of changes in Population and Prices across MSAs.

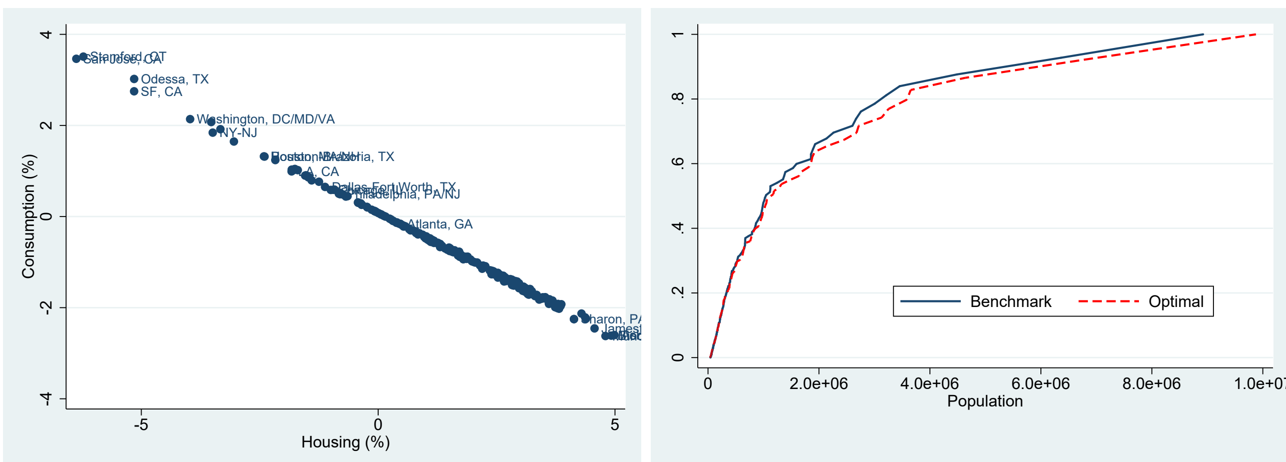

Figure 11: Implied changes of implementing the optimal policy $\tau^{\star}$: A. Substitution between c and h; B. Cumulative Distribution of city sizes.

7.65%, despite the fact that the top three are large in part because they also offer high amenities a. Most importantly, in the aggregate there is a reallocation of population from less productive, smaller cities to the more productive, larger cities. As a result there is first-order stochastic dominance in the population distribution, as is evident from Figure 11.B. Not surprisingly, aggregate housing prices go up by 4.85%. Due to higher prices, aggregate housing consumption declines by 1.90%.

Despite relatively large output gains, welfare gains are tiny. Given free mobility and a representative agent economy, all agents have the same utility level. After implementing the optimal policy, utility increases by only 0.05%. The reason for such tiny welfare gains is quite simple. Under the optimal spatial taxes, after-tax wages increase in cities that have initially high productivity. These cities, however, also get more crowded and housing prices go up. With higher prices, housing consumption in these cities

Aggregate Outcomes	
Welfare Gain (%)	0.0534
Output Gain (%)	1.00
Consumption (%)	0.95
Housing Consumption (%)	-1.90
Population Change, top 5 cities (%)	7.65
Fraction of Population that Moves (%)	3.31
Change in Average Housing Prices (%)	4.85

Table 3: Benchmark Economy, move from τ to $\tau^\star$

declines, and from the substitution of goods for housing, this generates higher goods consumption. However, the welfare gains associated with higher goods consumption get almost completely offset by lower housing consumption.

6 Understanding the Mechanism

The critical trade-off that determines the optimal τ in the model is between higher output and higher housing prices. Lower τ implies lower taxes in high-wage cities, making them more attractive. As more workers move to more productive cities, the planner has a larger tax base in these cities that more than compensates for lower revenue per person in these cities. This force towards lower τ is stronger if we allow for agglomeration externalities. On the other hand, as more workers move to high-wage cities, housing prices increase, reducing the utility gain from higher after-tax wages. Without changes in housing prices or other congestion effects, the planner would like to locate everyone in the most productive city. But higher housing prices, and the associated decline in housing consumption, limit how much the planner can lower taxes in high-wage cities. In this section, we highlight key model features that affect this trade-off.

6.1 Level of Government Spending

First, we show that the level of taxes collected matters. If the planner has to collect higher tax revenue, placing more people in high-wage cities, i.e., increasing the tax base, becomes relatively more attractive.

Based on the evidence for the US economy, we have chosen parameter values for λ and τ that are most plausible. The total tax revenue is given by $1 - \lambda$. Our value for the tax revenue of 15% ($\lambda = 0.856$) includes income taxes as well as social security taxes. Instead, if we exclude social security contributions, the tax revenues would be around 8% ($\lambda = 0.922$). Alternatively, if consider the whole tax revenue including corporate and other taxes not related to labor income, then the tax revenue is

	$\lambda = 0.922$	$\lambda = 0.856$ *Benchmark*	$\lambda = 0.828$
Optimal τ^*	0.0856	0.0123	-0.0218
Welfare Gain (%)	0.0078	0.0534	0.0858
Output Gain (%)	0.38	1.00	1.27
Consumption (%)	0.37	0.95	1.18
Housing Consumption (%)	-0.71	-1.90	-2.43
Population Change, top 5 cities (%)	2.92	7.65	9.67
Fraction of Population that Moves (%)	1.27	3.31	4.19
Change in Average Housing Prices (%)	1.81	4.85	6.19

Table 4: The Role of Government Revenue Requirements

18.2% ($\lambda = 0.828$).[24] As Table 4 shows when government spending increases (λ decreases), the optimal τ^* declines and taxes becomes relatively lower in bigger cities. Indeed, with $\lambda = 0.828$, the optimal spatial tax schedule is regressive with $\tau^* = -0.0218$, and workers in more productive cities pay lower average taxes compared to the ones in smaller cities. In contrast, when revenue requirements are smaller ($\lambda = 0.922$), government does not need so many people in productive cities and instead chooses a much more progressive schedule with $\tau^* = 0.0856$.

In contrast to λ, as we show in the Appendix, the initial level of τ does not change the level of optimal dramatically. It does have an important effect on output changes and as a result on welfare.

6.2 Absentee Landlords

Second, we focus on the ownership of land. Often in the urban economics literature as in our benchmark model, housing ownership is assumed to be 100% in the hands of workers. Alternatively, at the other extreme, some models assume 'absentee landlords', where all housing is owned by a zero-measure of agents who do not enter in the planner's social welfare function. In our benchmark economy, rents from land are distributed equally across all workers in the economy, i.e., as if each worker holds an equal share of a diversified portfolio of land across US MSAs. This particular ownership structure allows the planner to choose a low τ, reducing after-tax wages in highly-productive cities. After all, the increase in housing prices benefits all households through the redistribution of land rents. Hence even workers in a less productive town that loses population benefit from higher housing prices in NY and LA. Instead, when ownership of land is in the hands of absentee landlords, benefits of higher land prices in more productive cities do not stay in the economy and can limit the planner's willingness to lower τ.

The right panel in Figure 12 shows average taxes across locations in the benchmark economy (dark blue) together with the tax function that arises when the planner chooses τ to maximize welfare (red). The same figure also shows the tax function with absentee landlords (green dash line). In an economy

[24]Source: National Income and Product Accounts (NIPA) Table 3.2. - Federal Government Current Receipts and Expenditures.

with absentee landlords where workers do not receive any income from land ownership, the planner lowers τ^{US} from 0.0123 to $\tau^* = 0.0603$, so taxes for $w = 0.5, 2$ and 5 are 10.5%, 15% and 22.6%, respectively, while they were 6.5%, 21.6% and 30.2% in the benchmark. But the level of τ is considerably higher compared to the benchmark where the land rents were distributed equally to all workers (see Table 5). As a result, there is a smaller rise in output, and therefore also a smaller increase in housing prices. The welfare gains from choosing τ optimally are much higher when land rents are distributed among all households equally.

	Benchmark	Absentee Landlords	Agglomeration	Equal Land
Optimal τ	0.0123	0.0603	-0.0323	0.0064
Welfare Gain (%)	0.0534	0.0196	0.1430	0.0598
Output Gain (%)	1.00	0.60	2.09	1.07
Consumption (%)	0.95	0.55	1.98	1.00
Housing Consumption (%)	-1.90	-1.14	-3.72	-1.80
Population Change, top 5 cities (%)	7.65	4.60	14.27	8.01
Fraction of Population that Moves (%)	3.31	1.99	6.33	3.51
Change in Average Housing Prices (%)	4.85	2.85	9.59	4.79

Table 5: The Role of Absentee Landlords, Agglomeration and Land Availability

6.3 Agglomeration Economies

There is a large empirical literature in urban economics that documents the extent of agglomeration economies in cities. Rosenthal and Strange (2004), Duranton and Puga (2004), Combes, Duranton, and Gobillon (2018) and Combes and Gobillon (2015) provide reviews of the recent papers that find elasticities of city level productivity with respect to the city size that are of the order of 0.03 to 0.08.

In this section, we introduce agglomeration economies as an externality in the production function. We assume that the production is given by $F(l_j) = (A_j l_j^\gamma) l_j$, where the l_j^γ term captures the level of agglomeration economies. Competitive firms still choose l_j to maximize profits, taking as given the externality $(A_j l_j^\gamma)$. The resulting wage rate is now given by $w_j = A_j l_j^\gamma$, where γ is the elasticity of wages with respect to the city size. As above, we use data on wages and the size of the work force across MSAs to estimate A_j and v, and then repeat our main quantitative exercise.[25] Given l_j, we estimate A_j and γ to fit the observed wages, w_j, in each city. Therefore the benchmark allocations in the economy with agglomeration externalities are identical to ones in the benchmark economy. The planner problem, on the other hand, now takes into account the fact that a larger workforce in a given city has a positive effect on average wages there.

Table 5 shows the aggregate outcomes for an economy with agglomeration externalities. Since there

[25] The estimated value of γ is about 0.045.

is now an extra external benefit from allocating workers to productive cities, the planner chooses a regressive tax schedule with $\tau = -0.0323$. The resulting tax function is shown in the left panel of Figure 12 (light blue line). Taxes decline in city size, and for $w = 0.5, w = 2$ and $w = 5$, the tax rates are 16.2%, 12.3%, and 9.7%, respectively. As a result, the share of population in the largest five MSAs grows by more than 14% and the resulting reallocation of labor generates a significant output gain that is higher than 2%.[26]

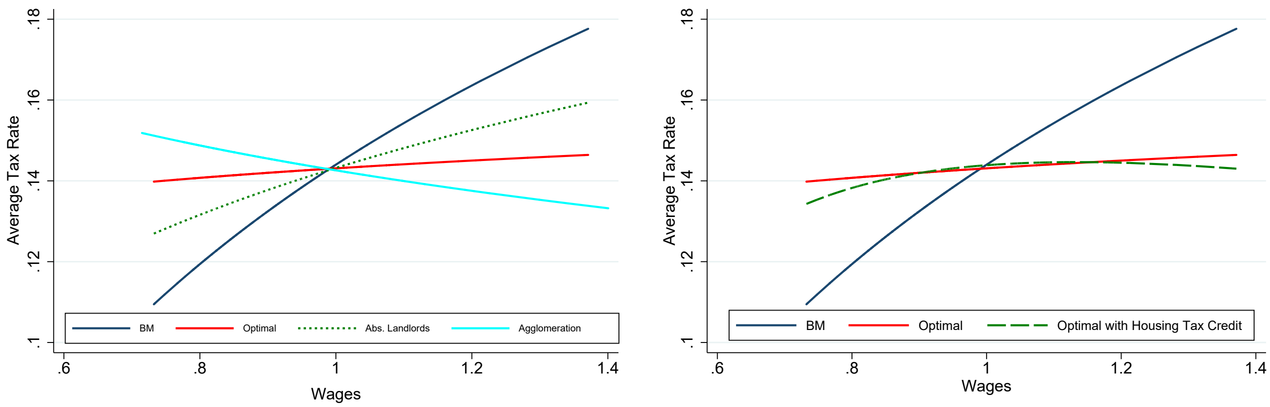

Figure 12: Changes in Optimal Spatial Tax Progressivity: A. Alternative Economies; B. Allowing for Tax Credits

6.4 An Economy with Identical Land Areas

We also study the effect of land distribution across MSAs. In the benchmark economy, more productive MSAs have less land; the correlation between wages and land size across MSAs is about -0.23. As a result, a lower τ makes these densely populated cities even more crowded, pushing housing prices up. Suppose NY or LA had as much available land as Anchorage, AK. Then, the higher population that results from lower τ would not influence housing prices as much. When we impose that each MSA has the same amount of land (the US average), the planner chooses an even less progressive tax schedule, which is almost proportional with a constant rate of around 14%, as shown in Table 5. The planner can now lower taxes in productive cities more than she did in the benchmark economy since the rise in housing prices is slightly more muted.

[26]Changing the congestion externality parameter has qualitatively similar effects as the introduction of agglomeration externalities since the functional form is the same. Congestion externalities however are tiny compared to the agglomeration externalities we find here. If we only shut down the congestion, i.e. set δ equal to zero, the optimal τ^* would be close to zero (-0.009). Hence, the planner would lower taxes in bigger cities more than she would do in the benchmark economy, which is not surprising.

6.5 Allowing for Tax Credits

Finally, we also study whether allowing the planner to use additional tax instruments can generate higher welfare gains. In particular, as the changes in house prices play a key role for the results, we consider refundable tax credits on housing expenditure. Hence, if in city i , households spend $p_i h_i$ on housing, the after tax wages are now given by

$$\tilde{w}_i = \lambda w_i^{1-\tau} + \chi p_i h_i$$

where $\chi \leq 1$ is the share of housing expenditure rebated to households. The planner chooses both τ and χ to maximize the welfare of the representative agent, and, as in other exercises, λ is adjusted so that the total tax collection is the same as in the benchmark economy. The resulting optimal spatial tax function is shown in in the right panel of Figure 12, together with the benchmark tax function and the optimal one obtained when the planner can only choose τ. The tax function with a tax credit is surprisingly close to one when the planner is constrained to use only τ. The tax credit allows the planner to lower taxes in low and high wage locations, creating a slightly hump-shaped tax function.[27] Yet, the welfare gains from using this additional instrument is small: welfare gains were 0.0534% when the planner can only choose τ, while they are only slightly higher, 0.0541%, when the planner can also choose χ, and the aggregate effects are very similar to the ones obtained in the benchmark experiment.

7 Conclusions

We have studied the role of federal income taxation on the misallocation of labor across geographical areas. More productive cities pay higher wages, and with redistributive taxes, workers in those cities also pay higher average taxes. Given perfect mobility, the tax schedule affects the incentives of where workers locate. Our objective has been to calculate the shape of the optimal spatial tax schedule in general equilibrium. When taxes change, citizens respond by relocating, but that in turn affects equilibrium prices. Those equilibrium effects determine both the optimal spatial tax schedule as well as the quantitative implications.

Our findings are, first, that the optimal spatial tax schedule is not flat and is sensitive to the level of government spending, to the presence of agglomeration externalities, and to the concentration of housing wealth. From a welfare viewpoint, what matters for the population allocation is the amount of government revenue and hence where it is best generated across differentially productive locations.

Second, quantitatively, the optimal spatial tax is less redistributive across space than the existing schedule in the data. Implementing the optimal schedule therefore favors the more productive cities.

[27] The optimal τ in this case is -0.1058, resulting in a regressive tax function. The optimal χ is close to 1, hence the planner wants to rebate the whole tax expenditure. This rebate lower the taxes paid more in low wage locations, so the resulting tax function turns out to be pretty flat, with a slight hump-shape.

In equilibrium this leads to output growth economy-wide, and population growth in the largest cities. The output growth is 1.42%. At the same time, there is first-order stochastic dominance in the city size distribution where the fraction of the population living in the five largest cities grows by 8%. The welfare effects however are small, 0.07%. Welfare obviously goes up, but in small amounts. This is due to the fact that the cost of living in the productive cities has increased commensurately. Our quantitative exercise also shows that the size of the government, the concentration of housing wealth, as well as the presence of agglomeration externalities play a critical role in determining the optimal spatial tax differences between large and small cities.

Appendix

Characterization of Equilibrium

Consider first the problem of construction firms. The First-Order Conditions are given by

$$\widetilde{p}_j B \frac{1}{\sigma}[(1-\theta)K_j^\sigma + \theta T_j^\sigma]^{\frac{1}{\sigma}-1}(1-\theta)\rho K_j^{\sigma-1} = 1, \tag{9}$$

and

$$\widetilde{p}_j B \frac{1}{\sigma}[(1-\theta)K_j^\sigma + \theta T_j^\sigma]^{\frac{1}{\sigma}-1}\theta\rho T_j^{\sigma-1} = r_j. \tag{10}$$

These conditions imply

$$K_j^\star = \left(\frac{1-\theta}{\theta} r_j\right)^{\frac{1}{1-\sigma}} T_j, \tag{11}$$

and

$$N_j = B\left[(1-\theta)\left(\frac{1-\theta}{\theta}r_j\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{1/\sigma} T_j. \tag{12}$$

The zero-profit condition then implies (after factoring out T_j and r_j):

$$\widetilde{p}_j = r_j \frac{\left(1 + \left(\frac{1-\theta}{\theta}\right)^{\frac{1}{1-\sigma}} r_j^{\frac{\sigma}{1-\sigma}}\right)}{B\left[(1-\theta)\left(\frac{1-\theta}{\theta}r_j\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{1/\sigma}}. \tag{13}$$

From the household problem we know that $p_j h_j = \alpha(\tilde{w}_j + R_j + TR)$. Since market clearing in the housing market requires that $h_j l_j = H_j$, this implies $\alpha(\tilde{w}_j + R + TR)l_j = p_j H_j$. Also, we know that $H_j = \frac{N_j}{\delta_h}$ and $p_j = \frac{r+\delta_h}{1+\rho}\tilde{p}_j$ which can be written as

$$p_j B\left[(1-\theta)\left(\frac{1-\theta}{\theta}r_j\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{1/\sigma} T_j = \alpha\delta_h l_j(\widetilde{w}_j + R + TR),$$

Substituting for other prices:

$$\frac{\rho+\delta_h}{1+\rho}\tilde{p}_j B\left[(1-\theta)\left(\frac{1-\theta}{\theta}r_j\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{1/\sigma} T_j = \alpha\delta_h l_j(\widetilde{w}_j + R + TR),$$

or, after substituting equation (13), rearranging and canceling terms:

$$r_j\left(1 + \left(\frac{1-\theta}{\theta}\right)^{\frac{1}{1-\sigma}} r_j^{\frac{\sigma}{1-\sigma}}\right) = \frac{\alpha\delta_h l_j(\widetilde{w}_j + R + TR)}{T_j}\frac{1+\rho}{v+\delta_h}. \tag{14}$$

Observe that this expression consists of one equation in one unknown, r_j. Given the solution for r_j, we can use equation (13) to find p_j.

In equilibrium each location has to give the same utility. Given equation (4), and normalizing $a_1 = 1$, we have

$$\frac{a_j}{a_1} = \frac{l_1^{\delta}(\tilde{w}_1 + R + TR_1)\left((\tilde{w}_1 + R + TR_1)l_1\right)^{-\alpha} H_1^{\alpha}}{l_j^{\delta}(\tilde{w}_j + R + TR_j)\left((\tilde{w}_j + R + TR_j)l_j\right)^{-\alpha} H_j^{\alpha}}.$$

Using the expression for H_j in (12) and since $N_j = H_j\delta_h$ we obtain

$$\frac{a_j}{a_1} = \frac{(\tilde{w}_1 + R + TR_j)^{1-\alpha} l_j^{\alpha-\delta}\left[(1-\theta)\left(\frac{1-\theta}{\theta}r_1\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{\alpha/\sigma} T_1^{\alpha}}{(\tilde{w}_j + R + TR_j)^{1-\alpha} l_1^{\alpha-\delta}\left[(1-\theta)\left(\frac{1-\theta}{\theta}r_j\right)^{\frac{\sigma}{1-\sigma}} + \theta\right]^{\alpha/\sigma} T_j^{\alpha}}.$$

The first order condition of production firms implies $w_j = A_j$, and $\tilde{w}_j$ is given by $(1-t_j)w_j$. Individuals own an equal share in a diversified portfolio of land holdings. Therefore R satisfies:

$$R = \frac{\sum_j r_j T_j}{\sum_j l_j}.$$

The population allocation must satisfy feasibility: $\sum_j l_j = \mathcal{L}$, and $\sum_{j=1}^{J} TR_j l_j = \phi G$. Hence, equation (**??**) can be used to pin down l_j for a given values of a_j.

Finally, in order to arrive at the aggregate resource constraint for this economy, we first aggregate the household budget constraints, $c_j + p_j h_j \leq \tilde{w}_j + R + TR$, across cities

$$\sum_{j=1}^{J} l_j c_j + \sum_{j=1}^{J} l_j p_j h_j = \sum_{j=1}^{J} l_j \tilde{w}_j + \sum_{j=1}^{J} l_j R + \sum_{j=1}^{J} l_j TR_j.$$

Since $l_j h_j = H_j$, $\sum_{j=1}^{J} l_j R = \sum_{j=1}^{J} l_j \frac{\sum_j r_j T_j}{\sum_j l_j} = \sum_j r_j T_j$, and $\sum_{j=1}^{J} l_j TR_j = \phi G$, we have

$$\sum_{j=1}^{J} l_j c_j + \sum_{j=1}^{J} p_j H_j = \sum_{j=1}^{J} l_j \tilde{w}_j + \sum_{j=1}^{J} r_j T_j + \phi G.$$

Adding and subtracting $\sum_j K_j$ to the right-hand side of this expression, we get

$$\sum_{j=1}^{J} l_j c_j + \sum_{j=1}^{J} p_j H_j = \sum_{j=1}^{J} l_j \tilde{w}_j + \sum_{j=1}^{J} r_j T_j + \sum_{j=1}^{J} K_j - \sum_{j=1}^{J} K_j + \phi G.$$

Since the housing production function is constant returns to scale, $\sum_{j=1}^{J} \tilde{p}_j N_j = \sum_j r_j T_j + \sum_j K_j$, $N_j = H_j \delta_h$, $\tilde{p}_j = \frac{1+r}{r+\delta_h} p_j$, which would mean:

$$\frac{\delta_h(1+r)}{r+\delta_h}\sum_{j=1}^{J} p_j H_j = \sum_j r_j T_j + \sum_j K_j$$

$$\sum_{j=1}^{J} p_j H_j = \frac{r+\delta_h}{\delta_h(1+r)}\left(\sum_j r_j T_j + \sum_j K_j\right)$$

Plug it in budget constraint:

$$\sum_{j=1}^{J} l_j c_j + \frac{r+\delta_h}{\delta_h(1+r)}\left(\sum_j r_j T_j + \sum_j K_j\right) = \sum_{j=1}^{J} l_j \tilde{w}_j + \sum_{j=1}^{J} r_j T_j + \phi G.$$

Factoring out terms:

$$\sum_{j=1}^{J} l_j c_j = \sum_{j=1}^{J} l_j \tilde{w}_j + \frac{r(\delta_h-1)}{\delta_h(1-r)}\sum_{j=1}^{J} r_j T_j - \frac{r+\delta_h}{\delta_h(1+r)}\sum_{j=1}^{J} K_j + \phi G.$$

Finally:

$$\sum_{j=1}^{J} l_j \tilde{w}_j = \sum_{j=1}^{J} l_j (1-t_j) w_j = \sum_{j=1}^{J} l_j w_j - \sum_{j=1}^{J} l_j t_j w_j$$

so that

$$\sum_{j=1}^{J} l_j c_j = \sum_{j=1}^{J} l_j w_j - \sum_{j=1}^{J} t_j l_j w_j + \phi\sum_{j=1}^{J} t_j w_j l_j + \frac{r(\delta_h-1)}{\delta_h(1-r)}\sum_{j=1}^{J} r_j T_j - \frac{r+\delta_h}{\delta_h(1+r)}\sum_{j=1}^{J} K_j,$$

which delivers the aggregate resource constraint for the economy:

$$\sum_{j=1}^{J} l_j c_j + \frac{r+\delta_h}{\delta_h(1+r)}\sum_{j=1}^{J} K_j + (1-\phi)\sum_{j=1}^{J} t_j w_j l_j = \sum_{j=1}^{J} l_j w_j + \frac{r(\delta_h-1)}{\delta_h(1-r)}\sum_{j=1}^{J} r_j T_j.$$

Proof of Proposition 1

The Ramsey planner's problem can be written as:

$$\max_{t_1,t_2} \; c_1^{1-\alpha}h_1^{\alpha}x + c_2^{1-\alpha}h_2^{\alpha}(1-x)$$
$$\text{s.t.} \; t_1w_1x + t_2w_2(1-x) = G$$
$$c_1^{1-\alpha}h_1^{\alpha} = c_2^{1-\alpha}h_2^{\alpha}$$
$$R = p_1h_1x + p_2h_2(1-x)$$
$$h_1x = 1; \; h_2(1-x) = 1$$
$$\max_{c_i,h_i} c_i^{1-\alpha}h_i^{\alpha}$$
$$\text{s.t.} \; c_i + p_ih_i = (1-t_i)w_i + R$$

The consumer's problem satisfies

$$h_i = \frac{\alpha((1-t_i)w_i + R)}{p_i} \quad \text{and} \quad c_i = (1-\alpha)((1-t_i)w_i + R).$$

Using $h_1 = \frac{1}{x}, \; h_2 = \frac{1}{1-x}$, prices can be written as:

$$p_1 = \alpha((1-t_1)w_1 + R)x \quad \text{and} \quad p_2 = \alpha((1-t_2)w_2 + R)(1-x).$$

Then R is equal to:

$$\begin{aligned} R &= p_1 + p_2 \\ &= \alpha((1-t_1)w_1 + R)x + \alpha((1-t_2)w_2 + R)(1-x) \\ &= \frac{\alpha}{1-\alpha}\left(x(1-t_1)w_1 + (1-t_2)(1-x)w_2\right) \end{aligned}$$

Utility equalization implies:

$$\frac{((1-t_1)w_1 + R)^{1-\alpha}}{x^{\alpha}} = \frac{((1-t_2)w_2 + R)^{1-\alpha}}{(1-x)^{\alpha}}$$

or

$$\left((1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}}\right)\left(\frac{\alpha(x(1-t_1)w_1 + (1-t_2)(1-x)w_2)}{1-\alpha}\right) + (1-x)^{\frac{\alpha}{1-\alpha}}(1-t_1)w_1 - $$
$$- x^{\frac{\alpha}{1-\alpha}}(1-t_2)w_2 = 0$$

or

$$\alpha((1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}})(x(1-t_1)w_1 + (1-t_2)(1-x)w_2) + (1-\alpha)((1-x)^{\frac{\alpha}{1-\alpha}}(1-t_1)w_1 - \\ -x^{\frac{\alpha}{1-\alpha}}(1-t_2)w_2) = 0$$

There is no explicit solution for general α, but for illustration, when $\alpha = 0.5$ we obtain two solutions:

$$x^*_{1,2} = \frac{2w_2(1-t_2) \pm \sqrt{2(w_2^2(1-t_2)^2 + w_1^2(1-t_1)^2)}}{2(w_2(1-t_2) - w_1(1-t_1))}$$

Since x and t_1, t_2 are constrained to be in the unit interval $[0, 1]$, the only feasible solution is x^*_2 (the negative root).

The planner problem can now be rewritten as:

$$\max_{t_1,t_2} \; ((1-\alpha)(1-t_1)w_1x + \alpha w_1x^2(1-t_1) + \alpha x(1-x)(1-t_2)w_2)^{1-\alpha} + ((1-\alpha)(1-x)(1-t_2)w_2 \\ +\alpha x(1-x)(1-t_1)w_1 + \alpha(1-x)^2(1-t_2)w_2)^{1-\alpha}$$

$$\text{s.t.} \;\; \alpha((1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}})(x(1-t_1)w_1 + (1-t_2)(1-x)w_2) + (1-\alpha)((1-x)^{\frac{\alpha}{1-\alpha}}(1-t_1)w_1 - \\ -x^{\frac{\alpha}{1-\alpha}}(1-t_2)w_2) = 0 \\ t_1 = \frac{G - t_2w_2(1-x)}{w_1x}$$

The Lagrangian:

$$\mathcal{L} = ((1-\alpha)(1-t_1)w_1x + \alpha w_1x^2(1-t_1) + \alpha x(1-x)(1-t_2)w_2)^{1-\alpha} + ((1-\alpha)(1-x)(1-t_2)w_2 \\ +\alpha x(1-x)(1-t_1)w_1 + \alpha(1-x)^2(1-t_2)w_2)^{1-\alpha} \\ -\lambda_G(t_1w_1x^*_2 + t_2w_2(1-x^*_2) - G) \\ -\lambda_x\Big(\alpha((1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}})(x(1-t_1)w_1 + (1-t_2)(1-x)w_2) + (1-\alpha)((1-x)^{\frac{\alpha}{1-\alpha}}(1-t_1)w_1 \\ -x^{\frac{\alpha}{1-\alpha}}(1-t_2)w_2)\Big)$$

We are not able to derive the analytical solution to the system of FOCs of this Lagrangian. Numerical solutions show that $t_1 = \frac{G}{w_1}$ and $t_2 = \frac{G}{w_2}$. We plug these solutions in the expression for x with general α:

$$\alpha((1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}})(x(1-\frac{G}{w_1})w_1 + (1-\frac{G}{w_2})(1-x)w_2) + (1-\alpha)((1-x)^{\frac{\alpha}{1-\alpha}}(1-\frac{G}{w_1})w_1 \\ -x^{\frac{\alpha}{1-\alpha}}(1-\frac{G}{w_2})w_2) = 0$$

$$\alpha((1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}})(x(w_1-G) + (w_2-G)(1-x)) + (1-\alpha)((1-x)^{\frac{\alpha}{1-\alpha}}(w_1-G) - x^{\frac{\alpha}{1-\alpha}}(w_2-G)) = 0 \quad (15)$$

If $G = 0$, the expression looks like:

$$\alpha((1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}})(xw_1 + w_2(1-x)) + (1-\alpha)((1-x)^{\frac{\alpha}{1-\alpha}}w_1 - x^{\frac{\alpha}{1-\alpha}}w_2) = 0$$

Comparative Statics:

$$\begin{aligned}
\frac{\partial t_1}{\partial G} &= \frac{1}{w_1} > 0 \\
\frac{\partial t_2}{\partial G} &= \frac{1}{w_2} > 0 \\
\frac{\partial (t_1 - t_2)}{\partial G} &= \frac{w_2 - w_1}{w_2 w_1} > 0 \\
\frac{dx}{dG} = \frac{(1-\alpha)x^{\frac{1-2\alpha}{1-\alpha}}(1-x)^{\frac{1-2\alpha}{1-\alpha}}\left(1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}}\right)}{\alpha\left((1-x)^{\frac{1-2\alpha}{1-\alpha}} + x^{\frac{1-2\alpha}{1-\alpha}}\right)(G - w_1 x - w_2(1-x))} &< 0
\end{aligned}$$

The last inequality follows from the fact that $G - w_1 x - w_2(1-x) < 0$ and all other expressions are positive.

This establishes the results in the Proposition regarding the change of taxes, population and output as G changes. Now, we show the last piece that establishes that when $G = 0$, the decentralized equilibrium allocation coincides with the Ramsey solution.

Decentralized allocation when $G = 0$. The household's problem is:

$$\max_{c_i, h_i} c_i^{1-\alpha} h_i^{\alpha}$$

subject to

$$c_i + p_i h_i = w_i + R,$$

and market clearing:

$$\begin{aligned}
x p_1 h_1 + (1-x) p_2 h_2 &= R \\
x h_1 &= 1 \\
(1-x) h_2 &= 1.
\end{aligned}$$

These market clearing equations jointly imply

$$p_1 + p_2 = R.$$

The FOC to the consumer maximization problem are:

$$
\begin{aligned}
(1-\alpha)c_i^{-\alpha}h_i^{\alpha} + \lambda_i &= 0 \\
\alpha h_i^{\alpha-1}c_i^{1-\alpha} + \lambda_i p_i &= 0
\end{aligned}
$$

or

$$
\frac{\alpha c_i}{(1-\alpha)h_i} = p_i
$$

Aggregate consumption satisfies:

$$
\begin{aligned}
x(c_1 + p_1h_1) + (1-x)(c_2 + p_2h_2) &= xw_1 + (1-x)w_2 + R \\
xc_1 + (1-x)c_2 &= xw_1 + (1-x)w_2
\end{aligned}
$$

The FOC together with feasibility in the housing market implies:

$$
\begin{aligned}
\frac{\alpha}{1-\alpha}c_1 x &= p_1 \\
\frac{\alpha}{1-\alpha}c_2(1-x) &= p_2
\end{aligned}
$$

then

$$
\frac{\alpha}{1-\alpha}c_1 x + \frac{\alpha}{1-\alpha}c_2(1-x) = p_1 + p_2 = R.
$$

Now we can write the budget constraint of an individual household as:

$$
c_i + p_ih_i = c_i + \frac{\alpha}{1-\alpha}c_1 = w_i + R \;\Rightarrow\; c_i = (1-\alpha)(w_i + R),
$$

and we obtain

$$
\begin{aligned}
c_i &= \ (1-\alpha)(w_i + R) \\
h_i &= \frac{\alpha(w_i + R)}{p_i}
\end{aligned}
$$

and

$$
\begin{aligned}
R &= \frac{\alpha}{1-\alpha}c_1 x + \frac{\alpha}{1-\alpha}c_2(1-x) \\
&= \alpha(x(w_1 + R) + (1-x)(w_2 + R)) \\
&= \frac{\alpha(w_1 x + w_2(1-x))}{1-\alpha}.
\end{aligned}
$$

Utility equalization in cities implies:

$$\frac{c_1^{1-\alpha}}{x^\alpha} = \frac{c_2^{1-\alpha}}{(1-x)^\alpha}$$

$$(w_1+R)^{1-\alpha}(1-x)^\alpha = (w_2+R)^{1-\alpha}x^\alpha$$

$$(1-x)^{\frac{\alpha}{1-\alpha}}\left(w_1+\frac{\alpha(xw_1+(1-x)w_2)}{1-\alpha}\right) = x^{\frac{\alpha}{1-\alpha}}\left(w_2+\frac{\alpha(xw_1+(1-x)w_2)}{1-\alpha}\right)$$

or equivalently

$$((1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}})\left(\frac{\alpha(xw_1+(1-x)w_2)}{1-\alpha}\right) + (1-x)^{\frac{\alpha}{1-\alpha}}w_1 - x^{\frac{\alpha}{1-\alpha}}w_2 = 0$$

$$\alpha((1-x)^{\frac{\alpha}{1-\alpha}} - x^{\frac{\alpha}{1-\alpha}})(xw_1+(1-x)w_2) + (1-\alpha)((1-x)^{\frac{\alpha}{1-\alpha}}w_1 - x^{\frac{\alpha}{1-\alpha}}w_2) = 0$$

Where the last expression gives solution for x^* which is the same as in Ramsey problem with $G=0$.

Wage and Population Distributions

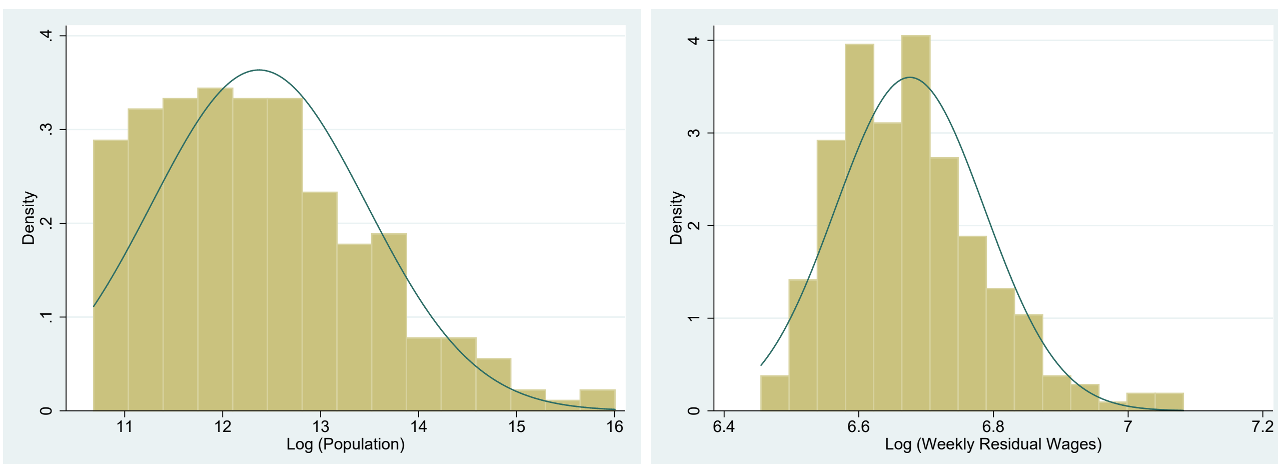

Figure 13: A. Histogram and Kernel density of labor force; B. Histogram and Kernel density of wages.

Estimating the Tax Functions

The OECD tax-benefit calculator provides the gross and net (after taxes and benefits) labor income at every percentage of average labor income on a range between 50% and 200% of average labor income, by year and family type. We simulate values for after and before taxes for increments of 25% of average labor income. As the OECD tax-benefit calculator only allows us to calculate wages up to 200% of average labor income, we use the procedure proposed by Guvenen, Burhan, and Ozkan (2014). In particular, let w denote average wage income before taxes as a multiple of mean wage income before taxes, and $t(w)$ and $\bar{t}(w)$ the marginal and average tax rates on wage income w. Also let t_{top} and w_{top}

be the top marginal tax rate and top marginal income tax bracket.[28] Suppose $w > 2$ and $w_{top} < 2$, i.e. top income bracket is less than 2. Then,

$$t(w) = \frac{(\bar{t}(2) \times 2 + t_{top} \times (w - 2))}{w}.$$

If $w_{top} > 2$ (which is the case for the US), we do not know the marginal tax rate between $w = 2$ and w_{top}. First set

$$t(2) = \frac{(\bar{t}(2) \times 2 - \bar{t}(1.75) \times 1.75)}{0.25}$$

and use linear interpolation between $t(2)$ and t_{top}

$$t(w) = \begin{cases} (t(2) + \frac{t_{top} - t(2)}{w_{top} - 2}(w - 2) & \text{if } 2 < w < w_{top} \\ t_{top} & \text{if } w > w_{top} \\ \end{cases}$$

Then average tax rate function for $w > 2$ is

$$\bar{t}(w) = \begin{cases} (\bar{t}(2) \times 2 + t(w) \times (w - 2))/w & \text{if } 2 < w < w_{top} \\ (\bar{t}(2) \times 2 + \frac{t_{top} + t(2)}{2}(w_{top} - 2) + t_{top} \times (w - w_{top}))/w & \text{if } w > w_{top} \end{cases}$$

Land Distribution across MSAs

The Figure 14 shows the distribution of land across MSAs.

Figure 15 shows the relation between weekly wages and available land across MSAs.

[28] Top marginal tax rate is taken from http://www.oecd.org/tax/tax-policy/oecdtaxdatabase.htm, Table I.7.

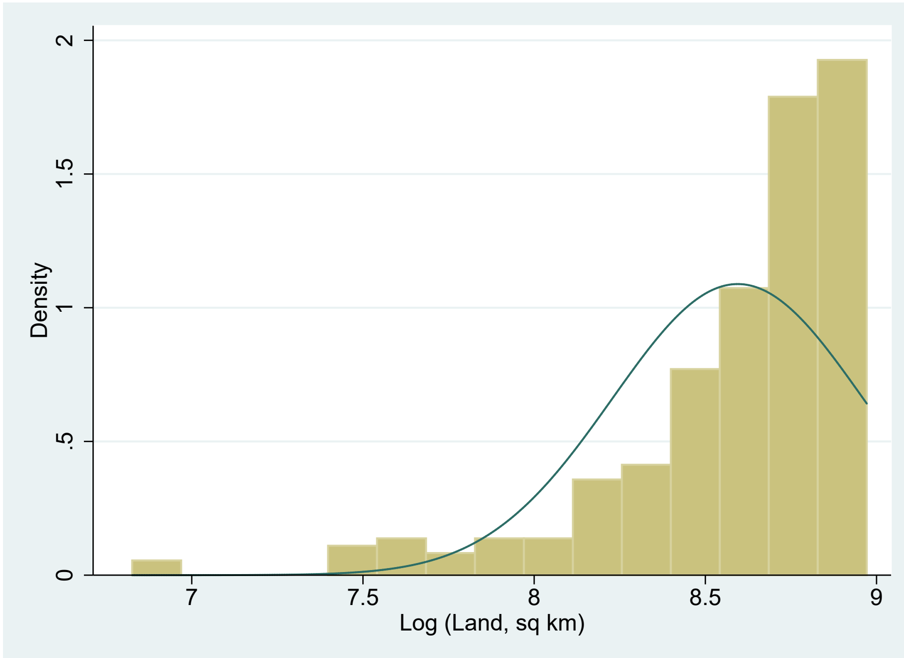

Figure 14: Land Distribution

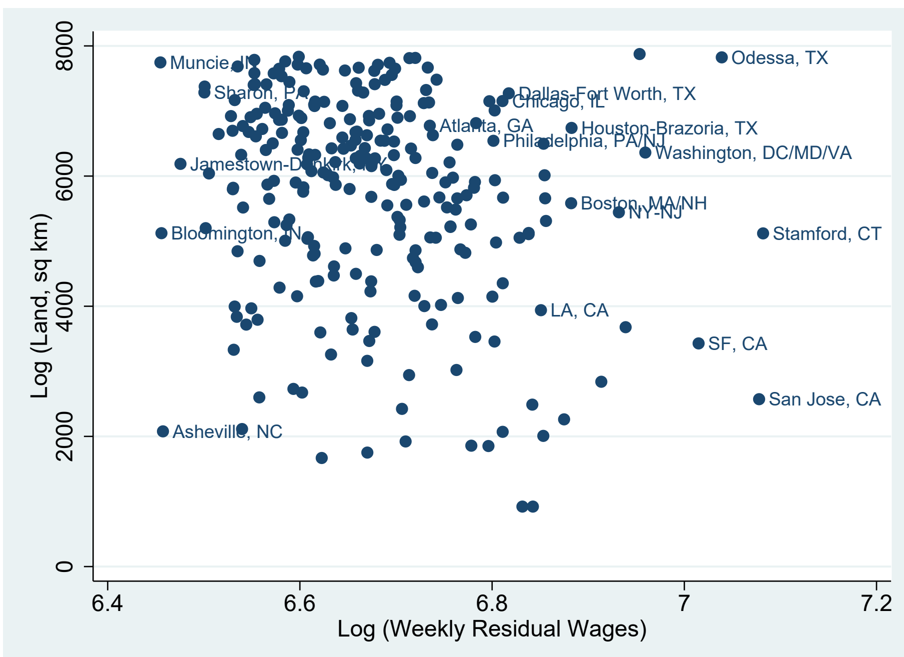

Figure 15: Land and Wages

City-Specific Transfers

The data in Figure 4 is from the US Bureau of Economic Analysis (BEA) Regional Accounts (https://www.bea.gov/re
Per capita transfers, excluding social security payments, are calculated, using data from Table CA35 Personal Current Transfer Receipts. The regression line is obtained from an OLS regression with per-

capita transfers as the dependent variable and per-worker wage income (productivity) as the explanatory variable.

We assume that city-specific transfers, TR_j, are a declining function of city-specific wages in the model, given by

$$TR_j = \eta_1 + \eta_2 w_j, \text{ with } \eta_1 > 0, \text{ and } \eta_2 < 0.$$

Then,

$$TR_{\min} = \eta_1 + \eta_2 w_{\max} \tag{16}$$

$$TR_{\max} = (1+\chi)TR_{\min} = \eta_1 + \eta_2 w_{\min}, \tag{17}$$

where $\chi = 0.857$ is calculated from the ratio of maximum to minimum transfers along the OLS line estimated in the model.

We want to find $TR_{\min}$, η_1 and η_2 in the model satisfies equations (16) and (17) and

$$\sum_i TRL_i = \eta_1 \sum L_i + \eta_2 \sum w_i L_i = G, \tag{18}$$

where L_i population of MSA i, and G is total government transfers.

From equations (16) and (17), we have

$$\chi TR_{\min} = \eta_2(w_{\min} - w_{\max}),$$

or

$$\eta_2 = -\frac{\chi TR_{\min}}{(w_{\max} - w_{\min})}.$$

Let $w_{\max} - w_{\min} = w_{gap}$, then equation (3) implies

$$\eta_1 \sum L_i - \frac{\chi \tau_{\min}}{w_{gap}} \sum w_i L_i = G.$$

Using $TR_{\min} = \eta_1 + \eta_2 w_{\max}$, we have

$$\eta_1 = TR_{\min} - \eta_2 w_{\max}.$$

As a result,

$$(\tau_{\min} - \eta_2 w_{\max}) \sum L_i - \frac{\lambda TR_{\min}}{w_{gap}} \sum w_i L_i = G,$$

or

$$(TR_{\min} - (-\frac{\lambda TR_{\min}}{w_{gap}}) w_{\max}) \sum L_i - \frac{\lambda TR_{\min}}{w_{gap}} \sum w_i L_i = G,$$

which is one equation in one unknown $\tau_{\min}$, given by

$$TR_{\min}\left[\sum L_i + \frac{\chi}{w_{gap}} w_{\max} \sum L_i - \frac{\chi}{w_{gap}} \sum w_i L_i\right] = G,$$

or

$$TR_{\min} = \frac{G}{\left[\sum L_i + \frac{\lambda w_{\max}}{w_{gap}} \sum L_i - \frac{\chi}{w_{gap}} \sum w_i L_i\right]}. \tag{19}$$

Once $TR_{\min}$ is determined, η_1 and η_2 can also be determined, from equations (16) and (17).

Note that we could also use $(1+\chi)TR_{\min} = \eta_1 + \eta_2 w_{\min}$ to set

$$\eta_1 = (1+\chi)TR_{\min} - \eta_2 w_{\min}.$$

Then,

$$[(1+\chi)TR_{\min} - \eta_2 w_{\min}] \sum L_i - \frac{\lambda TR_{\min}}{w_{gap}} \sum w_i L_i = G,$$

This is equivalent to equation (19) since

$$TR_{\min} - \eta_2 w_{\max} = (1+\chi)TR_{\min} - \eta_2 w_{\min}.$$

The Effect of Initial τ

The benchmark value of $\tau = 0.127$ reflects taxes on labor income based on the OECD tax calculator. Instead, we could have focused on total household income from the IRS micro data that includes income on assets. Considering both taxes paid and Earned Income Tax Credits (EITC) refunds received by the households, Guner, Kaygusuz, and Ventura (2014) estimate a lower $\tau = 0.053$, for all households. Their estimates for married households with children, who are much more likely to benefit EITC, imply a higher $\tau = 0.2$. Also, taking into account transfers, Heathcote, Storesletten, and Violante (2017) estimate $\tau = 0.18$. We repeat the same exercise for different initial values of τ, the results of which are reported in Table 6. The initial level of τ does not change the level of optimal dramatically, it does have an important effect on output and as a result welfare.

	$\tau = 0.053$	$\lambda = 0.856$ $\tau = 0.127$ *Benchmark*	$\tau = 0.2$
Optimal τ^*	0.0092	0.0123	0.0159
Welfare Gain (%)	0.0077	0.0534	0.1389
Output Gain (%)	0.38	1.00	1.62
Consumption (%)	0.36	0.95	1.54
Housing Consumption (%)	-0.71	-1.90	-3.12
Population Change, top 5 cities (%)	2.91	7.65	12.29
Fraction of Population that Moves (%)	1.26	3.31	5.33
Change in Average Housing Prices (%)	1.80	4.85	8.01

Table 6: The effect of different initial τ

References

Ahlfeldt, G. M., F. Bald, D. Roth, and T. Seidel (2024): "Measuring Quality of Life under Spatial Frictions," Discussion paper, Centre for Economic Performance, Discussion Paper No. 2061.

Albert, C., and J. Monras (2022): "Immigration and Spatial Equilibrium: The Role of Expenditures in the Country of Origin," *American Economic Review*, 112(11), 3763–3802.

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

Albouy, D., K. Behrens, F. Robert-Nicoud, and N. Seegert (2019): "The optimal distribution of population across cities," *Journal of Urban Economics*, 110, 102–113.

Albouy, D., and G. Ehrlich (2012): "Metropolitan Land Values and Housing Productivity," NBER Working Paper No. 18110.

Bénabou, R. (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 70(2), 481–517.

Borella, M., M. De Nardi, M. Pak, N. Russo, and F. Yang (2023): "FBBVA Lecture 2023. The Importance of Modeling Income Taxes over Time: U.S. Reforms and Outcomes," *Journal of the European Economic Association*, 21(6), 2237–2286.

Coen-Pirani, D. (2010): "Understanding gross worker flows across US states," *Journal of Monetary Economics*, 57(7), 769–784.

Colas, M., and K. Hutchinson (2021): "Heterogeneous Workers and Federal Income Taxes in a Spatial Equilibrium," *American Economic Journal: Economic Policy*, 13(2), 100–134.

Combes, P.-P., G. Duranton, and L. Gobillon (2018): "The Costs of Agglomeration: House and Land Prices in French Cities," *The Review of Economic Studies*, 86(4), 1556–1589.

Combes, P.-P., and L. Gobillon (2015): "The Empirics of Agglomeration Economies," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J. F. thisse, vol. 5, pp. 247–348. Elsevier B.V.

Davis, M., and J. Heathcote (2007): "The Price and Quantity of Residential Land in the United States," *Journal of Monetary Economics*, 54(8), 2595–2620.

Davis, M., and M. G. Palumbo (2008): "The Price of Residential Land in Large U.S. Cities," *Journal of Urban Economics*, 63(1), 352–384.

Davis, M. A., J. D. Fisher, and M. Veracierto (2021): “Migration and urban economic dynamics,” *Journal of Economic Dynamics and Control*, 133, 104234.

Davis, M. A., W. D. Larson, S. D. Oliner, and J. Shui (2021): “The Price of Residential Land for Counties, ZIP Codes, and Census Tracts in the United States,” *Journal of Monetary Economics*, 118, 413–431.

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

Desmet, K., and E. Rossi-Hansberg (2013): “Urban Accounting and Welfare,” *American Eco*, 103(6), 2296–2327.

Diamond, R. (2016): “The determinants and welfare implications of US workers’ diverging location choices by skill: 1980–2000,” *The American Economic Review*, 106(3), 479–524.

Duranton, G., and D. Puga (2004): “Micro-foundations of urban agglomeration economies,” *Handbook of regional and urban economics*, 4, 2063–2117.

Duranton, G., and M. A. Turner (2012): “Urban Growth and Transportation,” *The Review of Economic Studies*, 79(4), 1407–1440.

Dyrda, S., and B. Pugsley (2024): “The Rise of Pass-Throughs: an Empirical Investigation,” *The Economic Journal*, 135(665), 387–403.

Eeckhout, J. (2004): “Gibrat’s Law for (All) Cities,” *American Economic Review*, 5(94), 1429–1451.

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

Fajgelbaum, P. D., E. Morales, J. C. Suárez Serrato, and O. Zidar (2018): “State Taxes and Spatial Misallocation,” *The Review of Economic Studies*, 86(1), 333–376.

Fleck, J., J. Heathcote, K. Storesletten, and G. L. Violante (2025): “Fiscal Progressivity of the U.S. Federal and State Governments,” Discussion paper, NBER Working Paper 33385.

Floetotto, M., M. Kirker, and J. Stroebel (2016): “Government intervention in the housing market: Who wins, who loses?,” *Journal of Monetary Economics*, 80, 106–123.

Giannone, E., N. Paixão, X. Pang, and Q. Li (2023): “Unpacking Moving: A Spatial Equilibrium Model with Wealth,” Discussion paper, CEPR Discussion Paper No. 18072.

Glaeser, E. L. (1998): “Should transfer payments be indexed to local price levels?,” *Regional Science and Urban Economics*, 28(1), 1–20.

Guner, N., R. Kaygusuz, and G. Ventura (2014): "Income Taxation of U.S. Households: Facts and Parametric Estimates," *Review of Economic Dynamics*, 17(4), 559–581.

Guner, N., M. Lopez-Daneri, and G. Ventura (2016): "Heterogeneity and Government revenues: Higher taxes at the top?," *Journal of Monetary Economics*, 80, 69–85.

Guner, N., G. Ventura, and X. Yi (2008): "Macroeconomic Implications of Size-Dependent Policies," *Review of Economic Dynamics*, 11(4), 721–744.

Guvenen, F., K. Burhan, and S. Ozkan (2014): "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis," *Review of Economic Studies*, 81, 818–850.

Haughwout, A. F. (2002): "Public infrastructure investments, productivity and welfare in fixed geographic areas," *Journal of Public Economics*, 83(3), 405–428.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal Tax Progressivity: An Analytical Framework*," *The Quarterly Journal of Economics*, 132(4), 1693–1754.

Helpman, E., and D. Pines (1980): "Optimal Public Investment and Dispersion Policy in a System of Open Cities," *The American Economic Review*, 70(3), 507–514.

Herkenhoff, K. F., L. E. Ohanian, and E. C. Prescott (2018): "Tarnishing the golden and empire states: Land-use restrictions and the U.S. economic slowdown," *Journal of Monetary Economics*, 93, 89–109, Carnegie-Rochester-NYU Conference on Public Policy held at the Stern School of Business at New York University.

Hochman, O., and D. Pines (1993): "Federal Income Tax and Its Effects On Inter- and Intracity Resource Allocation," *Public Finance Quarterly*, 21(3), 276–304.

Hsieh, C.-T., and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124(4), 1403–1448.

Hsieh, C.-T., and E. Moretti (2015): "Why Do Cities Matter? Local Growth and Aggregate Growth," NBER Working Paper No. 21154.

Huggett, M., and W. Luo (2023): "Optimal income taxation: An urban economics perspective," *Review of Economic Dynamics*, 51, 847–866.

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

Kaplan, G., and S. Schulhofer-Wohl (2017): "Understanding the Long-Run Decline in Interstate Migration," *International Economic Review*, 58(1), 57–94.

Kaplow, L. (1995): "Regional Cost-of-Living Adjustments in Tax/Transfer Schemes," *Tax Law Review*, 51, 175–198.

Karahan, F., and S. Rhee (2014): "Population aging, migration spillovers, and the decline in interstate migration," NY Fed mimeo.

Kennan, J. (2013): "Open Borders," *Review of Economic Dynamics*, 16(2), L1–L13.

Kindermann, F., and D. Krueger (2022): "High Marginal Tax Rates on the Top 1 Percent? Lessons from a Life-Cycle Model with Idiosyncratic Income Risk," *American Economic Journal: Macroeconomics*, 14(2), 319–66.

Klein, P., and G. Ventura (2009): "Productivity Differences and The Dynamic Effects of Labor Movements," *Journal of Monetary Economics*, 56(8), 1059–1073.

Knoll, M. S., and T. D. Griffith (2003): "Taxing Sunny Days: Adjusting Taxes for Regional Living Costs and Amenities," *Harvard Law Review*, 116, 987–1023.

Lipman, B. J., et al. (2006): "A heavy load: The combined housing and transportation burdens of working families," John D. and Catherine T. MacArthur Foundation.

Luccioletti, C. (2024): "Should Governments Subsidize Homeownership? A Quantitative Analysis of Spatial Housing Policies," Discussion paper, Bank of Italy.

Lutz, C., and B. Sand (2019): "Highly Disaggregated Land Unavailability," Working Paper.

McKenzie, B., and M. Rapino (2011): "Commuting in the United States: 2009," American Community Survey Reports, ACS-15. U.S. Census Bureau.

Notowidigdo, M. J. (2020): "The Incidence of Local Labor Demand Shocks," *Journal of Labor Economics*, 38(3), 687–725.

of Commerce, U. D. (2015): "2015 Characteristics of New Housing," U.S. Department of Housing and Urban Development.

Oswald, F. (2019): "The effect of homeownership on the option value of regional migration," *Quantitative Economics*, 3(3), 1453–1493.

Parkhomenko, A. (2023): "Local causes and aggregate implications of land use regulation," *Journal of Urban Economics*, 138, 103605.

Restuccia, D., and R. Rogerson (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11(4), 707–720.

Roberto, E. (2008): "Commuting to Opportunity: The Working Poor and Commuting in the United States," The Brookings Institution Series on Transportation Reform.

Roca, J. D. L., and D. Puga (2016): "Learning by Working in Big Cities," *The Review of Economic Studies*, 84(1), 106–142.

Rosenthal, S. S., and W. C. Strange (2004): "Evidence on the Nature and Sources of Agglomeration Economies," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J. F. thisse, vol. 4, pp. 2119–2171. Elsevier B.V.

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

Sommer, K., and P. Sullivan (2018): "Implications of US Tax Policy for House Prices, Rents, and Homeownership," *American Economic Review*, 108(2), 241–274.

Tiebout, C. (1956): "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64(5), 416–424.

Topa, G., A. Sahin, and G. Violante (2014): "Mismatch Unemployment," *American Economic Review*, 104(11), 3529–3564.

Wildasin, D. E. (1980): "Locational efficiency in a federal system," *Regional Science and Urban Economics*, 10(4), 453–471.