Optimal Spatial Policies, Geography and Sorting[*]

Pablo D. Fajgelbaum
UCLA and NBER

Cecile Gaubert
UC Berkeley, NBER and CEPR

May 2018

Abstract

We study optimal spatial policies in quantitative trade and geography frameworks with spillovers and sorting of heterogeneous workers. We characterize the spatial transfers that must hold in efficient allocations, as well as labor subsidies that would implement them. Assuming homogeneous workers and constant-elasticity spillovers, a constant labor tax over space restores efficiency regardless of micro heterogeneity in fundamentals. Place-specific subsidies are needed to attain optimal sorting if there are spillovers across different types of workers. We show how to quantify optimal spatial transfers, and apply the framework to data across U.S. cities using existing estimates of the spillover elasticities. The results suggest that the U.S. economy features too much spatial sorting by skill and wage inequality in larger cities relative to efficient allocations.

1   Introduction

A long tradition in economics argues that the concentration of economic activity leads to spillovers. For instance, agglomeration economies drive productive efficiency while congestion spillovers may arise in dense cities. These spillovers shape the distribution of city size and productivity. Groups of workers with different skills arguably vary in how much they contribute to these spillovers and in how much they are impacted by them, so that these forces also shape spatial sorting of heterogeneous workers across cities. Being external in nature, spillovers likely lead to inefficient spatial outcomes. In this paper, we ask: is the observed spatial distribution of economic activity inefficient? If so, what policies would restore efficiency and what would be their welfare impact? Would an optimal spatial distribution feature stronger, or weaker, spatial disparities and sorting by skill than what is observed?

To answer these questions, we develop and implement a new approach. Our framework nests two recent strands of quantitative spatial research with spillovers: the location choice model with sorting of heterogeneous workers in Diamond (2016), and the economic geography models in Allen and Arkolakis (2014) and Redding (2016). Crucially, we generalize these models to allow for arbitrary transfers across agents and regions. We show that first-best allocations can only be attained under a particular set of transfers, which we characterize alongside labor subsidies that would implement them. We then combine the framework with data across metropolitan statistical areas (MSAs) in the United States, and quantitatively explore the impact of optimal spatial transfers on the spatial sorting by skill, wage inequality within cities, and welfare relative to the observed equilibrium. Under existing estimates of the spillover elasticities, the results suggest that the U.S. economy features too much concentration of high-skill workers and wage inequality in larger cities relative to efficient allocations.

The framework incorporates many key determinants of the spatial distribution of economic activity. Firms produce differentiated tradeable commodities and non-tradeables such as housing using labor, intermediate inputs, and land. Locations may differ in fundamental components of productivity and amenities, bilateral trade frictions, and housing supply elasticities. Productivity and amenities are endogenous through agglomeration and congestion spillovers that may depend on the composition of the workforce.[1] Different types of workers vary in how productive they are in each location, in their ownership of other factors of production, in their preference for each location, and in the efficiency and amenity spillovers they generate on other workers. In the market allocation, government policies may arbitrarily redistribute income across agents and regions.[2]

---

[1] As summarized by Duranton and Puga (2004), efficiency spillovers may result from several forces such as knowledge externalities, labor market pooling, or scale economies in the production of tradeable commodities. Amenity spillovers may result from congestion through traffic or environmental factors such as noise or pollution; availability of public services such as education, health, and public transport; availability of public amenities such as parks and recreation; or specialization thanks to scale effects in the provision of urban amenities such as restaurants or entertainment. Duranton and Venables (2018) discuss some of these congestion forces.

[2] A wide range of government policies lead to spatial transfers. Some of these are explicit "place-based policies", such as tax relief schemes targeted at distressed areas (e.g. New Markets Tax Credit, or Enterprise Zones) or direct public investment in specific areas (e.g. Tennessee Valley Authority). Other policies are not explicitly spatial, but

In the model, the spillovers have complex general-equilibrium ramifications through factor mobility and trade linkages. However, in the spirit of the "principle of targeting" pointed out by Dixit (1985), the first-best allocation can be implemented by policies acting only upon inefficient margins. Here, these margins consist of labor supply and demand decisions: workers do not internalize the impact of their location choice on city-level amenities, and firms do not internalize the impact of their hiring decisions on city-level productivity. We derive a necessary efficiency condition on the joint distribution of expenditures and employment across worker types and regions. Using this condition we then characterize the transfers that must hold in an efficient allocation. Furthermore, we identify a condition on the distributions of spillover and housing supply elasticities under which these optimal transfers are also sufficient to implement the efficient allocation.

This characterization generalizes a standard efficiency requirement from non-spatial environments such as Hsieh and Klenow (2009), whereby the marginal product of labor should be equalized across productive units. Here, the optimal spatial allocation balances the net benefit of agglomeration (through production or amenities) against the opportunity cost of attracting workers-consumers to each location. Because the location and consumption decisions are not separable, these opportunity costs are measured in terms of local consumption expenditures, and they vary across locations due to the compensating differentials born of geographic forces (congestion, amenities, trade costs, and non-traded goods). An important lesson is that diagnosing whether an observed allocation is efficient therefore requires information about expenditure per capita across locations, alongside the standard requirement of observing wages and employment.

We then impose specific functional forms in order to inspect policies that lead to the optimal transfers in well-known models nested in our framework.[3] We consider labor subsidies that tackle the spatial inefficiencies due to spillovers, whereas lump-sum transfers take care of distributional concerns across types. As a first benchmark, we apply our results to a model with constant-elasticity spillovers with respect to city size and no spillovers across worker types. Under homogeneous workers and no trade costs, this case corresponds to the canonical Roback (1982) framework extended with spillovers in both amenities and productivity. A common intuition running through several studies of place-based policies is that, unless externalities are different across cities, policies that reallocate workers cannot lead to welfare gains, because increases in externalities in locations that grow are exactly offset by decreases in ones that shrink.[4] However, our results imply that this intuition does not hold up. The allocation is inefficient even if spillover elasticities are constant. In that case, a constant labor tax over space coupled with a lump sum subsidy generates net transfers that distort the allocation of workers across cities and restores efficiency. Somewhat paradoxically, in this case the optimal "place-based" policy turns out to be a policy that is independent across space.

---

end up redistributing income to specific places (e.g., nominal income taxes and credits, state and local tax deductions, or sectoral subsidies). Neumark et al. (2015) and Austin et al. (2018) review some of these policies.

[3]Since our focus is on spatial inefficiencies from spillovers, our analysis assumes that distributional concerns across types can be tackled through type-specific lump-sum taxes.

[4]This notion is discussed for instance in Glaeser and Gottlieb (2008), Kline and Moretti (2014a) and the review of the literature on place-based policies by Neumark et al. (2015).

We then apply our results to establish the normative properties of well-known economic geography models, such as Allen and Arkolakis (2014), Helpman (1998) and Redding (2016). These models correspond to special cases of our framework provided we assume inelastic housing supply, a single worker type, constant elasticity spillovers and no intermediate inputs. In this context, global efficiency is characterized by the distribution of trade imbalances between regions. This distribution can be implemented by a simple transfer rule that depends on spillover elasticities and the non-traded expenditure share, but is independent from the distribution of fundamentals or trade costs. We show that, because these models make different assumptions about transfers in the *laissez-faire* allocation, they have different implications for whether the optimal government intervention should redistribute income from high- to low-wage regions, or the opposite.

Our theory thus implies that under constant-elasticity spillovers and either homogeneous workers or no spillovers across different types of workers, constant labor taxes over space implement the efficient allocation. In the general case with different worker types and heterogeneous spillovers, spatial variation in optimal policies arises to attain efficient sorting. For example, if the productivity of low skill workers is positively affected by the presence of high skill workers, the decentralized pattern of sorting by skill may be too strong. Spatial efficiency may then require subsidies to high skill workers in cities where they are relatively scarce under the *laissez-faire* allocation, increasing the degree of mixing of heterogeneous workers across locations relative to the competitive allocation.

We then show how to use the main framework for quantitative analysis. We identify the set of sufficient statistics in the data which, in addition to the spillover elasticities and standard production and utility function parameters, are needed to compute the optimal allocation. The exercise requires information on expenditures per capita across worker types and regions, on top of the standard data requirements in economic geography models such as the distributions of wages, employment and trade flows.

We apply the model to data on the distribution of economic activity across MSAs in the United States. We allow for two skills groups, high skill (college) and low skill (non college) workers. To fulfill the previous data requirements, we combine data on labor and non-labor income, taxes and transfers at the city level from the BEA, with Census data that allows us to break down these MSA-level totals by skill group within cities. To parametrize the spillover elasticities we rely on existing values from the literature that, also in the U.S. context, has estimated equations that are consistent with our model. Specifically, we draw estimates of the amenity spillovers and of the heterogeneity in spillovers across workers from Diamond (2016), and of the city-level elasticity of labor productivity with respect to employment density from Ciccone and Hall (1996). We also explore results under several different specifications around these estimates.

We first implement the quantification assuming homogeneous workers. In that case, we find negligible welfare gains from optimal reallocation, as well as optimal transfers that are close to the data. To demonstrate that this result is not a feature of our model but a feature of the data, we apply our approach to counterfactual data corresponding to alternative fundamentals, and show that it can deliver substantial welfare gains even with homogeneous workers.

The full model with high and low skill workers yields welfare gains from roughly 4% to 8% across specifications of the spillover elasticities. These gains materialize through a stronger redistribution towards low-wage cities than what is observed empirically. These transfers imply higher labor income taxes in high-wage cities for both low and high skill workers. In the case of low skill workers, the higher taxes in high-wage cities arise because these workers generate congestion and small productivity spillovers. In contrast, for high skill workers, they arise because these workers generate positive spillovers onto low skill workers, who are more prevalent in low-wage cities. This second force offsets the strong positive spillovers that high skill workers generate among each other, which by itself would call for a subsidy to these workers in high-wage cities. As a result of the reallocations induced by these transfers, within-city wage inequality decreases in initially larger and more unequal cities. These qualitative patterns are consistent across specifications that vary the spillover elasticities around the benchmark by a factor of two.

These results suggest that nudging U.S. policies towards generating a greater mixing of high and low skill workers in currently low-wage cities would be welfare improving. To further identify the key spillovers driving this result, in a final exercise we invert the logic of the approach: instead of calibrating spillover elasticities to existing estimates and inspecting optimal transfers for a range of values around them, we assume that the observed equilibrium is efficient and use our optimal transfers formulas to infer the corresponding spillover elasticities. This procedure implies strong congestion spillovers that are similar, on average, to the elasticities drawn from existing estimates, as well as positive efficiency spillovers generated by high skill workers. However, it also implies negative amenity spillovers of similar magnitude for both skill groups, unlike in the calibration where low skill and high skill workers generate spillovers of opposite signs. In this sense, we identify a key role for the heterogeneous amenity spillovers across skill types.[5]

The rest of the paper is structured as follows. Section 2 connects the paper to the related literature. Section 3 presents the general environment and defines the decentralized allocation and the planner's problem. Section 4 characterizes the optimal policies, teases out their implications in specific cases of the theory corresponding to the models from the literature, and determines the data that suffices to implement the model. Section 5 describes the data and the calibration. Section 6 presents the quantitative implementation and Section 7 concludes. Proofs, additional derivations and data construction are detailed in the appendix.

2 Relation to the Literature

The economic geography models nested in our framework introduce labor mobility and spillovers in general-equilibrium models with trade frictions such as Eaton and Kortum (2002) and Anderson

---

[5]In our parametrization, these spillovers rely on numbers from Diamond (2016), who estimates a positive response of an urban amenity index (including congestion in transport, crime, environmental indicators, supply per capita of different public services, and variety of retail stores) to the relative supply of college workers, as well as a higher marginal valuation for these amenities for college than for non-college workers.

and van Wincoop (2003).[6] These models are the basis of a growing body of research, summarized by Redding and Turner (2015) and Redding and Rossi-Hansberg (2017), that studies the counterfactual implications of particular shocks to economic fundamentals or trade costs in spatial settings. Existing applications of these frameworks, such as Caliendo et al., (forthcoming), impose restrictions on transfers across locations through assumptions on how the income generated by fixed factors is distributed across locations. We show that the optimal allocation must be implemented through a particular set of spatial transfers, and characterize the distribution of imbalances that ensures spatial efficiency. We focus our theoretical analysis and quantitative application on models with local spillovers, where efficiency and amenities are a function of the distribution of workers within a location. In an extension, we also derive optimal transfers in the case where economic activity in one location may generate spillovers in other locations in the spirit of Rossi-Hansberg (2005) and recent quantitative studies such as Ahlfeldt et al. (2015) and Desmet and Rossi-Hansberg (2014).[7]

Our framework also nests the model of Diamond (2016), who allows for asymmetric amenity and efficiency spillovers across workers with different skills in a Rosen (1979)-Roback (1982) model and takes the model to the data using the U.S. as an empirical setting. Her analysis shows that these forces partly determine the spatial distribution of skill shares and skill wage premia, including the recent "great divergence" in these outcomes across cities in the U.S. pointed out by Moretti (2012), among others.[8] Giannone (2017) argues that productivity spillovers across workers are relevant to understand the spatial effects of skill biased technological change.

In this paper, we study on the normative implications of these forces. In doing so, we are also motivated by the well known empirical evidence that larger cities feature higher wages, higher share of skilled workers, and higher skill premium, as documented among others by Combes et al. (2008). We ask whether these observed patterns of spatial disparities are too strong from the perspective of spatial efficiency. Recent papers such as Eeckhout et al. (2014), Behrens et al. (2014), and Davis and Dingel (2012) include spatial sorting of heterogeneous individuals to rationalize some of these patterns, but do not focus on optimal policies. Our results suggest that spatially efficient allocations would typically feature weaker sorting and lower wage dispersion within larger cities than what is currently observed.

Recent papers study spatial misallocation due to wedges in factor or housing markets, such as Brandt et al. (2013), Desmet and Rossi-Hansberg (2013) and Hsieh and Moretti (2017), or due to specific spatial policies such as state or federal income taxes or firm subsidies, such as Albouy (2009), Colas and Hutchinson (2017), Fajgelbaum et al. (2015) and Ossa (2018).[9] These studies

---

[6]These trade models belong to the class considered by Arkolakis et al. (2012), who show that the welfare implications of changes in trade costs depend on the elasticity of trade with respect to trade costs. Our theoretical analysis does not impose the constant-elasticity form for trade flows that is typically assumed in the gravity literature. The optimal spatial tax turns out to be independent from the trade elasticity.

[7]Some features studied in recent applications of quantitative spatial models that are not included in our study are multiple traded sectors with input-output linkages as in Caliendo et al., (forthcoming) and commuting as in Monte et al. (2015), and dynamics as in Caliendo et al. (2015) and Desmet et al., (forthcoming).

[8]Baum-Snow and Pavan (2013) document that a considerable share of the increase in wage in inequality in the U.S. is due to the increase in wage dispersion within larger cities after controlling for worker characteristics.

[9]Hsieh and Moretti (2017) inspect the role of changes in housing supply elasticities in driving changes in output

compare the actual allocation to counterfactual ones that typically remove the dispersion of wedges or policies. We focus instead on inefficiencies created by spillovers, and we implement the globally optimal spatial allocation through transfers. In addition, these models usually proceed under the assumption of a single worker type, whereas we incorporate heterogeneous workers with asymmetric spillovers, allowing us to study optimal sorting patterns. Relative to the magnitude of the welfare impact of removing dispersion in spatial wedges or policies, our results suggest significant welfare gains from moving to an optimal allocation arising from dispersion of spillover elasticities across heterogeneous workers. Gaubert (2015) studies a model without trade frictions, homogeneous workers, heterogeneous firms and a complementarity between city size and firm productivity, and characterizes the optimal allocation in that setup.[10]

Starting with Henderson (1974), a theoretical literature in urban economics studies optimal city sizes. In environments that feature homogeneous cities or one-dimensional heterogeneity, these models make predictions for whether cities are too large, and for whether there are too few cities.[11] Henderson (1977) and Helpman and Pines (1980) are early studies of optimal city sizes in models with heterogeneous cities in either amenity or productivity. Albouy et al. (2017) and Eeckhout and Guner (2017) suggest that large cities are too small under the assumptions of homogeneous workers, frictionless trade, and efficiency spillovers only. Relative to that literature, we characterize and quantify optimal spatial transfers in a framework with several dimensions of heterogeneity across cities (amenities, productivity, housing supply elasticities, and bilateral trade frictions), spillovers in both production efficiency and amenities, and sorting of heterogeneous workers with asymmetric spillovers. Due to these multiple sources of heterogeneity, and the fact that we allow for an arbitrary distribution of observed initial transfers, initial city size alone is a poor predictor of whether a city is too large or too small in the observed allocation. We abstract from analyzing the extensive margin of city creation, but we note that this margin does not affect our characterization of optimal transfers among populated locations.

While we focus on place-based policies that tax and redistribute income across regions and workers, our paper also relates to analyses of place-based policies in the form of public investments. A large empirical literature, summarized by Kline and Moretti (2014b), Neumark et al. (2015) and Duranton and Venables (2018), estimates the local impacts of policies such as government investments in public goods or infrastructure, often in the presence of agglomeration economies. For instance, Kline and Moretti (2014a) study the effect of a large regional development program. Glaeser and Gottlieb (2008) and Kline (2010) draw some of the aggregate implications of these policies. We characterize and implement first-best place-based subsidies in a quantitative spatial framework accounting for multiple indirect effects across the economy and heterogeneous spillovers across workers.[12] Finally, a literature in fiscal federalism focuses on the optimal financing

---

in a welfare. We incorporate heterogeneity in housing supply elasticities and keep them constant throughout our analysis, treating them as a technological constraint in the planner's problem.

[10]Other recent papers studying different spatial policies include Allen et al. (2015) who consider zoning restrictions within a city and Fajgelbaum and Schaal (2017) who consider transport network investments.

[11]See Abdel-Rahman and Anas (2004) for a review.

[12]In his review of the policy implications of empirical economic-geography studies, Combes (2011) notes the lack

of local public goods, showing that an efficient arrangement may involve a fiscal union with inter-governmental grants to correct distortions caused by local taxes.[13] In an extension, we show that incorporating public goods with inter-governmental transfers does not affect the optimal policies designed to deal with the spillovers.

3 Economic Geography Model with Worker Sorting and Spillovers

3.1 Environment

We start by describing the physical environment, then the competitive equilibrium and finally the planner's allocation. We consider a closed economy with a discrete number J of locations typically indexed by j or i. Workers are heterogeneous. Each worker belongs to one of Θ different types. The type indexes each worker's preference and productivity in each location, as well as each worker's capacity to generate and absorb productivity and amenity spillovers. The utility of a worker of type θ in location j is

$$u_j^\theta = a_j^\theta \left(L_j^1, .., L_j^\Theta \right) U \left(c_j^\theta, h_j^\theta \right). \tag{1}$$

The function $a_j^\theta \left(\cdot \right)$ depends on the number L_j^θ of each type-θ worker located at j, and it captures the valuation of a worker of type θ for location j's amenities. Workers may vary in how much they value fundamental amenities associated with exogenous features of each location. They may also vary in how much they value amenity spillovers created by each type; for example, a demographic group may prefer living in locations with higher density of their own demographic group, or may value some urban amenities generated or congested by specific groups. Workers also derive utility from a bundle of differentiated tradeable commodities (c_j^θ) and from non-tradeable services including housing (h_j^θ).[14] The utility function $U(c, h)$ is homogeneous of degree 1.[15]

In both the competitive and optimal equilibria we consider, the number L_j^θ of type-θ workers

---

[13]Early papers making this point include Flatters et al. (1974) and Wildasin (1980). Another concern in that literature is to analyze the responses of local governments to taxes set in other jurisdictions, as in for example Gordon (1983), and the role of a fiscal union in correcting the potential distortions from fiscal competition. We take the observed transfers as given, without taking a stand on how they are generated.

[14]There is no heterogeneity across workers within each type. However, since the number of types is unrestricted, this is not a restriction on the theory. For example, two skilled workers with different preferences for location can be represented as two different types. A standard formulation in the empirical literature is to allow for heterogeneity in preferences for location across workers distributed according to an extreme value distribution. In Section 4.4 we show how to incorporate this margin.

[15]The analysis can be easily generalized to homothetic utility functions which are not homogeneous of degree 1 at the cost of more burdensome notation. E.g., we can accommodate functions of the form $F_0 \left(U \left(c_j^\theta, h_j^\theta \right) \right)$ at the cost of keeping track of the elasticity of F with respect to U. Such a formulation would capture decreasing marginal utility from the consumption bundle of each location.

---

located at j is endogenous. National labor market clearing is

$$\sum_j L_j^\theta = L^\theta, \tag{2}$$

where L^θ is the fixed aggregate supply of group θ.

Every location produces traded and non-traded goods. Tradeable output uses the aggregate technology $Y_j\left(N_j^Y, I_j^Y\right)$ requiring services of labor N_j^Y and intermediates I_j^Y. Similarly, production in the non-traded sector is $H_j\left(N_j^H, I_j^H\right)$. The functions Y_j and H_j may be city-specific and feature constant or decreasing returns to scale. They can also be constants if the good is supplied inelastically. The decreasing returns arise due to the use of fixed factors in production, such as land. Therefore, the framework allows for different housing supply elasticities across cities through the city specific decreasing returns to scale in $H_j\left(\cdot\right)$.[16]

The feasibility constraint in the non-traded sector in j is

$$H_j\left(N_j^H, I_j^H\right) = \sum_\theta L_j^\theta h_j^\theta. \tag{3}$$

Goods in the traded sector can be shipped domestically or to other locations. The country's geography is captured by iceberg trade frictions $d_{ji} \geq 1$. These frictions mean that $d_{ji}Q_{ji}$ units must be shipped from location j to i for Q_{ji} units to arrive in location i. The feasibility constraint dictates that the quantity of traded goods produced at j must equal the quantities shipped:

$$Y_j\left(N_j^Y, I_j^Y\right) = \sum_i d_{ji}Q_{ji}. \tag{4}$$

Traded goods may be differentiated by origin, reflecting either industrial specialization at the regional level or variety specialization at the plant level.[17] The traded goods arriving in i are combined through the homothetic and concave aggregator $Q\left(Q_{1i}, .., Q_{Ji}\right)$. This bundle of traded commodities may be used for the final consumption of the local population or as an intermediate input in the production of traded and non-traded sectors:

$$Q\left(Q_{1i}, .., Q_{Ji}\right) = \sum_\theta L_i^\theta c_i^\theta + I_i^Y + I_i^H. \tag{5}$$

All workers supply one unit of labor with efficiency that may vary by worker type and location.

---

[16]I.e., if $H_j\left(N_j^H, I_j^H\right)$ is homothetic, so that it can be expressed as $H_j\left(N_j^H, I_j^H\right) = H_j^1\left(H_j^0\left(N_j^H, I_j^H\right)\right)$ for some $H_j^0\left(N_j^H, I_j^H\right)$ that is homogeneous of degree 1, then in the competitive allocation the elasticity of H_j^1 is the housing supply elasticity. To keep notations light, we do not feature land explicitly as a variable input and we assume that the fixed factors in production cannot be reallocated between the traded and non-traded sectors. This assumption can also be easily relaxed.

[17]Section 4.3 shows that our analysis encompasses a market structure with free entry of producers of differentiated varieties under monopolistic competition.

Each type-θ worker in location j supplies

$$z_j^\theta = z_j^\theta \left(L_j^1, .., L_j^\Theta \right) \tag{6}$$

efficiency units. The function z_j^θ captures exogenous differences in productivity between locations and skill groups, as well as productivity spillovers across workers. As with amenities, these spillovers may depend on the distribution of types. For example, high-skill workers may benefit more than low-skill workers from being employed in the same city as other high-skill workers, or in more densely populated areas. In both traded and non-traded sectors, the services $z_j^\theta L_j^\theta$ of the various types of labor are combined through the constant-returns aggregator $N \left(z_j^1 L_j^1, .., z_j^\Theta L_j^\Theta \right)$.[18] This aggregator also captures imperfect substitution across workers.[19] Feasibility in the use of labor services dictates that

$$N_j^Y + N_j^H = N \left(z_j^1 L_j^1, .., z_j^\Theta L_j^\Theta \right). \tag{7}$$

We highlight two key features relative to an otherwise standard neoclassical framework with a representative worker/consumer. First, the location of a worker drives both her marginal product (productivity is place specific) and her marginal utility of consumption (through local amenities). Therefore, production and consumption decisions are not separable. Second, the framework features two potential sources of non-convexities through the amenity and productivity spillover functions. The utility of each agent may increase with the number of other agents in the same location through a_j^θ, and the labor aggregator $N \left(\cdot \right)$ may feature increasing returns to the number of workers in a particular group through $z_j^\theta \left(L_j^1, .., L_j^\Theta \right) L_j^\theta$. Both sources of non-convexities operate though the same margin, L_j^θ.

3.2   Competitive Allocation

In the decentralized equilibrium each worker chooses location and consumption to maximize utility, while competitive producers hire labor and buy intermediate inputs to maximize profits. Being atomistic, these agents do not take into account the impact of their choices on the spillover functions $a_j^\theta \left(L_j^1, .., L_j^\Theta \right)$ and $z_j^\theta \left(L_j^1, .., L_j^\Theta \right)$.

Workers   Conditional on living in j, a type-θ worker with expenditure level x_j^θ solves

$$\max_{c_j^\theta, h_j^\theta} U \left(c_j^\theta, h_j^\theta \right) \quad s.t. \quad P_j c_j^\theta + R_j h_j^\theta = x_j^\theta, \tag{8}$$

---

[18]This aggregator can more generally be homothetic, in which case the expressions for optimal transfers would include an additional term with the aggregate curvature of this function. We limit the analysis to the CRS case to save notation.

[19]This formulation implies that different types of workers supply the same amount of efficiency units to every sector, ruling out comparative advantages of different types across sectors. Including this force would be feasible in both the theory and the quantification, subject to the availability of the proper production function and sector-specific spillover parameters. Our current formulation allows us to calibrate the model to match existing values of the spillover elasticities estimated at the city level with respect to each type of worker.

where P_j is the price of the bundle of traded goods and R_j is the unit price in the nontraded sector. As a result, the common component of utility from (1) becomes

$$u_j^\theta = a_j^\theta \left(L_j^1, .., L_j^\Theta \right) \frac{x_j^\theta}{\psi \left(P_j, R_j \right)}, \tag{9}$$

where $\psi \left(P, R \right)$ is the price index associated with U. Since workers are homogeneous within each type, their utility is equalized. In equilibrium, all type-θ workers attain the same utility u^θ. Workers' location choice implies that

$$u_j^\theta \le u^\theta, \tag{10}$$

with equality if $L_j^\theta > 0$.

Firms Producers of traded and non-traded commodities solve:

$$\Pi_j^Y = \max_{N_j^Y, I_j^Y} p_j Y_j \left(N_j^Y, I_j^Y \right) - W_j N_j^Y - P_j I_j^Y, \tag{11}$$

$$\Pi_j^H = \max_{N_j^H, I_j^H} R_j H_j \left(N_j^H, I_j^H \right) - W_j N_j^H - P_j I_j^H, \tag{12}$$

where p_j is the domestic price of the tradeable commodity produced in j and W_j is the wage per efficiency unit of labor. Given a distribution of wages per worker $\left\{ w_j^\theta \right\}$, the wage per efficiency unit is the solution to a standard cost-minimization problem,

$$W_j = \min_{\left\{ L_j^\theta \right\}} \sum_j w_j^\theta L_j^\theta \quad s.t. \quad N \left(z_j^1 L_j^1, .., z_j^\Theta L_j^\Theta \right) = 1.$$

This labor hiring decision takes the efficiency z_j^θ as given, ignoring its dependence on the employment distribution. As a result, the wage of type-θ workers in location j equals the value of its marginal product given the efficiency distribution $\left\{ z_j^\theta \right\}$:

$$w_j^\theta = W_j \frac{\partial N \left(z_j^1 L_j^1, .., z_j^\Theta L_j^\Theta \right)}{\partial L_j^\theta}. \tag{13}$$

We assume a no-arbitrage condition, so that the price in location i of the traded good from j equals $d_{ji} p_j$. Free entry of intermediaries who can buy and resell goods between regions ensures this condition holds. Given these prices, the trade flows are:

$$P_i \frac{\partial Q \left(Q_{1i}, .., Q_{Ji} \right)}{\partial Q_{ji}} = d_{ji} p_j. \tag{14}$$

In the competitive equilibrium the prices of consumption goods, P_j and R_j, adjust so that the corresponding goods markets clear.

Expenditure Per Worker The only component of the competitive allocation left to define is the per capita expenditure for a type-θ worker who lives in j, x_j^θ. By definition, x_j^θ includes labor and non-labor income as well as, potentially, government taxes and transfers. Each type-θ worker in location j earns the wage w_j^θ and owns a fraction b^θ of the national returns to fixed factors $\Pi \equiv \sum_j \Pi_j^Y + \Pi_j^H$, so that $\sum b^\theta L^\theta = 1$. We assume that each individual's portfolio is type- but not place-specific, meaning that individuals of a given group hold the same portfolio regardless of where they locate.[20] In addition, we allow for flexible government policies that tax and transfer income across locations. As a result, expenditure per capita is

$$x_j^\theta = w_j^\theta + b^\theta \Pi + t_j^\theta, \tag{15}$$

where t_j^θ is the net transfer to a type-θ worker living in j. Net government transfers equal zero, implying that expenditure equals net income:

$$\sum_j \sum_\theta L_j^\theta x_j^\theta = \sum_j \sum_\theta L_j^\theta w_j^\theta + \Pi. \tag{16}$$

Definition 1. *A competitive allocation consists of quantities $c_j^\theta, h_j^\theta, L_j^\theta, Q_{ij}, N_j^Y, I_j^Y, N_j^H, I_j^H$, utility levels u^θ, prices P_j, R_j, p_j, wages per efficiency unit W_j, and wages per worker w_j^θ such that*

(i) the consumption choices c_j^θ, h_j^θ are a solution to (8) for expenditures x_i^θ satisfying (15), and employment L_j^θ is consistent with the spatial mobility constraint (10);

(ii) the labor and intermediate input choices $N_j^Y, I_j^Y, N_j^H, I_j^H$ are a solution to (11) and (12), labor demand is given by (13), and trade flows Q_{ji} are given by (14);

(iii) the government budget constraint is satisfied; i.e. (16) holds, and

(iv) all markets clear, i.e. (2) to (7) hold.

3.3 Planning Problem

We characterize a planning problem where the planner chooses the endogenous margins of the competitive allocation: a distribution of workers over locations and types $\left\{ L_j^\theta \right\}$, consumption of traded and non-traded commodities $\left\{ c_j^\theta, h_j^\theta \right\}$, trade flows $\{Q_{ij}\}$, and an allocation of efficiency units and intermediate inputs, $\left\{ N_j^Y, I_j^Y, N_j^H, I_j^H \right\}$. The planner implements policies that treat in the same way all individuals within a type θ, and is bound by the spatial mobility constraint (10). Along with that constraint, the market clearing conditions (2) to (7) define a set $\mathcal{U}$ of attainable

---

[20]This assumption rules out cases considered in the trade and economic geography literature in which individuals are assumed to own part of the fixed factors where they choose to live, as in Caliendo et al., (forthcoming). These cases introduce an additional distortion in the competitive allocation, therefore creating an additional role for optimal policy besides correcting spillovers. Empirically, that assumption has been made to match trade imbalances across regions. In our model, that distribution can be fully explained by a combination of policies that tax and redistribute income across locations and sorting of heterogeneous workers types with type-specific ownership. In Section 4.3 we formally show the implications of assuming local ownership.

utility levels $\{u^\theta\}$. The optimal planning problem consists of reaching the frontier of that set,

$$\max u^\theta$$
$$\text{s.t.: } u^{\theta'} = \underline{u}^{\theta'} \quad \text{for } \theta' \neq \theta$$
$$u^{\theta'} \in \mathcal{U} \quad \text{for all } \theta'$$

The set of solutions of this problem given an arbitrary θ for all feasible values of $\underline{u}^{\theta'} \in \mathcal{U}$ for $\theta' \neq \theta$ defines the utility frontier. Competitive equilibria according to Definition 1 may not correspond to a point on the frontier due to spatial inefficiencies: workers do not internalize the impact of their location choice on amenities through a_j^θ and firms do not internalize the impact of their hiring decisions on efficiency through z_j^θ. In the next section we turn to the solution of the planning problem.

4 Optimal Transfers

To understand what policies may remedy the market inefficiencies we first characterize a necessary efficiency condition, and then use it to determine the set of transfers that must hold in any efficient allocation. We then move to inspecting the labor subsidies that would implement these transfers in well-known models nested in our framework. Furthermore, we demonstrate how to extend our approach to accommodate additional forces considered in recent studies. We conclude this section by showing how to bring the framework to the data. Formal derivations are detailed in Appendix A.1.

4.1 Efficiency Condition and Optimal Transfers

To characterize efficiency, it is useful to note that the competitive allocation can be determined given an arbitrarily chosen expenditure distribution of $x_j^\theta \geq 0$ satisfying (16), without imposing the requirement that x_j^θ satisfies its definition (15). In other words, we can modify Definition 1 to treat expenditure per capita x_j^θ as primitive, leaving aside how it is determined. Using the wages and profits arising at the resulting equilibrium allocation, we can then compute the transfers t_j^θ consistent with (15) given the arbitrarily chosen x_j^θ.

This logic suggests that we can first obtain a condition over the expenditure distribution x_j^θ that must hold in any efficient allocation, regardless of what particular policy tools are used to achieve it. Comparing a decentralized allocation given expenditures x_j^θ to the outcomes of the planning problem, detailed in Definition 2 of Appendix A.1, we reach the following result.

Proposition 1. *If a competitive equilibrium is efficient, then*

$$W_j \frac{dN_j}{dL_j^\theta} + \sum_{\theta'} \frac{x_j^{\theta'} L_j^{\theta'}}{a_j^{\theta'}} \frac{\partial a_j^{\theta'}}{\partial L_j^\theta} = x_j^\theta + E^\theta \quad \textit{if } L_j^\theta > 0, \tag{17}$$

for all j and θ and some constants $\{E^\theta\}$. If the planner's problem is globally concave and (17) holds for some specific $\{E^\theta\}$, then the competitive equilibrium is efficient.

Condition (17) defines a relationship between expenditure per capita and the labor allocation that must hold in any efficient allocation. This condition shows the equalization of the marginal benefits and costs of type-θ workers across all inhabited locations. The first term on the left is the value of the marginal product of labor, including both the direct output effect and the productivity spillovers.[21] The second term is the marginal benefit (or costs if negative) through amenity spillovers on each group of workers living in j, measured in expenditure equivalent terms.

These marginal benefits from allocating a type θ worker to location j are equated to the marginal costs on the right. The first term, x_j^θ, results from the non-separability between a worker's location and consumption: each type-θ worker in j requires x_j^θ units of expenditures in that particular location. From a social planning perspective this is a cost, because each additional worker in j translates into lower consumption of traded and non-traded commodities for other workers in that location. The last term, E^θ, measures the opportunity cost of employing a worker elsewhere, and is not place-specific.[22]

We can draw several useful implications from this result. First, it implies that market inefficiencies can be fully tackled through policies acting only on x_j^θ to reach (17). This is because the set of equations defining the competitive allocation coincides with the set defining the planner's allocation, except potentially for the expenditure distribution.[23] Therefore, asking whether the spatial allocation is efficient is equivalent to asking whether the expenditure distribution in the market allocation lines up with (17). Despite the multiple general-equilibrium ramifications of the spillovers, no other margin but the expenditure distribution requires policy intervention.[24] We construct policies that implement the optimal expenditure distribution in the next section.

Second, Proposition 1 extends a familiar efficiency condition from the misallocation literature to spatial environments. In our economy, "space" enters through trade costs, non-traded goods, congestion and amenities. In the absence of these forces,[25] there would be no compensating differentials across locations and, as a result, the equilibrium would exhibit the same expenditure per capita x_j^θ for each type θ across locations. In that case, the model would be equivalent to

---

[21]$\frac{dN_j}{dL_j^\theta}$ denotes the total differential of $N_j = N\left(z_j^1\left(L_j^1,..,L_j^\Theta\right)L_j^1,..,z_j^\Theta\left(L_j^1,..,L_j^\Theta\right)L_j^\Theta\right)$ with respect to L_j^θ.

[22]The constant E^θ is the multiplier of the national market clearing constraint (2) in the planner's problem. It is the marginal welfare impact of adding an additional type-θ worker in the economy.

[23]In the planner's allocation, x_j^θ coincides with the shadow value of the local goods assigned to each θ-worker in j. The market allocation of efficiency units $\{N_j^H, N_j^Y\}$, intermediates $\{I_j^H, I_j^Y\}$, trade $\{Q_{ij}\}$ and relative consumption per capita $\left\{\frac{c_j^\theta}{h_j^\theta}\right\}$ is efficient if the equilibrium prices $\{W_j, R_j, P_j, p_j\}$ equal the multipliers of the constraints on efficiency units, non-traded goods, traded goods, and exports faced by the planner, respectively. See the proof of Proposition 1.

[24]This compartmentalization of the inefficiencies in few margins reflects a broader "principle of targeting" noted by Bhagwati and Johnson (1960) in trade-policy contexts and by Sandmo (1975) and Dixit (1985) in economies with external effects.

[25]I.e., assuming away trade frictions ($d_{ij} = 1$, implying $P_j = P$), non-traded goods ($U\left(c_j^\theta, h_j^\theta\right) = c_j^\theta$), congestion ($\frac{\partial a_j^\theta}{\partial L_j^{\theta'}} = 0$) and preferences for location ($a_j^\theta = a_j$).

one describing the allocation of workers across firms selling differentiated products as in Hsieh and Klenow (2009), and (17) would collapse to the familiar condition that the marginal value-product of labor, $W_j\left(dN_j/dL_j^\theta\right)$, is equalized across locations.

Third, information about the distribution of expenditure per capita x_j^θ is needed to assess the economy's efficiency. In studies of misallocation across firms, the absence of compensating differentials justifies the common practice of inferring allocative inefficiencies from differences in income and number of workers. In our spatial environment with compensating differentials, the non-separability of consumption and production means that the net marginal benefit of reallocating a worker includes the local expenditure of that worker. As a result, assessing the efficiency of the allocation requires data on the distribution of expenditure per capita x_j^θ. This property motivates an efficiency check of the observed allocation derived in the next section, and it drives the data requirements for the quantitative implementation of the model.

Finally, we note that (17) is a necessary but not sufficient condition for efficiency. Even if it holds, inefficient market equilibria could exist. However, the inefficient allocations consistent with (17) can be ruled out if the planner's problem is globally concave. In that case, there is only one allocation that satisfies the first order conditions of the planner's problem. Because these first order conditions correspond to the definition of a competitive equilibrium given expenditures, it must be that only one competitive equilibrium exists and is efficient. In Section 4.5 we introduce conditions for global concavity of the planner's problem.

We move on now to deriving the set of transfers implied by the efficiency condition presented above. These transfers will deliver the efficient allocation if the market equilibrium is unique, an issue we come back to in Section 4.5. To that end, we define the productivity spillover elasticity

$$\gamma_{\theta,\theta'}^{P,j} \equiv \frac{L_j^\theta}{z_j^{\theta'}}\frac{\partial z_j^{\theta'}}{\partial L_j^\theta}, \tag{18}$$

and the amenity spillover elasticity

$$\gamma_{\theta,\theta'}^{A,j} \equiv \frac{L_j^\theta}{a_j^{\theta'}}\frac{\partial a_j^{\theta'}}{\partial L_j^\theta}. \tag{19}$$

These elasticities capture the marginal spillover of a type θ worker on the efficiency and utility of each type θ' worker in city j. The case without spillovers corresponds to $\gamma_{\theta,\theta'}^{P,j} = \gamma_{\theta,\theta'}^{A,j} = 0$. So far we have not imposed functional forms, so that these elasticities can be variable. Combining these definitions with (1) and labor demand (13), we obtain the optimal transfers.

Proposition 2. *The optimal allocation can be implemented by the transfers*

$$t_j^{\theta*} = \sum_{\theta'}\left(\gamma_{\theta,\theta'}^{P,j}w_j^{\theta'*} + \gamma_{\theta,\theta'}^{A,j}x_j^{\theta'*}\right)\frac{L_j^{\theta'*}}{L_j^{\theta*}} - \left(b^\theta\Pi^* + E^\theta\right), \tag{20}$$

where the terms $\left(x_j^{\theta*}, w_j^{\theta*}, L_j^{\theta*}, \Pi^*\right)$ are the outcomes at the efficient allocation, and $\{E^\theta\}$ are constants equal to the multipliers on the resource constraint of each type in the planner's allocation.

We highlight that the optimal transfers $t_j^{\theta*}$ take care of both inefficiencies in the market allocation due to spillovers and of distributional concerns across types. In the absence of spillovers ($\gamma_{\theta,\theta'}^{P,j} = \gamma_{\theta,\theta'}^{A,j} = 0$ for all θ, θ', j), we would still have $t_j^{\theta*} = -\left(b^\theta \Pi^* + E^\theta\right)$, so that these transfers would still be present for redistribution across types (as implied by the second welfare theorem). Therefore, the burden of dealing with the spatial inefficiencies falls on the spatial component of the optimal transfers, corresponding to the first term in (20). Along with the definition of expenditures in (15), condition (20) can be used to diagnose the efficiency of an observed allocation using only data on wages, expenditures and employment, in addition to the spillover elasticities.[26]

4.2 Optimal Subsidies under Constant Elasticity Spillovers

So far we have not imposed functional forms, but a predominant assumption in the related literature is that the spillover functions $z_j^\theta(\cdot)$ and $a_j^\theta(\cdot)$ adopt the constant-elasticity forms, $\gamma_{\theta,\theta'}^{P,j} = \gamma_{\theta,\theta'}^P$ and $\gamma_{\theta',\theta}^{A,j} = \gamma_{\theta',\theta}^A$. From now on we adopt these constant-elasticity formulations. We will also adopt them for the quantitative application (as defined in (37) and (38) below). Under these assumptions, the optimal transfers simplify to $t_j^\theta = s_j^\theta w_j + T^\theta$, where

$$s_j^\theta = \frac{\gamma_{\theta,\theta}^P + \gamma_{\theta,\theta}^A}{1 - \gamma_{\theta,\theta}^A} + \sum_{\theta' \neq \theta} \frac{\gamma_{\theta,\theta'}^P w_j^{\theta'} + \gamma_{\theta,\theta'}^A x_j^{\theta'}}{1 - \gamma_{\theta,\theta}^A} \frac{L_j^{\theta'}}{w_j^\theta L_j^\theta} \tag{21}$$

and $T^\theta = b^\theta \Pi + \frac{E^\theta}{1 - \gamma_{\theta,\theta}^A}$. This representation readily implies that the optimal transfers can be implemented by labor income subsidies $s^\theta(\cdot)$ coupled with lump-sum transfers T^θ. The labor income subsidy function $s^\theta(\cdot)$ may be type specific and vary over space as a function of outcomes such as wages or population.[27] These policies are defined such that the labor subsidies tackle spatial inefficiencies due to spillovers, and the lump-sum transfers take care of distributional concerns. We now draw the implications of this formula in special cases.

No Spillover Across Types We consider first a case with several worker types, but with $\gamma_{\theta',\theta}^P = \gamma_{\theta',\theta}^A = 0$ for $\theta' \neq \theta$, so that there are no spillovers across types. The optimal subsidy (21) becomes:

$$s^\theta = \frac{\gamma_{\theta,\theta}^P + \gamma_{\theta,\theta}^A}{1 - \gamma_{\theta,\theta}^A}. \tag{22}$$

In the special case of a single worker type, the policy is further simplified to (s, T) with $s = \frac{\gamma^P + \gamma^A}{1 - \gamma^A}$ and T which ensures that the aggregate budget condition (16) holds. This case corresponds to the

---

[26]Specifically, an observed allocation is efficient only if the the function of these variables in the left hand side of (A.14) in Appendix A.1 is constant over space for each type θ.

[27]The policy rules $\left(s^\theta(\cdot), T^\theta\right)$ are type-specific but defined independently from space. Alternatively, optimal transfers could also be implemented with type-independent but place-specific labor income subsidies $s_j(w)$ defined as step functions satisfying $s_j\left(w_j^{\theta*}\right) \equiv \frac{t_j^{\theta*}}{w_j^{\theta*}}$, except in knife-edge cases where different types earn identical wages in a given location.

majority of studies on place-based policies and quantitative economic geography reviewed in the introduction. Hence, our discussion in this section applies to that case as well.

We point out a few important implications of this result. First, the optimal policy scheme $\left(s^\theta, T^\theta\right)$ does not vary over space. So, somewhat paradoxically, optimal transfers can be implemented by a "place-based" policy that turns out to be independent from space. In spite of this, the policy does impact the spatial allocation. To see why, we must consider the role of the compensating differentials implied by the spatial dimensions (trade costs, non-traded goods, and amenities). From the mobility constraint (10), indifference across populated locations j and j' implies:

$$\frac{\psi\left(P_{j'}, R_{j'}\right)/a_{j'}^\theta\left(L_{j'}\right)}{\psi\left(P_j, R_j\right)/a_j^\theta\left(L_j\right)} = \frac{\left(1+s^\theta\right)W_{j'}z_{j'}^\theta\left(L_{j'}\right)+T^\theta+b^\theta\Pi}{\left(1+s^\theta\right)W_j z_j^\theta\left(L_j\right)+T^\theta+b^\theta\Pi}. \tag{23}$$

The left hand side is the relative compensating differential (amenity-adjusted cost of living) and the right hand side is the relative expenditure (equal to relative after-tax income) between locations j' and j for type θ. In the presence of amenities, non-traded goods and trade costs, the relative compensating differentials vary across space. As a result, changes to the policy scheme (s^θ, T^θ) lead to changes in the employment distribution of type θ. In the absence of these compensating differentials, these policies would cease to impact the spatial allocation of workers of type θ.

Second, the formula (22) for the optimal subsidy has a simple interpretation. Under negative congestion spillovers for type θ ($\gamma_{\theta,\theta}^A < 0$), if the agglomeration spillover of that type is not too strong ($\gamma_{\theta,\theta}^P < -\gamma_{\theta,\theta}^A$), then all workers of type θ should pay as tax the same fraction of their income everywhere (a negative subsidy, $s^\theta < 0$). In this case, the net transfer t_j^θ received by type-θ workers is smaller, and potentially negative, in cities where the wage of these workers is higher.

Third, the decentralized allocation without intervention is not efficient (i.e., optimal transfers are non-zero) even though the spillover elasticities are constant. This would be true even if workers were homogeneous, in which case there would exist aggregate gains from implementing the policy scheme (s, T) relative to a *laissez-faire* allocation. This result runs a against the notion, suggested by several studies of placed-based policies, that there can be no gains from policy interventions that reallocate workers across space, unless externalities are different across cities. This notion is present for instance in Glaeser and Gottlieb (2008), Kline and Moretti (2014a) and the review of the literature on place-based policies by Neumark et al. (2015). Intuitively, when the spillover elasticities are the same everywhere, externalities in locations that grow following the implementation of a policy could be expected to be exactly offset by decreases in locations that shrink. However, our previous results show that this intuition does not hold up.

We emphasize that compensating differentials due to spatial forces are the key margin leading to inefficiencies when the spillover elasticities are the same everywhere. To highlight this property, in Appendix A.2 we inspect an even more restricted case with homogeneous workers, no trade costs, no intermediate inputs in production, exogenous amenities and no valuation for non-traded goods. In this simpler case, differences in amenities are the only source of compensating differentials, the *laissez-faire* allocation is generically inefficient and there are gains from implementing the

16

optimal subsidy-cum-lump sum transfer scheme. However, if differences in amenities across cities are eliminated, the policies (s, T) do not distort the allocation and the *laissez-faire* allocation turns out to be efficient.

Spillovers Across Types Spillovers across types create a rationale for place-specific labor subsidies for each labor type. To see how, consider a polar case without any amenity spillover ($\gamma_{\theta,\theta'}^A = 0$ for all θ, θ') and without efficiency spillover on one's own type ($\gamma_{\theta,\theta}^P = 0$). Assume, furthermore, that there are only two types, $\theta = U, S$. Then, the optimal t subsidy to type-θ workers located in j simplifies to $s_j^\theta = s^\theta \left(\left\{ w_j^\theta L_j^\theta \right\}_{\theta=U,S} \right)$, where

$$s^\theta \left(\left\{ w^\theta L^\theta \right\}_{\theta=U,S} \right) = \gamma_{\theta,\theta'}^P \left(\frac{w^{\theta'} L^{\theta'}}{w^\theta L^\theta} \right). \tag{24}$$

In this special case, the optimal subsidy for workers in group θ varies across locations according to the distribution of relative wage bills, $w_j^\theta L_j^\theta$. A positive efficiency spillover of type θ on type θ' ($\gamma_{\theta,\theta'}^P > 0$) implies a higher marginal gain from attracting type-θ workers to locations where the economic size of the population of type θ' workers is relatively larger. The result is a higher optimal subsidy for θ-workers where they are more scarce. Relative to a *laissez-faire* equilibrium, this policy is likely to temper the degree of sorting across cities.[28] Condition (24) also implies $\frac{ds_j^S}{ds_j^U} < 0 \longleftrightarrow \gamma_{S,U}^P \gamma_{U,S}^P > 0$, so that subsidies of high and low skill workers are negatively correlated across cities if both types generate positive efficiency spillovers. In Section 6 we will return to these basic intuitions, as they will help us rationalize the quantitative findings about the spatial efficiency of the current transfer scheme across cities and workers with different levels of skill in the U.S. economy.

4.3 Economic Geography Frameworks and Monopolistic Competition

The environment laid out in Section 3 includes standard economic geography models, such as Helpman (1998), Allen and Arkolakis (2014) and Redding (2016) including either perfect competition or monopolistic competition with free entry. Our presentation so far has assumed that each location sells a different product under perfect competition. In Appendix A.3 we show that the normative implications would be the same assuming free entry of producers of differentiated varieties under monopolistic competition as in the standard Krugman (1980) model.[29]

These models are the backbone of a growing body of quantitative research studying the spatial implications of regional shocks such as infrastructure investments and regional productivity changes,

---

[28]For example, if only skilled workers generate positive productivity spillovers on the other type, with a CES aggregator of skills and $\gamma_{S,U}^P < 1$ then the optimal subsidy to high skill workers is higher where the population of low skill workers is larger.

[29]In particular, under a single worker type, the normative implications of a differentiated-products model with elasticity of substitution $\kappa > 1$ across varieties and no productivity spillovers are equivalent to those of a homogeneous-product model with spillover elasticity equal to $\gamma^P = \frac{1}{\kappa-1}$. Therefore, our normative analysis encompasses the models in Helpman (1998) and Redding (2016), who assume free entry of monopolistically competitive firms.

summarized by Redding and Turner (2015) and Redding and Rossi-Hansberg (2017). However, to the best of our knowledge, their normative implications have barely been explored. We now apply the previous results to shed light on the optimal transfers and policies in these models.

To specialize the physical environment to these models we need to impose a number of assumptions. First, we assume a single worker type with Cobb-Douglas preferences with weight α_C on the traded goods and with a constant amenity spillover elasticity γ^A. The common component of utility (1) then becomes $u_j = A_j L_j^{1+\gamma^A} c_j^{\alpha_C} h_j^{1-\alpha_C}$. Production only uses labor and the efficiency spillover has a constant elasticity γ^P, so that tradeable output in region j is $Y_j = Z_j L_j^{1+\gamma^P}$. Supply of non-traded goods in location j is inelastic and equal to H_j. In a competitive allocation, workers in j receive a gross wage w_j equal to tradeable output per worker. We still remain more general than in the previous economic geography models in that the trade-flows aggregator $Q\left(Q_{1i}, .., Q_{Ji}\right)$ is not restricted to the constant-elasticity of substitution (CES) form. Therefore, we do not impose a constant trade elasticity.

Applying these functional form assumptions to Proposition 1, we find that the expenditure of workers in j in the optimal allocation must be:

$$x_j = (1 - \eta)w_j + \eta\bar{w}, \tag{25}$$

where $\bar{w}$ is the average wage across workers in the economy and $\eta \equiv 1 - \frac{\alpha_C\left(1+\gamma^P\right)}{1-\gamma^A}$ is a constant combining the spillover elasticities and the expenditure share in traded goods. So, we find that a simple linear relationship between expenditure and wages implements the efficient allocation. The corresponding optimal transfers are simply linear in wages: $t_j = \eta\left(\bar{w} - w_j\right)$.[30] Micro heterogeneity in amenities, productivity, and trade costs does not affect the shape of this relationship. Moreover, barring knife-edge cases on the parameters ($\eta = 0$) or the fundamentals (such that $w_j = w$), the efficient allocation generically features trade imbalances equal to $L_j t_j$. In particular, under the empirically consistent case of $\eta < 0$, efficiency requires net trade deficits in high-wage regions.

Should the optimal policy that implements (25) redistribute towards or away from high-wage locations? The answer depends on the distribution of income in the market allocation, absent policy intervention. This distribution is shaped by how the returns to the fixed factor H_j are distributed across workers in different locations. The existing applications in this literature make different assumptions in this regard, ranging from no transfers to equal ownership. To nest these cases we can assume, as in Caliendo et al., (forthcoming), that a fraction ω of the returns to H_j is distributed locally to the L_j workers in j and the remainder is evenly split across all workers.[31] Then, the optimal policy can again be expressed as a constant labor subsidy s coupled with a lump-sum

---

[30]Net transfers equal the difference between value added and expenditure per capita of a region, which in turn equal the difference between value added and expenditure in the tradeable sector.

[31]After imposing these assumptions on the framework laid out in Section 3, and further imposing that the aggregator $Q\left(Q_{1i}, .., Q_{Ji}\right)$ is CES, the model becomes equivalent to Helpman (1998) when $\gamma^A = \omega = 0 < \gamma^P$, to Allen and Arkolakis (2014) when $\alpha_C = 1$, and to Redding (2016) when $\gamma^A = 0 < \gamma^P$ and $\omega = 1$. The pattern of ownership assumed in (15) in our benchmark model is equivalent in this context to $\omega = 0$.

payment T. The optimal subsidy is common across locations and equal to

$$s = \frac{1 + \gamma^P}{1 - \gamma^A} \left[1 - (1 - \alpha_C)\, \omega \right] - 1, \tag{26}$$

with lump-sum transfer of $T = -s\bar{w}$. The optimal policy redistributes income away from low-wage regions under a labor subsidy ($s > 0$), and into low-wage regions under a labor tax ($s < 0$).

We can now identify the optimal policy in different models. We consider again the case with $\eta < 0$. In Helpman (1998), who assumes common ownership of the national portfolio ($\omega = 0$), the default transfers to low-income regions in the competitive equilibrium are too large from the perspective of spatial efficiency; therefore, the optimal policy redistributes income to regions with above-average wage ($s > 0$). In contrast, in Allen and Arkolakis (2014) and Redding (2016), who assume away trade imbalances, the default transfers to low-income regions in the competitive equilibrium are too small; as result, the optimal policy redistributes income to low-wage regions ($s < 0$).

In sum, in this application of our more general model, we have shown that several details of the microeconomic structure and the country's economic geography (represented by bilateral trade costs) do not impact the relationship between optimal trade imbalances and wages in economic geography models, nor the policies that implement them. We have also shown how the specific assumptions about ownership of fixed factors, and therefore about trade imbalances, in the *laissez-faire* allocation of these models impact whether the optimal policies should redistribute income towards or away from high-wage regions.

4.4 Extensions

Spillovers Across Locations   So far we assumed that efficiency and amenities are a function of the distribution of workers within each location. In the spirit of Lucas and Rossi-Hansberg (2002) and Rossi-Hansberg (2005), recent quantitative studies in urban economics, such as Ahlfeldt et al. (2015), and in economic geography, such as Desmet and Rossi-Hansberg (2014), emphasize spillovers at different spatial scales, such that economic activity in one location may generate spillovers in other locations. We now derive the optimal transfers in this case. To simplify the exposition, we consider a special case of our model with homogeneous workers and constant-elasticity spillovers in amenities. However, we now extend our model to allow for the efficiency of location j to be an arbitrary function of the number of workers in every location: $z_j = z_j\left(\{L_{j'}\}\right)$. This formulation accommodates a commonly used specification where spillovers decay with distance between spatial units.[32] We define the efficiency spillover elasticity across locations,

$$\gamma^{P,j,j'} = \frac{\partial z_{j'}}{\partial L_j} \frac{L_j}{z_{j'}}, \tag{27}$$

---

[32]For instance, this nests the formulation in Ahlfeldt et al. (2015) where $z_{j'} = \left(\sum_j L_j e^{-\delta t_{jj'}}\right)^{\alpha}$ where $t_{jj'}$ is travel time between j and j' and δ is a decay parameter.

as the elasticity of the efficiency of workers at j' with respect to the number of workers located in j. Following similar steps to propositions 1 and 2, the optimal transfers now are:

$$t_j = \frac{\gamma^{P,j,j} + \gamma^A}{1 - \gamma^A} w_j + \sum_{j' \neq j} \frac{\gamma^{P,j,j'}}{1 - \gamma^A} \frac{L_{j'} w_{j'}}{L_j} + T, \tag{28}$$

where T is a constant that ensures goods market clearing. Hence we find that, as before, the optimal transfers can be characterized as a function of spillover elasticities and outcomes such as wages and employment, regardless of micro heterogeneity in fundamentals. In particular, non-localized spillovers lead to the intuitive implication that the optimal transfers should be higher in places that generate strong spillovers to larger locations, as measured by their total wage bill.

Preference Draws within Types A standard assumption in the urban economics literature is that workers differ in their individual tastes for location. Our theory accommodates this margin, as θ may index preferences. For example, two workers who are identical in every dimension except for their preference for location can be represented as belonging to two different θ's. However, for broad definitions of the types (e.g., if θ defines high and low skill workers), preference heterogeneity within types may be relevant.

This margin can be incorporated into our analysis. Letting $\mathcal{L}_\theta$ index the set of type-θ workers, the utility of a worker $l \in \mathcal{L}_\theta$ is now $u_j^\theta \epsilon_j^l$, for u_j^θ defined in (1). Now, ϵ_j^l captures the worker's idiosyncratic preference for location j. We make the standard assumption that the preference draw ϵ_j^l of a type-θ worker is i.i.d. across workers and locations according to a Fréchet distribution, $\Pr\left(\epsilon_j^l < x\right) = e^{-x^{-1/\sigma_\theta}}$. The parameter $\sigma_\theta \in [0, 1]$ indexes workers' immobility. The individual preference draws are eliminated when $\sigma_\theta = 0$, in which case we return to the original formulation of the model. Under this assumption, every aspect of the model environment and competitive allocation described in sections 3.1 and 3.2 remains unchanged except for the spatial mobility constraint (10), which is is now replaced with the following labor-supply equation:

$$\frac{L_j^\theta}{L^\theta} = \left(\frac{u_j^\theta}{u^\theta}\right)^{1/\sigma_\theta}. \tag{29}$$

We can compute the optimal allocation and construct optimal policies using the same definition of the planner's problem as in 3.3, replacing constraint (10) with (29). In this formulation, the planner still conditions outcomes on location j and type θ, but not on the individual preference draws ε_j^θ. I.e., the planner's objective function weighs symmetrically every individual within a group, without knowledge of the draw received by each particular individual. Propositions 1 to 4 then go through with minor modifications. In particular, the optimal transfers shown in Proposition 2 take the same form as before, with only one difference: instead of $\gamma_{\theta,\theta}^{A,j}$, the relevant amenity spillover elasticity on the own type becomes $\tilde{\gamma}_{\theta,\theta}^{A,j} \equiv \gamma_{\theta,\theta}^{A,j} - \sigma_\theta$.

Preference draws that are unobserved by the planner create a reason for intervention. Specif-

20

ically, even without spillovers ($\gamma^{A,j}_{\theta',\theta} = \gamma^{P,j}_{\theta',\theta} = 0$), this force leads to a labor tax $s^\theta = -\frac{\sigma_\theta}{1+\sigma_\theta} < 0$ coupled with a lump-sum subsidy $T^\theta > 0$, resulting in redistribution towards low-wage locations. Hence, under $\sigma_\theta > 0$, a utilitarian planner has incentives to redistribute income across cities for reasons other than spillovers. These incentives arise because different individuals within a group θ receive the same planner's weight, but vary in their marginal utilities from expenditures in traded and non-traded commodities depending on where they sort.[33] In particular, individuals have higher draws on average conditional on sorting into lower-wage locations, creating the incentive to redistribute income towards low-wage locations.

Public Spending So far we have focused the spatial policies on pure transfers. However, place-based policies often take the form of funding of government spending in public goods that are valued by consumers and firms. Our previous results can be easily generalized to this case. Relative to the previous framework we can allow for government spending valued by workers (G^U_j) to enter as an additional shifter in the amenity valuation for location j, $a^\theta_j = a^\theta_j \left(G^U_j, L^1_j, .., L^\Theta_j \right)$, and government spending in public goods that impact productivity (G^Y_j) to enter as an additional shifter in the efficiency units of labor, $N_j = N_j \left(G^Y_j, z^1_j L^1_j, .., z^\Theta_j L^\Theta_j \right)$. We can also allow for a production function of public goods $G_j = G_j \left(I^G_j, H^G_j \right)$ that takes as inputs the bundle of traded commodities (I^G_j) and non-traded services (H^G_j). This formulation encompasses standard representations of public spending, including classic studies such as Flatters et al. (1974) and recent quantitative analyses such as Fajgelbaum et al. (2015).

Following the same steps as in Section 3.2, we define a decentralized allocation under the assumption that, for any given G_j, the government uses inputs efficiently. Focusing for simplicity on the case with a single type of worker, the efficiency condition (17) here becomes:

$$x_j = \left[\left(1 + \gamma^{A,G} \right) \frac{1 + \gamma^P}{1 + \gamma^A} + \frac{\varepsilon_{N,G}}{\varepsilon_{N,L}} \right] w_j - \frac{1 + \gamma^{A,G}}{1 + \gamma^A} E, \tag{30}$$

where $x_j \equiv x^P_j + x^G_j$ is the sum of public and private expenditure per capita. In addition to the standard spillover elasticities, (30) includes the elasticities of labor efficiency N_j with respect to government spending ($\varepsilon_{N,G}$) and the services of labor ($\varepsilon_{N,L}$), as well as the elasticity of amenities with respect to spending ($\gamma^{A,G}$), all of which are here assumed to be constant. Therefore, here too we obtain that a simple linear relationship between total expenditure per capita and wages must hold in an efficient allocation. In addition to this necessary condition, efficiency now also requires an optimal breakdown of total expenditure into its private and public components:

$$x^G_j = \gamma^{A,G} x^P_j + \frac{\varepsilon_{N,G}}{\varepsilon_{N,L}} w_j. \tag{31}$$

---

[33]More precisely, the conditional distributions of ϵ^l_j are not the same across locations after sorting: $\Pr \left[\epsilon^l_j < x \mid \mathrm{j=arg\,max}_{j'} u^\theta_{j'} \epsilon^l_{j'} \right]$ will vary by j. Shourideh and Hosseini (2018) study optimal redistribution when workers have idiosyncratic utility shocks for working in different sectors in a small open economy.

The efficiency conditions (30) and (31) can be implemented as before through a policy scheme of the form (s, T), now coupled with inter-governmental transfers. Importantly, under any such scheme, the optimal subsidy s remains the same as in the absence of public spending, as implied by (22). The investment in public goods does not modify the planner's incentives to assign workers according to the trade-off between marginal product and private expenditure cost. As a result, the same labor subsidy s derived before targets the spillovers. Simultaneously, the optimal government expenditure in j is attained by an inter-governmental grant equal to $x_j^G L_j$. Hence, as long as transfers across regions are flexible, bringing in government spending to the picture does not change the optimal spatial policy designed to deal with the spillovers.

4.5 Quantitative Implementation

We now return to the main model and show how to bring it to the data. First, we impose the functional-form assumptions that will be used in the quantitative implementation. Second, under those functional forms, we identify conditions that guarantee concavity of the planner's problem and, in the light of Proposition 1, uniqueness of the competitive allocation under the optimal spatial policies. Then, we identify a set of data that suffices to identify the fundamentals of the model and compute the allocation under the optimal spatial policies.

Functional Forms On the demand side, we assume that preferences for traded and non-traded goods are Cobb-Douglas:

$$U(c, h) = c^{\alpha_C} h^{1-\alpha_C}, \tag{32}$$

while the aggregator of traded commodities is CES,

$$Q(Q_{1i}, .., Q_{Ji}) = \left(\sum_i Q_{ji}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{33}$$

where $\sigma > 0$ is the elasticity of substitution across products from different origins.

On the supply side, the production functions of traded and non-traded goods are Cobb-Douglas in labor and intermediate inputs:

$$Y_j\left(N_j^Y, I_j^Y\right) = z_j^Y \left(N_j^Y\right)^{1-b_{Y,j}^I} \left(I_j^Y\right)^{b_{Y,j}^I}, \tag{34}$$

$$H_j\left(N_j^H, I_j^H\right) = z_j^H \left(\left(N_j^H\right)^{1-b_{H,j}^I} \left(I_j^H\right)^{b_{H,j}^I} \right)^{\frac{1}{1+d_{H,j}}}, \tag{35}$$

where $d_{H,j} \geq 0$ and $\left\{ z_j^Y, z_j^H \right\}$ are TFP shifters. We note that traded goods are produced under constant returns to scale, but we allow for decreasing returns in the housing sector. In particular, the coefficient $d_{H,j}$ plays the role of the inverse housing supply elasticity of location j, and is allowed to vary across regions.[34]

---

[34] I.e., optimization of producers of non-traded goods implies that housing supply is $\ln H_j \propto \frac{1}{d_{H,j}} \ln R_j$.

The aggregator of labor types is CES,

$$N_j \left(z_j^1 L_j^1 .., z_j^\Theta L_j^\Theta \right) = \left[\sum_\theta \left(z_j^\theta L_j^\theta \right)^\rho \right]^{\frac{1}{\rho}}, \tag{36}$$

where $\frac{1}{1-\rho} > 0$ is the elasticity of substitution across types of workers.

Finally, we impose constant-elasticity forms for the spillovers:

$$z_j^\theta \left(L_j^1, .., L_j^\Theta \right) = Z_j^\theta \prod_{\theta'} \left(L_j^{\theta'} \right)^{\gamma_{\theta',\theta}^P}, \tag{37}$$

$$a_j^\theta \left(L_j^1, .., L_j^\Theta \right) \equiv A_j^\theta \prod_{\theta'} \left(L_j^{\theta'} \right)^{\gamma_{\theta',\theta}^A}. \tag{38}$$

These functional forms are consistent with studies that estimate spillover elasticities, allowing us to draw from existing estimates. The Z_j^θ capture exogenous comparative advantages in production across types and A_j^θ capture preferences for location across types. We refer to $\left\{ Z_j^\theta, A_j^\theta \right\}$ as fundamental components of productivity or amenities in location j.

Concavity Condition To ease the notation, we introduce the following composite elasticities of efficiency and congestion spillovers:

$$\Gamma^P = \max_\theta \left\{ \sum_{\theta'} \gamma_{\theta',\theta}^P \right\},$$

$$\Gamma^A = \min_\theta \left\{ - \sum_{\theta'} \gamma_{\theta',\theta}^A \right\}.$$

Also, we let $D = \min_j \{d_{H,j}\}$ be the lowest inverse elasticity of housing supply. Under the functional form assumptions (32) to (38) we have the following property.

Proposition 3. *The planning problem is concave if*

$$\Gamma^A > \Gamma^P, \tag{39}$$

$\Gamma^A \geq 0$ *and* $\gamma_{\theta,\theta'}^A > 0$ *for* $\theta \neq \theta'$. *Under a single worker type* ($\Theta = 1$), *the planning problem is quasi-concave if:*

$$1 - \gamma^A > \left(1 + \gamma^P \right) \left(\frac{1 - \alpha_C}{1 + D} + \alpha_C \right). \tag{40}$$

Condition (39) ensures that congestion forces are at least as large as agglomeration forces. Specifically, the congestion from the type that generates the weakest congestion, measured by Γ^A, dominates the agglomeration from the type that generates the strongest agglomeration, measured by Γ^P. Condition (39) may also apply when there is a single type, in which case it simplifies to $-\gamma^A > \gamma^P$. With a single type, further assuming Cobb-Douglas preferences over traded and non-

23

traded goods we obtain the weaker restriction (40). These cases apply to the economic-geography models we reviewed in the introduction, which typically assume a single worker type.[35]

Proposition 3 establishes conditions under which the market allocation is unique given the optimal spatial policies. It extends existing uniqueness results in two dimensions. First, we complement results that characterize uniqueness of the spatial equilibrium under no policy intervention and trade balance (Allen et al., 2014). Second, our result holds in a context with heterogeneous workers and cross-groups spillovers. We note that our uniqueness condition applies under the optimal expenditure distribution. Multiplicity is still possible under sub-optimal policies or no policy intervention, but this poses no limitation for our approach.

Implementation in Changes and Data Requirements We bring the model to the data by assuming that the observed allocation is generated by a decentralized equilibrium consistent with Definition 1 under the functional form assumptions (32) to (38). To compute the optimal allocation, we solve for a planner's problem that optimizes over the the expenditure of each type of worker relative to the observed equilibrium, $\hat{x}_j^\theta$, as well as over the remaining endogenous variables $\left\{\hat{P}_i, \hat{p}_i, \hat{Y}_i, \hat{W}_i, \hat{N}_j, \hat{L}_j^\theta, \hat{R}_i, \hat{u}^\theta\right\}$ relative to the observed equilibrium.[36] The planner controls these variables to maximize the welfare of one arbitrarily chosen group, $\hat{u}^\theta$, relative to the observed allocation, subject to lower bounds for the welfare changes of the remaining groups. We can then trace the utility frontier by solving the problem under different lower bounds for the remaining groups.

The solution to this problem in changes is equivalent to the solution to the problem in levels from Section (3.3), assuming that the same parameters and fundamentals underlie the observed equilibrium and the planning problem in levels. In the spirit of Dekle et al. (2008), as shown in Section A.7, this strategy allows us to identify the data that is sufficient to back out fundamentals and solve for the optimal allocation, as stated in the following proposition.

Proposition 4. *Assume that the observed data is generated by a competitive equilibrium consistent with Definition 1 under the functional forms* (32) *to* (38)*. Then, relative to the initial equilibrium, the optimal allocation can be fully characterized as function of:*

i) the distributions of wages, employment and expenditures across labor types and locations;

ii) the distribution of bilateral import and export shares across locations;

iii) the utility and production function parameters $\left\{\alpha_C, \sigma, \rho, b_{Y,j}^I, b_{H,j}^I, d_{H,j}\right\}$; and

iv) the spillover elasticities $\left\{\gamma_{\theta',\theta}^A, \gamma_{\theta',\theta}^P\right\}$.

---

[35]The CES restriction (33) on the aggregator of trade flows $Q(\cdot)$ is not needed for any of these results. Therefore, these condition holds regardless of product differentiation across locations. We also note that these conditions are sufficient but not necessary for uniqueness.

[36]As shown in Appendix A.7, given an arbitrary change in expenditure per capita $\left\{\hat{x}_j^\theta\right\}$, an equilibrium in changes relative to the observed allocation consists of $\left\{\hat{P}_i, \hat{p}_i, \hat{Y}_i, \hat{W}_i, \hat{N}_j, \hat{L}_j^\theta, \hat{R}_i, \hat{u}^\theta\right\}$ such that the system of equations (A.51) to (A.60) hold.

The proposition establishes data requirements that are sufficient, in the light of our model, to characterize the optimal allocation. In particular, it is sufficient to observe the data in i) and ii) and the elasticities in iii) and iv).

Importantly, this implementation does not impose restrictions on the distributional policies across locations in the observed equilibrium. These policies, entering the model through the transfers t_j^θ defined in (15), manifest themselves empirically through the expenditure distribution x_j^θ. As a result, we do not impose that the observed allocation is inefficient: the efficiency of the observed allocation depends on whether the distribution of expenditures lines up with condition (17) in Proposition 1. It could be that the empirical relationship between expenditures, wages and employment is not far from that relationship, in which case our implementation of the planner's problem would predict small welfare gains from implementing optimal policies.

We also highlight that, by following this approach, we ensure that the model is initially parametrized in a way that exactly matches the observed outcomes enumerated in items i) and ii) of the proposition. In particular, the fundamentals $\left\{Z_j^\theta, A_j^\theta\right\}$, TFP shifters $\left\{z_j^Y, z_j^H\right\}$, and bilateral trade costs $\{d_{ij}\}$ are chosen such that the equilibrium from Definition 1 generates these outcomes.

5   Data and Calibration

To take the model to the data, we use as an empirical setting the distribution of economic activity across Metropolitan Statistical Areas (MSAs) in the United States in the year 2007. We focus on the spatial allocation of two skill groups, high skill (college) and low skill (non college) workers. In section 5.1 we first describe the data sources used to fulfill the data requirements in points i) and ii) of Proposition 4. Then, in Section 5.2 we describe the strategy to set the parameters required in points iii) and iv) of the proposition. The broad steps are described here while more details are given in Appendix B. We conclude with basic stylized facts that set the stage for the quantitative analysis in the next section.

5.1   Data

As established in point i) of Proposition 4, we need data on income and expenditures by group and MSA. To that end, we rely on the BEA's Regional Accounts, which report labor income, capital income and welfare transfers by MSA. A complementary BEA dataset for the years 2000 to 2007 reports total taxes paid by individuals and MSA (Dunbar, 2009). Taken together, these sources give us a dataset at the MSA level. We then apportion each of these MSA-level totals into two labor groups: high skill, defined as workers who have completed at least four years of college, and low skill, defined as every other working age individual. To implement this apportionment, we use shares of labor income, capital income, taxes and transfers corresponding to each group in each MSA from the March supplement of the Current Population Survey (IPUMS-CPS, Flood et al., 2017) collected by the Census. Since the March CPS supplement is not available for the smallest

MSAs, our dataset covers 245 MSAs in the continental US.[37]

An important concern when measuring these variables is that the model does not include heterogeneity across individuals within each group of skill θ, whereas in reality these groups are heterogeneous across cities. If we did not control for this heterogeneity, our procedure to implement the model would interpret the observed variation in net individual transfers across MSAs within a group as place-based transfers, when they reflect, in part, differences in the types of workers within each group across MSAs. In principle, this concern can be mitigated by allowing for several θ groups corresponding to the fine individual characteristics observed in the CPS. While potentially feasible, such an approach would increase the dimension of the problem and the number of elasticities to calibrate. Alternatively, we choose to purge the observed measures of income, expenditure, taxes and transfers by skill and MSA from compositional effects using a set of socio-demographic controls at the MSA-group level built from individual level Census data (IPUMS) on age, educational attainment, sector of activity, race, and labor force participation status of individuals in a given MSA-group. In the quantification we then use measures of income, expenditures, taxes and transfers that are net of variation in socio-demographic composition within groups across MSAs.[38] We discuss the details of this step in Appendix B.

We use the variables above to construct expenditure per capita, x_i^θ, using its definition (15) as labor plus capital income net of taxes and transfers, which also corresponds to the BEA's definition of disposable income. In the model we assume no variation in capital income across cities for each type. Therefore, we use a group-specific measure of capital income consistent with the fact that 59% of non-labor income is owned by high skill workers according to the BEA/CPS data.[39]

As implied by ii) of Proposition 4, quantifying the model also requires data on trade flows between MSAs. The Commodity Flow Survey (CFS) reports the flow of manufacturing goods shipped between CFS zones in the US every five years. The CFS zones correspond to larger geographic units than our unit of observation, the MSA. To overcome this data limitation, we adapt the approach in Allen and Arkolakis (2014) and Monte et al. (2015), who use estimates of trade frictions as function of geography to project CFS-level flows to the MSA level. In our context, we use the gravity equation predicted by the model to find the unique estimates of trade flows between MSAs that are consistent with actual distance between MSAs, existing estimates of trade frictions with respect to distance, and observed trade imbalances, computed as the difference between income in the traded sector and expenditure on traded goods (for both final and intermediate use) in each MSA.

Finally, to calibrate the labor shares in production in part iii) of Proposition 4, we use CPS data on employment in traded and non-traded sectors by MSA.[40] We also adjust this measure to

---

[37]These areas correspond to 97% of the population and 98% of income of all US urban areas. Urban areas in the US in turn cover 83% of the population, and 87% of personal income.

[38]Following this procedure, we have a dataset of 225 MSAs for which information on socio-demographic characteristic is available.

[39]This step involves setting a national share of profits in GDP consistent with the general equilibrium of the model. See Appendix B for details.

[40]We define employment in the following NAICS sectors as corresponding to the non-traded sector in the model:

26

remove variation from compositional effects following a similar approach to the one described above for income, expenditure, taxes and transfers.

5.2 Calibration with Heterogeneous Workers

Our model is consistent with Diamond (2016) and generates similar estimating equations to those used in her analysis. We use the same definition of geographic units (MSA) and skill groups (College and Non College), and we rely on similar data sources for quantification. Therefore, her estimates constitute a natural benchmark to parametrize the model. In what follows, we discuss these elasticities and several alternative specifications that are also used in the quantitative section.

Utility and Production Function Parameters $\left\{\alpha_C, \sigma, \rho, b_{Y,j}^I, b_{H,j}^I, d_{H,j}\right\}$ We use the Diamond (2016) estimate of the Cobb-Douglas share of traded goods in expenditure ($\alpha_C = 0.38$), of the inverse housing supply elasticity ($d_{H,j}$ in (35)) for each MSA,[41] and of the elasticity of substitution between high and low skill, estimated at 1.6 and implying $\rho = 0.392$.

We calibrate the Cobb-Douglas share of intermediates in traded good production ($b_{Y,j}^I = 0.468$ for all j in (34)) using the share of material intermediates in all private good industries production in 2007 from the U.S. KLEMS data. Having calibrated the previous parameters, the Cobb-Douglas share of labor in non-traded production in each city ($1 - b_{H,j}$ in (35)) can be chosen to match the share of workers in the non-traded sector of each MSA, as detailed in Section B.2. We assume an elasticity of substitution σ among traded varieties in (33) equal to 5, corresponding to a central value of the estimates reported by Head and Mayer (2014).

Efficiency Spillovers $\left\{\gamma_{\theta',\theta}^P\right\}$ Previous empirical studies, such as Ciccone and Hall (1996), Combes et al. (2008), and Kline and Moretti (2014a), estimate elasticities of labor productivity with respect to employment density. Across specifications, these studies find elasticities in the range of $(0.02, 0.2)$.[42] Hence, we set a a properly weighted average of the elasticities $\gamma_{\theta',\theta}^P$, corresponding to what the empirical specifications of these previous studies would recover in data generated by our model, to match the benchmark value for the U.S. economy of 0.06 from Ciccone and Hall (1996). In addition, Diamond (2016) estimates an elasticity of MSA wages with respect to population by skill group. As detailed in Appendix B.2, under the previous normalization, these estimates can be mapped to the relative values of our $\gamma_{\theta,\theta'}^P$ parameters using the wage equation (13) and the elasticity of substitution between skilled and unskilled workers ρ.

This approach has the advantage of preserving an aggregate elasticity of labor productivity with respect to density that is consistent with standard estimates from the literature, as well as heterogeneity in spillover elasticities across skill types that is consistent with Diamond (2016). As a result we obtain $\left(\gamma_{UU}^P, \gamma_{SU}^P, \gamma_{US}^P, \gamma_{SS}^P\right) = (.003, .044, .02, .053)$. These parameters imply stronger

---

retail, real estate, construction, education, health, entertainment, hotels and restaurants.

[41]For MSAs that we cannot match to Diamond (2016) we use the average housing supply elasticity across MSAs.

[42]Most of the related studies reviewed by Combes and Gobillon (2015) and Melo et al. (2009) also fall in this range.

efficiency spillovers generated by high skill workers, and close to zero spillovers from low skill workers.

Amenity Spillovers $\left\{ \gamma_{\theta',\theta}^{A} \right\}$ Diamond (2016) estimates elasticities of labor supply by skill group with respect to real wages and with respect to an amenity index that summarizes forces such as congestion in transport, crime, environmental indicators, supply per capita of different public services, and variety of retail stores. She estimates a positive elasticity for the the supply of this MSA level amenity index with respect to the relative supply of college workers, as well as a higher marginal valuation for these amenities for college than for non-college workers. As detailed in Appendix B.2, we can directly map her estimates to both the level and distribution of our amenity spillovers $\gamma_{\theta',\theta}^{A}$ using the labor-supply equation implied by the spatial mobility constraint (10). As a result we obtain $\left(\gamma_{UU}^{A}, \gamma_{SU}^{A}, \gamma_{US}^{A}, \gamma_{SS}^{A} \right) = (-.43, .18, -1.24, .77)$.[43] These parameters imply positive amenity spillovers generated by high skill workers, and negative spillovers generated by low skill workers.[44]

Alternative Parametrizations of the Spillover Elasticities We implement all our counterfactuals under different parametrizations of the spillover elasticities. The alternatives deviate from the benchmark described so far in terms of the efficiency or amenity spillover elasticities. In particular, we implement the model under: i) a more conservative parametrization that scales down the amenity spillover elasticities $\gamma_{\theta,\theta'}^{A}$ by 50% (referred to as the "Low amenity spillover" parametrization); ii) mappings of the amenity spillovers $\gamma_{\theta,\theta'}^{A}$ assuming values of the elasticity of city amenities to the share of college workers that are either one standard deviation above or below Diamond (2016) point estimates (referred to "High cross amenity spillover" and "Low cross amenity spillover" parametrizations, respectively); and iii) a less conservative parametrization that scales up the efficiency spillover elasticities $\gamma_{\theta,\theta'}^{P}$ to 0.12, i.e., twice the benchmark of 0.06 from Ciccone and Hall (1996) (referred to "High efficiency spillover" parametrization). The values of these alternative parametrizations are reported in Appendix B.2.

5.3 Calibration with Homogeneous Workers

We will start the quantitative section with a special version of our model that assumes a single worker type. Proposition 4 establishes the data requirements of that restricted model as well. To

---

[43]If our model was specified with Fréchet-distributed idiosyncratic preference draws for MSAs across individuals within types, as discussed in Section 4.4, the welfare-relevant coefficient $\tilde{\gamma}_{\theta,\theta}^{A}$ derived in that section would be parametrized to be the same as $\gamma_{\theta,\theta}^{A}$ in a specification without preference draws. Therefore, given the definition of the planner's problem in Section 4.4, whether or not idiosyncratic preference draws are included in our model does not matter for the quantification, once we condition on the moments from the data reported by Diamond (2016). See Appendix B.2 for details.

[44]At these values, all but one of the concavity conditions implied by Proposition 3 are satisfied. Specifically, the conditions that $\Gamma^{A} > \Gamma^{P}$, $\Gamma^{A} > 0$, and $\gamma_{\theta,\theta'}^{P} > 0$ for $\theta \neq \theta'$ are all satisfied, as well as the condition that $\gamma_{SU}^{A} > 0$. However, our parametrization sets $\gamma_{US}^{A} < 0$. In principle, therefore, concavity of the planner's problem is not guaranteed. However, in the quantitative exercise we check for the possibility of multiple local maxima by repeating the welfare maximization algorithm starting from 100 spatial allocations taken at random. Reassuringly, we fail to find any alternative local maximum.

implement that version, we use the aggregate MSA-level variables constructed above. To determine the spillover elasticities, we compute one-group elasticities that are consistent with the two-groups spillovers parametrized in the previous section. Specifically, we set one-group elasticities $\left(\gamma^A, \gamma^P\right)$ to the value that would be estimated through the lens of the labor supply and demand equations of the single-group model, if one were to use an MSA-level dataset generated by the model with heterogeneous groups and elasticities $\left\{\gamma^P_{\theta',\theta}\right\}$ and $\left\{\gamma^P_{\theta',\theta}\right\}$ calibrated above. This procedure by construction delivers $\gamma^P = 0.06$, equal to the value drawn from Ciccone and Hall (1996). The last subsection of Appendix B.2 shows that this procedure delivers an aggregate amenity elasticity of $\gamma^A = -0.19$.

5.4 Stylized Facts

Figure 1 revisits standard stylized facts on spatial disparities and sorting in the data, as well as a relatively less known fact on the spatial structure of trade imbalances. These facts will serve as a benchmark to evaluate the impact of optimal spatial policies.

Panels A to C show the standard facts about spatial disparities and sorting as function of city size, or "urban premia". Panel A documents the urban wage premium, defined as the increase in average nominal wages with city size. The elasticity of wages to city size is 7.4%. Panel B shows spatial sorting, in terms of the share of high-skill workers. The semi-elasticity of the share of high skill workers with respect to city size is 2.2%. I.e., doubling population increases the skill share by 2.2 percentage points.[45] Panel C shows the urban skill premium, defined as the increase in the ratio of high- to low-skill wage as city size increases. The slope of 0.11 means that larger cities feature a more unequal nominal wage distribution. The first fact suggests differences in productivity and cost of living across cities, while the last two suggest complementarities between city size and skill.

Panel D shows a somewhat less known fact, the relationship between city size and net imbalances. For each city we construct the net imbalance as the difference between expenditures and total income (from labor and non-labor sources). The graph shows net imbalance relative to labor income at the MSA level across MSAs. Given our construction of the expenditure variable, these differences in imbalances across cities result purely from the government policies that we measure (taxes and transfers). The negative slope reflects that government policies redistribute income from larger, high wage, high skill cities to smaller, low wage, low skill cities. These transfers are net of compositional effects according to detailed demographic characteristics in IPUMS, as mentioned above. Therefore, distributive government policies that vary with these characteristics across individuals do not underlie these patterns across cities.

---

[45]This spatial sorting has been increasing over time. Cities with initially higher shares of high skill workers have seen that increase. This trend, often referred to as "Great Divergence", is at the heart of the analyses of Moretti (2013) and Diamond (2016) on real wage inequality.

Figure 1: Urban Premia

(a) Urban Wage Premium

(b) Sorting

(c) Urban Skill Premium

(d) Imbalances

Note: each figure shows data across MSAs. All the city level outcomes reported on the vertical axis are adjusted by socio-demographic characteristics of each city, as detailed in Appendix B.1.

6 Optimal Spatial Policies in the U.S. Economy

We now inspect the optimal spatial policies using the parametrized model. To tease out the role of skill heterogeneity in driving results, we first solve for the optimal allocation in a version of the model that assumes homogeneous workers. Then, we move to the main case with skilled and unskilled workers. As a final exercise, to understand which key elasticity drive our results, we use the model to infer what spillover elasticities would rationalize the observed data as an efficient equilibrium, and compare them to our benchmark calibration.

6.1 Homogeneous Workers

We first solve the planner's problem in relative changes described in Section 4.5 assuming that the economy is populated by a homogeneous group of workers, under the parametrization described in Section 5.3. In this case the welfare gains from implementing the optimal allocation are negligible, equal to 0.06%.[46]

Why are the gains from optimal spatial policies with homogeneous workers tiny? This result follows from the optimal transfers being close to the measured ones. Figure A.1 in Appendix C shows transfers relative to wages across MSAs in the optimal allocation and in the data. As implied by our discussion in Section 4.2, under homogeneous workers the optimal net transfer per worker relative to the wage in city j takes the form $\frac{t_j}{w_j} = s + \frac{T}{w_j}$ for $s = \frac{\gamma^P + \gamma^A}{1 - \gamma^A}$. The figure shows that the best empirical fit to this relationship is quite close to the optimal.

We can also demonstrate that this result is not a feature of our model but a feature of the data. Varying the productivity and amenity spillover elasticities $\left(\gamma^A, \gamma^P\right)$ over a range from 0 to 4 times their baseline values, the welfare gains from optimal reallocations are below 0.2%. If we instead randomly modify the observed net transfers relative to wages keeping their mean and variance at the level in the data, the welfare gains vary between 0.02% and 0.5%.[47] Perturbing the data in other dimensions leads to even larger welfare gains. We simulate data corresponding to equilibria without government policies under counterfactual vectors of exogenous fundamentals such that the variance of wages across MSAs is larger than in the observed U.S. data equilibrium. Figure A.2 in Appendix C shows that the welfare gains can be substantial given artificial data with high wage dispersion.

Hence, through the lens of a model with homogeneous workers, the observed spatial allocation and transfers in the U.S. economy seem to be close to efficient. However, the model does not deliver small gains by design, as larger gains are possible under counterfactual configurations of transfers or the fundamentals.

6.2 Heterogenous Workers

We now explore the implications of optimal spatial policies using the full model with two groups of workers, skilled (S) and unskilled (U). Here again we solve the planner's problem in changes relative to the observed equilibrium. Specifically, we solve the problem of maximizing the change in utility of skilled workers, $\hat{u}^S$, subject to a lower bound for the change in utility of unskilled worker, $\hat{u}^U$. Varying this lower bound allows us to trace the Pareto frontier.

Aggregate Welfare Gains The left panel of Figure 2 shows the utility frontier of the U.S. economy in the benchmark parametrization, expressed in changes relative to the observed equilib-

---

[46]Eeckhout and Guner (2017) also find very small welfare gains in a numerical optimization over the income tax schedule.

[47]Specifically, we compute the welfare gains for 100 random draws of the distribution of transfer rates $(x_i - w_i)/w_i$ across cities, where the draws are taken from a normal distribution that matches the mean and variance of the distribution of transfer rates in the data.

Figure 2: Utility Frontier of the U.S. Economy between High and Low Skill Workers

(a) Benchmark

(b) Alternative Parametrizations
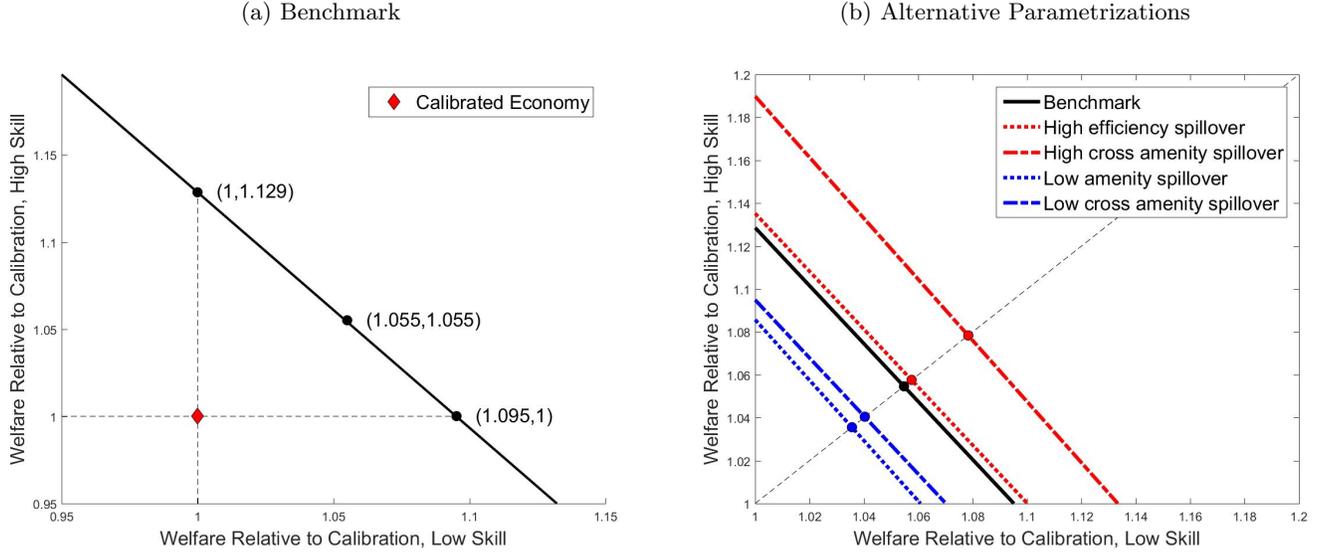
The figure shows the optimal welfare changes $\left(\hat{u^L}, \hat{u^H}\right)$ between the optimal and observed allocation, corresponding to the solution of the planner's problem in relative changes described in Appendix A.7. Each point corresponds to a maximization of $\hat{u^H}$ subject to a different lower bound on $\hat{u^L}$. The benchmark parametrization on the left panel corresponds to the black line on the right panel. The circles in the right panel represent intersections with the 45 degree line where the welfare of skilled and unskilled workers increase by the same amount.

rium. The point (1,1) represented with a red diamond corresponds to allocations where the welfare of skilled and unskilled workers is unchanged compared to the calibrated equilibrium. When the welfare gain of unskilled and skilled workers is restricted to be the same, optimal transfers lead to a 5.5% welfare gain for both types of workers. When only the welfare of one group is maximized subject to a constant level of welfare for the other group, we find gains of 12.9% for high skill workers and of 9.5% for low skill workers.

The right panel of Figure 2 shows the utility frontier under each of the alternative parametrizations discussed at the bottom of Section 5.2, including the benchmark. Under these parametriza-

Table 1: Welfare gains under different levels of the spillovers

Spillovers	Welfare Gain (%)
Benchmark	5.5
High efficiency spillover	5.9
Low amenity spillover	3.6
High cross-amenity spillover	7.8
Low cross-amenity spillover	4.0

The table reports the common welfare gains for skilled and unskilled workers under alternative parametrizations described in Section 5.2. The second row corresponds to $\gamma_{\theta'\theta}^{P}$ that are twice as large as in the benchmark. The third row corresponds to $\gamma_{\theta',\theta}^{A}$ that are 50% lower than the benchmark. The last two rows correspond to configurations that assume higher or lower cross-amenity spillovers corresponding to the plus or less one standard deviation of the estimates in Diamond (2016). See Appendix B.2 for details on these parametrizations.

tions, the frontier shifts up and down with little change in its slope. The welfare gains from implementing optimal policies are larger in the two frontiers in red, corresponding to high efficiency spillovers and high amenity spillovers across workers. The gains are lower under low amenity spillovers. Table 1 shows the welfare gains corresponding to the points at the intersection between these frontiers and the 45 degree line, such that skilled and unskilled workers gain the same. Across these specifications, the common welfare gains range from roughly 4% to 8%. Lowering the amenity spillover by 50% brings the common welfare gain down to 3.6%, while multiplying the efficiency spillovers by 2 increases the gain to 5.9%. Hence, in contrast to the one-group case, we now find sizable welfare gains from the optimal spatial reallocation.

Actual versus Optimal Transfers   The optimal allocation is implemented through net transfers by worker type and city. How does the optimal spatial redistribution implied by the quantified model compare to the data? To guide the answer to this question we can return to the optimal policies. Let t_j^θ be the optimal transfers received by type θ according to (20) in Proposition 2. Figure 3 shows the net transfers per capita relative to (gross) wages t_j^θ/w_j^θ by MSA and worker type on the vertical axis, against the average wage w_j^θ of each MSA in both the data (blue circles) and the optimal allocation (red diamonds), for low skill workers (hollow markers) and high skill workers (solid markers). The optimal allocation corresponds to the point on the Pareto frontier in the left panel of Figure 2 where welfare gains are equal for both types of workers.

The transfers in the data present a clear pattern of redistribution from high skill workers and high-wage cities towards low skill workers and low-wage cities. Net average transfers are positive for low skill workers and negative for high skill workers in most MSAs. Within skill groups, net transfers decrease with the average wage of the MSA.[48] The observations in red show the efficient allocation, which satisfies the optimality condition from Proposition 1. Across cities, the optimal transfers relative to labor income decrease more steeply with wages than in the data for both labor types, implying a stronger redistribution towards low-wage cities than what is observed empirically. Moreover, this increase in redistribution is sharper for high-skill workers, as seen in the greater difference between the slopes in the data and in the optimal allocation for that group.

The optimal transfers are implemented with a labor tax (a negative subsidy, $s_j^\theta < 0$) for both groups, coupled with lump-sum transfers T^θ. Moreover, the tax is higher in larger, high-wage cities. To understand the source of this spatial variation, we return to the expression for optimal subsidies (21). The first term of (21) is driven by own spillovers, while the second term is shaped by cross spillovers. In our parametrization of spillovers for low skill workers, both of these terms are negative. The negative cross-spillovers through amenities lead to the higher tax of low skill workers in large, high-wage cities where a larger share of expenditures accrues to high skill workers. The logic that rationalizes a higher labor tax in high-wage cities is different for high skill workers. In

---

[48]On average across MSAs, they equal 1.9 thousand dollars for low-skill workers, or 13% of their average wage. For high skill workers, the corresponding numbers are -3 thousand dollars or -7% of the average wage. In cities where high skill workers earn on average more than $50k per year, net transfers of high skill workers are -6.9 thousand dollars or -12% of wages.

Figure 3: Per Capita Transfers by Skill Level and MSA, Data and Optimal Allocation
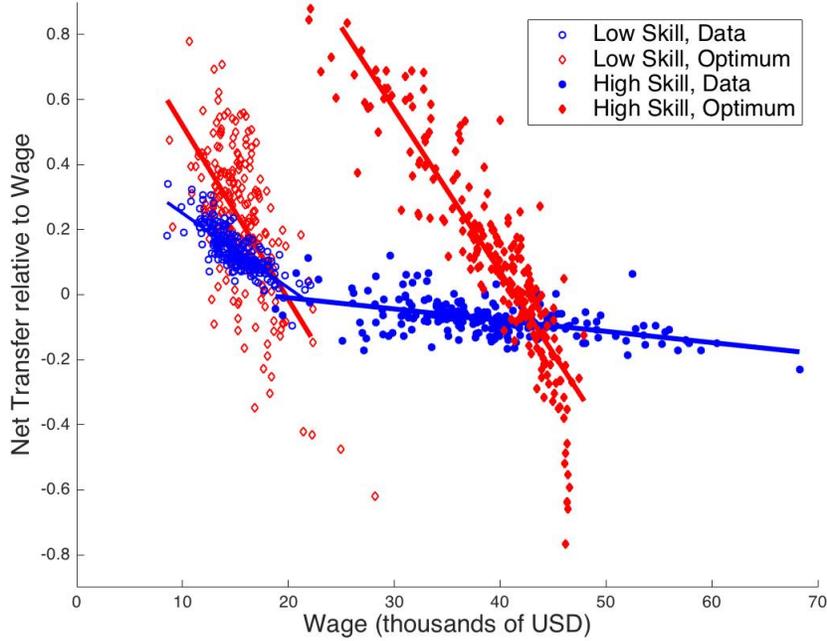
Note: each point in the figure corresponds to an MSA-skill group combination. The vertical axis show the difference between average expenditure and wage relative to wage and the horizontal axis shows the average wage. For details of how the data is constructed see Appendix B. The slopes of each linear fit (with SE) are: Low Skill, Data: -0.02 (0.001); Low Skill, Optimum: -0.052 (0.005); High Skill, Data: -0.003 (0.0004); High Skill, Optimum: -0.05 (0.001). The figure corresponds to planner's weights such that both types of workers experience the same welfare gain in Figure 2.

our parametrization, high skill workers generate positive own spillovers. According to the first term in (21), these positive spillovers would call for a labor income *subsidy*. However, this force is more than offset by strong positive cross spillovers onto low skill workers. A higher tax in high-wage cities directs skilled workers into small, low-wage cities where low skill workers are relatively more prevalent.[49]

Optimal Reallocation and Sorting The optimal transfers change the spatial distribution of economic activity compared to the data. By changing the location incentives of workers, they affect spatial sorting and the city size distribution. These reallocations in turn impact the spatial distributions of labor productivity and wages through agglomeration spillovers, and the distribution of urban amenities through amenity spillovers. These effects feed back to location choices, changing the spatial pattern of skill premia and inequality. We now describe the spatial equilibrium resulting from this process.

Figure 4 shows the pattern of reallocation. Panel (a) shows the initial total population of each

---

[49]Figure A.3 in Appendix C replicates the figure for alternative Pareto weights for high and low skill workers. The main impact of a different Pareto weight is to shift the transfer schedules up and down depending on the Planner's preference for each group, without changing the qualitative patterns we have discussed.

Figure 4: Changes in Population, Skill Shares, and Skill Premium across MSAs
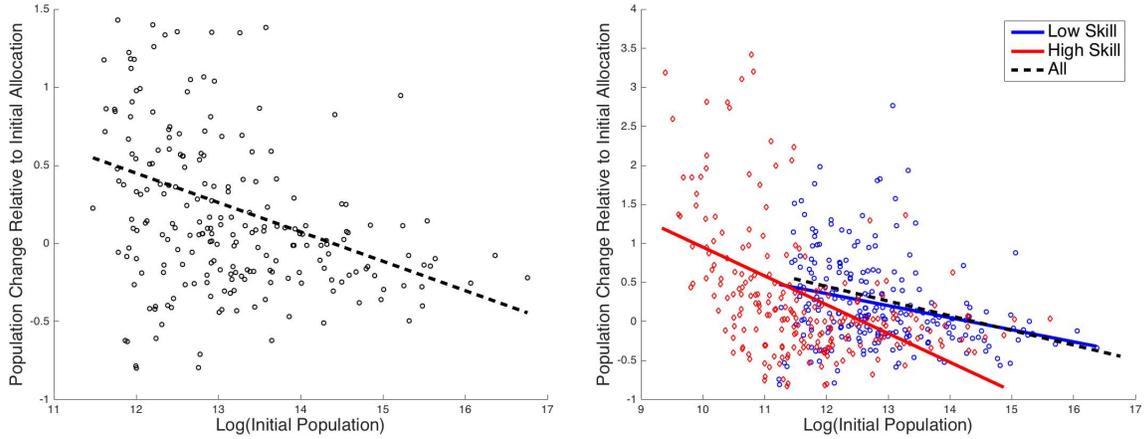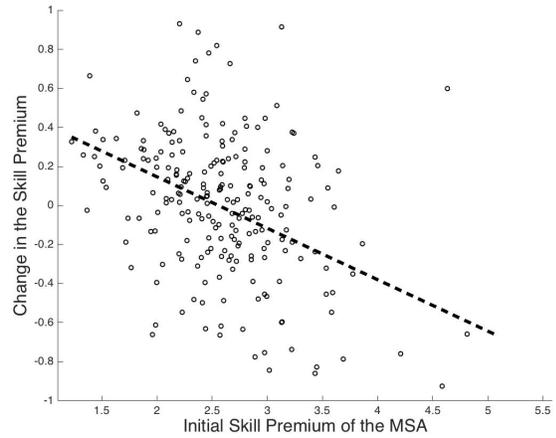

(a) Change in Total Population

(b) Change in Population by Skill Group
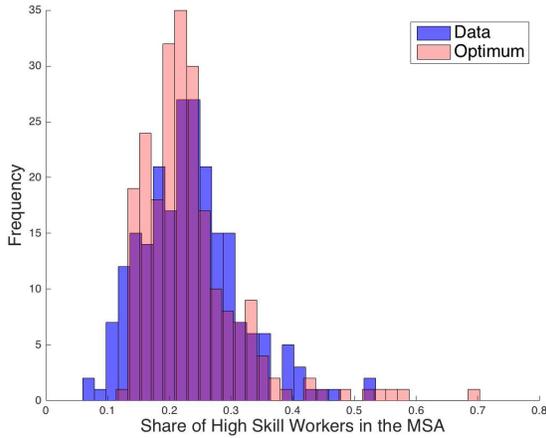
(c) Histogram of High-Skill Shares across MSAs

(d) Change in Skill Premium

Note: Panel (a) shows the change in population between the optimal allocation and the initially observed equilibrium and the linear fit. Slope (SE): -0.19 (0.04); R^2=0.1. Panel (b) displays the same outcomes for high and low skill workers. Slopes (with SE): High Skill: -0.22 (0.03); Low Skill: -0.11 (0.03). Panel (d) displays in the vertical axis the difference in the skill premium between the optimal and initial allocation. Slope (SE): -0.25 (0.04). The figures correspond to planner's weights such that both types of workers experience the same welfare gain in Figure 2.

MSA on the horizontal axis and the change in population implied by the optimal allocation relative to the initial allocation on the vertical axis, defined as $\hat{L}_j - 1$. The stronger redistribution to low-wage locations discussed in the previous section implies that, on average, there is reallocation from large to small cities. However, there is also considerable heterogeneity in growth rates over the size distribution, including middle- and small-MSAs that shrink alongside large MSA's that grow. In other words, due to the multiple sources of spatial heterogeneity and spillovers, initial city size is a poor predictor of whether a city is too large or too small in the observed allocation (the R^2 of the linear regression is 10%). Panel (b) shows the same figure within skill groups alongside the linear fit from Panel (a). The sharper increase in tax progressivity for high skill workers implies that the

reallocations towards initially smaller places is also stronger within that group.

These differential patterns of reallocation by skill lead to less segregation of worker by skill across cities. Panel (c) shows the histogram of skill shares across MSAs in the initial and optimal allocation, implying a concentration of skill shares. An additional implication of these reallocations, shown in panel (d), is that the skill premium tends to decrease in cities that are initially more unequal. The loss of agglomeration spillovers in productivity among high skill workers implies that the optimal spatial transfers reduce inequality within cities that are more unequal in the observed allocation.

Figure 5: Urban Premia, Data and Optimal Allocation

(a) Urban Wage Premium

(b) Sorting

(c) Urban Skill Premium

(d) Imbalances
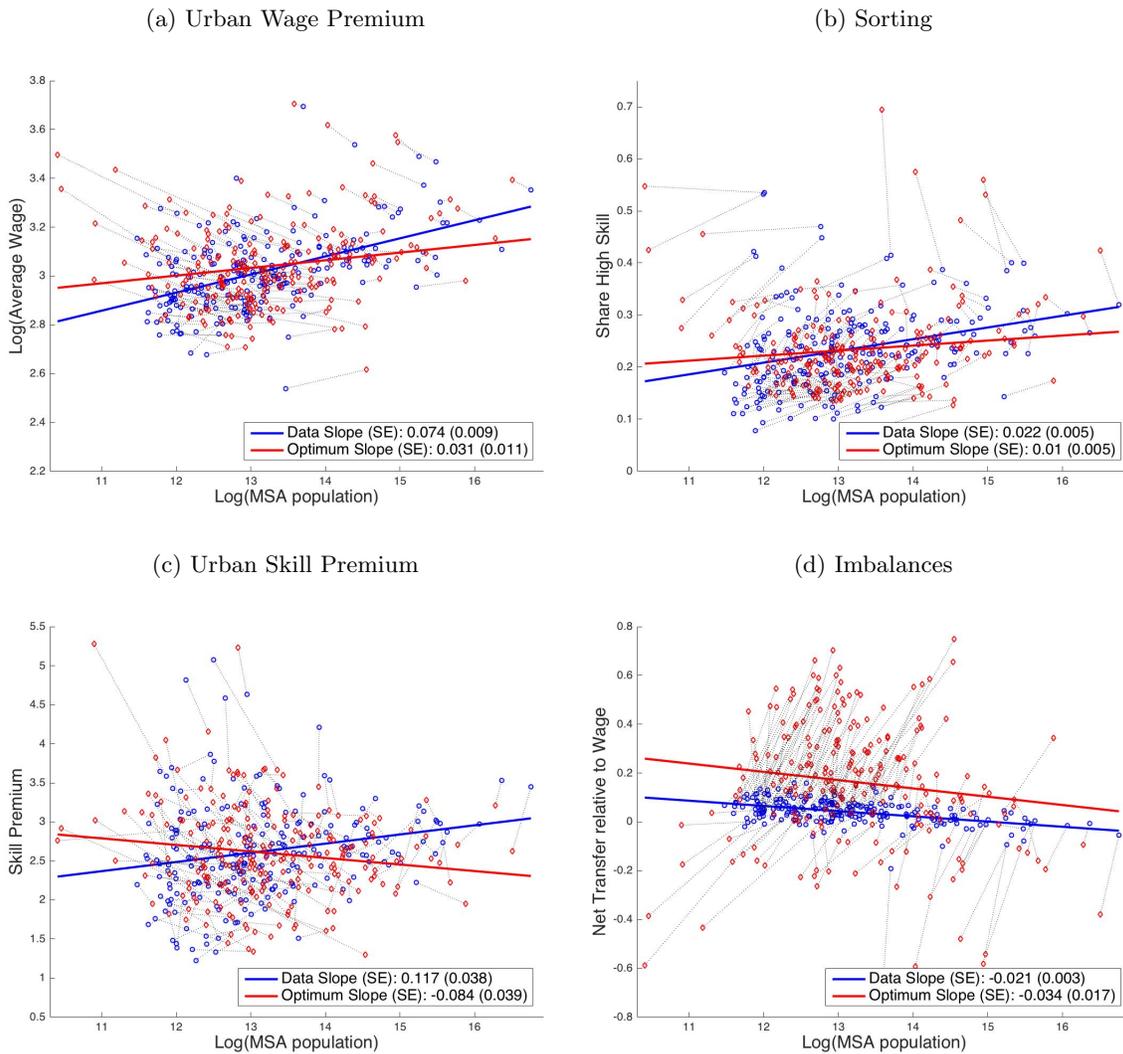
Note: each panel reports outcomes across MSAs in the data and in the optimal allocation. Each linked pair of observations corresponds to the same MSA.

The Urban Premia in the Optimal Allocation   We now revisit the stylized facts about urban premia presented in Section 5. Figure 5 reproduces Figure 1. Each pair of linked observations corresponds to the same MSA, in the data and in the optimal allocation. As expected from the previous discussion, the optimal allocation features a higher absolute value of the imbalances at the city level (panel (d)), since redistribution to smaller MSAs is stronger in the optimal allocation than in the data. The optimal allocation features more mixing of high and low skills in large cities (panel (b)), resulting from the larger increase in implicit taxes for high skill workers than for low skill workers discussed above. The urban skill premium vanishes (panel (c)), implying that the sorting pattern from panel (b) ends up being detached from the urban skill premium. Instead, it is driven by stronger preferences for urban amenities among high skill workers. As seen in panel (a), the wage premium in the large cities is still noticeable, but lower than in the data. It is driven by an average productivity advantage across both skill groups in larger cities, rather than by a relatively higher productivity of high-skill workers in these places.

In addition to the average patterns of urban premia, these figures also demonstrate the rich patterns of optimal reallocation resulting from heterogeneity in the many different fundamentals across cities. In panels (b) and (d) we encounter cities that may either increase or decrease their share of high skill workers and skill premium conditional on either growing or shrinking in size. In particular, some of the smallest cities in the optimal allocation are also relatively small in the initial allocation, and they shrink along with a reduction in the skill share. Similarly, we can find that some initially large cities should grow even more and increase their skill share. Therefore, our model does not predict that all currently small cities should grow and increase their skill share at the expense of large cities.

6.3   Optimal Urban Premia under Alternative Parametrizations

Pareto Weights   Throughout our discussion in the previous section we compared the data to an optimal allocation corresponding to the same welfare gains to all workers (i.e., the welfare gains $\left(\hat{u}^U, \hat{u}^S\right) = (1.05, 1.05)$ in Figure 2). Different planner's weights are implemented through the lump-sum transfers across types, which turn out to have little effect on the optimal urban premia. To highlight this property, Figure A.4 in Appendix C shows the same outcomes as in Figure 5 but assuming planner weights that are either 10 times larger for high skill workers (case labeled "low weight on U") or 10 times larger for low skill workers (case labeled "high weight on U"). As we should expect, the optimal welfare changes relative to the observed allocation under these weights are different from the benchmark case. However, we find similar patterns of optimal urban premia.

Amenity and Efficiency Spillover Elasticities   We revisit the main results on spatial real-location under different values of the spillovers. In particular we consider two of the alternative parametrizations from Table 1: low amenity spillovers, where $\gamma^A_{\theta,\theta'}$ is 50% below the benchmark parametrization; and high efficiency spillover, where we set the efficiency spillovers $\gamma^P_{\theta,\theta'}$ to twice the benchmark. Figure A.5 replicates the optimal patterns of urban premia from Figure 5 under

the benchmark and under these alternative parametrizations. The key qualitative patterns of optimal reallocation and spatial premia discussed in the benchmark are similarly present under these alternative parametrizations.

6.4   Contrast of the Results under Homogeneous and Heterogeneous Workers

Why are the welfare gains so much larger under multiple worker types relative to a single worker type? To answer this question, we note that the quantification with heterogeneous workers would deliver the exact same MSA-level outcomes and welfare gains (common to all workers) as the quantification with homogeneous workers if, under heterogeneous workers: (i) the data exhibited no spatial sorting by skill (a constant share of workers of type θ over space), no urban skill premium (a constant skill wage premium w_j^S/w_j^U over space), and no relative differences in expenditures (a constant expenditure ratio x_i^S/x_i^U over space); (ii) high and low skill workers were perfectly substitutable in production ($\rho = 1$); and (iii) spillovers across types were absent, with own spillovers equal to the calibration in the one-group case, i.e.: $\gamma_{\theta,\theta'}^A = \gamma_{\theta,\theta'}^P = 0$ if $\theta \neq \theta'$, $\gamma_{\theta,\theta}^A = \gamma^A$, and $\gamma_{\theta,\theta}^P = \gamma^P$. We can investigate the role played by different forces by releasing each of the restrictions (i)-(iii). As we do so, we move from a "one-group calibration" of the model with heterogeneous workers to the full two-groups case calibration we have previously implemented.[50]

Table 2 reports the results. Each of the columns allows for a different relaxation of (i), (ii), or (iii). The first row corresponds to the welfare gain under homogeneous workers reported in Section 6.1, and the last row corresponds to the welfare gains from model with heterogeneous workers in Section 6.2, when the gains are common to all workers.

Starting from a case with perfect substitution across workers, bringing in the full matrix of heterogeneous spillovers considerably increases the welfare gains, both without sorting (comparing rows 1 and 2) and with it (comparing rows 4 and 5). Therefore, heterogeneous spillovers across workers play a key role. The sorting observed in the data also plays a relevant role, and particularly so under heterogeneous spillovers. We conclude that accounting for heterogeneity in skills is important for the implied benefits of implementing optimal spatial policies, and that dispersion in spillovers across different types of workers play a substantial role.

Table 2: Welfare gains under different model specifications

	(i) Data	(ii) Elasticity ρ	(iii) Heterogeneous spillovers	Welfare Gain (%)
(1)	No Sorting	1	No	0.06
(2)	No Sorting	1	Yes	0.36
(3)	No Sorting	0.39	Yes	0.25
(4)	Sorting	1	No	0.29
(5)	Sorting	1	Yes	6.21
(6)	Sorting	0.39	Yes	5.47

---

[50]This "one-group calibration" is not the calibration of the model with homogeneous workers, but rather a calibration of a model with heterogeneous workers which, by design, delivers the same welfare effects for all workers and MSA-level reallocation as the model with homogeneous workers.

6.5 Inferring the Spillover Elasticities assuming Efficiency in the Data

Our logic so far was to discipline the model with existing estimates of the spillover elasticities, and then use it to compute the efficient allocation. In this final section we invert this logic, and instead ask: what spillover elasticities would be consistent with assuming that the observed spatial allocation is efficient? By comparing these inferred spillover elasticities with those used in the calibration, this exercise allows us to identify the key elasticities behind our results.

Assuming that the observed equilibrium with transfers is optimal, the condition on optimal transfers (20) must hold. Combined with the definition of expenditure per worker in (15), we obtain an optimal relationship between transfers, wages, expenditures, and employment of the form

$$t_j^\theta = a_0^\theta + a_1^\theta w_j^\theta + a_2^\theta \left(\frac{w_j^{\theta' \neq \theta} L_j^{\theta' \neq \theta}}{L_j^\theta} \right) + a_3^\theta \left(\frac{x_j^{\theta' \neq \theta} L_j^{\theta' \neq \theta}}{L_j^\theta} \right) + \varepsilon_j^\theta \tag{41}$$

for $\theta \in \{U, S\}$, where ε_j^θ is a measurement error term, and the reduced-form parameters have the following structural interpretations: $a_0^\theta \equiv -b^\theta \Pi^* - \frac{E^\theta}{1-\gamma_{\theta,\theta}^A}$, $a_1^\theta \equiv \frac{\gamma_{\theta,\theta}^P + \gamma_{\theta,\theta}^A}{1-\gamma_{\theta,\theta}^A}$, $a_2^\theta \equiv \frac{\gamma_{\theta,\theta'}^P}{1-\gamma_{\theta,\theta}^A}$, and $a_3^\theta = \frac{\gamma_{\theta,\theta'}^A}{1-\gamma_{\theta,\theta}^A}$. We estimate the parameters $\{a_i^\theta\}$ by running (41) as a regression in the cross-section, and then infer the spillover elasticities $\{\gamma_{\theta,\theta'}^A, \gamma_{\theta,\theta'}^P\}$ up to a normalization for each type.[51] We normalize the own-spillover elasticity for productivity to the benchmark level for the U.S. used in Section 5.2.

This exercise yields $(\gamma_{UU}^A, \gamma_{SU}^A, \gamma_{US}^A, \gamma_{SS}^A) = (-.26, -.08, -.09, -.33)$ and $(\gamma_{UU}^P, \gamma_{SU}^P, \gamma_{US}^P, \gamma_{SS}^P) = (.003, .11, .07, .053)$. The average level of both types of spillovers are similar to the parameters implied by the empirical estimates used in the calibration. In both these inferred elasticities and the calibrated ones, the amenity spillovers are larger than the agglomeration spillovers, and high-skill workers generate stronger efficiency spillovers than low-skill workers. However, the assumption that the observed allocation is optimal implies negative amenity spillovers both across and within skill groups, whereas the calibrated elasticities imply positive amenity spillovers generated by high skilled workers. Therefore, we find that heterogeneity in spillovers across groups is key for spatial policies, as previously suggested by the contrast between the quantified model under homogeneous and heterogeneous workers.

7 Conclusion

In this paper we study optimal spatial policies in quantitative geography frameworks with spillovers and sorting of heterogeneous workers. Our framework nests recent strands of quantitative spatial research and includes many key determinants of the spatial distribution of economic activity

---

[51]This normalization is needed because from (41) the own-spillover elasticities for productivity and amenities are not separately identified. Assuming values for $\gamma_{\theta,\theta}^P$ we can then infer the remaining elasticities as follows: $\gamma_{\theta,\theta}^A = \frac{a_1^\theta - \gamma_{\theta,\theta}^P}{1+a_1^\theta}$, $\gamma_{\theta,\theta'}^P = a_2^\theta \left(1 - \gamma_{\theta,\theta}^A\right)$, and $\gamma_{\theta,\theta'}^A = a_3^\theta \left(1 - \gamma_{\theta,\theta}^A\right)$.

such as geographic frictions and asymmetric spillovers across heterogeneous workers. We generalize these models to allow for arbitrary policies that transfer resources across agents and regions.

We first derive the set of transfers across workers and regions that must hold in an efficient allocation. Furthermore, we provide conditions on the distributions of spillover and housing supply elasticities such that these transfers are also sufficient for efficiency. Under constant-elasticity spillovers and either homogeneous workers or no spillovers across different types of workers, constant labor subsidies and lump-sum transfers over space implement the efficient allocation, regardless of micro heterogeneity in fundamentals. When workers are heterogeneous and there are spillovers across different types of workers, spatial efficiency may require place-specific taxes or subsidies to attain optimal sorting.

We then show how to use the framework to assess the efficiency of the observed spatial allocation. The distributions of fundamentals needed to compute the optimal allocation can be backed out from data on the distribution of wages, employment, and expenditures across worker types and regions. We apply the model to data on the distribution of economic activity across MSAs using spillover elasticities in a range around existing estimates. Our results suggest that spatial efficiency calls for more redistribution to low-wage cities and weaker sorting by skill relative to the data, leading to lower wage inequality in larger cities. Under the assumption of homogeneous workers, our approach delivers negligible gains. However, under heterogeneous workers, the gains are considerably larger.

Overall, we find that accounting for skill heterogeneity and spillovers across different types of workers is important for the design and aggregate welfare effects of spatial policies. The results suggest that nudging current U.S. policies towards generating a greater mixing of high and low skill workers in low-wage cities could be socially desirable.

References

Abdel-Rahman, H. and M. Fujita (1990). Product variety, marshallian externalities, and city sizes. *Journal of regional science 30*(2), 165–183.

Abdel-Rahman, H. M. and A. Anas (2004). Theories of systems of cities. *Handbook of regional and urban economics 4*, 2293–2339.

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

Albouy, D. (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 (2017). The optimal distribution of population across cities. Technical report, National Bureau of Economic Research.

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

Allen, T., C. Arkolakis, and X. Li (2015). Optimal city structure. *Yale University, mimeograph*.

Allen, T., C. Arkolakis, and Y. Takahashi (2014). Universal gravity. Technical report, National Bureau of Economic Research.

Anderson, J. E. and E. van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review 93*(1), 170–192.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (2012). New trade models, same old gains? *The American Economic Review 102*(1), 94–130.

Austin, B., E. Glaeser, and L. Summers (2018). Saving the heartland: Place-based policies in 21st century america. *Brookings Papers on Economic Activity Conference Draft*.

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

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

Bhagwati, J. and H. G. Johnson (1960). Notes on some controversies in the theory of international trade. *The Economic Journal 70*(277), 74–93.

Brandt, L., T. Tombe, and X. Zhu (2013). Factor market distortions across time, space and sectors in China. *Review of Economic Dynamics 16*(1), 39–58.

Caliendo, L., M. Dvorkin, and F. Parro (2015). The impact of trade on labor market dynamics. Technical report, National Bureau of Economic Research.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (forthcoming). The impact of regional and sectoral productivity changes on the us economy. *The Review of economic studies*.

Ciccone, A. and R. E. Hall (1996). Productivity and the density of economic activity. *The American Economic Review*, 54–70.

Colas, M. and K. Hutchinson (2017). Heterogeneous workers and federal income taxes in a spatial equilibrium. *University of Wisconsin-Madison*.

Combes, P.-P. (2011). The empirics of economic geography: how to draw policy implications? *Review of World Economics 147*(3), 567–592.

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

Combes, P.-P. and L. Gobillon (2015). The empirics of agglomeration economies. In *Handbook of regional and urban economics*, Volume 5, pp. 247–348. Elsevier.

Davis, D. R. and J. I. Dingel (2012). A spatial knowledge economy. Technical report, National Bureau of Economic Research.

Dekle, R., J. Eaton, and S. Kortum (2008). Global rebalancing with gravity: Measuring the burden of adjustment. Technical Report 3, International Monetary Fund.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (forthcoming). The geography of development: Evaluating migration restrictions and coastal flooding. *Journal of Political Economy*.

Desmet, K. and E. Rossi-Hansberg (2013). Urban accounting and welfare. *American Economic Review 103*(6), 2296–2327.

Desmet, K. and E. Rossi-Hansberg (2014). Spatial development. *American Economic Review 104*(4), 1211–43.

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.

Dixit, A. (1985). Tax policy in open economies. *Handbook of public economics 1*, 313–374.

Dunbar, A. E. (2009). Metropolitan area disposable personal income: Methodology and results for 2001-2007.

Duranton, G. and D. Puga (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics*, Volume 4, pp. 2063–2117. Elsevier.

Duranton, G. and A. J. Venables (2018). Pace-based policies for development. Technical report, National Bureau of Economic Research.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Eeckhout, J. and N. Guner (2017). Optimal spatial taxation: Are big cities too small?

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

Fajgelbaum, P. D., E. Morales, J. C. S. Serrato, and O. M. Zidar (2015). State taxes and spatial misallocation. Technical report, National Bureau of Economic Research.

Fajgelbaum, P. D. and E. Schaal (2017). Optimal transport networks in spatial equilibrium. Technical report, National Bureau of Economic Research.

Flatters, F., V. Henderson, and P. Mieszkowski (1974). Public goods, efficiency, and regional fiscal equalization. *Journal of Public Economics 3*(2), 99–112.

Flood, S., M. King, S. Ruggles, and J. R. Warren (2017). Integrated public use microdata series, current population survey: Version 5.0.[dataset]. minneapolis: University of minnesota, 2017.

Gaubert, C. (2015). Firm sorting and agglomeration. *University of California, Berkeley*.

Giannone, E. (2017). Skilled-biased technical change and regional convergence. Technical report, University of Chicago Working Paper, available at: http://home. uchicago. edu/˜ elisagiannone/files/JMP_ElisaG. pdf.

Glaeser, E. L. and J. D. Gottlieb (2008). The economics of place-making policies. *Brookings Papers on Economic Activity 39*(1 (Spring)), 155–253.

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

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. *Handbook of International Economics, Vol. 4*.

Helpman, E. (1998). The size of regions: transport and housing as factors in agglomeration. In D. Pines, E. Sadka, and I. Zilcha (Eds.), *Topics in Public Economics*, pp. 33–54. Cambridge University Press Cambridge.

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.

Henderson, J. V. (1974). The sizes and types of cities. *The American Economic Review*, 640–656.

Henderson, J. V. (1977). *Economic theory and the cities*. Academic Press.

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

Hsieh, C.-T. and E. Moretti (2017). Housing constraints and spatial misallocation. Technical report.

Khajavirad, A., J. J. Michalek, and N. V. Sahinidis (2014). Relaxations of factorable functions with convex-transformable intermediates. *Mathematical Programming 144*(1-2), 107–140.

Kline, P. (2010). Place based policies, heterogeneity, and agglomeration. *The American Economic Review 100*(2), 383–387.

Kline, P. and E. Moretti (2014a). Local economic development, agglomeration economies and the big push: 100 years of evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics*.

Kline, P. and E. Moretti (2014b). People, places, and public policy: Some simple welfare economics of local economic development programs. *Annual Review of Economics 6*(1), 629–662.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review*, 950–959.

Lucas, R. E. and E. Rossi-Hansberg (2002). On the internal structure of cities. *Econometrica 70*(4), 1445–1476.

Melo, P. C., D. J. Graham, and R. B. Noland (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional science and urban Economics 39*(3), 332–342.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2015). Commuting, migration and local employment elasticities. Technical report, National Bureau of Economic Research.

Moretti, E. (2012). *The new geography of jobs.* Houghton Mifflin Harcourt.

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

Neumark, D., H. Simpson, et al. (2015). Place-based policies. *Handbook of Regional and Urban Economics 5*, 1197–1287.

Ossa, R. (2018). A quantitative analysis of subsidy competition in the us. Technical report, National Bureau of Economic Research.

Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics 101*, 148–167.

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

Redding, S. J. and M. A. Turner (2015). Transportation costs and the spatial organization of economic activity. *Handbook of Regional and Urban Economics 5*, 1339–1398.

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

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

Rossi-Hansberg, E. (2005). A spatial theory of trade. *American Economic Review 95*(5), 1464–1491.

Sandmo, A. (1975). Optimal taxation in the presence of externalities. *The Swedish Journal of Economics*, 86–98.

Shourideh, A. and R. Hosseini (2018). Inequality, Redistribution, and Optimal Trade Policy: A Public Finance Approach. Technical report, Carnegie Mellon University.

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

A   Appendix to Section 4

A.1   Planning Problem and Proofs of Propositions 1 to 3

The planning problem can be described as follows.

Definition 2. *The planning problem is*

$$\max L^\theta u^\theta$$

subject to (i) the spatial mobility constraints

$$L_j^\theta u^\theta \leq L_j^\theta a_j^\theta \left(L_j^1, .., L_j^\Theta \right) U \left(c_j^\theta, h_j^\theta \right) \text{ for all } j;$$
$$L_j^{\theta'} \underline{u}^{\theta'} \leq L_j^{\theta'} a_j^\theta \left(L_j^1, .., L_j^\Theta \right) U \left(c_j^\theta, h_j^\theta \right) \text{ for all } j \text{ and } \theta' \neq \theta;$$

(ii) the tradable and non-tradable goods feasibility constraints

$$\sum_i d_{ji} Q_{ji} \leq Y_j \left(N_j^Y, I_j^Y \right) \text{ for all } j, i;$$

$$\sum_\theta L_j^\theta c_j^\theta + I_j^Y + I_j^H \leq Q \left(Q_{1j}, .., Q_{Jj} \right) \text{ for all } j;$$

$$\sum_\theta L_j^\theta h_j^\theta \leq H_j \left(N_j^H, I_j^H \right) \text{ for all } j;$$

(iii) local and national labor-market clearing,

$$N_j^Y + N_j^H = N \left(z_1^\theta \left(L_j^1, .., L_j^\Theta \right) L_j^1, .., z_j^\Theta \left(L_j^1, .., L_j^\Theta \right) L_j^\Theta \right) \text{ for all } j;$$

$$\sum_j L_j^\theta = L^\theta \text{ for all } \theta; \text{ and}$$

(iv) non-negativity constraints on consumption, trade flows, intermediate inputs, and labor.

Proposition 1. *If a competitive equilibrium is efficient, then*

$$W_j \frac{dN_j}{dL_j^\theta} + \sum_{\theta'} \frac{x_j^{\theta'} L_j^{\theta'}}{a_j^{\theta'}} \frac{\partial a_j^{\theta'}}{\partial L_j^\theta} = x_j^\theta + E^\theta \quad \text{if } L_j^\theta > 0, \tag{A.1}$$

for all j and θ and some constants $\{E^\theta\}$. If the planner's problem is globally concave and (A.1) holds for some specific $\{E^\theta\}$, then the competitive equilibrium is efficient.

Proof. First we present the system of necessary first order conditions in the planner's problem. Then we contrast it with the market allocation. The Lagrangian of the planning problem is:

$$\begin{aligned}
\mathscr{L} = u^\theta \\
&- \sum_j \omega_j^\theta L_j^\theta \left(u^\theta - a_j^\theta \left(L_j^1, .., L_j^\Theta \right) U \left(c_j^\theta, h_j^\theta \right) \right) \\
&- \sum_{\theta' \neq \theta} \sum_j \omega_j^{\theta'} L_j^{\theta'} \left(u^{\theta'} - a_j^{\theta'} \left(L_j^1, .., L_j^\Theta \right) U \left(c_j^{\theta'}, h_j^{\theta'} \right) \right) \\
&- \sum_j p_j^* \left(\sum_i d_{ji} Q_{ji} - Y_j \left(N_j^Y, I_j^Y \right) \right) \\
&- \sum_j P_j^* \left(\sum_\theta L_j^\theta c_j^\theta + I_j^Y + I_j^H - Q \left(Q_{1j}, .., Q_{Jj} \right) \right) - \sum_j R_j^* \left(\sum_\theta L_j^\theta h_j^\theta - H_j \left(N_j^H, I_j^H \right) \right) \\
&- \sum_j W_j^* \left(N_j^I + N_j^H - N \left(z_j^1 \left(L_j^1, .., L_j^\Theta \right) L_j^1, .., z_j^\Theta \left(L_j^1, .., L_j^\Theta \right) L_i^\Theta \right) \right) \\
&- \sum_\theta E^\theta \left(\sum_j L_j^\theta - L^\theta \right) + ...
\end{aligned} \tag{A.2}$$

where we omit notation for the non-negativity constraints. The first-order conditions with respect to trade flows, labor services and intermediate inputs are:

$$[Q_{ji}] \qquad P_i^* \frac{\partial Q \left(Q_{1i}, .., Q_{Ji} \right)}{\partial Q_{ji}} \leq p_j^* \tau_{ji}, \tag{A.3}$$

$$\left[N_j^Y, N_j^H \right] \qquad p_j^* \frac{\partial Y_j}{\partial N_j^Y} \leq W_j^*; R_j^* \frac{\partial H_j}{\partial N_j^H} \leq W_j^*, \tag{A.4}$$

$$\left[I_j^Y, I_j^H \right] \qquad p_j^* \frac{\partial Y_j}{\partial I_j^Y} \leq P_j^*; R_j^* \frac{\partial H_j}{\partial I_j^H} \leq P_j^*, \tag{A.5}$$

each holding with equality in an interior solution. The first-order conditions with respect to individual consumption of traded and non-traded goods can be written:

$$\left[c_j^\theta\right] \qquad \omega_j^\theta a_j^\theta \frac{\partial U\left(c_j^\theta, h_j^\theta\right)}{\partial c_j^\theta} c_j^\theta = P_j^* c_j^\theta$$

$$\left[h_j^\theta\right] \qquad \omega_j^\theta a_j^\theta \frac{\partial U\left(c_j^\theta, h_j^\theta\right)}{\partial h_j^\theta} h_j^\theta = R_j^* h_j^\theta$$

Adding up the last two expressions and using degree-1 homogeneity of U gives

$$\omega_j^\theta a_j^\theta U\left(c_j^\theta, h_j^\theta\right) = x_j^{\theta*}, \tag{A.6}$$

where

$$x_j^{\theta*} \equiv R_j^* h_j^\theta + P_j^* c_j^\theta. \tag{A.7}$$

Therefore, we can write

$$\left[c_j^\theta\right] \qquad c_j^\theta = \frac{\alpha_C\left(c_j^\theta, h_j^\theta\right)}{P_j^*} x_j^{\theta*} \tag{A.8}$$

$$\left[h_j^\theta\right] \qquad h_j^\theta = \frac{1 - \alpha_C\left(c_j^\theta, h_j^\theta\right)}{R_j^*} x_j^{\theta*} \tag{A.9}$$

where $\alpha_C\left(c, h\right) \equiv \frac{\partial U(c,h)}{\partial c} \frac{c}{U(c,h)}$ is the elasticity of U with respect to c.

Using (A.7) and the slackness condition on the spatial mobility constraint, the first-order condition of the planning problem with respect to L_j^θ is:

$$\sum_{\theta'} \omega_j^{\theta'} L_j^{\theta'} \frac{\partial a_j^{\theta'}\left(L_j^1, .., L_j^\Theta\right)}{\partial L_j^\theta} U\left(c_j^{\theta'}, h_j^{\theta'}\right) + W_j^* \frac{dN_j}{dL_j^\theta} \leq x_j^{\theta*} + E^\theta, \tag{A.10}$$

with equality if $L_j^\theta > 0$. Further using (A.6), if $L_j^\theta > 0$ then:

$$W_j^* \frac{dN_j}{dL_j^\theta} + \sum_{\theta'} \frac{\left(x_j^{\theta*}\right)' L_j^{\theta'}}{a_j^{\theta'}} \frac{\partial a_j^{\theta'}}{\partial L_j^\theta} = x_j^{\theta*} + E^\theta. \tag{A.11}$$

An optimal allocation is given by quantities $\left\{Q_{ji}, N_j^Y, N_j^H, I_j^Y, I_j^H, c_j^\theta, h_j^\theta, L_j^\theta, u^\theta\right\}$ and multipliers $\left\{P_j^*, p_j^*, R_j^*, W_j^*, \omega_j^\theta\right\}$ such that the first-order conditions (A.3)-(A.11) and the constraints enumerated in (i) to (iii) in Definition 2 hold.

It is straightforward to show that (A.3) to (A.9) coincide with the optimality conditions of producers and consumers (i) and (ii) in the competitive equilibrium from Definition 1 given competitive prices $\left\{P_j, p_j, R_j, W_j\right\}$ equal to the multipliers $\left\{P_j^*, p_j^*, R_j^*, W_j^*\right\}$ and decentralized expenditure x_j^θ equal to $x_j^{\theta*}$. In addition, the restrictions (i) to (iii) from definition 2 of the planning problem are the same as restriction (iii) from the competitive equilibrium. Therefore, the system characterizing the competitive solution for $\left\{Q_{ji}, N_j^Y, N_j^H, I_j^Y, I_j^H, c_j^\theta, h_j^\theta, L_j^\theta\right\}$ given the prices $\left\{P_j, p_j, R_j, W_j\right\}$ and the expenditure x_j^θ is the same as the system characterizing the planner allocation for those same quantities given the multipliers $\left\{P_j^*, p_j^*, R_j^*, W_j^*\right\}$ and $x_j^{\theta*}$. As a result, if the competitive allocation is efficient, then $x_j^\theta = x_j^{\theta*}$ where $x_j^{\theta*}$ is given by (A.11). Conversely, if $x_j^\theta = x_j^{\theta*}$ for $x_j^{\theta*}$ defined in (A.1) given the W^θ that solves the planner's problem, there is a solution for the competitive allocation such that $\left\{P_j, p_j, R_j, W_j\right\} = \left\{P_j^*, p_j^*, R_j^*, W_j^*\right\}$. If the planning problem is concave then there is a unique solution to the system characterizing the planner's allocation, in which case $\left\{P_j, p_j, R_j, W_j\right\} = \left\{P_j^*, p_j^*, R_j^*, W_j^*\right\}$ is the only competitive equilibrium.

$\square$

Proposition 2. *The optimal allocation can be implemented by the transfers*

$$t_j^{\theta*} = \sum_{\theta'} \left(\gamma_{\theta,\theta'}^{P,j} w_j^{\theta'*} + \gamma_{\theta,\theta'}^{A,j} x_j^{\theta'*}\right) \frac{L_j^{\theta'*}}{L_j^{\theta*}} - \left(b^\theta \Pi^* + E^\theta\right), \tag{A.12}$$

where the terms $\left(x_j^{\theta*}, w_j^{\theta*}, L_j^{\theta*}, \Pi^*\right)$ are the outcomes at the efficient allocation, and $\left\{E^\theta\right\}$ are constants equal to the multipliers on the resource constraint of each type in the planner's allocation.

Proof. The result follows from combining (1) with the expressions for labor demand (13) and expenditure per capita (15). Using the labor demand condition (13) and the fact that the aggregator of efficiency units $N(\cdot)$ has constant returns to scale, we obtain that the value of the marginal product of labor can be written as function of wages, employment and elasticities:

$$W_j \frac{dN_j}{dL_j^\theta} = w_j^\theta \left(1 + \gamma_{\theta,\theta}^{P,j}\right) + \sum_{\theta' \neq \theta} w_j^{\theta'} \left(\frac{L_j^{\theta'}}{L_j^\theta}\right) \gamma_{\theta,\theta'}^{P,j}. \tag{A.13}$$

Combining this expression with the definition of $\gamma_{\theta,\theta'}^{A,j}$, we can re-write (17) as follows:

$$w_j^\theta - x_j^\theta + \sum_{\forall \theta'} \left(\gamma_{\theta,\theta'}^{P,j} w_j^{\theta'} + \gamma_{\theta,\theta'}^{A,j} x_j^{\theta'}\right) \frac{L_j^{\theta'}}{L_j^\theta} = E^\theta. \tag{A.14}$$

Combining this last expression with (15) gives the result.

$\square$

Proposition 3. *The planning problem is concave if $\Gamma^A > \Gamma^P$, $\Gamma^A \geq 0$ and $\gamma_{\theta,\theta'}^A > 0$ for $\theta \neq \theta'$. Under a single worker type ($\Theta = 1$), the planning problem is quasi-concave if $1 + \gamma^A > \left(1 + \gamma^P\right)\left[\frac{1-\alpha_C}{1+D} + \alpha_C\right]$.*

Proof. We consider the following planning problem defined in section 3.3:

$$\max u^\theta$$
$$\text{s.t.: } u^{\theta'} = \underline{u}^{\theta'} \quad \text{for } \theta' \neq \theta$$
$$u^{\theta'} \in \mathcal{U} \quad \text{for all } \theta'$$

where θ is a given type, $\mathcal{U}$ is the set of attainable utility levels $\left\{u^\theta\right\}$ and $\underline{u}^{\theta'}$ for $\theta' \neq \theta$ is an arbitrary attainable utility level for group θ'. $\mathcal{U}$ is characterized by a set of feasibility constraints which are defined in the main text, and which we come back to below. We show here that this problem, noted $\mathcal{P}$, can be recast as a concave problem, under the condition stated in proposition 2. Therefore, a local maximum of $\mathcal{P}$ is necessarily its unique global maximum. The planning problem $\mathcal{P}$ can be recast as the following equivalent problem $\mathcal{P}'$, after simple algebraic manipulations:

$$\max_{\left\{v^\theta, U_j^\theta, C_j^\theta, H_j^\theta, \tilde{L}_j^\theta, \tilde{N}_j^k, I_j^k, Q_{ij}, M_j, S_j\right\}} v^\theta \tag{A.15}$$

subject to the set of constraints $\mathcal{C}$:

$$v^\theta - \mathcal{F}\left(\frac{U_j^\theta \prod_{\theta' \neq \theta} \left(\tilde{L}_j^{\theta'}\right)^{\frac{\gamma_{\theta',\theta}^A}{1+\Gamma^P}}}{\left(\tilde{L}_j^\theta\right)^{\frac{1-\gamma_{\theta,\theta}^A}{1+\Gamma^P}}}\right) \leq 0 \text{ for all } j; \tag{A.16}$$

$$\underline{v}^{\theta'} - \mathcal{F}\left(\frac{U_j^{\theta'} \prod_{\theta'' \neq \theta'} \left(\tilde{L}_j^{\theta''}\right)^{\frac{\gamma_{\theta'',\theta'}^A}{1+\Gamma^P}}}{\left(\tilde{L}_j^{\theta'}\right)^{\frac{1-\gamma_{\theta',\theta'}^A}{1+\Gamma^P}}}\right) \leq 0 \text{ for all } j \text{ and } \theta' \neq \theta; \tag{A.17}$$

$$U_j^\theta - U\left(C_j^\theta, H_j^\theta\right) \leq 0 \tag{A.18}$$

$$\sum_i d_{ji} Q_{ji} - \left(b_Y^N \left(N_j^Y\right)^{\beta_Y} + b_Y^I \left(I_j^Y\right)^{\beta_Y}\right)^{\frac{1}{\beta_Y}} \leq 0 \text{ for all } j, i; \tag{A.19}$$

$$\sum_\theta C_j^\theta + \left(I_j^Y\right) + \left(I_j^H\right) - Q\left(Q_{1j}, .., Q_{Jj}\right) \leq 0 \text{ for all } j; \tag{A.20}$$

46

$$\sum_\theta H_j^\theta - \left(b_H^N \left(N_j^H \right)^{\beta_H} + b_H^I \left(I_j^H \right)^{\beta_H} \right)^{\frac{1}{\beta_H}} \leq 0 \tag{A.21}$$

$$M_j - \left[\sum_\theta \left(Z_j^\theta \prod_{\theta'} \left(\widetilde{L}_j^{\theta'} \right)^{\frac{\gamma_{\theta',\theta}^P}{1+\Gamma^P}} \left(\widetilde{L}_j^\theta \right)^{\frac{1}{1+\Gamma^P}} \cdot \right)^\rho \right]^{\frac{1}{\rho}} \leq 0 \text{ for all } j; \tag{A.22}$$

$$N_j^Y + N_j^H - M_j \leq 0 \tag{A.23}$$

$$\sum_j \left(\widetilde{L}_j^\theta \right)^{\frac{1}{1+\Gamma^P}} - L^\theta = 0 \text{ for all } \theta \tag{A.24}$$

To reach these expressions, we have introduced the auxiliary variables M_j, S_j, and U_j^θ we have used the following change of variables

$$v^\theta = \mathcal{F}\left(u^\theta \right)$$
$$H_j^\theta = L_j^\theta h_j^\theta$$
$$C_j^\theta = L_j^\theta c_j^\theta$$
$$\widetilde{L}_j^\theta = \left(L_j^\theta \right)^{1+\Gamma^P} \text{ for all } j \in 1, ..., N \text{ and } \theta,$$

where the function $\mathcal{F}(.)$ is defined by $\mathcal{F}(x) = -x^b$ for $b = \frac{1+\Gamma^P}{\Gamma^P - \Gamma^A}$. Problems $\mathcal{P}$ and $\mathcal{P}'$ are equivalent: any solution to $\mathcal{P}'$ is a solution to $\mathcal{P}$ and vice-versa. We then consider the relaxed problem $\mathcal{P}''$ that is identical to $\mathcal{P}'$ except that the last constraint of $\mathcal{P}'$ is relaxed into an inequality constraint:

$$L^\theta - \sum_j \left(\widetilde{L}_j^\theta \right)^{\frac{1}{1+\Gamma^P}} \leq 0 \text{ for all } \theta. \tag{A.25}$$

We now show that problem $\mathcal{P}''$ has a concave objective and convex constraints under the assumptions of proposition 2. To that end, we show that under these assumptions, each constraint of $\mathcal{P}''$ is convex.

Consider first constraints (A.16) and (A.17), and examine specifically the expression:

$$f_j^\theta(U_j^\theta, \left\{ L^\theta \right\}, \left\{ L^{\theta'} \right\}) = \frac{U_j^\theta \prod_{\theta' \neq \theta} \left(\widetilde{L}_j^{\theta'} \right)^{\frac{\gamma_{\theta',\theta}^A}{1+\Gamma^P}}}{\left(\widetilde{L}_j^\theta \right)^{\frac{1-\gamma_{\theta,\theta}^A}{1+\Gamma^P}}} \tag{A.26}$$

This expression is a multivariate function of the form $f(y,z) = \frac{\prod_{i=1}^k y_i^{a_i}}{z^b}$ where $a_i > 0$, $b > 0$ and $\sum_{i=1}^k a_i < b$. By proposition 11 of Khajavirad et al. (2014), such functions are G-concave, meaning that the function $G(f(y,z))$ is concave in (y,z), for functions $G(x)$ that are concave transforms of $-x^{\frac{1}{\sum a_i - b}}$. Assumptions made on parameter values in Proposition 3 ensure that $\gamma_{\theta',\theta}^A \geq 0$ for all $\theta' \neq \theta$ and $1 + \frac{\gamma_{\theta',\theta}^A}{1+\Gamma^P} < \frac{1-\gamma_{\theta,\theta}^A}{1+\Gamma^P}$, which follows from $\Gamma^A > \Gamma^P$. Therefore, by Proposition 11 of Khajavirad et al. (2014), the transformation $G_\theta(x) = -x^{\frac{1+\Gamma^P}{\Gamma^P - \left(\sigma_\theta + \sum_{\theta'} \gamma_{\theta',\theta}^A \right)}}$ ensures that $G_\theta(f_j^\theta(.))$ is concave. Finally, given the definition of Γ^A, $\mathcal{F}(.)$ is a concave transform of $G_\theta(.)$. Therefore, (A.16) and (A.17) are convex for all θ.

Second, functions of the form $f(x_1, ..., x_n) = \left[\sum a_i x_i^\beta \right]^\rho$ are concave whenever $\beta \in (0,1)$ and $\rho\beta \leq 1$. Therefore, constraints (A.19), (A.21) and (A.25) are convex.

The constraint (A.18) is convex because $U(.)$ is concave. The constraint (A.20) is convex because the aggregator $Q(.)$ is concave.

Next, consider the constraint (A.22). The second term is the negative of a composition of an increasing CES function with exponent $\rho \leq 1$, which is concave, and a series of functions of the form

$$f(x_1, ..., x_\Theta) = \prod_{\theta'} \left(x^{\theta'} \right)^{\frac{\gamma_{\theta',\theta}^P}{1+\Gamma^P}} \left(x^\theta \right)^{\frac{1}{1+\Gamma^P}} .$$

As concave transforms of a geometric mean, these functions are concave, whenever $\frac{1+\sum_{\theta'} \gamma_{\theta',\theta}^P}{1+\Gamma^P} \in (0,1)$. This restriction holds by definition of Γ^P. We finally invoke that the vector composition of a concave function that is increasing in all its elements with a concave function is concave. Therefore, constraint (A.22) is convex. Finally, constraint (A.23) is linear hence convex.

It follows that the relaxed problem $\mathcal{P}''$ is a maximization problem with concave objective and convex inequality constraints. It admits at most one global maximum, and a vector satisfying its first order conditions is necessarily the global maximum. If at this unique optimal point for $\mathcal{P}''$ the relaxed constraint (A.25) binds, so that (A.24) holds, we guarantee that the solution to $\mathcal{P}''$ is also the unique global maximizer of $\mathcal{P}'$ and the unique global maximizer of the equivalent problem $\mathcal{P}$.[52]

We now specialize to the case of a single type of workers ($\Theta = 1$) where the decreasing returns to scale in the production of housing help make the problem concave. The relaxed planner's problem $\mathcal{P}''$ can be further simplified in this case to:

$$\max_{\left\{v^\theta, U_j^\theta, C_j^\theta, H_j^\theta, \widetilde{L}_j^\theta, \widetilde{N}_j^k, I_j^k, Q_{ij}, M_j, S_j\right\}} \min_j \frac{\left(C_j^\theta\right)^{\alpha_C} \left(\widetilde{H}_j^\theta\right)^{\frac{1-\alpha_C}{1+d_{H,j}}}}{\left(\widetilde{L}_j^\theta\right)^{\frac{1-\gamma_{\theta,\theta}^A}{1+\Gamma^P}}}$$

subject to the constraints (A.19), (A.20), (A.22), (A.23) and (A.25), which are unchanged except that they now hold for only one group. The modified constraint for housing production is:

$$\widetilde{H}_j^\theta - \left(b_H^N \left(\widetilde{N_j^H}\right)^{\beta_H(1+D)} + b_H^I \left(\widetilde{I_j^H}\right)^{(1+D)\beta_H}\right)^{\frac{1}{\beta_H} \frac{1}{1+D}} \leq 0 \tag{A.27}$$

where we have used the following change of variable $\widetilde{H}_j^\theta = \left(H_j^\theta\right)^{1+d'_{H,j}}$. The modified housing market constraint (A.27) is convex. The objective of the planner is quasi-concave as the minimum of a ratio of a concave and a convex function, as long as $(1-\alpha_C)\frac{1}{1+d'_{H,j}} + \alpha_C \leq \frac{1-\gamma_{\theta,\theta}^A}{1+\Gamma^P}$ in each city. The constraints are convex. Therefore, the problem is a quasi-concave maximization problem as long as the parameter restriction in (ii) holds.

$\square$

A.2 Inefficiency in a Simple Case with Homogeneous Workers and Constant Elasticities

We consider a special case of our model with homogeneous workers ($|\Theta| = 1$), homogeneous products ($Q(\cdot) = \sum_j Q_{ji}$), no trade costs ($d_{ij} = 1$), no intermediate inputs in production ($Y_j = F(N_j^Y)$), exogenous amenities (a_j exogenously given), and no valuation for non-traded goods ($u(c,h) = c$). We assume further that productivity spillovers have constant elasticity: $z_j(L_j) = L_j^{\gamma^P}$ so that $N_j^Y = z_j(L_j)L_j$. Finally, we assume that tradable good production features decreasing returns to scale: $F(N_j^Y) = (N_j^Y)^{1-\beta}$ where $0 \leq \beta \leq 1$. In equilibrium, all efficiency units of labor are employed in the production of tradeable goods, so $Y_j = (z_j(L_j)L_j)^{1-\beta}$.

Consider first the market allocation, possibly with a constant labor subsidy s and lump sum transfer T, so that the transfer to workers in location j is $t_j = sw_j + T$. The definition of expenditure (A.7) gives expenditure per worker in location j equal to $x_j = w_j + t_j + \pi$, where $\pi = \frac{\Pi}{L}$ is the share of profits owned by each worker. In the market allocation, $a_j x_j = u$ for all populated locations. Normalizing the price of the tradeable good to 1, the wage in (13) is $w_j = (1-\beta)\frac{Y_j}{L_j}$. As a result, a market allocation is given by $\{\{L_j\}, u, \pi\}$ such that:

$$a_j\left((1+s)(1-\beta)\frac{Y_j}{L_j} + T + \pi\right) = u \text{ for all } j \tag{A.28}$$

and the goods and labor markets clear. In turn, the planner's problem can be stated as

$$\max u = \frac{Y\left(\{L_j\}\right)}{\sum \frac{L_j}{a_j}}, \tag{A.29}$$

where $Y\left(\{L_j\}\right) \equiv \sum Y_j\left(L_j\right)$ is aggregate output, subject to the labor market clearing constraint with multiplier λ.[53] The solution to the planner's problem consists of $\{\{L_j\}, u, \lambda\}$ such that

$$a_j \left(\left(1 + \gamma^P\right)(1 - \beta) \frac{Y_j}{L_j} - \lambda \frac{Y}{u} \right) = u \text{ for all } j \tag{A.30}$$

and the goods and labor markets clear.

Comparing (A.28) and (A.30), it follows that the market allocation is generically inefficient if there is dispersion in amenities and non-zero productivity spillovers α, and efficient otherwise. To see why, note that in the absence of government policies ($s = T = 0$), equating (A.28) and (A.30) leads to the condition that $\pi + \lambda \frac{Y}{u} = (1 - \beta) \gamma^P \frac{Y_j}{L_j}$, i.e. the planner and market solution can only coincide only if either there are no spillovers ($\gamma^P = 0$) or if income per worker $\frac{Y_j}{L_j}$ is equalized across locations. But if $s = T = 0$, a constant income per worker cannot be a solution to (A.30) as long as there is dispersion in amenities. However, a constant policy scheme over space (s, T) with $s = \gamma^P$ and $T = -\pi - \lambda \frac{Y}{u}$ restores efficiency.[54] Without differences in amenities, a constant income per worker across locations is a solution to both the planner's and the market allocation, in which case the decentralized allocation is efficient.

A.3   Equivalence with Monopolistic Competition

Consider the economic geography environment from Section 4.3. As a reminder, that environment starts from the general model from Section 3 and imposes only one labor type, inelastic housing supply ($H_j\left(N_j^H, I_j^H\right) = H_j$ is a constant), and only labor used in production of traded goods ($Y_j\left(N_j^Y, I_j^Y\right) = N_j^Y = N_j = z_j\left(L_j\right)L_j$). Now suppose that, in addition, the production structure in the traded sector is the same as in Krugman (1980): in each location j, M_j homogeneous plants produce differentiated varieties with constant elasticity of substitution κ across them, and setting up a plant in location j requires F_j units of labor. The resulting environment corresponds to the Redding (2016) or to Helpman (1998) in the absence of individual preference shocks ($\sigma = 0$).

We now show that the competitive allocation of such an extended model, as well as their normative implications, are equivalent to the model with homogeneous products analyzed in Section 4.3 under an aggregate production function equal to:

$$\widetilde{Y_j}\left(L_j\right) = K_j \left(z_j\left(L_j\right) L_j\right)^{\frac{\kappa}{\kappa-1}}, \tag{A.31}$$

where $K_j \equiv \frac{\kappa-1}{\kappa}\left(\kappa F_j\right)^{\frac{1}{1-\kappa}}$ is a constant. Therefore, a monopolistic competition model with no productivity spillovers is equivalent to a homogeneous-product model with perfect competition and spillover elasticity equal to $\gamma^P = \frac{1}{\kappa-1}$. This property relates to the result, dating back to at least Abdel-Rahman and Fujita (1990) and also shown by Allen and Arkolakis (2014) within their model, that CES product differentiation with monopolistic competition has the same aggregate implications as a constant-elasticity aggregate production function with increasing returns. In our context, we must also demonstrate that the equivalence extends to the welfare implications summarized in Proposition 1.

The key reason why this equivalence holds is that under CES preferences the number of producers M_j and the bilateral trade flows are efficient given the allocation of labor $\{L_j\}$, as it is well known. Therefore, the labor allocation remains the only inefficient margin and our propositions and results from Section 4.3 go through. We note that these properties would not go through under monopolistic competition outside of CES. In that case, the entry and bilateral pricing decisions would be inefficient.

---

[53]The planner maximizes u subject to $u = a_j c_j$ for all j, the goods market clearing constraint $\sum L_j c_j = \sum Y_j$, and the labor market clearing constraint. Combining these constraints leads to $u = \frac{\sum Y_j}{\sum \frac{L_j}{a_j}}$, corresponding to (A.29).

[54]Note that this formula is a direct application of proposition 2.

Environment We start by describing how the physical environment of this model differs from the environments from Section 3. Now, the input to the aggregator $Q(\{Q_{ji}\})$ is $Q_{ji} = M_j^{\frac{\kappa}{\kappa-1}} q_{ji}$, where M_j is the number of plants in j and q_{ji} is the quantity exported by each of these from j to i. The feasibility constraint for traded goods (4) becomes $z_j(L_j) L_j = M_j \left(\sum_i \tau_{ji} q_{ji} + F_j \right)$ to account for the use of labor in setting up plants. Combining these two expressions, that constraint can be further expressed:

$$M_j^{\frac{1}{\kappa-1}} (z_j(L_j) L_j - F_j M_j) = \sum_i \tau_{ji} Q_{ji}. \tag{A.32}$$

Competitive Equilibrium Now we describe how the market allocation differs from the baseline environments. First, the optimization conditions (11) is replaced by:

$$\max \sum_i (\tilde{p_{ji}} - \tau_{ji} W_j) q_{ji} \tag{A.33}$$

subject to $q_{ji} = Q_{ji} \left(\frac{\tilde{p_{ji}}}{p_{ji}} \right)^{-\kappa}$, where $p_{ji} = M_j^{\frac{1}{1-\kappa}} \tilde{p_{ji}}$ is the price index corresponding to the exports from j to i and $\tilde{p_{ji}}$ is the price at which each firm from j sells in i. The solution to this problem yields the standard constant markup rule, $\tilde{p_{ji}} = \tau_{ji} \frac{\kappa}{\kappa-1} W_j$. We have as before that the price in location i of the aggregate traded good from j, p_{ji}, can be expressed according to the "mill pricing" rule as $\tau_{ji} p_j$, where now the price index corresponding to the domestic sales of traded goods in j is:

$$p_j \equiv M_j^{\frac{1}{1-\kappa}} \frac{\kappa}{\kappa - 1} W_j. \tag{A.34}$$

As a result, condition (14) still determines the flows in the competitive equilibrium. Combining these pricing rules with (A.33), imposing zero profits and using (4) we obtain the number of producers in a competitive allocation:

$$M_j = \frac{z_j(L_j) L_j}{\kappa F_j}. \tag{A.35}$$

And further combining with (A.32), we can write

$$\widetilde{Y_j}(L_j) = \sum_j \tau_{ij} Q_{ij} \tag{A.36}$$

for $\widetilde{Y_j}$ given in (A.31).

We conclude that the competitive allocation can be represented as in the model without product differentiation from Definition 1 under the restrictions from Section 4.3 and assuming the aggregate production function $\widetilde{Y_j}(L_j)$. I.e., it is given by quantities $\{c_j, h_j, L_j, Q_{ij}, L_j\}$ and prices P_j, R_j, p_j, such that: (i) consumers optimize (i.e., c_j, h_j are a solution to (8) given expenditures x_i); (ii) trade flows are given by (14); (iii) employment L_j is consistent with the spatial mobility constraint (10); and (iv) all markets clear, i.e. (2), (3) and (A.36) hold.[55]

Planning Problem In turn, the planning problem from Definition (2) is now associated with the Lagrangian

$$\mathscr{L} = u - \sum_j \omega_j \left(u - a_j(L_j) U(c_j, h_j) \left(\frac{L_j}{L} \right)^{-\sigma} \right)$$

$$- \sum_j p_j^* \left(\sum_i d_{ji} Q_{ji} - M_j^{\frac{1}{\kappa-1}} (z_j(L_j) L_j - F_j M_j) \right)$$

$$- \sum_j P_j^* (L_j c_j - Q(Q_{1j}, .., Q_{Jj})) - \sum_j R_j^* (L_j h_j - H_j) - W \left(\sum_j L_j - L \right) + ... \tag{A.37}$$

---

[55]Note that the definition of the competitive allocation can dispense with the wage W_j, which can be determined residually from (A.34).

Relative to Definition 2, now the planner also chooses the number of firms M_j in each location and faces the constraint (A.32) instead of (4). We readily see that entry is efficient by noting that the first-order condition with respect to M_j implies (A.35). As a result, the market clearing constraint in the second line of (A.37) can be replaced by (A.36). The resulting planning problem is equivalent to Definition 2 applied to the economic geography model in Section 4.3 under the production function $\widetilde{Y_j}(L_j)$.

A.4   Spillovers Across Locations

The Lagrangian of the planning problem described in the extension to spillovers across locations in Section 4.4 is a special case of (A.2), except that now the the supply of efficiency units in j is $N_j(\{L_{j'}\}) = z_j(\{L_{j'}\}) L_j$. Compared to our derivation of Proposition 1, the only difference is the first-order condition with respect to employment. Now, instead of (A.11) we reach:

$$\sum_{j'} W_{j'}^* \frac{dN_{j'}}{dL_j} + x_j^* \frac{L_j}{a_j} \frac{\partial a_j}{\partial L_j} = x_j^* + E. \tag{A.38}$$

In addition, instead of (A.13) we now have:

$$W_j \frac{dN_j}{dL_j} = \begin{cases} w_{j'} \frac{L_{j'}}{L_j} \gamma^{P,j,j'} & \text{if } j' \neq j, \\ w_j \left(\gamma^{P,j,j} + 1\right) & \text{if } j' = j. \end{cases} \tag{A.39}$$

Combining the last two expressions with (15) gives (28).

A.5   Preference Heterogeneity within Groups

The Lagrangian of planning problem described in the extension to preference draws in Section 4.4 is a special case of (A.2), except that now the spillover function $a_j^{\theta'}\left(L_j^1,..,L_j^\Theta\right)$ is replaced by $a_i^{\theta'}\left(L_i^\theta\right)^{-\sigma_\theta}$. Following the same steps as in the proof of Proposition 1, we find that condition (17) is extended to

$$W_j \frac{dN_j}{dL_j^\theta} + \sum_{\theta'} \frac{x_j^{\theta'} L_j^{\theta'}}{a_j^{\theta'}} \frac{\partial a_j^{\theta'}}{\partial L_j^\theta} = x_j^\theta \left(1 + \sigma_\theta\right) + E^\theta \quad \text{if } L_j^\theta > 0. \tag{A.40}$$

Following the same steps as in the proof of Proposition 2, we find that (20) is extended to

$$t_j^\theta = \gamma_{\theta,\theta}^{P,j} + \left(\gamma_{\theta,\theta}^{A,j} - \sigma_\theta\right) + \sum_{\theta' \neq \theta} \left(\gamma_{\theta,\theta'}^{P,j} w_j^{\theta'*} + \gamma_{\theta,\theta'}^{A,j} x_j^{\theta'*}\right) \frac{L_j^{\theta'*}}{L_j^{\theta*}} - \left(b^\theta \Pi^* + E^\theta\right). \tag{A.41}$$

The general-equilibrium structure underlying propositions 3 and 4 under the assumptions of the quantitative model can be expressed exactly as in the proof of Proposition 3 and as in the planning problem in relative changes from Section A.7 below, the only modification being that the term $\gamma_{\theta,\theta}^A$ is replaced by $\gamma_{\theta,\theta}^A - \sigma_\theta$.

A.6 Public Spending

The Lagrangian of the planning problem described in the extension to public spending in Section 4.4 is

$$
\begin{aligned}
\mathscr{L} = u &- \sum_j \omega_j L_j \left(u - a_j \left(G_j^U, L_j \right) U \left(c_j, h_j \right) \right) - \sum_j p_j^* \left(\sum_i d_{ji} Q_{ji} - Y_j \left(N_j^Y, I_j^Y \right) \right) \\
&- \sum_j P_j^* \left(I_j^G + L_j c_j + I_j^Y + I_j^H - Q \left(Q_{1j}, .., Q_{Jj} \right) \right) - \sum_j R_j^* \left(H_j^G + L_j h_j - H_j \left(N_j^H, I_j^H \right) \right) \\
&- \sum_j P_j^{G*} \left(G_j^Y + G_j^U - G_j \left(I_j^G, H_j^G \right) \right) - \sum_j W_j^* \left(N_j^I + N_j^H - N_j \left(G_j^Y, z_j \left(L_j \right) L_j \right) \right) \\
&- E \left(\sum_j L_j - L \right) + ...
\end{aligned}
\tag{A.42}
$$

Letting $x_j^P \equiv P_j^* c_j + R_j^* h_j$ be private expenditure, following the same steps as in the proof of Proposition 1 we find

$$
x_j^P = \frac{1 + \gamma^P}{1 - \gamma^A} w_j - \frac{E}{1 - \gamma^A}.
\tag{A.43}
$$

Combining the first order condition over c_j with optimization over government spending gives

$$
\left[G_j^U \right] : \qquad L_j x_j^P \gamma^{A,G} = P_j^{G*} G_j^U
$$

$$
\left[G_j^Y \right] : \qquad W_j^* N_j \varepsilon_{N,G} = P_j^{G*} G_j^Y
$$

where $\gamma^{A,G} \equiv \frac{\partial a_j}{\partial G_j} \frac{G_j}{a_j}$ and $\varepsilon_{N,G} = \frac{\partial N}{\partial G_j} \frac{G_j}{N_j}$ are elasticities of amenities and labor efficiency to government spending. Adding up the two previous equations, using the first-order conditions over I_j^G and H_j^G, and applying homogeneity of the the production function of government spending $G_j \left(I_j^G, H_j^G \right)$ we reach condition (31) in the text. Further combining (31) with (A.43) gives (30) in the text. In terms of the efficient implementation, under constant labor subsidies with a lump sum transfer individuals earn $x_j^P = w_j \left(1 + s \right) + \frac{\Pi}{L} + T$. Hence, the optimal subsidy must still be $s = \frac{\gamma^P + \gamma^A}{1 - \gamma^A}$, as discussed in Section 4.2 when there is no government spending. Given the value of s, the lump-sum transfer T is such that $\sum_j \left[\left(w_j s + x_j^G \right) \frac{L_j}{L} + T \right] = 0$, which ensures that each local government can finance spending per capita x_j^G.

A.7 Planning Problem in Relative Changes and Proof of Proposition 4

We show how to express the solution for the competitive allocation under an optimal new policy relative to an initial equilibrium consistent with Definition 1, and then define the planning problem that optimizes over the policy space.

Preliminaries We adopt the functional forms from Section 4.5. From the optimization problems (11) and (12) and market clearing in the housing market we obtain the following conditions:

$$
W_i N_i^Y = \left(1 - b_{Y,i}^I \right) p_i Y_i,
\tag{A.44}
$$

$$
W_i N_i^H = \frac{1 - b_{H,i}^I}{1 + d_{H,i}} \left(1 - \alpha_C \right) X_i.
\tag{A.45}
$$

These terms imply the non-traded labor share, $\frac{N_i^H}{N_i}$, as function of the share of gross expenditures over tradeable income $\frac{X_i}{p_i Y_i}$:

$$
\frac{N_i^H}{N_i} = \frac{\frac{1 - b_{H,i}^I}{1 + d_{H,i}} \frac{1 - \alpha_C}{1 - b_{Y,i}^I} \left(\frac{X_i}{p_i Y_i} \right)}{\frac{1 - b_H^I}{1 + d_{H,i}} \frac{1 - \alpha_C}{1 - b_{Y,i}^I} \left(\frac{X_i}{p_i Y_i} \right) + 1}.
\tag{A.46}
$$

Using (A.44) and (A.45) along with labor-market clearing (A.4), we can further express gross expenditures over tradeable income as a function the shares of wages in expenditures:

$$\frac{X_i}{p_i Y_i} = \frac{1 - b_{Y,i}^I}{\frac{W_i N_i}{X_i} - \frac{1 - b_{H,i}^I}{1 + d_{H,i}} (1 - \alpha_C)}. \tag{A.47}$$

We now re-formulate some of the equilibrium from Definition 1 conditions to include prices. Consider first the market clearing condition (5). Multiplying both sides by the price of the traded bundle P_j, letting $E_j^Y = P_j Q_j = P_j Q (Q_{1j}, .., Q_{Jj})$ be the gross expenditures in tradeable goods in j (used both as intermediate and for final consumption), using equilibrium in the housing market, and using optimality condition for the choice of intermediate inputs in the traded sector, we can re-write that condition as

$$E_j^Y = \left(\alpha_C + (1 - \alpha_C) \frac{b_{H,j}^I}{d_{H,j} + 1} \right) X_j + b_Y^I (p_j Y_j), \tag{A.48}$$

where $X_j = \sum_{\theta'} L_j^{\theta'} x_j^{\theta'}$ are the aggregate expenditures in region j. This condition says that aggregate expenditures in traded goods results from the aggregation of expenditures by consumers and final producers. Second, consider the market condition (4) for traded commodities. Multiplying both sides by the price of traded commodities at j, p_j, this condition is equivalent to

$$\sum_i s_{ji}^X = 1, \tag{A.49}$$

where $s_{ji}^X \equiv \left(\frac{E_i}{p_j Y_j} \right) s_{ji}^M$ is region i's share of j's sales of tradeable goods (i.e., the export share of i in j) and $s_{ji}^M \equiv \frac{p_{ji} Q_{ji}}{E_i}$ is region $j's$ share of i's purchases of tradeable goods (i.e., the import share of region j in i). Finally, aggregating the budget constraints of individual consumers gives

$$\sum_j s_{ji}^M \equiv 1. \tag{A.50}$$

Equilibrium in Relative Changes We now express the solution for the competitive allocation from Definition 1 under the new policy relative to an initial equilibrium. Consider a policy change that affects the equilibrium expenditure distribution $\{x_i^\theta\}$. We now show that the outcomes in the new equilibrium relative to the initial equilibrium are given by a set of changes in prices $\{\hat{P}_i, \hat{p}_i, \hat{R}_i\}$, wages $\{\hat{W}_i\}$, employment by group $\{\hat{L}_i^\theta\}$, supply of efficiency units $\{\hat{N}_i\}$, production of tradeable goods $\{\hat{Y}_i\}$, and utility levels $\{\hat{u}^\theta\}$ that satisfy a set of conditions given the change in expenditure per capita by group and location $\{\hat{x}_i^\theta\}$. The planner's problem in relative changes will then choose the optimal $\{\hat{x}_i^\theta\}$.

From the previous expressions we obtain the following system in relative changes:

$$\sum_j s_{ij}^X \left(\frac{\hat{p}_i}{\hat{P}_j} \right)^{1-\sigma} \hat{E}_j^Y = \hat{p}_i \hat{Y}_i \text{ for all } i, \tag{A.51}$$

$$\sum_j s_{ji}^M \left(\frac{\hat{p}_j}{\hat{P}_i} \right)^{1-\sigma} = 1 \text{ for all } i, \tag{A.52}$$

$$\left(1 - \frac{N_i^H}{N_i} \right) \hat{p}_i \hat{Y}_i + \frac{N_i^H}{N_i} \hat{X}_i = \hat{W}_i \hat{N}_i \text{ for all } i, \tag{A.53}$$

$$\hat{W}_i^{1 - b_{Y,i}^I} \hat{P}_i^{b_{Y,i}^I} = \hat{p}_i \text{ for all } i, \tag{A.54}$$

where $\hat{X}_j = \sum_\theta s_j^{X,\theta} \hat{x}_j^\theta \hat{L}_j^\theta$ is the change in aggregate expenditures by region and $s_j^{X,\theta}$ is group θ's share in the consumer expenditures in j in the initial equilibrium. Equations (A.51) and (A.52) follow from expressing (A.49) and (A.50) in relative changes and using the CES functional for $Q(\cdot)$ form (33). In condition (A.51), using (A.48)

implies that the change in expenditures in tradeable commodities is:

$$\hat{E}_j^Y = \hat{X}_j + \frac{b_Y^I}{\left(\alpha_C + (1 - \alpha_C) \frac{b_H^I}{d_{H,j}+1}\right) \left(\frac{X_j}{p_j Y_j}\right) + b_Y^I} \hat{p_j} \hat{Y}_j. \tag{A.55}$$

Condition (A.53) follows from expressing labor-market clearing (7) in relative changes together with (A.44) and (A.45), where the non-traded labor share $\frac{N_i^H}{N_i}$ is defined in (A.46). Condition (A.54) follows from optimization of producers of tradeable commodities.

The system (A.51) to (A.54) defines a solution for $\left\{\hat{P}_j, \hat{p}_j, \hat{Y}_j, \hat{W}_j\right\}$ given the change in the number of efficiency units $\hat{N}_i$ and expenditures in each region $\hat{X}_i$, and independently from heterogeneity across groups or spillovers. Heterogeneous groups and spillovers enter through $\hat{N}_i$.

To reach an explicitly expression for $\hat{N}_i$, we first note that the labor demand expression in the market allocation (13) allows us to back out the efficiency of each group:

$$z_i^\theta = \frac{w_i^\theta}{W_i} \left(\frac{L_i^\theta}{N_i}\right)^{1-\rho}. \tag{A.56}$$

Expressing the CES functional form for the aggregation of labor types in (36) in relative changes and using (A.56) we obtain:

$$\hat{N}_i = \left(\sum_\theta s_i^{W,\theta} \left(\hat{L}_i^\theta\right)^\rho \prod_{\theta'} \left(\hat{L}_i^{\theta'}\right)^{\gamma_{\theta',\theta}^P}\right)^{\frac{1}{\rho}}, \tag{A.57}$$

where $s_j^{W,\theta} = \frac{w_j^\theta L_j^\theta}{\sum_{\theta'} w_j^{\theta'} L_j^{\theta'}}$ is group θ share of wages in city j. This expression relates the total change in efficiency units in a location to the distribution of wage bills in the observed allocation, the changes in employment by group, and the production function and spillover elasticity parameters.

The change in the number of workers $\left\{\hat{L}_i^\theta\right\}$ of each type in every location that is initially populated must also be consistent with the spatial mobility constraint, (10),

$$\hat{u}^\theta = \prod_{\theta'} \left(\hat{L}_i^{\theta'}\right)^{\gamma_{\theta',\theta}^A} \frac{\hat{x}_i^\theta}{\hat{P}_i^{\alpha_C} \hat{R}_i^{1-\alpha_C}}. \tag{A.58}$$

In this expression, $\hat{R}_i$ is the change in the price of non-traded goods in location i. Solving for the equilibrium in the market for non-traded goods, after some manipulations this relative price can be expressed as solely a function of the price of the own traded good, the price index of traded commodities, of aggregate expenditures in i:

$$\hat{R}_i = \left(\hat{p}_i^{\frac{1-b_{H,i}^I}{1-b_{Y,i}^I}} \hat{P}_i^{b_{H,i}^I - b_{Y,i}^I \frac{1-b_{H,i}^I}{1-b_{Y,i}^I}} \hat{X}_i^{d_{H,i}}\right)^{\frac{1}{1+d_{H,i}}}. \tag{A.59}$$

Finally, the national labor market must clear for each labor type is

$$\sum_j s_j^{L,\theta} \hat{L}_j^\theta = 1 \text{ for all } \theta, \tag{A.60}$$

where $s_j^{L,\theta} = \frac{L_j^\theta}{\sum_{\theta'} L_j^{\theta'}}$ is group θ's share of employment in city j.

In sum, an equilibrium in changes given a change in expenditure per capita $\left\{\hat{x}_j^\theta\right\}$ consists of $\left\{\hat{P}_i, \hat{p}_i, \hat{Y}_i, \hat{W}_i, \hat{N}_j, \hat{L}_j^\theta, \hat{R}_i, \hat{u}^\theta\right\}$ such that equations (A.51) to (A.60) hold. These equations conform a system of $5J + \Theta J + \Theta$ equations in equal number of unknowns, where J is the number of locations and Θ is the number of types.

Planner's Problem in Relative Changes In the implementation, we solve an optimization over $\left\{\hat{x}_j^\theta\right\}$ subject to $\left\{\hat{P}_i, \hat{p}_i, \hat{Y}_i, \hat{W}_i, \hat{N}_j, \hat{L}_j^\theta, \hat{R}_i, \hat{u}^\theta\right\}$ consistent with (A.51) to (A.60) in order to maximize the utility of a given

group θ, $\hat{u}^\theta$, subject to a lower bound for the change in utility of the other groups ($u^{\hat{\theta}'} \geq \underline{u}^{\hat{\theta}'}$ for $\theta' \neq \theta$). This problem (call it $\mathcal{P}_2''$) differs formally from the baseline problem in Definition 2 (call it $\mathcal{P}_2$) for two reasons. First, it features prices, expenditures and incomes rather than being expressed in terms of quantities alone, as in conditions (A.44) to (A.50). We denote by $\mathcal{P}_2'$ an intermediary problem expressed in terms of income and expenditure rather than quantities, but still in levels. Second, $\mathcal{P}_2''$ is expressed in changes relative to an initial equilibrium rather than in levels. We show here that the two problems are nevertheless equivalent. Therefore, the problem that we implement has a unique maximizer under the conditions of Proposition 2.

To see that the two problems have the same solutions, we first focus on the first order conditions of problem $\mathcal{P}_2$ and compare them to the problem in levels $\mathcal{P}_2'$ expressed in income and expenditures terms rather than in quantities. Conditions (A.3) and (A.5) define the Lagrange multipliers corresponding to good and factor prices for $\mathcal{P}_2$. They are identical to the price index definition constraint of problem $\mathcal{P}_2'$. Furthermore, manipulating these equations together with the constraints expressed in quantities leads to the constraints expressed in terms of income and expenditure. Therefore, a vector satisfies the first order conditions for $\mathcal{P}_2$ if and only if it satisfies the first order conditions for $\mathcal{P}_2$'. Then, note that the problem in relative changes stated here is simply the problem $\mathcal{P}_2'$ modified through the changes of variable $x \to x_o \hat{x}$ for all variables, where x_o is a constant corresponding to the observed data and $\hat{x}$ the optimization variable in $\mathcal{P}_2'$. The problem in relative changes considered here and the problem $\mathcal{P}_2'$, and in turn problem $\mathcal{P}_2$, have therefore the same solutions, subject to the appropriate change of variables. In particular, a point that satisfies the first order conditions under the conditions of Proposition 3 is the (unique) global maximizer for both problems.

Proof of Proposition 4 Proposition 4 follows from inspecting (A.51) to (A.60) under the planner's problem in relative changes defined above. Note that, given the elasticities $\{\alpha_C, \rho, b_{Y,j}^I, b_{H,j}^I, d_{H,j}\}$, and as long as $b_Y^I > 0$, computing the change in tradeable expenditures requires information about gross expenditures over tradeable income, $\frac{X_j}{p_j Y_j}$. This information is also needed to compute the non-traded labor share $\frac{N_i^H}{N_i}$ in (A.53). However, as shown in (A.46) and (A.47), $\frac{X_j}{p_j Y_j}$ can be constructed from the elasticities $\{\alpha_C, b_Y^I, b_H^I, d_{H,j}\}$ and the share of wages in gross expenditures, $\frac{W_i N_i}{X_i}$.

B   Data Appendix

We detail the construction of the variables used to implement the counterfactuals. We rely on three primary data sources: i) BEA regional economic accounts, CA4 Personal Income and Employment by Major Component (https://www.bea.gov/regional/downloadzip.cfm); ii) estimates of disposable income by MSA from Dunbar (2009) based on BEA regional economic accounts;[56] iii) March CPS based on the IPUMS-CPS, ASEC 2007 sample.[57]

B.1   Appendix to Section 5.1 (Data)

MSA-Level Outcomes We first extract from Dunbar (2009) the following information: *population*, *personal income*, and *personal taxes* paid by MSA, in 2007. To split personal income by source of income, we merge this data with the BEA Regional Economic Accounts. We compute the share of personal income corresponding to each possible source: labor income, capital income, and transfers. Specifically, we measure labor income as BEA's *earning by place of work*;[58] capital income as the sum of *proprietor's income*, and *dividends, interests and rents*; and transfers as *current transfer receipts*.[59] Combining these shares with the total personal income and taxes by MSA from Dunbar (2009) provides us with a measure of labor income, capital income, transfers and taxes at the MSA level.

---

[56]https://https://www.bea.gov/papers/xls/dpi_msa_working_paper_2001_2007_results.xls

[57]https://cps.ipums.org/cps/index.shtml

[58]The BEA's earning by place of work is comprised of: wages and salaries, supplements to wages and salaries, proprietor's income, net of contributions for government social insurance, plus adjustment for residence.

[59]Current transfer receipts is defined as the sum of government social benefits and net current transfer receipts from business (https://www.bea.gov/glossary/glossary_p.htm).

Break-Down By Skill Group We split these totals at the MSA level into two groups, high skill and low skill. To that end, we use March CPS data, part of the Integrated Public Use Microdata Series (Flood et al., 2017). The March CPS reports, at the individual level, income by source (labor, capital and transfers), MSA of residence and level of education. Consistent with Diamond (2016), we define as high skill those workers who have completed 4 years of college, or more; and as low skill those who have completed less than 4 years of college, or have not gone to college. We aggregate individual level data from the March CPS at the MSA-group level, to get an MSA-level estimate of capital income, labor income, transfers and taxes by group, as well as the population of both groups.[60] We do not use this information directly, as the MSA aggregates from individual-level data might be noisy, in particular for smaller MSAs. Instead, we use this information to construct the shares of the MSA-level outcomes from the BEA described above corresponding to each group of workers. That is, for each MSA i, we compute $s_i^L = \frac{\tilde{X}_i^L}{\tilde{X}_i^L + \tilde{X}_i^H}$ where $\tilde{X}_i^\theta$ denotes total MSA i-group θ level capital income, labor income, transfers, taxes or population in the march CPS data. We use this share s_i^L, together with the MSA-level dataset for income described above, to build our measure $X_i^\theta = X_i s_i^\theta$ of MSA-group level population, labor income, capital income, transfers and taxes. We also compute the corresponding per-capita measures for each MSA-group: $x_i^\theta = \frac{X_i^\theta}{L_i^\theta}$.

Controlling for Heterogeneity within Groups Before applying it in the quantification we purge the raw data described above from compositional effects across MSAs. We use Census data (IPUMS) to obtain the share of individuals with the following characteristics for each MSA-skill group: age by bins :<20, 20-40, 40-60, >60; detailed level of educational attainment: less that 8th grade, grade 9-12, some college (those are relevant for the low skill group) and bachelor, masters or professional degree (for the high skill group); share black; share male ; share unemployed, share out of the labor force; and share working in manufacturing, services, agriculture. We then proceed as follows: denoting by x_i^θ the per-capita measure $\frac{X_i^\theta}{L_i^\theta}$ (where X_i^θ can stand for labor income, capital income, transfers or taxes) in MSA i and group θ, we run the following MSA level regression, separately for each group θ:

$$x_i^\theta = x_0^\theta + \sum_j \beta_j^\theta DEM_{ij}^\theta + \varepsilon_i^\theta, \tag{A.61}$$

where DEM_{ij}^θ is the demographic variable j enumerated above in MSA i and for group θ. The coefficients β_j^θ measures how demographic characteristic j correlates with x_i^θ across cities within group θ. We then adjust the observed x_i^θ from compositional differences across cities measured as deviations from the population mean:

$$\hat{x}_i^\theta \equiv x_i^\theta - \sum_j \hat{\beta}_j^\theta \left(DEM_{ij}^\theta - \overline{DEM_j^\theta} \right) \tag{A.62}$$

where $\hat{\beta}_j^\theta$ is the estimate from (A.61) and $\overline{DEM_j^\theta} \equiv \frac{1}{I} \sum_i DEM_{ij}^\theta$.[61] The corresponding MSA-level variable is $\hat{X}_i^\theta = \hat{x}_i^\theta L_i^\theta$. The resulting data $(\hat{X}_i^\theta, \hat{x}_i^\theta, L_i^\theta)$ is our MSA-group level dataset, where X stands for labor income, capital income, transfers and taxes.

Expenditure per Capita We construct expenditure by group and by MSA, x_i^θ in the model, as disposable income by group. Disposable income is

$$x_i^\theta = w_i^\theta(1 + \tau_i^\theta) + T_i^\theta + b^\theta \Pi^H. \tag{A.63}$$

---

[60]Specifically, we aggregate the following categories to measure capital income: income from interest, from dividends, from rents. We aggregate the following categories to measure labor income: wage and salary income, non-farm business income, farm income, income from worker's compensation, alimony and child support. We aggregate the following categories to measure transfers: welfare income, social security income, income from SSI, income from unemployment benefits, income from veteran's, survivor's, disability benefit, income from educational assistance. We aggregate the following categories to measure taxes paid: federal income tax liability, after all credits, and state income tax liability, after all credits.

[61]I.e., we define $\hat{x}_i^\theta \equiv \hat{x}_0^\theta + \sum_j \hat{\beta}_j^\theta \overline{DEM_j^\theta} + \hat{\varepsilon}_i^\theta$, where $\hat{\varepsilon}_i^\theta$ is the estimated residual from (A.61).

The variables $\{w_i^\theta, \tau_i^\theta, T_i^\theta\}$, respectively labor income per capita, tax paid per capita (where $\tau_i^\theta < 0$ is a tax), and transfer received per capita, are directly taken from the BEA/CPS dataset constructed above. We measure b^θ as the average fraction of national capital income owned by each type θ worker in BEA/CPS dataset. This step gives $b^S L^S = 0.59$ and $b^U L^U = 0.41$. Finally, we set a value for national profits and returns to land Π^H that is consistent with the general equilibrium of the model. Using profit maximization in the housing sector and market clearing in the non-tradeable sector we obtain the following expression for Π^H as function of calibrated elasticities and observable outcomes:

$$\Pi^H = \frac{(1 - \alpha_C) \sum_i \frac{d_{H,i}}{d_{H,i}+1} \sum_\theta L_i^\theta \left(w_i^\theta (1 + \tau_i^\theta) + T_i^\theta \right)}{1 - (1 - \alpha_C) \sum_i \frac{d_{H,i}}{d_{H,i}+1} \sum_\theta b^\theta L_i^\theta}. \tag{A.64}$$

Using x_i^θ we then construct X_i (aggregate expenditure by MSA) as $X_i = \sum L_i^\theta x_i^\theta$ and $s_j^{X,\theta}$ (share of expenditures by type within MSA). Following these adjustments, we need to ensure that the sum of transfers paid by the government equal the sum of taxes levied. To that end, we scale all transfers uniformly so that they add up to the sum of taxes. This ensures that the government budget constraint holds.[62]

Traded and Non-Traded Sectors We need data on the relative size of the non-traded sector in each city to calibrate the labor shares by sector. The CPS data also reports the sector of activity of workers. We measure at the MSA level the share of workers who work in the non traded sector by counting all workers in the following NAICS sectors: retail, real estate, construction, education, health, entertainment, hotels and restaurants. This measure is not group-specific. To remove unmodeled heterogeneity in this measure, we compute a series of MSA-level socio-demographic characteristics, as above, and regress the share of workers in the non-traded sector on these demographic characteristics. We compute, as above, the predicted share of workers in the non-traded sector in each city, assuming that demographic characteristics of the city are at the nationwide mean.

Trade Shares We need data on trade shares between MSAs, s_{ij}^M and s_{ij}^X (import and export shares). These flows are observed in the CFS data, but not at the finer geographic level that we consider here (MSA). Therefore, we adapt the procedure in Allen and Arkolakis (2014) and Monte et al. (2015), whereby the import shares from the CFS data are used to parametrize the elasticity of trade with respect to distance. In particular, the model implies the following expression for share of location i's imports originating from j:

$$s_{ji}^M = \left(\frac{d_{ji}}{P_i} \frac{W_j^{1-b_Y^I} P_j^{b_Y^I}}{z_j} \right)^{1-\sigma} \equiv \left(d_{ji} \delta_j^D \delta_i^O \right)^{1-\sigma}, \tag{A.65}$$

where δ_i^O and δ_j^D are origin and destination fixed effects. We assume that trade costs have the form $\ln d_{ji} = \psi \ln dist_{ji} + e_{ji}$, where $dist_{ji}$ is the great circle distance between MSAs j and i. We the use Allen and Arkolakis (2014) estimate for ψ and set trade costs to $d_{ji} = dist_{ji}^\psi$. We then construct the smoothed import shares s_{ji}^M between MSAs using (A.65). To that end we must obtain the values of $\{\delta_j^D, \delta_i^O\}$, which are uniquely pinned down, up to a normalization, by considering the identity that sales equals income,

$$p_j Y_j = \sum_i s_{ji}^M E_i, \tag{A.66}$$

together with equation (A.65) and the definition of the price index, leading to:

$$\left(\delta_i^O \right)^{\sigma - 1} = \sum_j \left(d_{ji} \delta_j^D \right)^{1-\sigma}. \tag{A.67}$$

---

[62]This step implies that transfers are uniformly scaled down by 35%. The fact that total taxes and transfers do not match in our dataset comes in part from having removed heterogeneity that is not place-specific from the data and from our treatment of capital to be consistent with the sources of capital income (profits from housing rents), which scales down its share in income relative to the data.

Plugging (A.65) and (A.67) in (A.66), we get a system N equations in N unknowns, which we solve to recover $\{\delta_j^D, \delta_i^O\}$ and in turn s_{ji}^M. The export shares are then constructed using $s_{ji}^X \equiv \left(\frac{E_i}{p_j Y_j}\right) s_{ji}^M$, where spending E_i and traded income $p_j Y_j$.

B.2   Appendix to Section 5.2 (Calibration)

Intermediate Input Shares   We provide details about the calibration of the intermediate input share in non-traded goods. We use the following equilibrium relationship from the market clearing condition in the non traded sector in city j:

$$1 - b_{H,j}^I = \frac{W_j N_j^H}{(1 - \alpha_C) X_j} \left(1 + d_{H,j}\right). \tag{A.68}$$

We compute this expression using the observed wage bill of workers in non-traded sectors $W_j N_j^H$ and total expenditure X_j described in the previous subsection, and our calibrated values for α_C and $d_{H,j}$ described in Section 5.2. In practice, this step requires observing $\frac{W_j N_j^H}{(1-\alpha_C) X_j} \left(1 + d_{H,j}\right) \in [0,1)$ in all cities. However, for some cities we find $\frac{W_j N_j^H}{(1-\alpha_C) X_j} \left(1 + d_{H,j}\right) > 1$. In these cases, we recalibrate the housing supply elasticity in the non-traded sector $d_{H,j}$ such that $\frac{W_j N_j^H}{(1-\alpha_C) X_j} \left(1 + d_{H,j}\right) = 1$, and set $b_{H,j}^I = 0$.[63] For four MSAs where $\frac{W_j N_j^H}{(1-\alpha_C) X_j} > 1$, we set $W_j N_j^H = (1 - \alpha_C) X_j$. This corresponds in practice to scaling down $W_j N_j^H$ by less than 2 % in these cities.

Efficiency Spillover Elasticities   The standard estimate of city-level spillovers reviewed by Combes and Gobillon (2015) are obtained from a regression of average city wages w_j on city population L_j. In log-changes, such an equation would take the form: $\hat{w}_j = \gamma^P \hat{L}_j + \psi_j$, where ψ_j is a city effect and γ^P is the city-level spillover elasticity. In our environment, city-level wages are $w_j L_j = N_j W_j$. Under the assumptions of the quantitative model, applying (A.57), an exogenous shift in the total population of city j keeping its composition across groups constant would then imply:

$$\hat{w}_j = \left[s_j^{W,S} \left(\gamma_{S,S}^P + \gamma_{U,S}^P \right) + \left(1 - s_j^{W,S}\right) \left(\gamma_{S,U}^P + \gamma_{U,U}^P \right) \right] \hat{L}_j + \hat{W}_j, \tag{A.69}$$

where $s_j^{W,S}$ is the share of skilled workers in wages in city j. Hence, through the lens of our model, the coefficient γ^P estimated at the city level in the empirical literature would correspond to $\overline{s^{W,S}} \left(\gamma_{S,S}^P + \gamma_{U,S}^P \right) + \left(1 - \overline{s^{W,S}}\right) \left(\gamma_{S,U}^P + \gamma_{U,U}^P \right)$, where $\overline{s^{W,S}}$ is the average skilled worker share across cities. Therefore, we uniformly normalize the distribution of the $\gamma_{\theta,\theta'}^P$ coefficients such that, under their scaled values, $\overline{s^{W,S}} \left(\gamma_{S,S}^P + \gamma_{U,S}^P \right) + \left(1 - \overline{s^{W,S}}\right) \left(\gamma_{S,U}^P + \gamma_{U,U}^P \right) = \gamma^P$. We set $\gamma^P = 0.06$, which is consistent with the standard estimate for the U.S. from Ciccone and Hall (1996), and $\overline{s^{W,S}} = 0.49$ as observed in our data.

Having chosen the level of the $\gamma_{\theta,\theta'}^P$ coefficients, we must still choose their distribution. Under the assumptions of the quantitative model, the labor demand condition (13) gives the following expression for the log wage of type-θ worker:

$$\ln w_j^\theta = \left[\rho \left(1 + \gamma_{\theta,\theta}^P\right) - 1 \right] \ln \left(L_j^\theta \right) + \rho \gamma_{\theta',\theta}^P \ln \left(L_j^{\theta'} \right) + \ln W_j - (\rho - 1) \ln N_j + \ln \varepsilon_j^\theta, \tag{A.70}$$

where $\ln \varepsilon_j^\theta = \rho \ln Z_j^\theta$ captures productivity shocks at the worker-city level. In data generated by this model and expressed in differences over time, we would have

$$\Delta \ln w_j^\theta = \left[\rho \left(1 + \gamma_{\theta,\theta}^P\right) - 1 \right] \Delta \ln \left(L_j^\theta \right) + \rho \gamma_{\theta',\theta}^P \Delta \ln \left(L_j^{\theta'} \right) + \Delta \kappa_j + \Delta \ln \varepsilon_j^\theta, \tag{A.71}$$

where $\Delta \kappa_j = \Delta \ln W_j - (\rho - 1) \Delta \ln N_j$ is a city effect. We can use (A.71) to map estimates from Diamond (2016). Specifically, she estimates equations (27) and (28) in her paper using Bartik shocks as instruments. The only difference between these equations in her paper and (A.71) is the fixed effect $\Delta \kappa_j$ here. Assuming that the inclusion of the fixed

---

[63]This step affects the profits generated in the economy, in practice reducing slightly the implied values of Π^H and X_j whose construction we described in the previous subsection of the appendix, which in turn modifies (A.68). We iterate to find the unique fixed point of the procedure.

effect $\Delta\kappa_j$ would not alter Diamond (2016) estimates, we can directly map her estimates from Column 3 of Table 5, i.e. $\rho\left(1+\gamma_{S,S}^P\right)-1 = 0.229$, $\rho\gamma_{U,S}^P = 0.312$, $\rho\left(1+\gamma_{U,U}^P\right)-1 = -0.552$, $\rho\gamma_{S,U}^P = 0.697$.

The elasticities resulting from this procedure are reported in the first row of Table A.1. The second row reports the coefficients from an alternative parametrization used in the quantitative section where we target $\gamma^P = 0.12$ instead of $\gamma^P = 0.06$.

Parametrization	γ_{UU}^P	γ_{SU}^P	γ_{US}^P	γ_{SS}^P
Benchmark	0.003	0.044	0.020	0.053
High Efficiency Spillover	0.007	0.087	0.039	0.106

Table A.1: Alternative Parametrizations of Efficiency Spillovers

Amenity Spillover Elasticities Diamond (2016) reports estimates for equation (31) in her paper, which (using our notation for the variables in common with her analysis) has the form:

$$\Delta \ln L_j^\theta = a_0^\theta \Delta \ln\left(\frac{w_j^\theta}{P_j}\right) + a_1^\theta \Delta \ln\left(\frac{R_j}{P_j}\right) + a_2^\theta \Delta \ln\left(a_j^D\right) + \Delta\xi_j^\theta, \tag{A.72}$$

where $a_j^D \equiv \left(L_j^S/L_j^U\right)^{\gamma^a}$ is the endogenous component of amenities in her analysis[64] and (a_0, a_1, a_2) are estimated coefficients. Column (3) of Table 5 of Diamond (2016) reports the following estimates: $\left(a_0^U, a_0^S, a_2^U, a_2^S, \gamma^a\right) = (4.026, 2.116, 0.274, 1.012, 2.6)$. We generate equation (A.72) in our setup and match the coefficients from our model to these estimates. For generality, we do so allowing for idiosyncratic preference draws within each type as in Section 4.4 (i.e., assuming $\sigma_\theta > 0$). The labor-supply equation implied by (29) is

$$\sigma_\theta \ln L_j^\theta = \ln\left(\frac{x_j^\theta}{P_j}\right) - (1-\alpha_C)\ln\left(\frac{R_j}{P_j}\right) + \ln\left(a_j^\theta\right) + \left(\sigma_\theta \ln L^\theta - \ln u^\theta\right). \tag{A.73}$$

Let $\zeta^{A,S} = \gamma^a$ and $\zeta^{A,U} = -\gamma^a$, and then redefine our amenity index a_j^θ for $\theta = U, S$ in (38) as a function of the amenity index a_j^D from Diamond (2016) as follows: $a_j^\theta = A_j^\theta\left(L_j^\theta\right)^{\gamma_{\theta,\theta}^A - \beta^{a,\theta}\zeta^{A,\theta}} a_j^D$, where $\beta^{a,\theta} \equiv \frac{\gamma_{\theta',\theta}^A}{\zeta^{A,\theta'}}$ is by construction constant over θ'. Using this equivalence in (A.73), re-arranging and expressing that equation in changes we obtain

$$\Delta \ln L_j^\theta = \frac{1}{\left(\sigma_\theta - \gamma_{\theta,\theta}^A\right) + \beta^{a,\theta}\zeta^{A,\theta}} \Delta \ln\left(\frac{x_j^\theta}{P_j}\right) - \frac{1-\alpha_C}{\left(\sigma_\theta - \gamma_{\theta,\theta}^A\right) + \beta^{a,\theta}\zeta^{A,\theta}} \Delta \ln\left(\frac{R_j}{P_j}\right)$$
$$+ \frac{\beta^{a,\theta}}{\left(\sigma_\theta - \gamma_{\theta,\theta}^A\right) + \beta^{a,\theta}\zeta^{A,\theta}} \Delta \ln\left(a_j^D\right) + \Delta\xi_j^\theta, \tag{A.74}$$

where $\Delta\xi_j^\theta \equiv \frac{1}{\left(\sigma_\theta-\gamma_{\theta,\theta}^A\right)+\beta^{a,\theta}\zeta^{A,\theta}}\left(\ln A_j^\theta + \sigma_\theta \ln L^\theta - \ln u^\theta\right)$. Comparing (A.72) with (A.74) readily allows us to map Diamond (2016) estimates to our parameters as follows:

$$\gamma_{\theta,\theta}^A - \sigma_\theta = \frac{a_2^\theta}{a_0^\theta}\zeta^{A,\theta} - \frac{1}{a_0^\theta}, \tag{A.75}$$

$$\gamma_{\theta',\theta}^A = \frac{a_2^\theta}{a_0^\theta}\zeta^{A,\theta'} \tag{A.76}$$

for $\theta = U, S$. Conditional the estimates of $\left(a_0^U, a_0^S, a_2^U, a_2^S, \gamma^a\right)$, we back out the value of $\gamma_{\theta,\theta}^A - \sigma_\theta$ but are unable to

---

[64]This index captures congestion in transport, crime, environmental indicators, supply per capita of different public services, and variety of retail stores. See Table 4 of Diamond (2016).

distinguish $\gamma^A_{\theta,\theta}$ from $-\sigma_\theta$. Our benchmark model is presented assuming $\sigma_\theta = 0$. However, as discussed in Section 4.4, $\gamma^A_{\theta,\theta} - \sigma_\theta$ is the relevant combination of parameters to characterize optimal allocations and policies under the definition of the planner problem with idiosyncratic preference draws defined in that section.

The resulting numbers are reported in the first row of Table A.2. The second row reports the coefficients from an alternative parametrization used in the quantitative section where we scale all amenity spillovers down by 50% relative to the benchmark. The third and fourth rows report parametrizations that, instead the coefficient $\gamma^a = 2.6$ reported in Column (3) of Table 5 of Diamond (2016), use that point estimate plus or minus the standard deviation reported in that table, respectively.

	γ^A_{UU}	γ^A_{SU}	γ^A_{US}	γ^A_{SS}
Benchmark	-0.43	0.18	-1.24	0.77
Low amenity spillover	-0.21	0.09	-0.62	0.38
High cross-amenity spillover	-0.46	0.22	-1.51	1.04
Low cross-amenity spillover	-0.39	0.14	-0.97	0.50

Table A.2: Alternative Parametrizations of Amenity Spillovers

Equivalence of Amenity Spillovers Between Multiple and Single Worker Types In the quantitative section we implement a version of the model with a single type of worker, using a single amenity spillover elasticity γ^A. To find this elasticity, we note that, under a single worker type, the labor-supply equation implied by (10) expressed in time differences becomes

$$\Delta \ln L_j = -\frac{1}{\gamma^A} \left(\Delta \ln \left(\frac{x_j}{P_j} \right) - (1 - \alpha_C) \Delta \ln \left(\frac{R_j}{P_j} \right) \right) + \Delta \xi_j \qquad (A.77)$$

where $\Delta \xi_j$ includes changes in aggregate labor supply and exogenous components of amenities, A_j. In turn, under multiple worker types, the labor supply equation at the city level results from aggregating the supply of multiple workers:

$$\Delta \ln L_j^\theta = -\sum_\theta \frac{s_j^{L,\theta}}{\gamma^A_{\theta,\theta}} \left(\Delta \ln \left(\frac{x_j^\theta}{P_j} \right) - (1 - \alpha_C) \Delta \ln \left(\frac{R_j}{P_j} \right) \right) - \sum_\theta s_j^{L,\theta} \sum_{\theta' \neq \theta} \frac{\gamma^A_{\theta',\theta}}{\gamma^A_{\theta,\theta}} \Delta \ln L_j^{\theta'} + \Delta \xi_j^\theta \qquad (A.78)$$

where $\Delta \xi_j^\theta$ includes changes in the labor supply of type-θ workers and in the exogenous component of amenities, A_j^θ. We can draw an equivalence between the aggregate elasticity that would be estimated assuming homogeneous workers (i.e., using (A.77)) when the true model includes heterogeneous workers, so that the data is generated by (A.78). In the latter, assuming a shock that exogenously changes population and expenditure per capita in the same proportion for every worker, aggregating the labor supplies by skill we obtain:

$$\hat{L}_j = \left(-\frac{\sum_\theta \frac{s_j^{L,\theta}}{\gamma^A_{\theta,\theta}}}{1 + \sum_\theta \sum_{\theta' \neq \theta} \frac{s_j^{L,\theta} \gamma^A_{\theta',\theta}}{\gamma^A_{\theta,\theta}}} \right) \left(\hat{x}_j - \hat{P}_j - (1 - \alpha_C) \left(\hat{R}_j - \hat{P}_j \right) \right) + \Delta \overline{\xi}_j. \qquad (A.79)$$

where $s_j^{L,\theta}$ is the share of type θ workers in j and $\Delta \overline{\xi}_j \equiv \sum_\theta s_j^{L,\theta} \Delta \xi_j^\theta$. Comparing (A.77) with (A.79), we obtain that, at the average share of type-θ workers in the economy $\overline{s^{L,\theta}} = \frac{1}{J} \sum_j s_j^{L,\theta}$, the coefficient that would be recovered is:

$$\gamma^A = \frac{1 + \sum_\theta \sum_{\theta' \neq \theta} \frac{s_j^{L,\theta} \gamma^A_{\theta',\theta}}{\gamma^A_{\theta,\theta}}}{\sum_\theta \frac{s_j^{L,\theta}}{\gamma^A_{\theta,\theta}}}. \qquad (A.80)$$

When implementing the model with a single worker type we use this expression to determine γ^A.

C Appendix Figures to Section 6

Figure A.1: Optimal Transfers and Reallocation under Homogeneous Workers
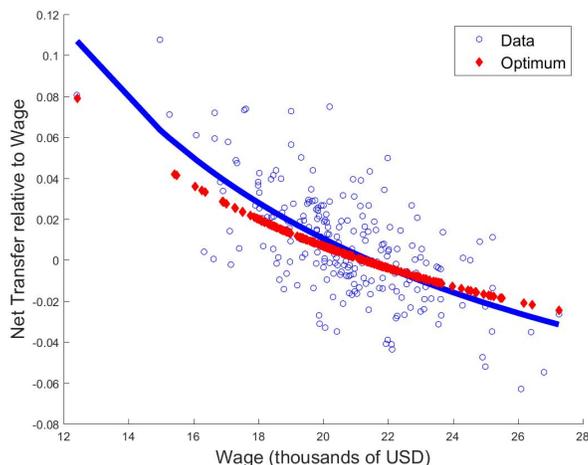
Note: This figure shows the transfer per worker relative to the wage in the optimal allocation and in the data. As implied by Section 4.2, the optimal net transfer relative to the wage takes the form $\frac{t_j}{w_j} = s + \frac{T}{w_j}$ for $s = \frac{\gamma^P + \gamma^A}{1 - \gamma^A}$. The solid lines shows the relationship $\frac{t_j}{w_j} = a + b\frac{1}{w_j}$ under parameters a and b that correspond to the best fit in an OLS regression.

Figure A.2: Gains from Optimal Policies given Different Initial Equilibria under Homogeneous Workers
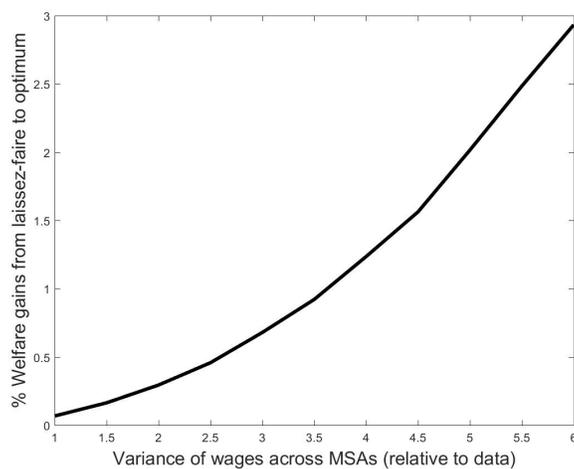
Note: We simulate laissez-faire equilibria with no government transfers under different fundamentals such that the joint distribution of wages and city sizes differs from the data in terms of the variance of the wage distribution across MSAs and the correlation between wages and city sizes across MSAs. In all the equilibria the distribution of city sizes has the same variance as in the data. Correlation and variances are reported in relative terms compared to the data. For each variance-correlation combination we draw 400 random distributions of wages and city sizes, and report the mean welfare gains from implementing optimal policies across these simulations.

Figure A.3: Transfers relative to Wage by Skill Level and MSA under Different Pareto Weights
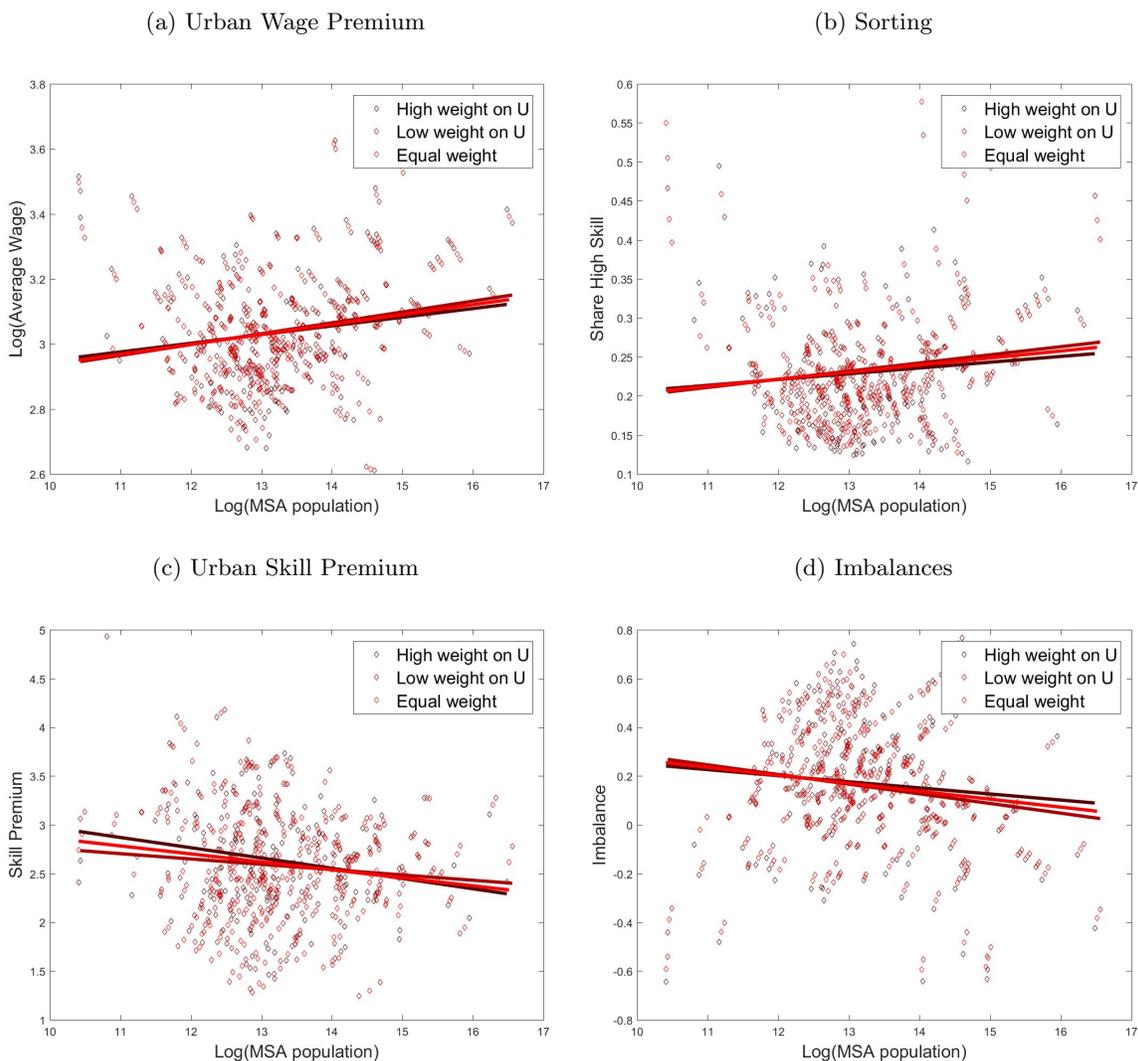
(a) High weight on U
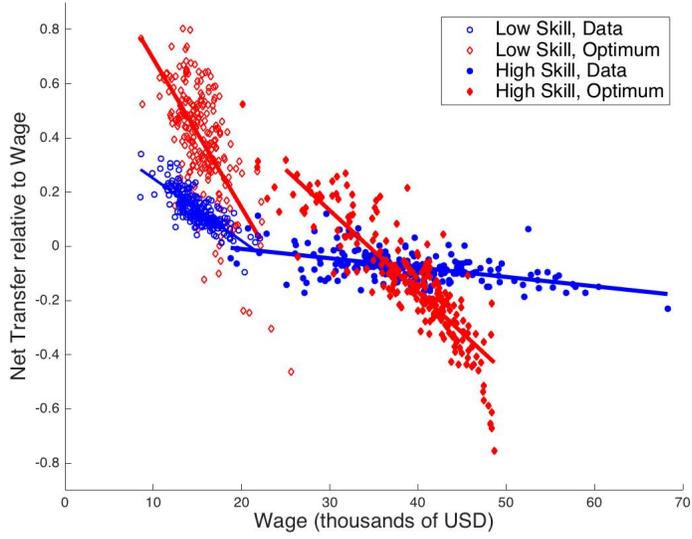
(b) Low Weight on U
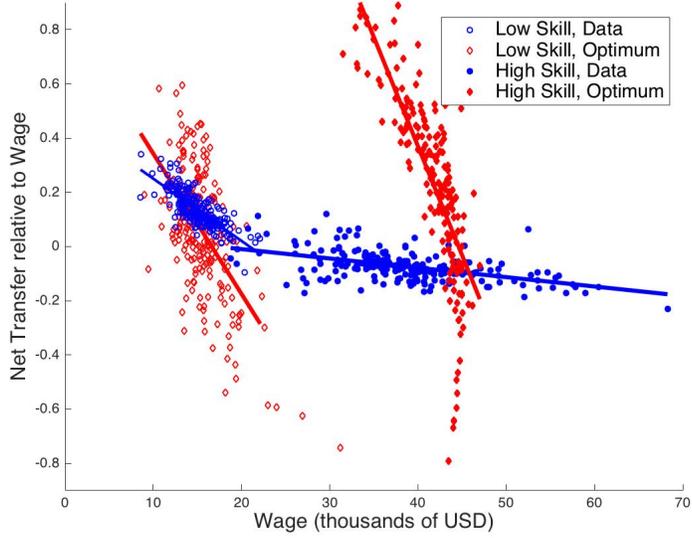
Note: The panels replicates Figure 3 under alternative weights for each skill group in the planning problem. In the case "high weight on U", the weight is 10 times larger for low skill workers and the welfare changes are $\left(\hat{u^U}, \hat{u^S}\right) = (1.18, 0.88)$. In the case "low weight on U", the weight is 10 times larger for high skill workers and the welfare changes are $\left(\hat{u^U}, \hat{u^S}\right) = (0.88, 1.28)$.
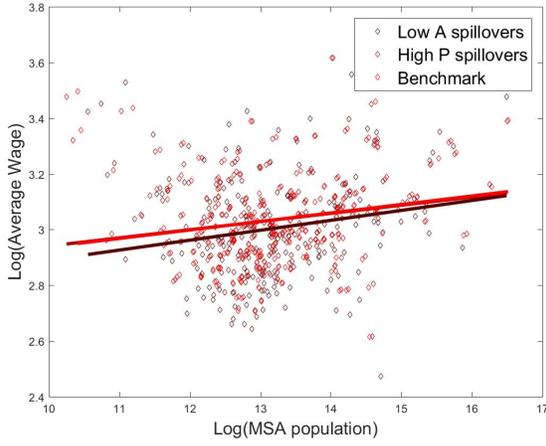
Figure A.4: Urban Premia under Different Pareto Weights

(a) Urban Wage Premium

(b) Sorting

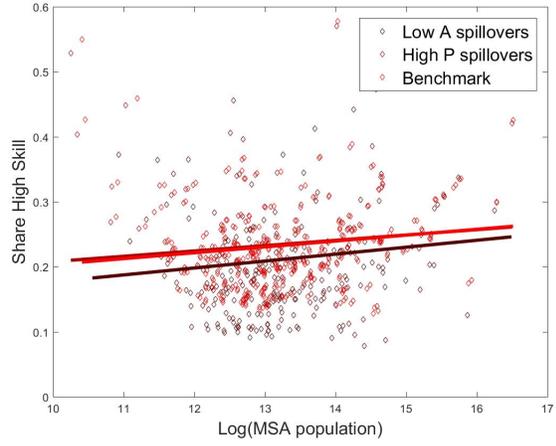(c) Urban Skill Premium

(d) Imbalances

Note: Each panel reports outcomes across MSAs in the optimal allocation under alternative weights for each skill group in the planning problem. In the case "high weight on U", the weight is 10 times larger for low skill workers and the welfare changes are $\left(\hat{u^U}, \hat{u^S}\right) = (1.18, 0.88)$. In the case "low weight on U", the weight is 10 times larger for high skill workers and the welfare changes are $\left(\hat{u^U}, \hat{u^S}\right) = (0.88, 1.28)$. The case "equal weight" corresponds to the optimal allocation in Figure 3.

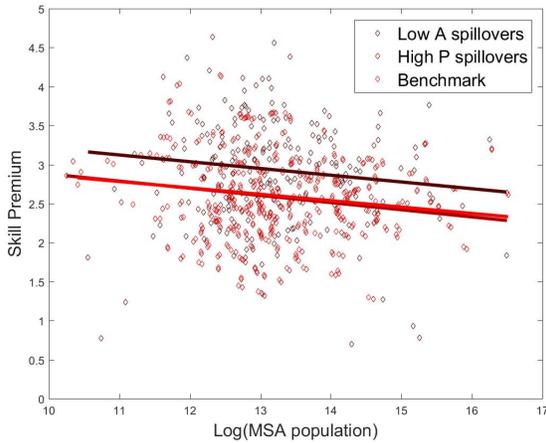Figure A.5: Urban Premia under Different Spillover Elasticities
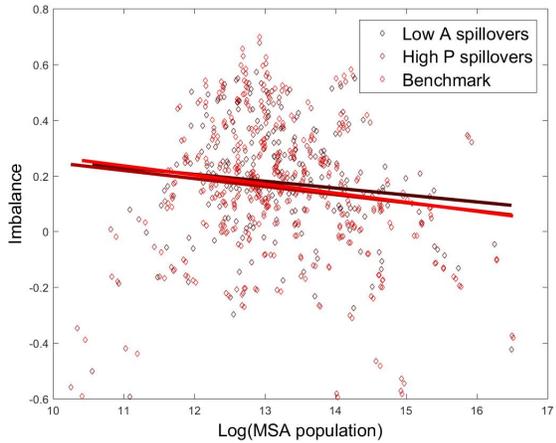
(a) Urban Wage Premium

(b) Sorting

(c) Urban Skill Premium

(d) Imbalances

Note: Each panel reports outcomes across MSAs in the optimal allocation under alternative parametrizations of the spillovers. "Low A spillovers" corresponds to the "Low Amenity Spillover" parametrization and "High P spillovers" corresponds to the "High Efficiency Spillover" parametrization. The welfare effects under these alternative parametrizations are reported in Table 1.