

Urban Accounting and Welfare*

Klaus Desmet

Esteban Rossi-Hansberg

Universidad Carlos III

Princeton University

September 30, 2010

Abstract

This paper provides a way of decomposing the characteristics that lead to the size distribution of cities into three main components: efficiency, amenities, and frictions. The theory consists of a multi-city neoclassical model where agents choose how much to work and where to live. Higher efficiency and better amenities lead to larger cities, but also to greater frictions through congestion and other negative effects of agglomeration. Using data on MSAs in the United States, we parametrize the model and empirically estimate efficiency, amenities, and frictions. Counter-factual exercises show that all three characteristics are important in that eliminating any of them leads to large population reallocations, though the welfare effects from these reallocations are small. Overall, we find that the gains from worker mobility across cities are modest. When allowing for externalities, we find an important city selection effect: eliminating differences in any of the city characteristics cause many cities to exit.

1. INTRODUCTION

Why do people live in particular cities? We can list many reasons, but two are undoubtedly relevant. Agents can enjoy the city or be more productive there. How many agents are attracted to a particular city due to its combination of life's amenities and the productivity level of its workforce determines the size of cities. The positive effects of attractive amenities and high productivity are capped by frictions arising from congestion. Depending on city governance and the flexibility of markets, these frictions can be more or less important. These city characteristics are in turn enhanced and amplified by the presence of urban externalities. Understanding the different forces that determine city sizes is crucial for answering a broad set of questions. What is the relative importance of these forces in determining the size distribution of cities? How much would we gain

*We thank Stephen Redding for helpful comments, and acknowledge the financial support of the International Growth Centre at LSE (Grant RA-2009-11-015).

or lose if cities had similar amenities, technology levels or frictions? How much reallocation would this cause? How do externalities interact with these forces to determine city selection?

In this paper we provide a simple way of decomposing the characteristics that lead to the size distribution of cities into three main components: efficiency, amenities, and excessive frictions. We use a simple urban theory to calculate these components and to do a wide set of counterfactual exercises that provide answers to the questions we asked above. The theory consists of a multi-city model with monocentric cities that produce a single good. Workers decide how much to work and where to live. Efficiency is modeled as TFP differences, amenities as preference shocks, and excessive frictions as the cost of providing urban infrastructure that is paid for with labor taxes. To measure “excessive frictions”, we use the concept of a “labor wedge” and decompose it into the standard cost effect of city size and the “excessive cost” of providing city services (see Chari, et al., 2007). We solve the general equilibrium model with and without externalities.

We use an empirical strategy to parametrize the model and to obtain the “excessive frictions” shocks and the efficiency shocks. We then use the model to determine the amenity shocks that make cities be of their actual sizes. The model therefore matches by construction the size distribution of cities in the U.S. The counterfactual exercises then allow us to analyze the effects of the different parameter values and of the different city characteristics. For many counterfactuals we find that the changes in utility (and in consumption) are modest in spite of massive population reallocations. For example, eliminating efficiency differences across cities lowers equilibrium utility levels by a mere 2.5%, and eliminating amenity differences reduces welfare by just 2.3%. When we account for externalities, these numbers decline even further. The welfare implications of redistributing agents across cities due to switching off any of the fundamental characteristics that account for the actual size distribution is never greater than a few percentage points.¹ This is surprising given that the differences in amenities and efficiency levels can be rather large and given that the implied population reallocations can be as large as 50%. Adding externalities has an important effect on the extensive margin in the counterfactual exercises, with many cities exiting and the urban population settling in the surviving cities. However, these externalities have only modest effects on welfare.

The paper relates to much of the literature on the size distribution of cities, but instead of taking a random growth approach in which city dynamics coming from productivity or preference shocks determine the size distribution (as in Gabaix, 1999a, b, Duranton, 2007, Rossi-Hansberg and Wright

¹This resembles the literature on business cycle accounting that found that eliminating business cycles would lead to trivial effects (as in Lucas, 1987, we do not have the necessary distributional cost to obtain larger losses as agents are identical, as emphasized by Storesletten, et al., 2001).

2007, and Cordoba, 2008), we use a model to decompose the individual city characteristics that lead to the cross-sectional distribution of city sizes. Since our model has no dynamic frictions or specific factors, agents move across cities as a response to any temporary shock. In that sense, city dynamics play no role in our decomposition. Of course, the measured levels of efficiency, amenities or frictions may still be the result of these dynamic mechanisms. To the extent that this is the case, our approach helps us assess the contribution of particular dynamic factors to the distribution of city sizes. More related to our work is Au and Henderson (2006) who use a model with agglomeration economies and congestion effects to analyze optimal city sizes in China. After structurally estimating their model, they calculate the welfare effects of migration constraints, and find that output per worker would increase substantially in some cities if labor were free to move. However, different from us, they limit their attention to efficiency, and do not focus on the other components of city size determination. Focusing on the U.S. and incorporating these other components result in much smaller welfare changes from urban reallocation.

Other work has emphasized the importance of efficiency, productivity and amenities in explaining the distribution of city sizes. Glaeser et al. (2001), Glaeser et al. (2005), Albouy (2008 and 2009), and Rappaport (2008 and 2009), for example, have underscored the importance of city amenities and institutional frictions. Others have emphasized the importance of the relative efficiency in production of the different urban areas (Holmes and Stevens, 2002 and 2004, Holmes, 2005, Duranton and Overman, 2008) or the geographic characteristics of the locations in which cities develop (Davis and Weinstein, 2002, Bleakley and Lin, 2010). Our approach relates to this work in that we also aim to identify important city characteristics that lead to urban agglomeration. We differ in that we have a fully specified general equilibrium model that we use to jointly account for efficiency, amenities and excessive frictions. The model is also flexible enough to incorporate externalities in efficiency levels and amenities. With the urban characteristics in hand we can perform a variety of counterfactual exercises and calculate the welfare implications of eliminating variation in any of these characteristics. Furthermore, the proposed methodology can be used to compare urban systems across countries, although we leave that for future research.

The rest of the paper is organized as follows. Section 2 introduces a simple urban model and explains the basic urban accounting exercise. Section 3 estimates a log-linear version of the structural equations using U.S. data between 2005 and 2008 and obtains the reduced form effects of the three main characteristics of cities on rents and city sizes. Section 4 performs counterfactual exercises using the empirical values of these city characteristics. Section 5 studies the effect of efficiency and

amenity shocks, and Section 6 concludes. Appendix A shows how the population sizes of individual cities are affected when certain characteristics change. Appendix B describes in detail the urban data set constructed.

2. THE MODEL

We use a standard urban model with elastic labor supply so that labor taxes create distortions. Agents work in cities with idiosyncratic productivities and amenities. They live in mono-centric cities that require commuting infrastructures that city governments provide by levying labor taxes. Large cities are more expensive to live in because of higher labor taxes and commuting costs, but are large because of high levels of efficiency or local amenities. City governments can be more or less efficient in the provision of the public infrastructure. We refer to this variation as a city’s “excessive frictions”. In later sections we augment the model to include local externalities in production and amenities.

2.1 Technology

Consider a model of a system of cities in an economy with N_t workers. Goods are produced in I mono-centric circular cities. Cities have a local level of productivity. Production in a city i in period t is given by

$$Y_{it} = A_{it}K_{it}^\theta H_{it}^{1-\theta}$$

where A_{it} denotes city productivity, K_{it} denotes total capital and H_{it} total hours worked in the city.² We denote the population size of city i by N_{it} . The standard first order conditions of this problem are

$$w_{it} = (1 - \theta) \frac{Y_{it}}{H_{it}} = (1 - \theta) \frac{y_{it}}{h_{it}} \tag{1}$$

and

$$r_t = \theta \frac{Y_{it}}{K_{it}} = \theta \frac{y_{it}}{k_{it}}$$

where small-cap letters denote per capita variables (e.g. $y_{it} = Y_{it}/N_{it}$). Note that capital is freely mobile across locations so there is a national interest rate r_t . Mobility patterns will not be determined solely by the wage, w_{it} , so there may be equilibrium differences in wages across cities at any point in time.

²It would be straightforward to generalize this model to include human capital. We experimented with this, and doing so did not substantially change any of the theoretical or empirical results.

We can then write down the “efficiency wedge” which is identical to the level of productivity, A_{it} , as

$$A_{it} = \frac{Y_{it}}{K_{it}^\theta H_{it}^{1-\theta}} = \frac{y_{it}}{k_{it}^\theta h_{it}^{1-\theta}}. \quad (2)$$

2.2 Preferences

Agents order consumption and hour sequences according to the following utility function

$$\sum_{t=0}^{\infty} \beta^t [\log c_{it} + \psi \log (1 - h_{it}) + \gamma_i]$$

where γ_i is a city specific amenity and ψ is a parameter that governs the relative preference for leisure. Each agent lives on one unit of land and commutes from his home to work. Commuting is costly in terms of goods.

The problem of an agent in city i_0 with capital k_0 is therefore

$$\max_{\{c_{it}, h_{it}, k_{it}, \tau_{it}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t [\log c_{it} + \psi \log (1 - h_{it}) + \gamma_i]$$

subject to

$$\begin{aligned} c_{it} + x_{it} &= r_t k_{it} + w_{it} h_{it} (1 - \tau_{it}) - R_{it} - T_{it} \\ k_{it+1} &= (1 - \delta) k_{it} + x_{it}, \end{aligned}$$

where x_{it} is investment, τ_{it} is a labor tax or friction associated with the cost of building the commuting infrastructure, R_{it} are land rents and T_{it} commuting costs (as we will see below, $R_{it} + T_{it}$ is constant in the city so the location of the agent’s home does not affect his choices).³

Throughout the paper we assume that we are in steady state so $k_{it+1} = k_{it}$ and $x_{it} = \delta k_{it}$. Furthermore, we assume k_{it} is such that $r_t = \delta$ (capital is at its Golden Rule level). The simplified budget constraint of the agent becomes

$$c_{it} = w_{it} h_{it} (1 - \tau_{it}) - R_{it} - T_{it}. \quad (3)$$

The first order conditions of this problem are given by $1/c_{it} = \lambda_{it}$, and

$$\psi \frac{1}{1 - h_t} = w_{it} (1 - \tau_{it}) \lambda_{it},$$

³Since agents can move across cities, the subscript i depends on t , as written under the maximization sign. To save on notation, we drop this additional subscript.

which imply

$$\psi \frac{c_{it}}{1 - h_{it}} = (1 - \tau_{it}) w_{it}. \quad (4)$$

Combining equations (1) and (4) we obtain

$$(1 - \tau_{it}) = \frac{\psi}{(1 - \theta)} \frac{c_{it}}{1 - h_{it}} \frac{h_{it}}{y_{it}}. \quad (5)$$

Agents can move freely across cities so utility in each period has to be determined by

$$\bar{u} = \log c_{it} + \psi \log (1 - h_{it}) + \gamma_i, \quad (6)$$

for all cities with $N_{it} > 0$, where \bar{u} is the economy-wide per period utility of living in a city.

2.3 Commuting Costs, Land Rents and City Equilibrium

Cities are mono-centric, all production happens at the center, and people live in surrounding areas characterized by their distance to the center, d . Cities are surrounded by a vast amount of agricultural land that can be freely converted into urban land. We normalize the price of agricultural land to zero. Since land rents are continuous in equilibrium (otherwise there would be arbitrage opportunities), this implies that at the boundary of a city, \bar{d} , land rents should be zero as well, namely, $R(\bar{d}) = 0$. Since all agents in a city are identical, in equilibrium they must be indifferent between where they live in the city, which implies that the total cost of rent plus commuting costs should be identical in all areas of the city. So

$$R_{it}(d) + T(d) = T(\bar{d}_{it}) = \kappa \bar{d}_{it}$$

since $T(d) = \kappa d$ where κ denotes commuting costs per mile.

Since everyone lives in one unit of land, $N_{it} = \bar{d}^2 \pi$, and so

$$\bar{d}_{it} = \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}},$$

we know that

$$R_{it}(d) + T(d) = \kappa \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}} \text{ all } d.$$

This implies that

$$R_{it}(d) = \kappa (\bar{d}_{it} - d)$$

and so total land rents in a city of size N_{it} are given by

$$TR_{it} = \int_0^{\bar{d}_{it}} (\kappa (\bar{d}_{it} - d) d 2\pi) dd = \frac{\pi^{\frac{1}{2}} \kappa}{3} N_{it}^{\frac{3}{2}}.$$

Hence, average land rents are equal to

$$AR_{it} = \frac{2\kappa}{3} \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}}.$$

Taking logs and rearranging terms, we obtain that

$$\ln(N_{it}) = o_1 + 2 \ln AR_{it}, \quad (7)$$

where o_1 is a constant.

We can also compute the total miles traveled by commuters in the city, which is given by

$$TC_{it} = \int_0^{\bar{d}_{it}} (d^2 2\pi) dd = \frac{2}{3} \pi^{-\frac{1}{2}} N_{it}^{\frac{3}{2}}. \quad (8)$$

2.4 Government Budget Constraint

The government levies a labor tax, τ_{it} , to pay for the transportation infrastructure. This tax can also be interpreted as a distortion in the labor market because of the congestion created in a large city. As in Chari et al. (2007), we will refer to τ_{it} as the “labor wedge”. Let government expenditure be a function of total commuting costs and wages such that

$$G(h_{it}w_{it}, TC_{it}) = g_{it}h_{it}w_{it}\kappa TC_{it} = g_{it}h_{it}w_{it}\kappa \frac{2}{3} \pi^{-\frac{1}{2}} N_{it}^{\frac{3}{2}}.$$

where g_{it} is a measure of government inefficiency. That is, the government requires κg_{it} workers per mile commuted to build and maintain urban infrastructure. The government budget constraint is then given by

$$\tau_{it}h_{it}N_{it}w_{it} = g_{it}h_{it}w_{it}\kappa \frac{2}{3} \pi^{-\frac{1}{2}} N_{it}^{\frac{3}{2}}$$

which implies that the “labor wedge” can be written as

$$\tau_{it} = g_{it}\kappa \frac{2}{3} \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}} \quad (9)$$

or

$$\ln \tau_{it} = o_2 + \ln g_{it} + \frac{1}{2} \ln N_{it}. \quad (10)$$

2.5 Equilibrium

The consumer budget constraint is given by

$$c_{it} = w_{it}h_{it}(1 - \tau_{it}) - R_{it} - T_{it} = (1 - \theta)(1 - \tau_{it})y_{it} - \kappa \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}}.$$

From (5) we know that

$$(1 - \tau_{it}) = \frac{\psi}{(1 - \theta)} \frac{c_{it}}{1 - h_{it}} \frac{h_{it}}{y_{it}}.$$

To determine output we know that the production function is given by $y_{it} = A_{it} k_{it}^\theta h_{it}^{1-\theta}$ and the decision of firms to rent capital implies that $r_t k_{it} = \theta y_{it}$. Hence,

$$y_{it} = A_{it} \left(\frac{\theta y_{it}}{r_t} \right)^\theta h_{it}^{1-\theta}$$

or

$$y_{it} = A_{it}^{\frac{1}{1-\theta}} \left(\frac{\theta}{r_t} \right)^{\frac{\theta}{1-\theta}} h_{it}.$$

Hence,

$$1 - h_{it} = c_{it} \frac{\psi}{(1 - \theta)} \frac{\left(\frac{r_t}{\theta} \right)^{\frac{\theta}{1-\theta}}}{(1 - \tau_{it}) A_{it}^{\frac{1}{1-\theta}}}$$

and so

$$h_{it} = \frac{1}{1 + \psi} \left(1 + \frac{\psi (R_{it} + T_{it})}{(1 - \theta) (1 - \tau_{it})} \frac{\left(\frac{r_t}{\theta} \right)^{\frac{\theta}{1-\theta}}}{A_{it}^{\frac{1}{1-\theta}}} \right).$$

which implies

$$1 - h_{it} = \frac{\psi}{1 + \psi} \left(1 - \frac{(R_{it} + T_{it})}{(1 - \theta) (1 - \tau_{it})} \frac{\left(\frac{r_t}{\theta} \right)^{\frac{\theta}{1-\theta}}}{A_{it}^{\frac{1}{1-\theta}}} \right)$$

and

$$c_{it} = \frac{(1 - \theta) (1 - \tau_{it}) A_{it}^{\frac{1}{1-\theta}} \left(\frac{\theta}{r_t} \right)^{\frac{\theta}{1-\theta}} - (R_{it} + T_{it})}{1 + \psi}.$$

The free mobility assumption in (6) implies that $\bar{u} = \log c_{it} + \psi \log(1 - h_{it}) + \gamma_{it}$ for some \bar{u}_t determined in general equilibrium so

$$\begin{aligned} & \bar{u}_t + (1 + \psi) \log(1 + \psi) - \psi \log \psi & (11) \\ = & \log \left((1 - \theta) (1 - \tau_{it}) \frac{A_{it}^{\frac{1}{1-\theta}}}{\left(\frac{r_t}{\theta} \right)^{\frac{\theta}{1-\theta}}} - (R_{it} + T_{it}) \right) \\ & + \psi \log \left(1 - \frac{(R_{it} + T_{it})}{(1 - \theta) (1 - \tau_{it})} \frac{\left(\frac{r_t}{\theta} \right)^{\frac{\theta}{1-\theta}}}{A_{it}^{\frac{1}{1-\theta}}} \right) + \gamma_i \\ = & \log \left((1 - \theta) \left(1 - \kappa g_{it} \frac{2}{3} \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}} \right) \frac{A_{it}^{\frac{1}{1-\theta}}}{\left(\frac{r_t}{\theta} \right)^{\frac{\theta}{1-\theta}}} - \kappa \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}} \right) \\ & + \psi \log \left(1 - \frac{\kappa \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}}}{(1 - \theta) \left(1 - \kappa g_{it} \frac{2}{3} \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}} \right)} \frac{\left(\frac{r_t}{\theta} \right)^{\frac{\theta}{1-\theta}}}{A_{it}^{\frac{1}{1-\theta}}} \right) + \gamma_i \end{aligned}$$

The last equation determines the size of the city N_{it} as an implicit function of city productivity, A_{it} , city amenities, γ_i , government inefficiency g_{it} , and economy-wide variables like r_t and \bar{u}_t . We can use this equation to derive the effect of the three city specific characteristics ($A_{it}, \gamma_{it}, g_{it}$) on N_{it} . First note that the LHS of (11) is decreasing in N_{it} . The LHS is also increasing in A_{it} and γ_i and decreasing in g_{it} . Hence, we can prove immediately that

$$\begin{aligned} \frac{dN_{it}}{dA_{it}} &> 0, \\ \frac{dN_{it}}{d\gamma_i} &> 0, \\ \frac{dN_{it}}{dg_{it}} &< 0. \end{aligned} \tag{12}$$

So population increases in a more productive city or a city with more amenities, but it decreases in a city with a less efficient government.

The economy-wide utility level \bar{u}_t is determined by the labor market clearing condition

$$\sum_{i=1}^I N_{it} = N_t. \tag{13}$$

This last equation clarifies that our urban system is closed; we do not consider urban-rural migration.

3. EVIDENCE OF EFFICIENCY, AMENITIES AND FRICTIONS

To lend validity to our theoretical model, we estimate the size of the three derivatives in (12) and estimate the effect of land rents on population as in (7). When doing so, the general equilibrium nature of the model will be key.

3.1 Empirical Approach

We first estimate the “labor wedge” in equation (5) and the “efficiency wedge” in equation (2). Note that the empirical measure of the “efficiency wedge” is related not just to productivity, but also to the relative price of city output. Although we have no way of disentangling these two terms, in a theory with multiple goods relative price effects across cities would have isomorphic effects to changes in productivity. Hence, we just equate productivity to our measure of the “efficiency wedge”.

The general equilibrium nature of the model is important. For example, if we regress the log of city size on the log of the labor wedge, we find a statistically significant positive effect (coefficient

of 1.2360 and p-value of 0.000). But it would be wrong to conclude from there that higher frictions lead to greater city size. Rather, according to the theory, this positive association would reflect more productive cities being larger, and larger cities experiencing greater commuting costs. That is, in as far as greater commuting costs are due to cities being more efficient, they will be positively associated with city size. Only frictions “in excess” of this basic tradeoff between efficiency and congestion will have a negative effect on city size. In what follows we propose a methodology that accounts for these general equilibrium links by decomposing these different effects.

We start by estimating the following equation

$$\ln N_{it} = \alpha_1 + \beta_1 \ln A_{it} + \varepsilon_{1it}. \quad (14)$$

The value of β_1 yields the effect of the “efficiency wedge” on city population. According to the model, $\beta_1 > 0$ by (12). Furthermore, $\ln \tilde{N}_{it}(A_{it}) = \beta_1 \ln A_{it}$ is the population size explained by the size of the “efficiency wedge”. In contrast, ε_{1it} is the part of the observed population in the city that is unrelated to productivity; according to the model it is related to both g_{it} and γ_{it} . We can thus define the function $\tilde{\varepsilon}_1(g_{it}, \gamma_{it}) \equiv \varepsilon_{1it}$.

Since the “efficiency wedge” increases population size, total commuting increases, which affects the “labor wedge” according to equation (10). This is the standard urban trade-off between productivity and agglomeration. We can estimate the effect of productivity on the “labor wedge” by using equation (10) and the decomposition of $\ln N_{it}$ into $\ln \tilde{N}_{it}(A_{it})$ and ε_{1it} provided by equation (14). Hence, we estimate

$$\ln \tau_{it} = \alpha_2 + \beta_2 \ln \tilde{N}_{it}(A_{it}) + \varepsilon_{2it}. \quad (15)$$

According to equation (10), $\beta_2 > 0$. That is, a city that is more productive and so has more population, will be more distorted. We denote the effect of efficiency on distortions by $\ln \widetilde{\tau}_{it} = \beta_2 \ln \tilde{N}_{it}(A_{it})$. Equation (10) also implies that the error term ε_{2it} is related to g_{it} and to $\tilde{\varepsilon}_1(g_{it}, \gamma_{it})$ (since the labor wedge depends on all factors affecting population and not just on $\ln \tilde{N}_{it}(A_{it})$). Hence, we define $\tilde{\varepsilon}_2(g_{it}, \tilde{\varepsilon}_1(g_{it}, \gamma_{it})) \equiv \varepsilon_{2it}$.

We now use equation (7) to decompose the effect from all three elements of $(A_{it}, \gamma_i, g_{it})$. To do so, we estimate

$$\ln (AR_{it}) = \alpha_3 + \beta_3 \ln \widetilde{\tau}_{it} + \beta_4 \varepsilon_{1it} + \beta_5 \varepsilon_{2it} + \varepsilon_{3it} \quad (16)$$

using median rents for AR_{it} . The model has clear predictions for β_3 , β_4 and β_5 . In particular, it implies $\beta_3 > 0$, since by equations (7) and (12) efficiency has a positive effect on population, which has a positive effect on the level of distortions and on average rents. This is the standard city size

effect. The effects of γ_{it} and g_{it} are determined by the estimates of β_4 and β_5 . Note that ε_{1it} and ε_{2it} depend on both γ_{it} and g_{it} . However, since $\varepsilon_{2it} = \tilde{\varepsilon}_2(g_{it}, \tilde{\varepsilon}_1(g_{it}, \gamma_{it}))$ depends only on γ_{it} through ε_{1it} and we are including ε_{1it} directly in the regression, β_5 will capture only the effect of changes in g_{it} on land rents. So, β_5 captures the effect of g_{it} on frictions and therefore average rents. Higher distortions imply a higher τ_{it} . Hence, the model implies that higher g_{it} , and therefore higher τ_{it} and ε_{2it} , implies lower population and lower rents (see (12)). Thus β_5 should be negative. Similarly, since we are controlling for the effect of g_{it} by including ε_{2it} , β_4 will capture the effect of ε_{1it} on land rents controlling for g_{it} , which is the effect of γ_{it} on land rents, since $\varepsilon_{1it} = \tilde{\varepsilon}_1(g_{it}, \gamma_{it})$. Hence, the model implies that β_4 should be positive by equations (7) and (12). Our model implies that rents are a non-linear function of $(A_{it}, \gamma_{it}, g_{it})$. In contrast, equation (16) assumes that it is a linear function. Adding higher degree polynomials and interaction terms to this relationship can in principle be important. We do so in our empirical implementation below, though this does not affect results in any substantial way.

Note that we can then use equation (7) to relate average rents and population sizes. So we estimate equation (7) as

$$\ln(N_{it}) = \alpha_4 + \beta_6 \ln AR_{it} + \varepsilon_{4it}. \quad (17)$$

According to the model, in a circular city, $\beta_6 = 2 > 0$.

3.2 Effects of Efficiency, Amenities and Frictions on City Size

To bring the model to the data, we need to construct a new data set on U.S. metropolitan statistical areas (MSAs) for the period 2005-2008. Apart from output and rental prices, few ready-to-use data are available at the MSA level. We rely on a combination of proxies previously used in the literature and micro-data to come up with measures for the other relevant variables, such as consumption, hours worked, and capital. Appendix B provides details on the construction of the data set. Computing the “labor wedge” and the “efficiency wedge” requires making assumptions on the values of some parameter values. Table 1 reports the parameter values we use and provides brief comments on the chosen values. Essentially, they make the model match key moments in the aggregate.

To implement the empirical exercise of the previous section, we pool the data for 2005-2008, and include time dummies in all regressions. One further difference is that we also include an interaction term $\varepsilon_{1it}\varepsilon_{2it}$ in equation (16), since we found it to be statistically highly significant. We denote the

coefficient associated with this interaction term β_7 . Standard errors for equations (16) and (15) are obtained by bootstrapping, since some of the regressors are estimated.⁴ The results are presented in Table 2.

Table 1		
Parameter	Value	Comments
ψ	1.4841	McGrattan and Prescott (2009)
θ	0.3358	McGrattan and Prescott (2009)
r	0.02	Standard number in the literature
δ	0.02	Theory requires $\delta = r$

As is clear from Table 2, all coefficients have the signs implied by the model and are highly significant. The estimation of equations (14), (15), (16), and (17) yield R^2 values of, respectively, 0.14, 0.37, 0.25, 0.18. The model implies that in a circular city $\beta_6 = 2$. The value we find is close to two and we fail to reject the hypothesis that it is equal to two at the 5% level.

Table 2				
j	β_j	s.e.	p-value	Theoretical Prediction
1	2.0964	0.3727	0.000	+
2	0.4127	0.0234	0.000	+
3	0.1283	0.0461	0.005	+
4	0.0959	0.0070	0.000	+
5	-0.2020	0.0420	0.000	-
6	2.1400	0.1652	0.000	2
7	-0.1841	0.0437	0.000	-
Number of observations: 768				

These results allow us to reach several conclusions. First, highly efficient cities are more populated. This is consistent with numerous empirical studies in the literature. Second, efficient cities are more distorted. Frictions are larger as a result of these cities being larger. The frictions that result from

⁴Correcting the standard errors for clustering by MSA does not qualitatively change any of the results, except for β_3 which is no longer statistically significant.

more efficient cities being larger are positively related to median rents as they are the result of the higher efficiency. Third, frictions that exceed the ones explained by efficiency have a negative effect on land rents and city size. Finally, cities that are larger due to amenities also exhibit larger median rents.

The model and the empirical exercise have allowed us to assess the impact of the three city characteristics (efficiency, excessive frictions, and amenities) on land rents and population size. It has also made the point that the general equilibrium effects are important. However, the empirical log-linear model that we have used does not inherit the entire structure of the model. For example, the derivatives in (12) need not be constant. It is therefore important to go beyond this simple empirical exercise to capture the full richness of the theoretical model. In the next section we propose a methodology to obtain the value of the three key city characteristics, and we use the model to perform counterfactual exercises. We show how the model can be made to account for all the variation in city sizes if we identify amenities as a residual from the theory.

4. COUNTERFACTUAL EXERCISES

4.1 Methodology and Shock Identification

The model provides a straightforward way of performing counterfactual exercises. Equation (11) implies that

$$\begin{aligned} & C_1(\bar{u}_t, \gamma_{it}) - \log(C_2(A_{it}, r_t)) \\ = & (1 + \psi) \log \left(1 - \left(\kappa \frac{2}{3} g_{it} + \frac{\kappa}{C_2(A_{it}, r_t)} \right) \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}} \right) - \psi \log \left(1 - \kappa \frac{2}{3} g_{it} \left(\frac{N_{it}}{\pi} \right)^{\frac{1}{2}} \right). \end{aligned}$$

If g_{it} and τ are small using the approximation $\log(1 - x) \approx -x$,⁵ we obtain

$$N_{it} = \frac{\pi}{\kappa^2} \left(\frac{\log(C_2(A_{it}, r_t)) - C_1(\bar{u}_t, \gamma_{it})}{\frac{(1+\psi)}{C_2(A_{it}, r_t)} + \frac{2}{3}g_{it}} \right)^2 \quad (18)$$

where

$$\begin{aligned} C_1(\bar{u}_t, \gamma_{it}) &= \bar{u}_t + (1 + \psi) \log(1 + \psi) - \psi \log \psi - \gamma_{it}, \\ C_2(A_{it}, r_t) &= (1 - \theta) \frac{A_{it}^{\frac{1}{1-\theta}}}{\left(\frac{r_t}{\theta}\right)^{\frac{\theta}{1-\theta}}}. \end{aligned}$$

⁵This approximation works best if τ_{it} and κ are small. In the exercise below the approximation error is likely very small.

This solves N_{it} as a function of $(A_{it}, \gamma_{it}, g_{it})$. Note that the approximation results in exactly the same derivatives with respect to $(A_{it}, \gamma_{it}, g_{it})$. Furthermore, $\partial N_{it} / \partial \bar{u} < 0$, namely, a higher equilibrium utility (smaller total population) makes concentration of workers in a given city less likely since concentration implies congestion costs.⁶

We can use the equation above to calculate N_{it} given the values of $(A_{it}, \gamma_{it}, g_{it})$ and other parameter values. We can also use these expressions to run counterfactual exercises. In particular we can calculate counterfactual distributions of city sizes assuming that all cities have similar values of any of the exogenous city characteristics $(A_{it}, \gamma_{it}, g_{it})$. Note that \bar{u}_t has to be selected such that the resulting city sizes satisfy the labor market clearing condition (13).

In order to perform any of these exercises we first need to develop a strategy to calculate $(A_{it}, \gamma_{it}, g_{it})$ for each city. A_{it} can be calculated directly from available data since

$$A_{it} = \frac{y_{it}}{k_{it}^\theta h_{it}^{1-\theta}}$$

and we have data on y_{it} , h_{it} and k_{it} . This is what we did in the empirical implementation above. An alternative way of calculating the relevant productivity and skill term without using k_{it} (which is potentially poorly measured in the data) is to use the prediction of the model on capital allocation. In particular the model implies that $r_t k_{it} = \theta y_{it}$ and so

$$k_{it} = \frac{\theta y_{it}}{r_t}.$$

We can then let

$$A_{it} = \frac{y_{it}^{1-\theta}}{\left(\frac{\theta}{r_t}\right)^\theta h_{it}^{1-\theta}}.$$

Equation (18) assumes that capital is determined in this way and so this method has the advantage of being theoretically more consistent (although it does not use the actual data on capital stocks). We have added capital in both ways, and found the results to be similar. The correlation of the model-based capital stock measure and the empirical capital stock measure is 0.9. Therefore, we omit here the exercise with the theoretical levels of capital and focus on the one where we use the

⁶Throughout this section we calculate an agents' utility based on his labor and capital income but not the income that agents obtain from land rents. Land is owned by absentee landlords and so rental income does not enter an agent's utility and does not affect their decision to move. We have calculated all of the results below if we use the alternative assumption that workers in a city own a diversified portfolio of land in the city and so obtain as income the average rents. The results under this assumption are essentially identical (utility differs only by less than 0.001) to the ones with absentee landlords, both in the case with and without externalities. The reason is that we are always normalizing the level of utility that reproduces the size distribution to $\bar{u} = 10$ and only relative utilities matter to determine location decisions.

empirical measure of the capital stock. (Figure A9 in the appendix shows the results when we use theory-based capital levels).

Obtaining values for the other two city characteristics is more complicated. First note that equation (10) can be used to estimate g_{it} . Based on this equation we can run the simple log-linear regression

$$\ln \tau_{it} = \alpha_5 + \beta_8 \ln N_{it} + \varepsilon_{5it}. \quad (19)$$

Equation (10) then implies that $\varepsilon_{5it} = \ln g_{it}$. Using data for 2005-2008 and adding time dummies, we estimate $\beta_8 = 0.08$ with a standard error of 0.009. The R^2 of the regression is 0.10. Note that equation (10) implies that $\beta_8 = 0.5$, so the estimated coefficient is significantly lower than the one predicted by the theory. The theoretical model can easily be modified to account for this mismatch if we allow for increasing returns in the provision of some forms of infrastructure. In particular if $\tilde{g}_{it}(N_{it}) = g_{it}N_{it}^{-\phi}$, substituting $\tilde{g}_{it}(\cdot)$ for g_{it} in equation (10) would yield the prediction that $\beta_8 = 0.5 - \phi$. Independently of how the coefficient β_8 maps with the theory, the equation allows us to recover the frictions not associated to population size, which constitute the natural logarithm of the city’s “excessive frictions”.

We still have to obtain the value of γ_{it} . There are a variety of ways to do this. The one that is most consistent with the theory is to use equation (18) and solve for the set of γ_{it} that make the model match city sizes exactly, given some normalization of \bar{u} (we set $\bar{u} = 10$). We can then fix γ_{it} and perform counterfactual exercises. Of course, this exercise depends on the value of all parameters in the model. As before, we use the parameters in Table 1, which match key moments in the aggregate. One important parameter in determining γ_{it} is κ , for which we have not assigned a value yet. To obtain a value for κ we also use equation (19) that regresses the labor wedge on city population size. In discussing the results of estimating this equation, we already pointed out that the theoretical value of the coefficient on population size (β_8) should be 0.5 according to the theory. So we fix it at this value and estimate the constant α_5 . Furthermore, equation (9) then implies that

$$\alpha_5 = \ln \left(\frac{2}{3} \right) + \ln \kappa - \frac{1}{2} \ln \pi$$

and so, given a value for α_5 , we can calculate κ . The estimation yields a constant equal to -7.3179 , which yields a value of $\kappa = 0.00176$. The time dummies we include are mostly not significant, but adding their values would yield a constant that is somewhat smaller which would imply a lower κ . Therefore, we use a value of $\kappa = 0.001$ and do many robustness checks for higher values of κ .

4.2 Counterfactuals

We are now ready to perform a number of counterfactual exercises. After analyzing the effect of commuting costs, the main focus will be on exploring the relative importance of different characteristics (efficiency, amenities and excessive frictions) in determining the city size distribution. In particular, we are interested in understanding how changes in city characteristics affect city sizes, welfare and the reallocation of people.

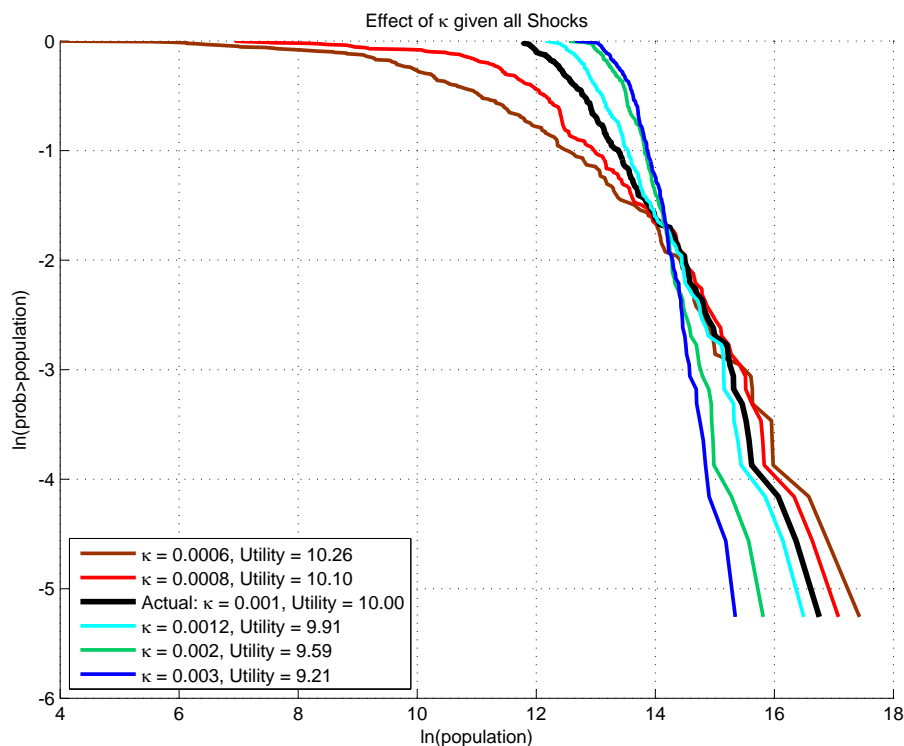


Figure 1: The City Size Distribution for Different Levels of Commuting Costs

Figure 1 shows the actual distribution of city sizes in the U.S. and counterfactual distributions of city sizes if we increase or decrease commuting costs κ , given the distribution of characteristics. The results are presented in the standard log population – log rank plots in which a Pareto distribution would be depicted as a line with slope equal to minus the Pareto coefficient. As is well known, the actual distribution is close to a Pareto distribution with coefficient one. By construction the model

matches the actual distribution exactly for $\kappa = 0.001$. In that case we normalize utility $\bar{u} = 10$.⁷ In all counterfactual exercises we solve for the value of \bar{u} for which the labor market clears, i.e., the sum of population across cities equals the actual total urban population. Note that percentage changes in utility are equivalent to percentage changes in consumption since we are using a log-utility specification. Figure 1 shows that larger commuting costs make the largest cities smaller and the smaller cities larger, leading to a less dispersed distribution of city sizes. The 20% increase in commuting costs decreases utility by about 0.9%. The decline in commuting costs increases dispersion and raises utility by 1%. Note that the smallest cities now become much smaller. The main advantage of some of these cities was their small size and their low commuting costs. As commuting costs decrease, this advantage becomes less important and their size decreases further. If we double or triple κ to .002 and 0.003 we obtain changes in utility of around 4% and 8%. Large cities now become much smaller, which implies that production is not done in the most productive and highest amenity cities which leads to substantial welfare losses.

Figure 2 shows three counterfactual exercises where we shut down each of the three shocks (efficiency, amenities, and excessive frictions) respectively. In all cases we eliminate a shock by setting its value to the population weighted average of the shock. We then calculate the utility level that clears the labor market, so total urban population is identical in all cases. Note that eliminating any of the shocks always leads to an increase in utility. Shocks create dispersion in the city size distribution and by equation (8) total commuting costs are convex. So utility in the model tends to increase if population is uniformly distributed in the 192 cities in our sample. If we eliminate all three shocks so that all sites are identical, welfare would increase by 3.26% and all cities would have a population of 1 million 68 thousand people. Of course, this increase in welfare does not constitute an upper bound on the importance of the shocks since the distribution of the shocks as well as their correlation matters for the final results.

The counterfactual exercises in Figure 2 show that eliminating differences in efficiency and amenities has a small effect on utility. In both case utility would increase by less than 2.5%. The limited effect on utility is due to several reasons. The most obvious one is that population can reallocate across cities. But there are others. For example, the effect of a negative shock to productivity on

⁷This normalization affects the size of the welfare changes due to the level taken by the amenity characteristics. On average we find that the share of utility coming from amenities is around 28%. We have no evidence that this number is reasonable. However, it is easy to adjust the numbers to be consistent with any share of amenities in utility. For example, if the share of utility coming from amenities is set to zero, all our utility numbers from counterfactuals would need to be multiplied by $1/(1 - .28) = 1.39$.

utility is also mitigated by people working less, by lowering the cost of providing city infrastructure, and by the fact that utility does not only depend on production but also on amenities.

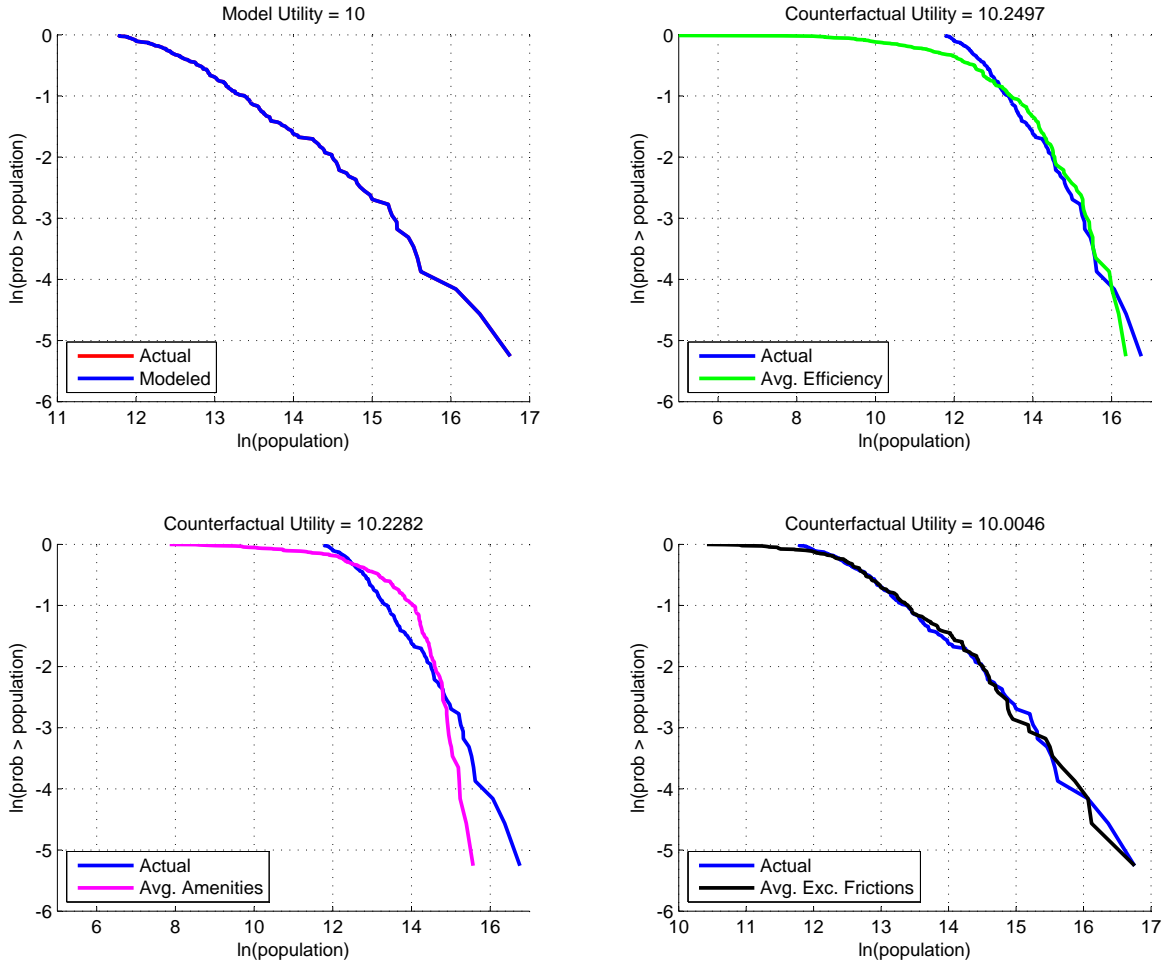


Figure 2: Counterfactuals Without one Shock

In spite of the small effect on utility, the effect on the size of individual cities is large. Many cities would substantially decline in size, especially the largest ones. For example, Los Angeles Long Beach-Santa Ana would lose 85% of its population if it had average amenities, and 35% if it had average efficiency. Similarly, the corresponding figures for New York-Northern New Jersey-Long Island are 79% and 53%. Maybe surprisingly, some of the small cities also decline. When this happens, it is because they lose their only comparative advantage. One such example is Santa Fe: if it had average amenities, it would lose 99% of its population. Intermediate-sized cities in general benefit as they

tend to experience a boost in productivity or amenities, and already are attractive enough in terms of other characteristics. These cities also grow because of the reallocation of population from larger cities. In contrast to amenities and efficiency, eliminating excessive frictions has only a small effect on the shape of the distribution, although again some individual cities would experience significant decreases in size, as they were large only because of their advantage in having low excessive frictions. Phoenix-Mesa-Scottsdale, for example, would lose 48% of its population if it had average excessive frictions.

Appendix A shows figures and maps with the percentage changes in population for individual cities when we set one of the shocks to its weighted average. In terms of the geographic distribution of city characteristics, we find that most cities in the coastal regions would lose population if we eliminated amenity differences. This is consistent with Rappaport and Sachs (2003) who argue that the concentration of population in coastal areas has to do increasingly with a quality of life effect. Central regions would tend to lose population if we eliminated efficiency differences as would most of the north-eastern regions. Perhaps the sharpest geographical pattern emerges when we eliminate excessive frictions. Most of the mid-west and north-east (the region that includes the “Rust Belt”) would gain population if we equalize frictions across cities: an indication that governance problems, as well as other labor market frictions, like unions, may be important in these regions.

The effect of the different shocks on the distribution of city sizes hides some of the implied population reallocation in these counterfactuals. That is, cities are changing ranking in the distribution even if the overall shape of the distribution does not exhibit large changes, as in the case of excessive frictions. We can calculate reallocation following Davis and Haltiwanger (1992) by adding the number of new workers in expanding cities as a proportion of total population when we change from the actual distribution to the counterfactual. This measure of reallocation is 51.79% when we eliminate differences in TFP, 42.54% when we eliminate amenities, and 9.9% when we eliminate excessive frictions: large numbers given the modest welfare gains. As a benchmark, the same reallocation number for the U.S. economy over a 5 year interval is around 2.1% (2.14% between 2003 and 2008 or 2001 and 2007).

Figure 3 shows the counterfactual distributions of city sizes when we shut down two of the three shocks. The distributions therefore show the heterogeneity in city sizes generated by a single shock or characteristic. Note that neither efficiency on its own nor amenities on their own can explain the relatively large sizes of the smallest cities in the actual distribution. This is because some of these small cities are attractive in terms of the other characteristics, making them larger than their

efficiency or their amenities would imply.

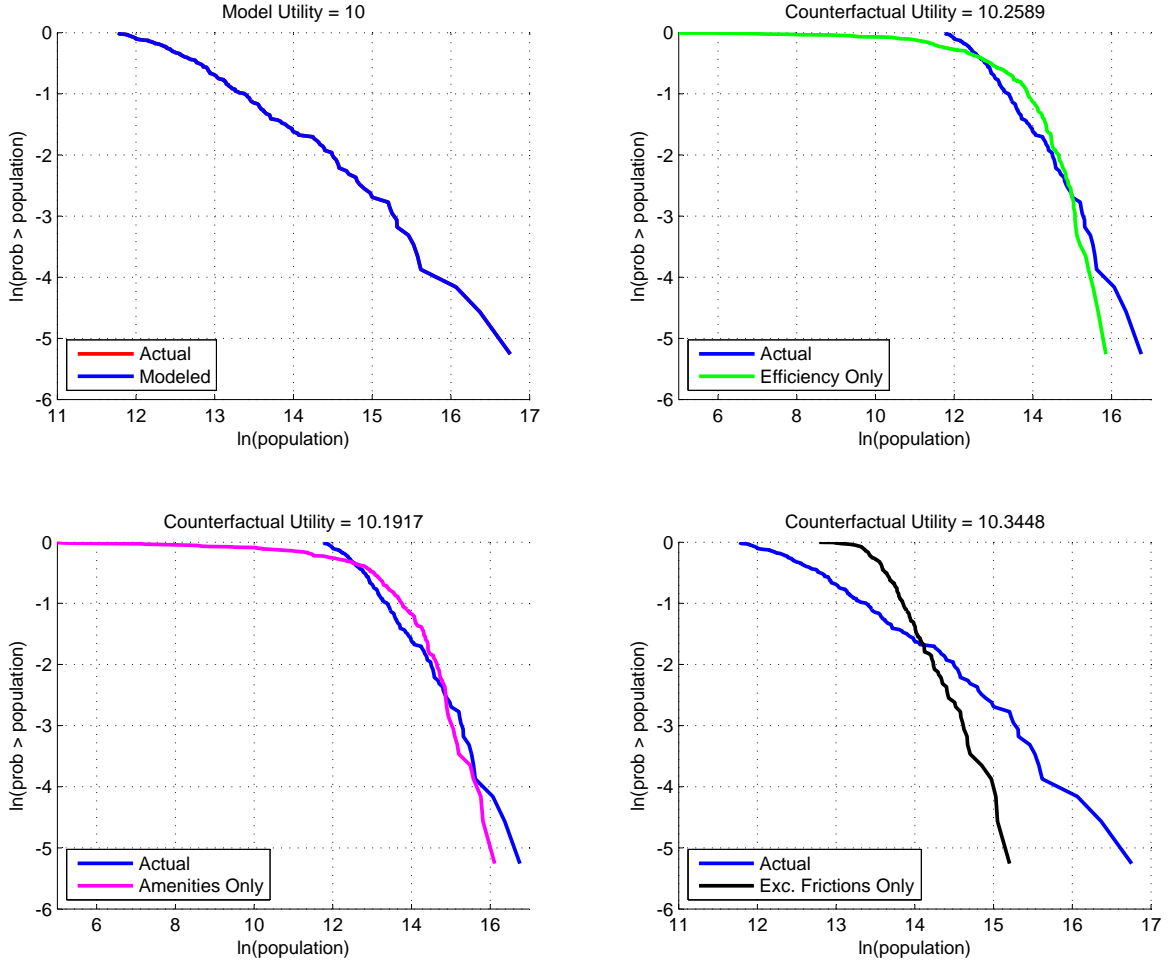


Figure 3: Counterfactuals with Only one Shock

The level of commuting costs κ , which we have estimated to be equal to 0.001, is a key parameter in determining the relative importance of each shock. The estimated value implies that agents spend about 2% of their wage income on commuting infrastructure.⁸ The larger κ , the smaller the relative importance of productivity shocks, since it becomes more costly to live in large productive cities and the people that live in them tend to work less since τ is larger. If we set $\kappa = 0.005$, a fivefold increase, the total reallocation if we equalize efficiency across locations drops from around 52% to 10%, with

⁸This number seems reasonable: estimates of government spending in transportation infrastructure as a share of GDP come up with figures around 1.5% (Congressional Budget Office, 2007).

a 1.5% increase in utility, a smaller effect than before. By decreasing the relative importance of productivity shocks, higher transport costs increase the implied dispersion in amenities across cities. As a result, with a fivefold increase in κ , reallocations increase from 43% to 45% when cities have average amenities, and utility now goes up by 14.3%, instead of by 2.3%. The reallocation if we set excessive frictions to their average level remains relatively constant at 10%. The changes in city sizes are highly correlated in the exercises with the two different values of κ . The correlation in the population changes if we keep amenities constant is 0.53, if we keep efficiency constant 0.95, and if we keep excessive frictions constant 0.98.

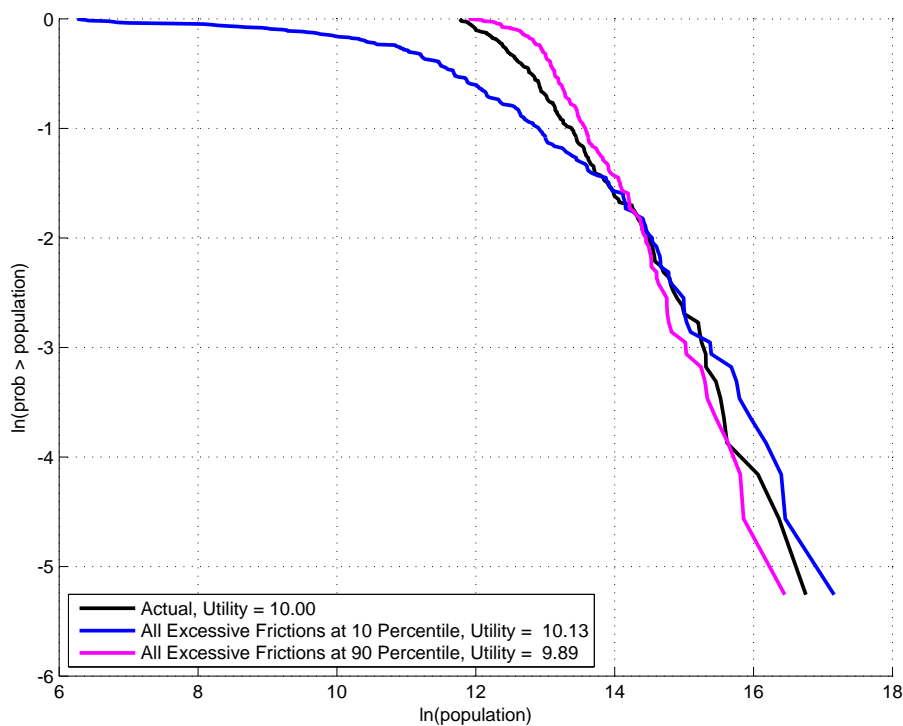


Figure 4: Changing the Level of Excessive Frictions

Figure 4 shows a counterfactual exercise when we set excessive frictions in all cities equal to the 10th or 90th percentile of the distribution of excessive frictions. The figure shows that reducing frictions in all cities to the 10th percentile increases the dispersion of city sizes. Large cities benefit the most from the change, and small cities suffer in terms of population. Utility increases by 1.3%. The opposite effect results from setting frictions to the 90th percentile, although the changes in the distribution are much more modest. In this case utility declines by 1.1%, even though the distribution

is less dispersed. Note that by construction $\ln g_{it}$ and $\ln N_{it}$ are orthogonal so this relationship is just the result of the fact that large cities use the commuting technology more intensively, as can be seen in Section 2.4.

4.1 Adding Production Externalities

So far we have taken productivity in a particular city to be exogenously given. We have assumed that the efficiency of a particular site is not affected by the level of economic activity at that site. That is, so far, efficiency has explained agglomeration, but we have assumed away the reverse link by which agglomeration explains efficiency. Of course, a standard view in urban economics suggests that agglomeration is, at least in part, created by an increase in productivity coming from a rise in the number of people living in a given city. Including these agglomeration effects in our calculations has the potential to change our results, as this will have an endogenous effect on the size of a city.

To incorporate these effects, we start with equation (18) but recognize that the term A_{it} , which captures the efficiency of city i , is a function of the size of the city N_{it} . In particular, we now let

$$A_{it} = \tilde{A}_{it} N_{it}^{\omega}. \quad (20)$$

That is, the level of productivity is now a function of an exogenous shock \tilde{A}_{it} , and city size, N_{it} , where the elasticity of the efficiency wedge with respect to population is given by ω . Note that externalities operate within cities, and not across cities. We can then use the previous calculation of efficiency wedges in the data, using equation (2), and divide by population raised to ω . The result is a set of new exogenous efficiency levels \tilde{A}_{it} . We then substitute (20) in (18) and solve for the γ_{it} 's that yield the city's exact population levels. Excessive frictions are calculated as before. With all the city shocks in hand, we can now perform the same set of counterfactual exercises as before. Note that equation (18) now includes N_{it} in the productivity terms and so cannot be solved analytically. However, we can solve the system of nonlinear equations numerically to obtain city sizes in the counterfactual exercises.

We still need to determine a suitable value for ω . Of course, the estimation of equation (14) is not useful to determine ω . In fact, this equation will fit exactly as in the data in our simulation of the actual economy. Instead, we rely on the literature, that suggests a fairly robust estimate of $\omega = 0.02$ (see, amongst others, Carlino, Chatterjee and Hunt, 2006, and Combes, et. al, 2009). We therefore start with an initial value of 0.02 and perform some robustness checks. We also set $\kappa = 0.001$ as estimated in the previous section.

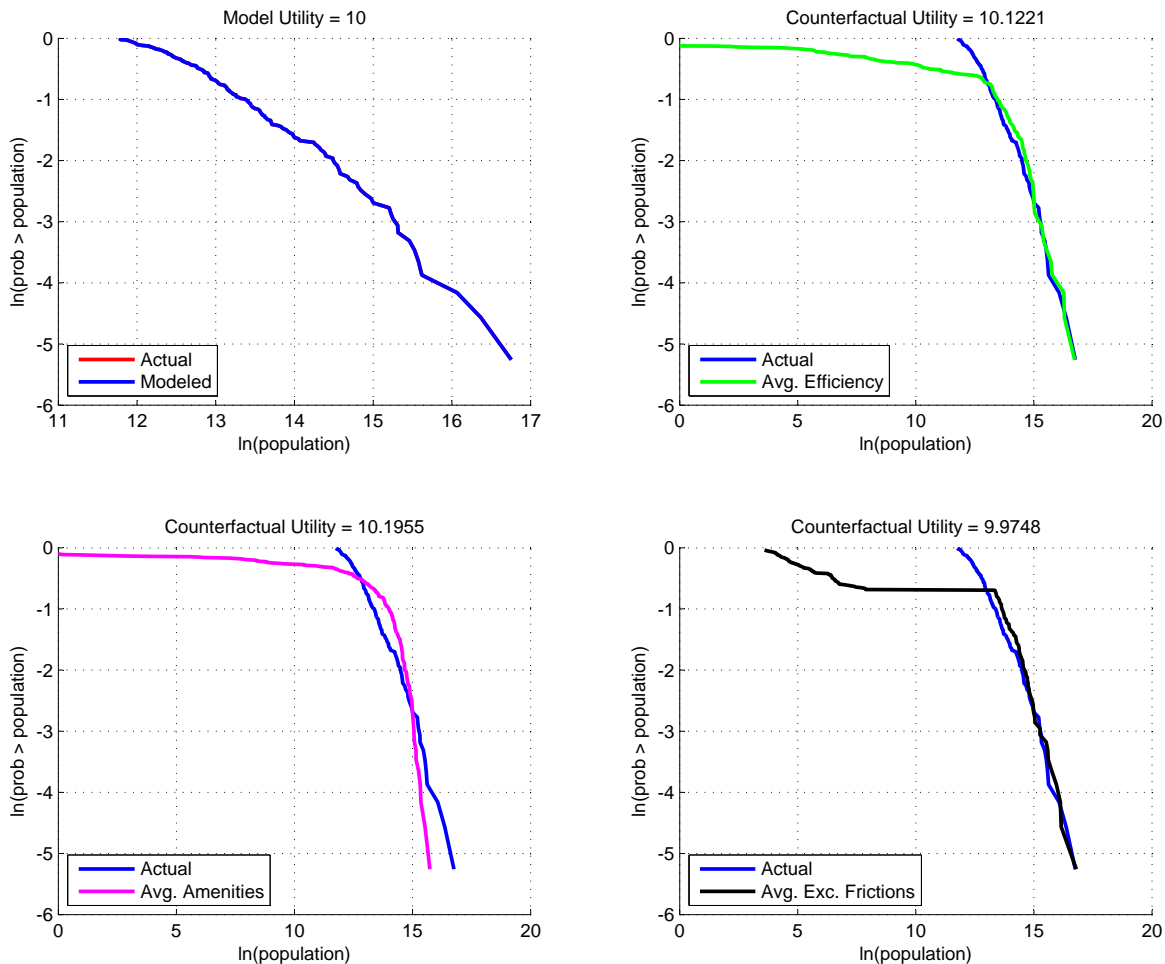


Figure 5: Counterfactuals Without one Shock and Externalities, $\kappa = 0.001$, $\omega = 0.02$

Figure 5 presents the exercise with externalities in the case where we eliminate each shock individually. First note that when we eliminate one of the shocks some cities no longer survive. In particular, 5 cities exit when we equalize \tilde{A}_{it} across cities to its population weighted mean, 6 cities exit when we set amenities to their average value, and 7 cities exit with average excessive frictions. As in the case without externalities, these are cities that lose their only comparative advantage. With externalities, this loss gets compounded, leading some small cities to exit. As for the large cities, they tend to become a lot smaller when eliminating differences in amenities, whereas their size does not change much when equalizing exogenous efficiency levels. This latter result can be explained by the smaller dispersion in exogenous efficiency. As for medium-sized cities, they generally benefit

from the drop in size in many other cities. Compared to the case without externalities, utility drops. With average efficiency, utility declines from 10.25 to 10.12, and for average amenities utility declines from 10.23 to 10.20. The utility levels with average excessive frictions declines from 10.005 to 9.97. The gains are smaller with externalities because even though the convex commuting costs lead to gains if cities become more alike, differences in city characteristics make cities exploit the gains from the external effects. As we increase the elasticity of efficiency to population size, the losses from the second channel dominate when we eliminate differences in one of the city characteristics.

Compared to the case without externalities, total reallocation required in the counterfactual tends to go down when we eliminate efficiency differences (41.54%). This is natural as now a fraction of the efficiency differences is explained by size. For the other two characteristics, reallocation increases slightly to 44.03% in the case of average amenities, and more significantly to 18.46% in the case of average excessive frictions. This happens because the changes introduced by the elimination of these shocks get compounded through the effect of changes in population on efficiency.

Figure 6 doubles the externality to $\omega = 0.04$. This is closer to the estimate of 0.05 reported by Behrens, Duranton and Robert-Nicoud (2010). In general, the effects discussed before get exacerbated. Many more cities either exit or become very small. The results suggest that selection of cities in the presence of externalities can be extremely important. Relative to the case without externalities, the increase in externalities lowers the utility gain obtained if we equalize one of the city characteristics. It also increases the decline in utility in the case of average excessive frictions where many cities decline in size substantially. By equalizing exogenous characteristics, some of the more attractive cities no longer benefit from agglomeration economies, and utility drops.

Adding externalities in production implies that the equilibrium allocation we compute is no longer efficient. In contrast with the exogenous productivity case, city planners could improve on the equilibrium allocation by subsidizing urban agglomeration. We can compute the optimal allocations in the case with production externalities by letting a representative firm internalize the external effect on productivity. Since the differences in welfare between the cases with and without externalities are so small, it is not surprising that the effect of these optimal urban policies is necessarily small as well. In fact, the gain in utility is only 0.06%. In the optimal allocation, some cities become very small and all the rest increase their size slightly. Furthermore, the informational requirements for these urban policies is extremely high so it is not clear that actual policy can achieve these small gains. We include a graph that compares the optimal and actual allocations in Appendix A.

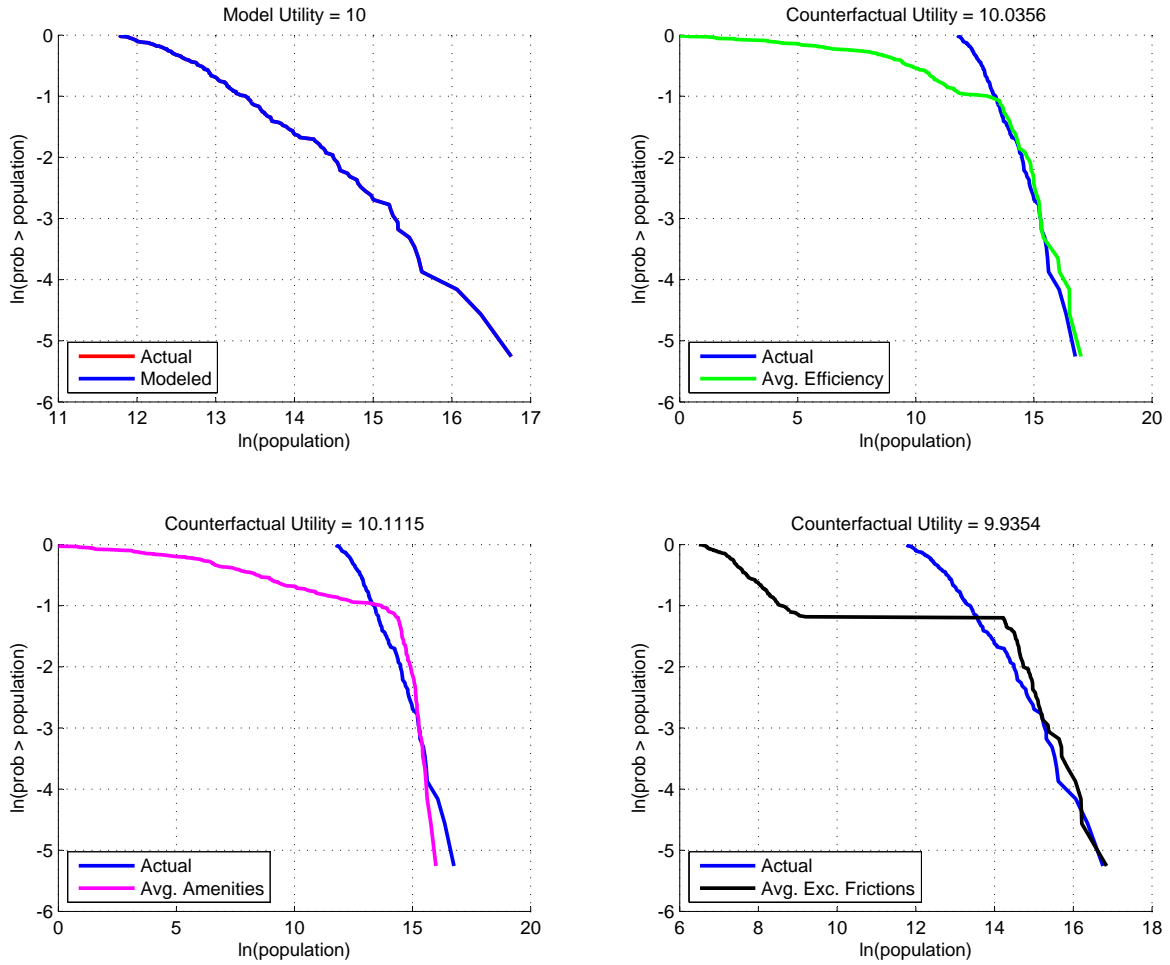


Figure 6: Counterfactuals Without one Shock and Externalities, $\kappa = 0.001$, $\omega = 0.04$

We should also mention here that the exercise with externalities leads to the possibility of multiple equilibria in the size of cities. For many cities it will be the case that, given the equilibrium utility level, there is only one equilibrium size. But for other cities it will be the case that there are several possible equilibrium sizes. Our theory does not provide a way of selecting between these equilibria so we always present the one that requires less reallocation. That is, we always initialize the search for a solution of the size of a city at its actual size.

4.2 Adding Externalities to Amenities

We can also add externalities in the amenities a city provides. That is, we can let the utility from living in a particular city depend on the size of the city directly. People live in New York because living around a large number of people leads to a scale that provides them with a variety of goods and services, and interactions with people, that they enjoy. We have modeled the preference to live in a particular city through the amenity shocks γ_{it} . So we can simply let

$$\gamma_{it} = \tilde{\gamma}_{it} N_{it}^{\zeta},$$

where now $\tilde{\gamma}_{it}$ is the exogenous amenity shock and ζ is the elasticity of amenities with respect to population size.

We repeat the exercise in Figure 5 but now we let $\zeta = 0.02$ as well. Figure 7 shows the results. The results are qualitatively similar but now we observe that more cities become extremely small. That is, the selection mechanism we emphasized above becomes stronger. Utility in all counterfactuals decreases relative to Figure 5. Including the externality in amenities results now in small losses in the case of identical exogenous productivity across cities, and 0.5% gains in the case where exogenous amenities are constant. The decrease in utility is natural given that with externalities in both production and amenities, equalizing city characteristics implies that externalities are not exploited as much.

Perhaps surprisingly, the effect on utility of eliminating the differences in any of our three characteristics are extremely small in magnitude even though the implied reallocation of agents are, again, fairly large. Eliminating efficiency differences reduces utility by less than 0.1% but implies that 42% of agents reallocate. The same reallocation statistic when we eliminate amenity differences is 40% and 32% for excessive frictions. Most of the reallocation come from the extensive margin. Many cities become extremely small: the city selection effect. Once again, by equalizing a given characteristic, some small cities lose their only comparative advantage. This loss is compounded by the existence of externalities, so that quite a few smaller cities become so small that they exit. However, the reallocation has small effects on agent's utility since even though small cities do not experience the benefits of large externalities they are not distorted through taxes since city infrastructure is cheap. The slope of the envelope of the value of living in different cities is extremely flat, so agents switching location leads to small utility gains.

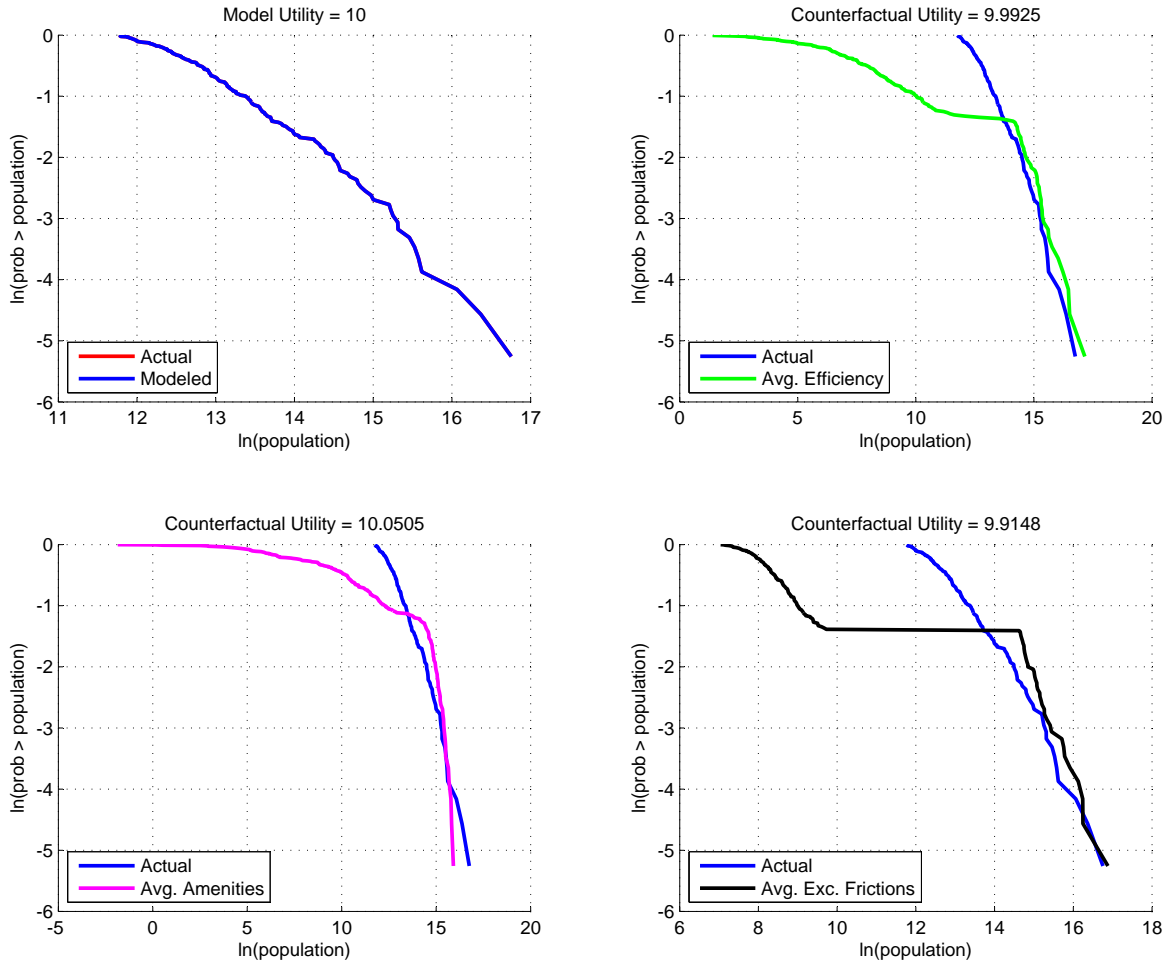


Figure 7: Counterfactuals Without one Shock and Externalities, $\kappa = 0.001$, $\omega = 0.02$, $\zeta = 0.02$

City selection can be most easily understood by studying what happens if we eliminate differences in all three city characteristics. In this case the urban structure has 61 cities with 3,360,745 agents and the other 131 cities essentially disappear and preserve a population of only 613 agents in each of them. Without any city characteristics, but with externalities, there are two city sizes that give agents identical utility levels and the number of cities in each size is determined by the market clearing condition so that all agents are housed in some city. So there is an equilibrium which specifies the number of cities of each type. The utility level in this case is 10.047. Thus, eliminating all shocks yield benefits to agents as most of them can live in larger cities and the ones that do not, live in small town that have no congestion or infrastructure costs. Note again that since there are no shocks we know that in this case there may be multiple equilibria. As before, in all cases we

compute the equilibrium with minimal reallocation of agents across cities, which also yields a level of utility closest to the one in the actual distribution, which we normalize to 10.

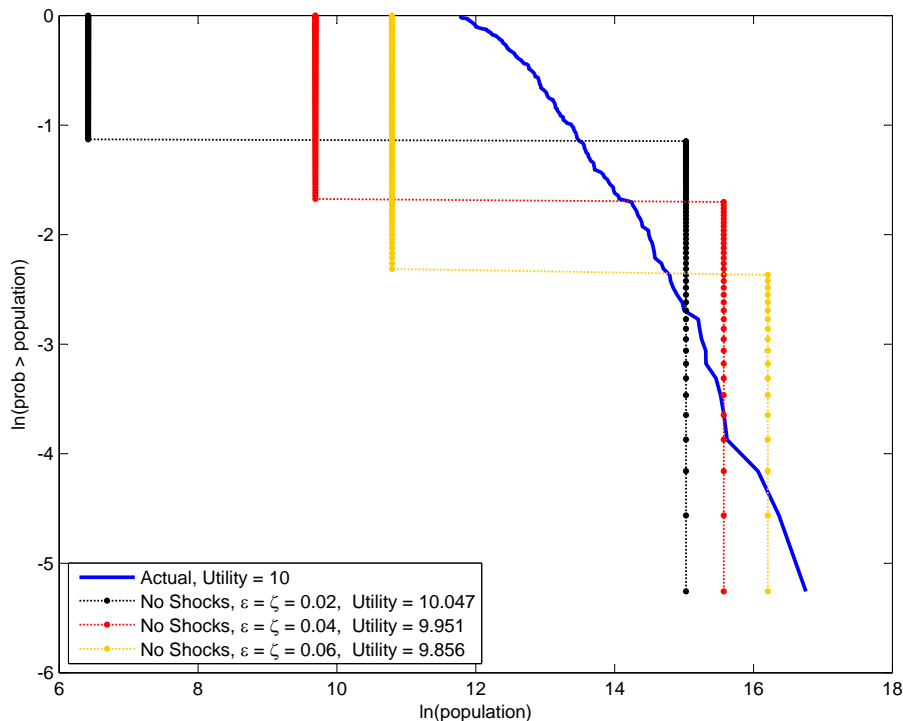


Figure 8: Counterfactuals Without Shocks, $\kappa = 0.001$

Figure 8 shows counterfactuals without shocks for different elasticities of city efficiency and amenities to population size. Clearly, as we increase the elasticity, and therefore the externality, we still have two sizes of cities, but the larger the externality, the larger and fewer the larger cities. So larger externalities make the larger and smaller cities larger, and increase the number of small cities. Furthermore, the larger the externality, the lower utility in the counterfactual without shocks. When externalities are large, differences across cities create agglomeration and result in benefits. Eliminating them yields lower utility.

5. THE EFFECT OF EFFICIENCY AND AMENITY SHOCKS

In this section we study the effect on welfare and labor reallocation of shocks to efficiency and amenities. The model we have proposed to analyze the urban hierarchy has no dynamics, since all

factors reallocate in a frictionless manner across cities. There are some dynamics in the process of accumulating capital but we neglect those and focus on the steady state (since these dynamics are just aggregate and have been well studied in the macroeconomics literature).

The purpose here is to understand what are the magnitudes of welfare losses if the economy is hit by, say, a 20% negative efficiency shock. We consider a proportional decrease in productivity in all cities and a random uniformly distributed proportional shock⁹ that amounts to different declines in average productivity. Figure 9 shows the effect on utility (top panel) and worker reallocation (bottom panel). In this exercise we calculate utility including the gains from changes in land rents in order to account exactly for agent’s welfare (although this has a negligible impact).

As can be seen in Figure 9, if every city experiences a proportional reduction in efficiency, the urban hierarchy remains the same, so there is no reallocation of agents. In this case, the value of commuting costs κ does not affect the slope of the welfare losses: the negative effect on welfare of an increase in the shock is independent of κ .¹⁰ The most important observation is that although shocks reduce utility, they do so much less than proportionally. The reason is threefold. First, agents can adjust their leisure as a result of the shock as well as the amount of capital. Second, agents obtain utility out of the amenities in the city and so consumption of goods is only one of the elements that determines an agent’s utility. The utility that agents obtain from city amenities amounts, on average, to around 28% of an agent’s utility. Third, commuting costs and the distortions created to build the related public infrastructure go down with productivity and wages.

If cities receive random, but on average negative, TFP shocks, the welfare costs are smaller, since agents can relocate to the cities that received relatively good shocks. For example, for shocks that reduce average TFP by 30%, welfare decreases by 4.6% rather than 5.3%. If we double commuting costs to $\kappa = 0.002$, the overall utility loss is slightly higher, as reallocation tends to go from smaller to larger cities. The impact of higher commuting costs on reallocation is more substantial, since higher congestion costs slow down the move towards larger cities. Overall, random shocks create significantly more reallocation, but the welfare implications are minor. The fact that all curves in the top panel are close to each other suggests that changes in commuting costs, as well as including externalities in production, have small effects on the welfare consequences of these shocks. The figure also shows that in all these cases the urban structure, and in particular migration between cities,

⁹We use a proportional shock to productivity of the form $s_{it} = 0.5 + x_{it}$, where $x_{it} \sim U(0, 1)$, and s_{it} is multiplied by 1 minus the size of the average shock we analyze.

¹⁰Of course, an increase in κ leads to lower utility as more output is lost on commuting, but this does not show up in the graphs, since in all cases the utility has been normalized to 10 when the shock is zero.

mitigates the impact of productivity shocks. It is also clear from the graph that the benefits from the implied population reallocations are extremely small. In general, adding externalities leads to more reallocation since many cities disappear. However, the welfare impact of these selection effects is, again, very small.

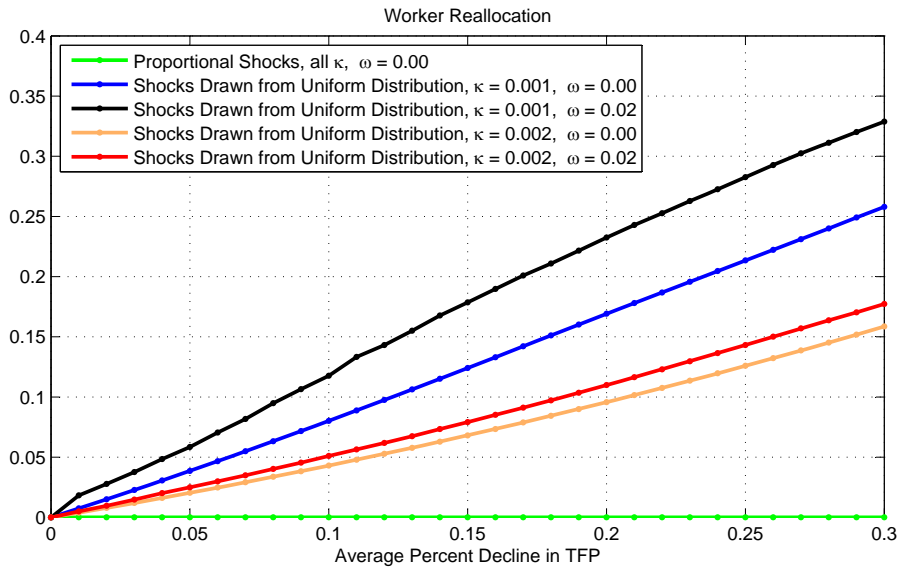
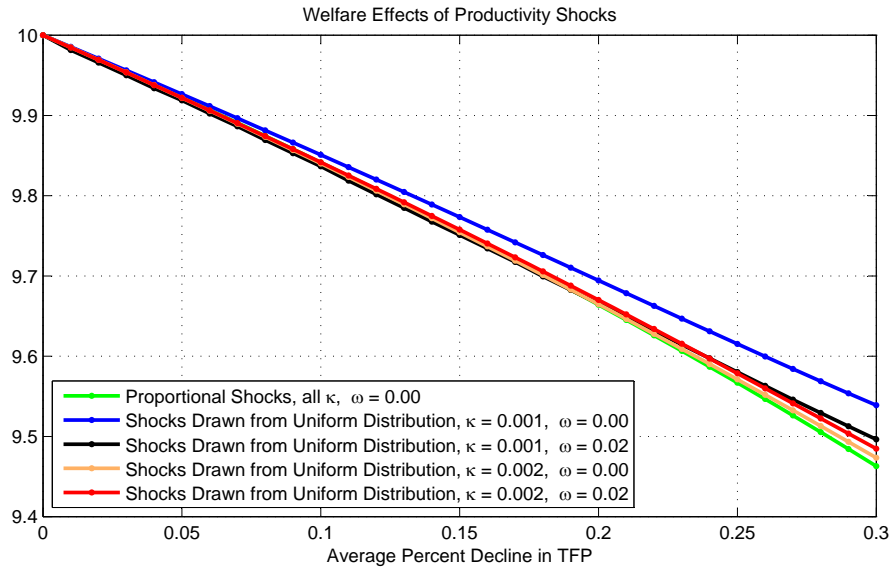


Figure 9: The Effect of Negative Productivity Shocks

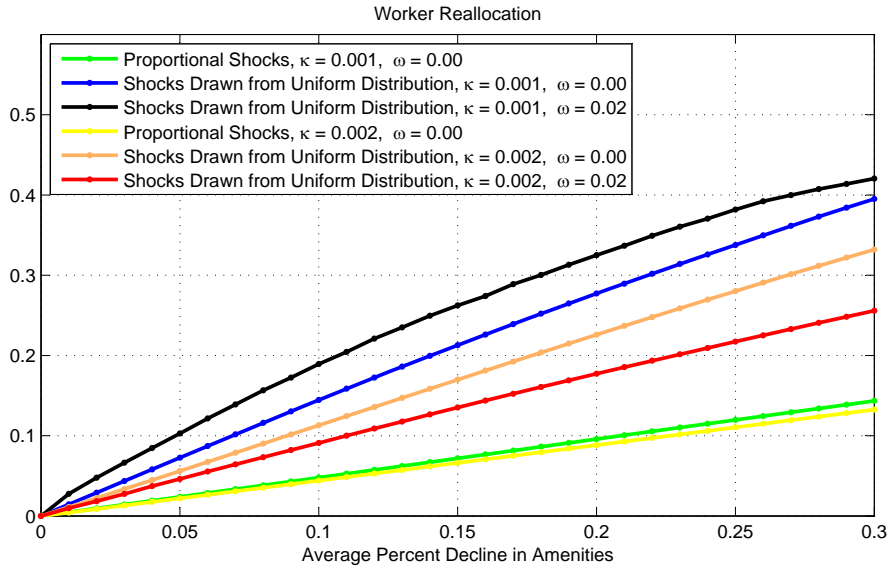
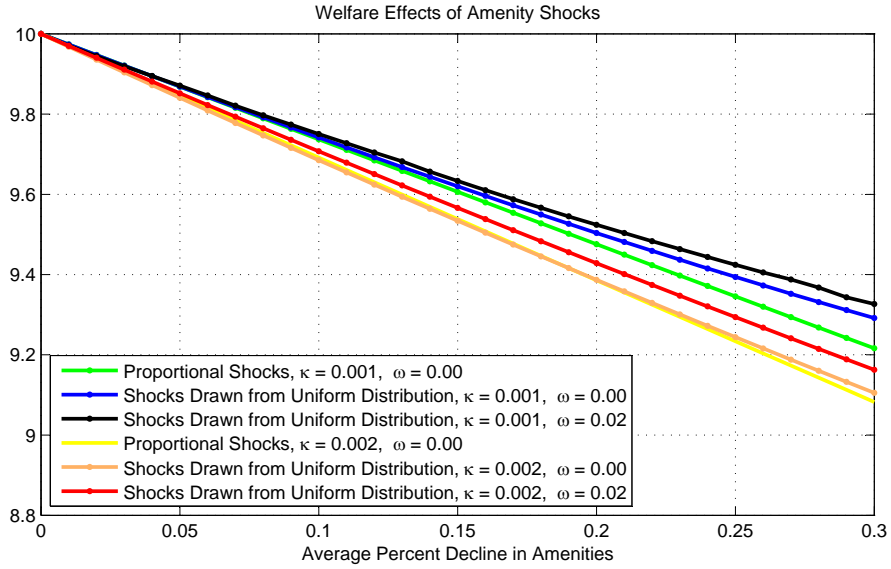


Figure 10: The Effect of Negative Productivity Shocks

We can also study the effect of amenity shocks. Figure 10 presents the results of a proportional (amenities change in all cities by the same proportion) and a random uniformly distributed proportional shock (the average of the shock amounts to a given proportion). When the amenity shock is proportional, there is reallocation of people across cities since the overall attractiveness of a city is

also formed by productivity and the effectiveness in providing urban infrastructure.

The welfare losses from these shocks are larger than in the case of productivity. Since we know that amenities amount to around 28% of average utility, we would expect welfare to go down by slightly more than one quarter of the original shock. The calculation is not exact since reallocation and other adjustments change the actual welfare losses, but it provides a helpful benchmark. In fact, the good accuracy of this back-of-the-envelope calculation shows that the benefits from reallocation are again very small. Adding externalities in this case always increases welfare, although not necessarily reallocation (the effect depends on the value of κ). In the case of amenity shocks, an increase in the value of κ has a greater negative effect on welfare, compared to the case of productivity shocks. This happens because amenity shocks imply more reallocation than productivity shocks. The conclusion is that shocks that yield on average a 30% reduction in amenities reduce welfare by around 7%.

6. CONCLUSION

In this paper we have decomposed the size distribution of cities into three main characteristics: efficiency, amenities, and excessive frictions. We find that each one of these components is important. Eliminating differences in any of them would imply large reallocations of people. However, the welfare gains or losses associated with particular distributions of these characteristics are very small. Eliminating any differences in characteristics across cities yields welfare gains of at most 3%. Dynamic models in which reallocation affects the growth rate could potentially generate larger effects. So changes in the particular shape of the size distribution, which is the result of the particular distribution of these characteristics across cities, seem to have negligible effects in terms of welfare. Note that the actual population movements required can be larger than 50%, so any small reallocation cost would turn these gains into losses. We also include externalities in both productivity and amenities. The welfare effects associated with eliminating particular characteristics of cities are even smaller in these cases, although we find a strong selection effect in the counterfactual distributions. Namely, many cities exit or become extremely small.

Though not directly related, the small effect on utility from reallocation is vaguely reminiscent of the limited welfare gains found in the literature on trade and capital market integration (see, for example, De Melo and Tarr, 1992, and Gourinchas and Jeanne, 2006). Dynamic models in which reallocation affects the growth rate could potentially generate larger effects. Welfare effects could also be larger if one were to add distributional effects in a model with heterogeneous agents, or if

the number of cities were smaller, making reallocation by moving to similar cities more difficult.

These results also suggest that a potential lack of mobility across locations (that could be caused, for example, by agents being “underwater” with their mortgages) can at most involve limited effects on welfare. We find that a 20% reduction in productivity leads to a reduction in welfare of between 3% and 4%, while a reduction of 20% in amenities leads to a reduction in welfare of between 5% and 6%. These reductions in welfare change only by a small fraction of a percent if we use idiosyncratic city shocks that average to the same decline. These results suggest that the implied reallocation of agents that results from idiosyncratic shocks has small welfare effects.

More generally, we have provided a simple methodology to study the determinants of the size distribution of cities. This methodology can be useful in comparing urban systems across countries. The data requirements to do the exercise are not extreme, and it could shed light on the sources of deviations in urban systems across countries. Such a comparison will be informative about the effectiveness and the welfare effects of different policies aimed at making the location of agents across cities more efficient.

REFERENCES

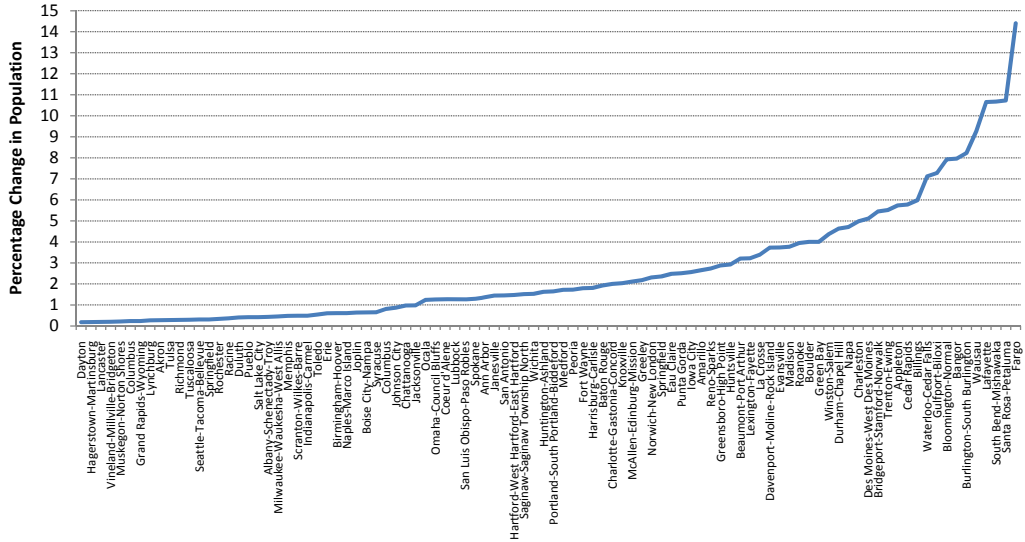
- [1] Albouy, D., 2008. "Are Big Cities Really Bad Places to Live? Improving Quality-of-Life Estimates across Cities," NBER Working Papers 14472.
- [2] Albouy, D., 2009. "What Are Cities Worth? Land Rents, Local Productivity, and the Capitalization of Amenity Values," NBER Working Papers 14981.
- [3] Asdrubali, P., Sorensen, B.E. and Yosha, O., 1996. "Channels of Interstate Risk Sharing: United States 1963-1990," *Quarterly Journal of Economics*, 111, 1081-1110.
- [4] Au, C.-C. and Henderson, J.V., 2006. "Are Chinese Cities Too Small?," *Review of Economic Studies*, 73, 549-576.
- [5] Behrens, K., Duranton, G. and Robert-Nicoud, F., 2010. "Productive Cities: Sorting, Selection, and Agglomeration," unpublished manuscript.
- [6] Bleakley, H. and Lin J., 2010. "Portage: Path Dependence and Increasing Returns in U.S. History," NBER Working Papers 16314.
- [7] Carlino, G., Chatterjee, S., and Hunt, R., 2006. "Urban Density and the Rate of Invention," Working Paper 06-14, Federal Reserve Bank of Philadelphia.
- [8] Chari, V. V., Kehoe, P., McGrattan, E., 2007. "Business Cycle Accounting," *Econometrica*, 75, 781-836.
- [9] Combes, P, Duranton, G., Gobillon, L., Puga, Diego and Roux, S., 2009. "The productivity advantages of large cities: Distinguishing agglomeration from firm selection," CEPR Discussion Papers 7191.
- [10] Congressional Budget Office, 2007. "Trends in Public Spending on Transportation and Water Infrastructure, 1956 to 2004", August.
- [11] Córdoba, Juan Carlos, 2008. "On the Distribution of City Sizes," *Journal of Urban Economics*, 63, 177-197.
- [12] Davis, S.J. and Haltiwanger, J., 1992. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics*, 107, 819-63.

- [13] Davis, D., and Weinstein, D., 2002. "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92, 1269-1289.
- [14] De Melo, J. and Tarr, D.G., 1992. *A General Equilibrium Analysis of US Foreign Trade Policy*, MIT Press: Cambridge, Mass.
- [15] Duranton, Gilles, 2007. "Urban Evolutions: The Fast, the Slow, and the Still," *American Economic Review*, 97, 197-221.
- [16] Duranton, Gilles and Henry Overman, 2008. "Exploring the Detailed Location Patterns of U.K. Manufacturing Industries Using Micro-Geographic Data," *Journal of Regional Science*, 48, 213-243.
- [17] Gabaix, X., 1999a. "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, 114, 739-767.
- [18] Gabaix, X., 1999b. "Zipf's Law and the Growth of Cities," *American Economic Review*, 89, 129-132.
- [19] Garofalo, G.A. and Yamarik, S., 2002. "Regional Convergence: Evidence from a New State-by-State Capital Stock Series," *Review of Economics and Statistics*, 84, 316-323.
- [20] Glaeser, E., Kolko, J., and Saiz, A., 2001. "Consumer City," *Journal of Economic Geography*, 1, 27-50.
- [21] Glaeser, E., Gyourko, J. and Saks, R., 2005. "Why Is Manhattan So Expensive? Regulation and the Rise in Housing Prices," *Journal of Law and Economics*, 48, 331-369.
- [22] Gourinchas, P.O. and Jeanne, O., 2006. "The Elusive Gains from International Financial Integration," *Review of Economic Studies*, 73, 715-741.
- [23] Hess, G.D. and Shin, K., 1997. "International and Intranational Business Cycles," *Oxford Review of Economic Policy*, 13, 93-109.
- [24] Holmes, T., 2005. "The Location of Sales Offices and the Attraction of Cities," *Journal of Political Economy*, 113, 551-581.
- [25] Holmes, T., and Stevens, J., 2002. "Geographic Concentration and Establishment Scale," *The Review of Economics and Statistics*, 84, 682-690.

- [26] Holmes, T., and Stevens, J., 2004. "Spatial distribution of economic activities in North America," *Handbook of Regional and Urban Economics*, in: J. V. Henderson & J. F. Thisse (ed.), V. 4, Chapter 63, 2797-2843, Elsevier.
- [27] Lucas, R. E., 1987. *Models of Business Cycles*, Basil Blackwell.
- [28] Lustig, H. and Van Nieuwerburgh, S., 2010. "How Much Does Household Collateral Constrain Regional Risk Sharing?", *Review of Economic Dynamics*, 13, 265-294.
- [29] McGrattan, E. and Prescott, E., 2009. "Unmeasured Investment and the Puzzling U.S. Boom in the 1990s," Research Department Staff Report 369, Federal Reserve Bank of Minneapolis.
- [30] Rappaport, J., 2008. "Consumption amenities and city population density," *Regional Science and Urban Economics*, 38, 533-552.
- [31] Rappaport, J., 2009. "The increasing importance of quality of life," *Journal of Economic Geography*, 9, 779-804.
- [32] Rappaport, J. and Sachs, J. D., 2003. "The United States as a Coastal Nation," *Journal of Economic Growth*, 8, 5-46.
- [33] Rossi-Hansberg, Esteban and Mark Wright, 2007. "Urban Structure and Growth," *Review of Economic Studies*, 74, 597-624.
- [34] Storesletten, K., Telmer, C. and Yaron, A., 2001. "The Welfare Cost of Business Cycles Revisited: Finite Lives and Cyclical Variation in Idiosyncratic Risk," *European Economic Review*, 45, 1311-1339. *Economic Review*, 45, 1311-1339.

APPENDIX A: THEORY-BASED COUNTERFACTUAL POPULATION CHANGES

Cities with Largest Percentage Change in City Population Without Amenity Differences



Cities with Smallest Percentage Change in City Population Without Amenity Differences

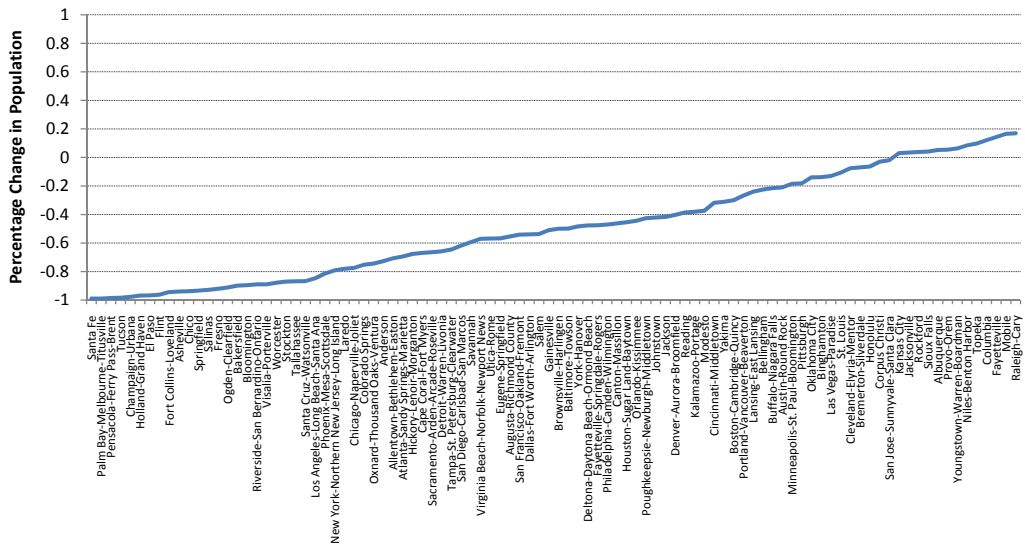
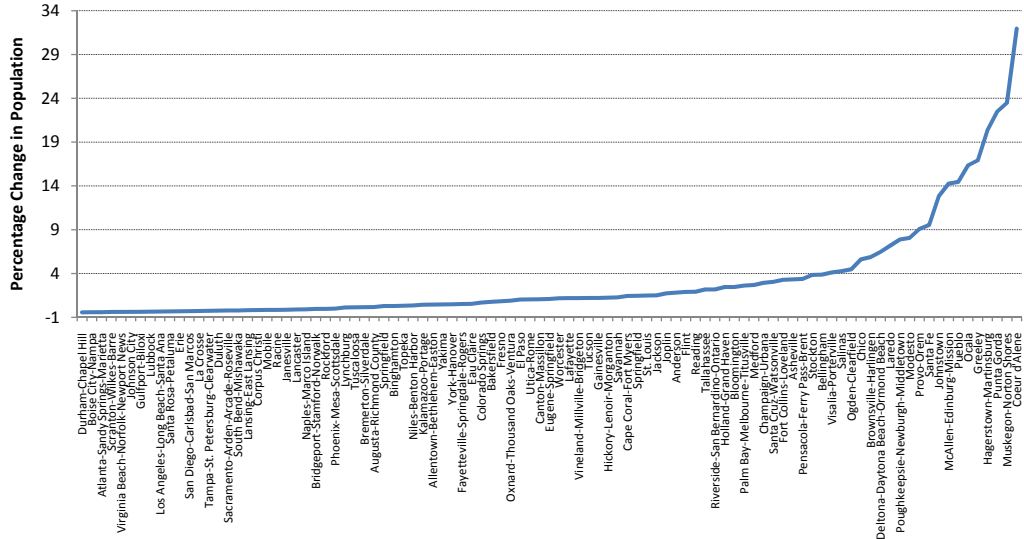


Figure A1: Changes in Population Sizes with Average Amenities

Cities with Largest Percentage Change in City Population Without Efficiency Differences



Cities with Smallest Percentage Change in City Population Without Efficiency Differences

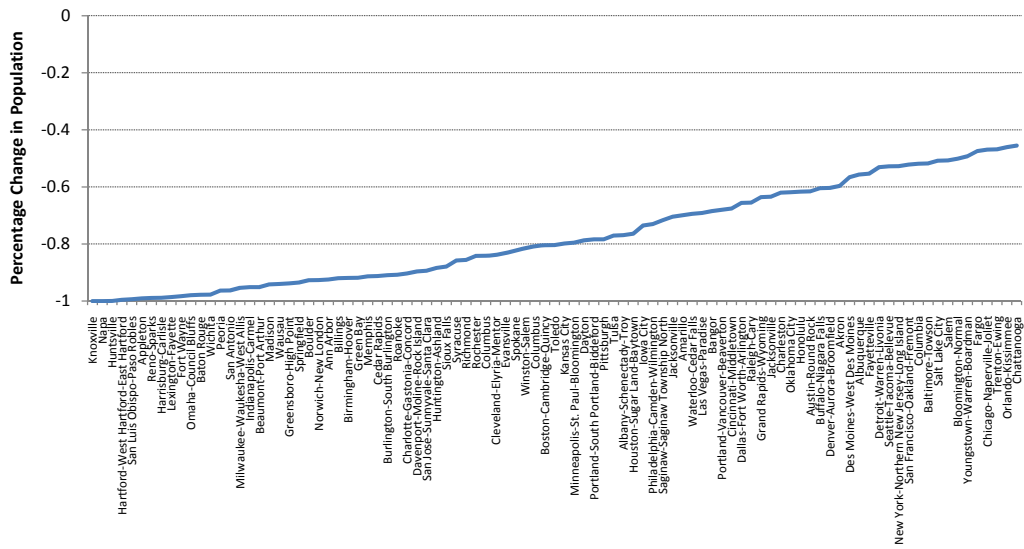
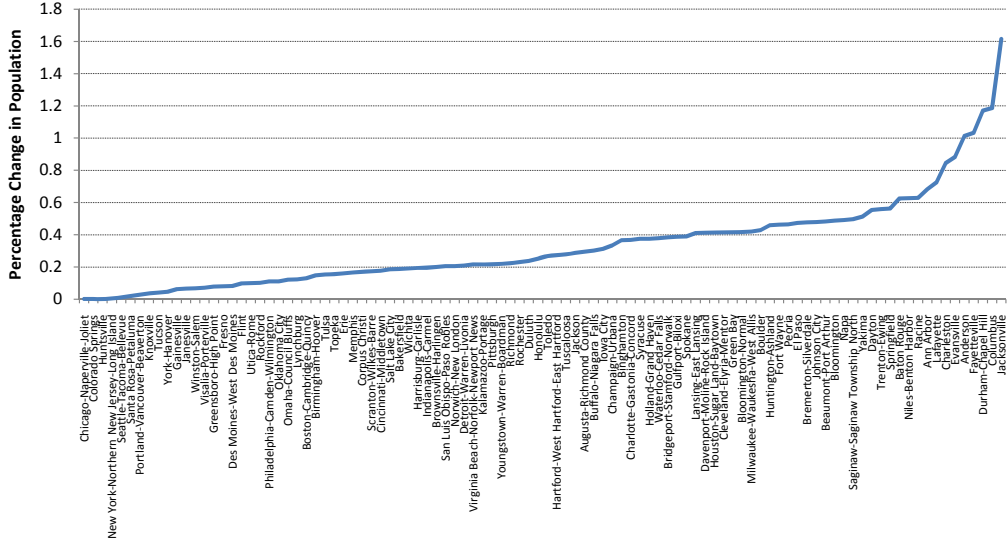
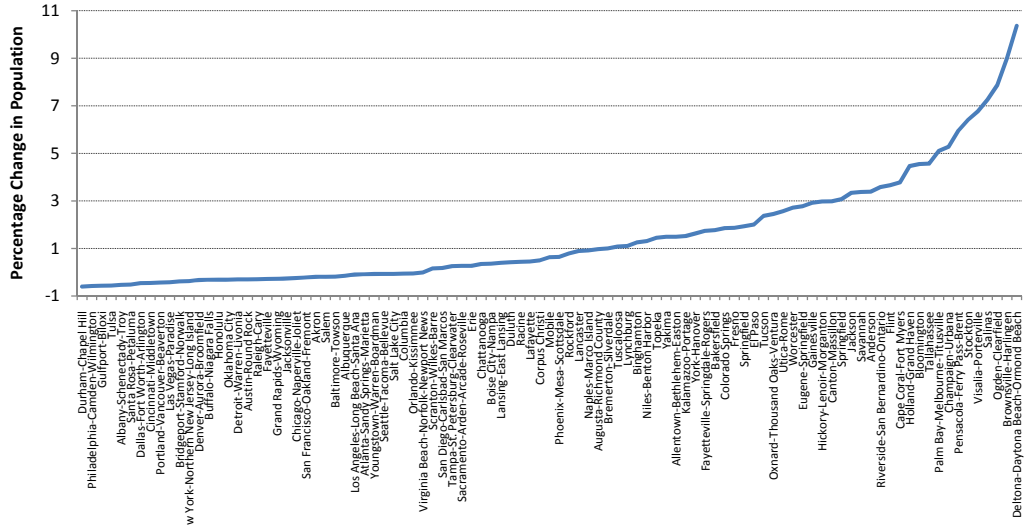


Figure A2: Changes in Population Sizes with Average Efficiency

Cities with Positive Percentage Change in City Population Without Excessive Friction Differences



Cities with Largest Percentage Change in City Population Without Efficiency Differences



Cities with Smallest Percentage Change in City Population Without Efficiency Differences

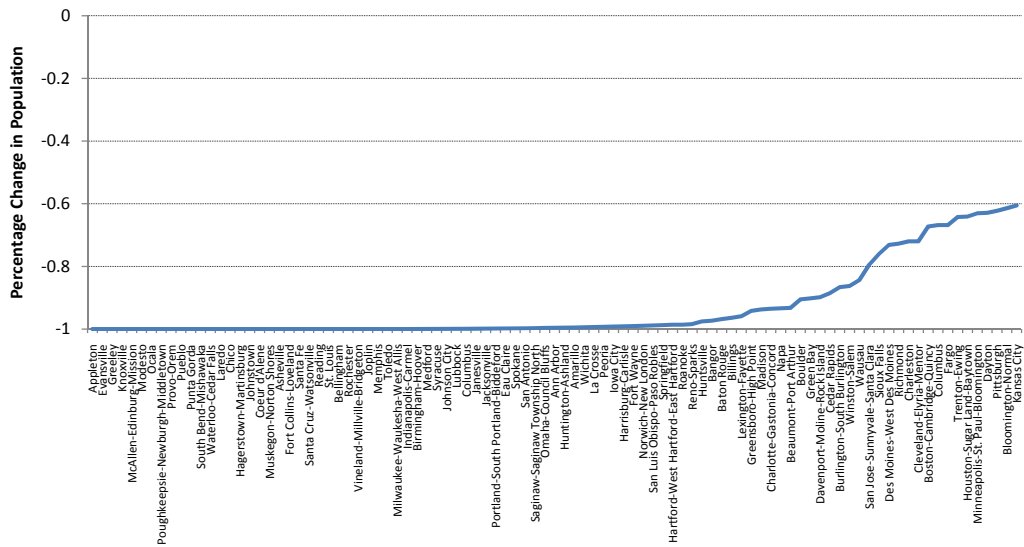
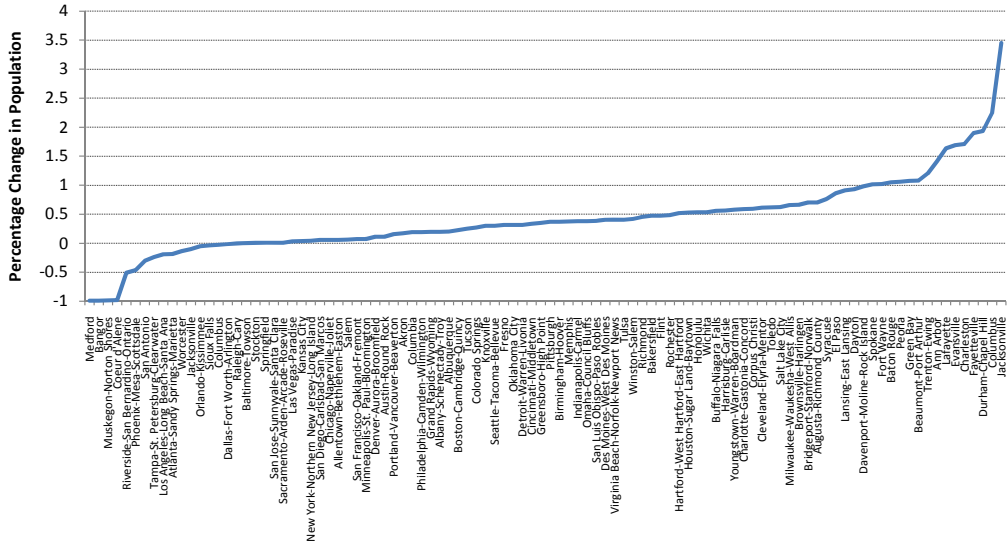


Figure A5: Changes in Population Sizes with Average Efficiency (with Externalities)

Cities with Largest Percentage Change in City Population Without Excessive Friction Differences



Cities with Smallest Percentage Change in City Population Without Excessive Friction Differences

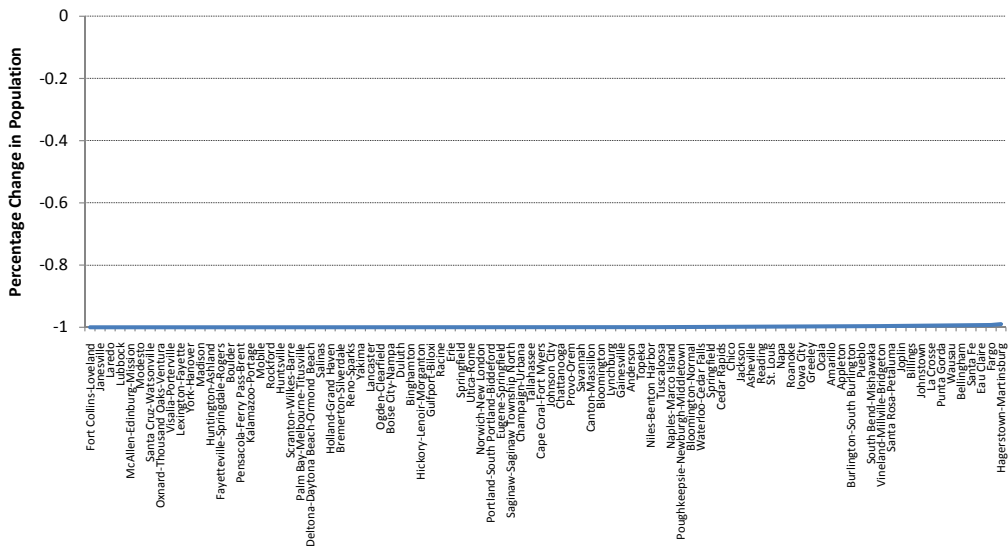
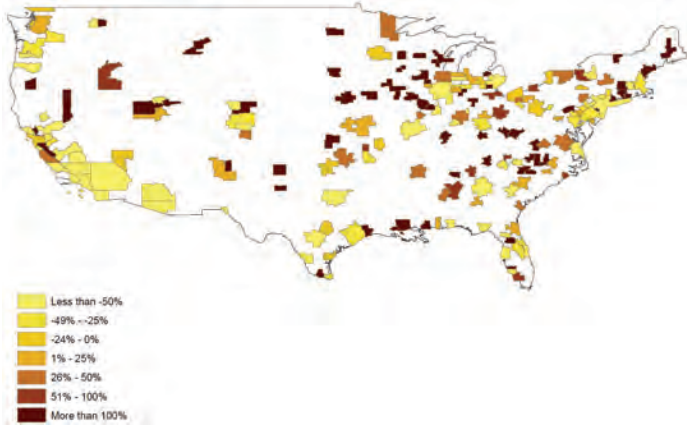
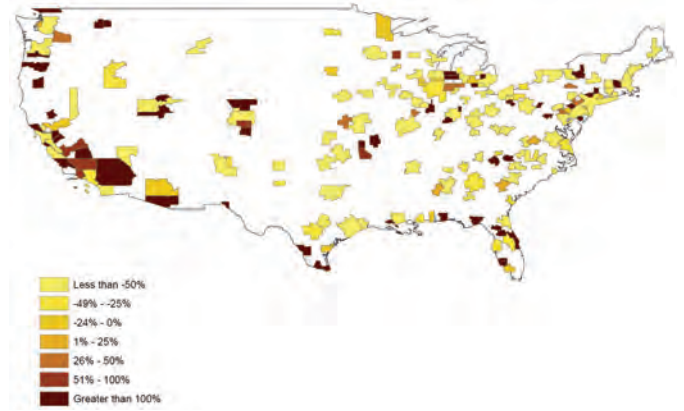


Figure A6: Changes in Population Sizes with Average Excessive Frictions (with Externalities)

Without Differences in Amenities:



Without Differences in Efficiency:



Without Differences in Excessive Frictions:

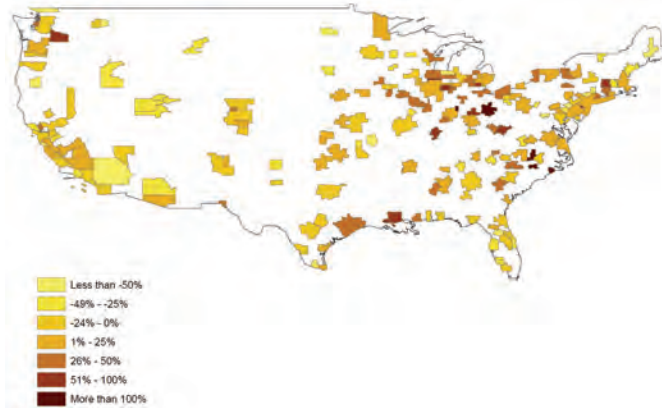
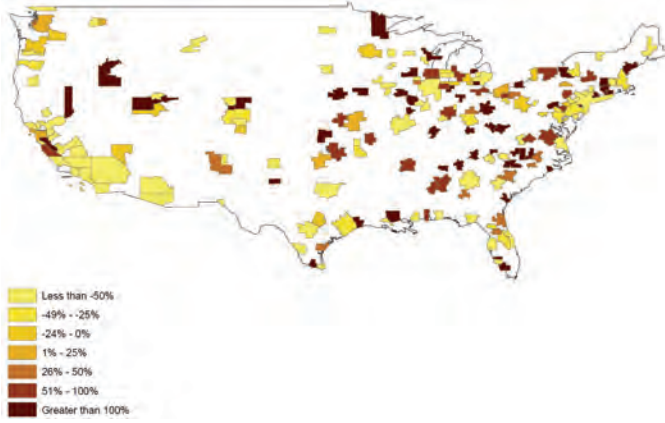
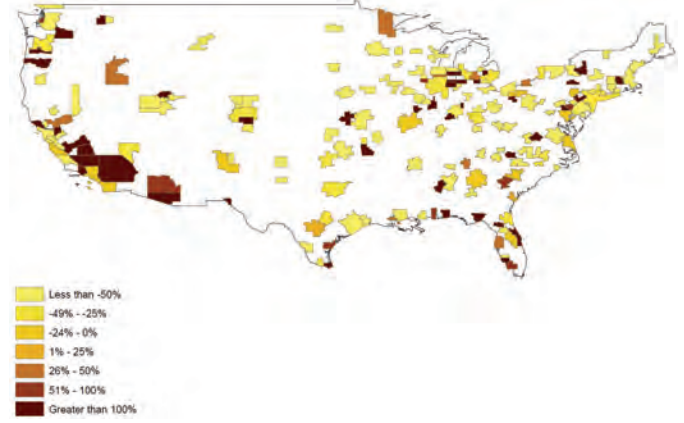


Figure A7: Maps of Changes in Population Sizes

Without Differences in Amenities:



Without Differences in Efficiency:



Without Differences in Excessive Frictions:

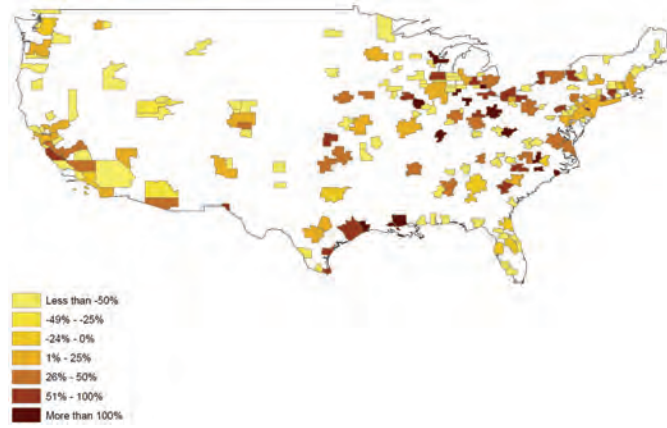


Figure A8: Maps of Changes in Population Sizes with Externalities, $\omega = 0.02$

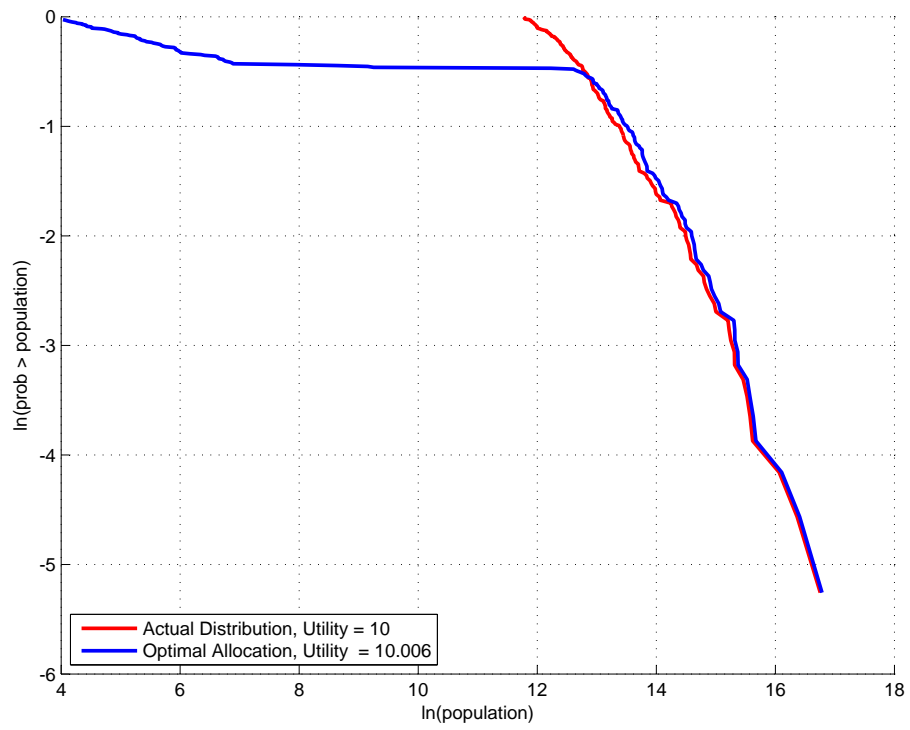


Figure A9: Optimal Allocation with Externalities, $\omega = 0.02$

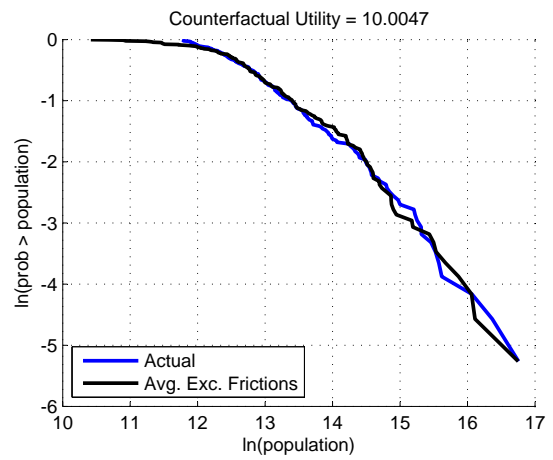
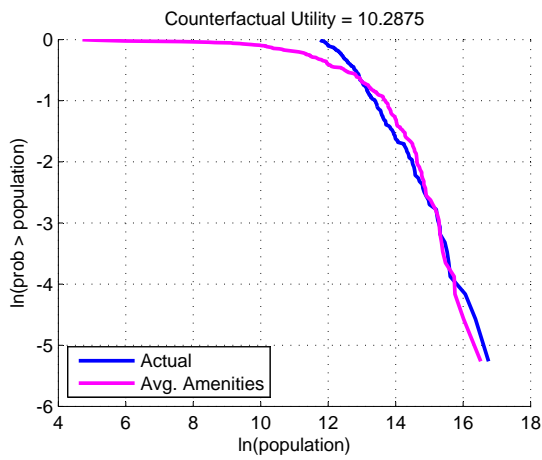
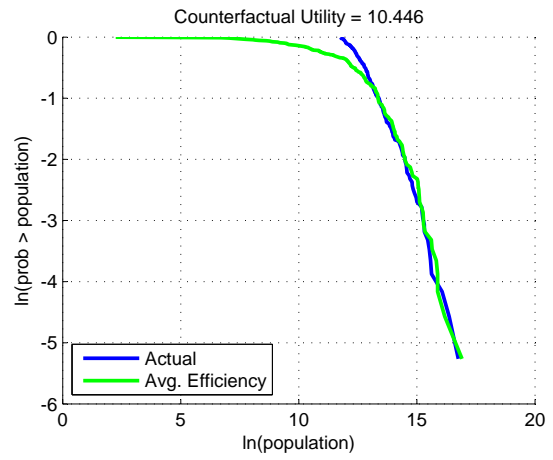
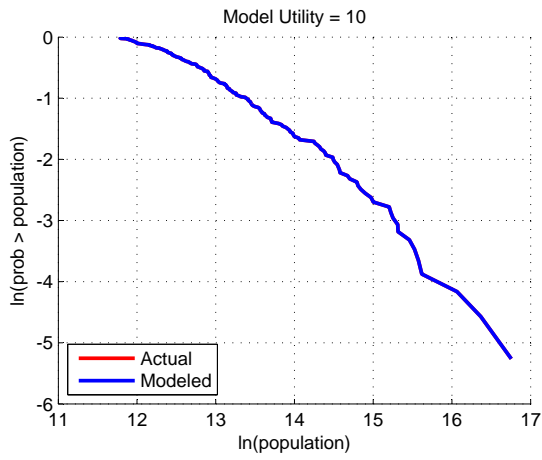


Figure A10: Counter-factuals with Theory-Based Capital Measures

APPENDIX B: DATA APPENDIX

This section provides a detailed description of the metropolitan data we use.

Unit of observation. The unit of observation is the metropolitan statistical area (MSA). A metropolitan statistical area is a collection of counties with at least one urbanized area of 50,000 or more inhabitants. We use data from 2005 to 2008. Going further back in time is complex, as the definition of MSAs changed in 2003, and there was a subsequent lag in the adoption of the new definitions. More recent data (for 2009) are not available yet for some of the relevant variables.

Production. Measured by Gross Domestic Product by Metropolitan Area. Source: Bureau of Economic Analysis, Regional Economic Accounts.

Private consumption. The measure of consumption used to compute the labor wedge is private consumption. There are no ready-to-use data on private consumption at the metropolitan area level. We start by decomposing private consumption into private consumption net of housing services and private consumption of housing services.

We proxy private consumption net of housing services by retail earnings. In particular, we use retail earnings (at the MSA level) multiplied by private consumption net of housing (at the U.S. level) divided by retail earnings (at the U.S. level). Data on retail earnings are defined as personal earnings from retail trade and come from the Bureau of Economic Analysis Regional Economic Accounts, Table CA05. Data on private consumption net of housing comes from the Bureau of Economic Analysis National Income and Product Accounts (NIPA), Table 2.3.5. In using this proxy, we follow the literature on interregional risk sharing which has used retail sales as a proxy for private consumption (Asdrubali et al., 1996, Hess and Shin, 1997, Lustig and Van Nieuwerburgh, 2009). Note that those papers use retail sales, rather than retail earnings, by using data from the Survey of Buying Power published by Sales & Marketing Management. However, that survey got interrupted during the period 2006-2008. Both proxies are very similar though. For 2005 and 2008 the correlation between retail earnings and retail sales at the MSA level is about 0.80. In addition, the correlation between private consumption and retail earnings at the U.S. level for the years 2001-2008 is 0.99.

We proxy private consumption of housing services by taking the sum of aggregate gross rent of renter-occupied housing and the rental value of the aggregate value of owner-occupied housing. Both

variables are available at the MSA level, and come from the annual American Community Survey run by the U.S. Census Bureau. To compute the rental value of owner-occupied housing, we assume that the aggregate value of owner-occupied housing is the discounted sum of future rental flows, taking into account depreciation. The rental value of owner-occupied housing is then computed as the aggregate value of owner-occupied housing multiplied by $r + \delta$ and divided by $1 + r$. For the benchmark calculation, we take $r = \delta = 0.02$, as in the rest of the paper (see Table 1).

Capital stock. Again, there are no ready-to-use capital stock data at the MSA level. We start by decomposing the capital stock into non-residential and residential capital.

We proxy non-residential capital at the MSA level by using sectoral non-residential capital stock data at the U.S. level and allocating it to the different MSAs in function of their sectoral weights. This is similar to the approach taken by Garofalo and Yamarik (2004) when estimating state capital stocks. Sectoral non-residential capital stock data come from the National Economic Accounts from the BEA. We take the sum of private and public capital. For private non-residential capital we take current-cost net stock of private fixed assets by industry (Table 3.1ES); for public non-residential capital we take current-cost net stock of government fixed assets at both the federal level and the state & local level (Table 7.1B). We then allocate the sectoral capital stock to the MSAs in function of their shares of sectoral earnings. In particular, capital stock in sector s in MSA i is computed as the capital stock in sector s in the U.S. multiplied by earnings in sector s in MSA i divided by earnings in sector s in the U.S. Data on sectoral earnings both at the MSA and the U.S. levels come from the Regional Economic Accounts of the Bureau of Economic Analysis (Table CA05N).

Non-residential capital at the MSA level is easier to come by. We take the sum of aggregate value of renter-occupied and of owner-occupied housing. In the case of owner-occupied housing, that information is available from the American Community Survey of the U.S. Census Bureau. In the case of renter-occupied housing, the same data source gives information on the aggregate gross rent. This allows us to compute the value of rental-occupied housing as the aggregate gross rent multiplied by $1 + r$ divided by $r + \delta$. This assumes, as before, that the value of housing is equal to the discounted sum of future rental streams, where future rental streams are the same as today's rental stream corrected for depreciation.

Hours worked. To compute average hours worked we take the total hours worked divided by the population 16 years old and above. We use data from the Current Population Survey (CPS) and

compute total hours worked at the MSA level by summing up the total hours worked by individuals (weighted by their representativeness in the sample) and then dividing them by all individuals aged 16 and above in the sample (weighted by their representativeness in the sample). To limit errors due to small sample problems, we leave out MSAs that have information on less than 50 individuals. The share of time worked is then equal to the average hours worked per day divided by 14.

Housing rental prices. As measure for housing rental prices, we take the median gross rent of rental-occupied housing. Data at the MSA level are available from the American Community Survey from the U.S. Census Bureau.