Pollution and city size: can cities be too small?*

Rainald Borck†
University of Potsdam, CESifo and DIW Berlin

Takatoshi Tabuchi‡
University of Tokyo and RIETI

This version: May 10, 2017

Abstract

We study optimal and equilibrium size of cities in a city system model with pollution. Pollution is a function of population size. With symmetric cities, if pollution is local or per capita pollution increases with population, equilibrium cities are too large. When pollution is global and per capita pollution declines with city size, cities may be too small. With asymmetric cities, the largest cities are too large when pollution is local or per capita pollution increases with population; when pollution is global and per capita pollution decreases with population, the largest cities are too small if the marginal damage of pollution is large enough. We calibrate the model to US cities and find that the largest cities may be undersized by 3-4%.

JEL classification: R12, Q54

Keywords: optimal city size distribution, agglomeration, pollution

*We thank Stefan Bauernschuster, Jan Brueckner, M. Morikawa, T. Morita, Michael Pflüger, and M. Yano as well as participants at MCC Berlin, Tinbergen Institute, University Duisburg-Essen, Free University Berlin, the UEA meeting in Minneapolis, ERSA/UEA meeting in Lisbon, the Verein für Socialpolitik meeting in Münster and its regional economics section in Dresden, the DP seminar at RIETI, and the Osaka spatial economics conference for comments and suggestions. The first author thanks the German Science Foundation (DFG) and the second author thanks RIETI for financial support.

†University of Potsdam, Faculty of Economic and Social Sciences, August-Bebel-Str. 89, 14482 Potsdam, Germany, e-mail: `rainald.borck@uni-potsdam.de`.

‡University of Tokyo, Faculty of Economics, Hongo 7-3-1, Bunkyo-ku, Tokyo, Japan, e-mail: `ttabuchi@e.u-tokyo.ac.jp`

1 Introduction

Urbanization is rapidly increasing, especially in developing countries. According to the UN Population Division, urbanization worldwide will increase from 51.6% in 2010 to 66.4% in 2050, and from 46.1% to 63.4% in the developing world. Some commentators are afraid that this urbanization may have adverse environmental consequences. For instance, Seto *et al.* (2012) argue that the projected urbanization until 2030 leads to significant loss of biodiversity and increased CO_2 emissions due to deforestation and land use changes. Urban economic activities such as manufacturing production, commuting, and residential energy use also contribute to pollution. Fig. 1 shows that over the last half century, urbanisation and CO_2 emissions have moved together. Of course, this may not be a causal relation.

In fact, some writers who claim that large, densely populated cities produce lower per capita emissions. Glaeser and Kahn (2010) show that in the US, inhabitants of large, densely populated cities such as New York City and San Francisco tend to produce lower CO_2 emissions from transport and residential energy use than those living in smaller and less densely populated cities, controlling for factors such as local weather. Glaeser (2011) writes about this *Triumph of the City* and in the subtitle succinctly states: "How our greatest invention makes us richer, smarter, *greener*, healthier, and happier" (our emphasis). This line of reasoning has prompted organizations such as the OECD and the World Bank to advocate high density urban development to mitigate environmental pollution.

Therefore, an important policy question is whether big cities are good or bad for the environment, especially in developing countries such as China, where new cities are springing up by the minute. While on the one hand, migrants flock to cities to take advantage of their economic opportunities, on the other hand, concern about congestion, environmental pollution and other side effects is mounting. So what is the optimal size of cities that are affected by environmental pollution? And what would be the unregulated equilibrium city size?

In this paper, we build a simple model of a city system to study how the equilibrium and optimal city size distributions are affected by environmental pollution. We use a standard monocentric city model, where people work, consume goods and housing in cities. Agglomeration externalities make workers more productive in big cities. Pollution is related to city size since it is a by-product of urban production, commuting and housing. In line with reality, we assume that externalities arising from pollution are not internalized. We distinguish between pollution which is purely local, such as certain kinds of emissions from traffic, and pollution which spills over between cities, such as greenhouse gas (GHG) emissions. When cities are symmetric, we find that with local pollution, equilibrium cities are too large and there are too few of them, mirroring the classic result of Henderson (1974). By contrast, when pollution is global, we find that equilibrium cities may be either too small or too big. The former case can occur when per capita pollution falls with city size. We also study the model with a given number of asymmetric cities. With local pollution, we find that the largest cities are too large and the smallest cities too small. With global pollution, if per capita pollution decreases with city size and the marginal damage of pollution is large enough, the largest cities are too small and the smallest too large.

We also quantify the extent to which cities may be undersized, using a calibrated version of the model. We use some standard parameter values from the literature, and, using data from Fragkias et al. (2013), we estimate the effect of the size of US metropolitan areas on CO_2 emissions. We find that doubling city size reduces per capita CO_2 emissions by 13 percent. In the symmetric city case, we find that cities might be undersized by up to 9 percent if pollution is global. With asymmetric cities, in the case of global pollution, the largest cities may be undersized by 3-4 percent while the smallest cities are oversized by 5-10 percent. If pollution is local, the largest cities are oversized, but by only about 0.3%. Finally, we use an estimate of the degree of pollution spillovers (so pollution is neither completely global nor completely local) and find that the largest city is undersized by 2.4% and the smallest is oversized by 6.6%.

Our paper is related to two strands of literature. First, the literature on city systems has studied equilibrium and optimal city sizes. Henderson (1974) first showed that in equilibrium, cities are too big. This finding also comes out of the models by Tolley (1974), Arnott (1979), and Abdel-Rahman (1988). Tolley (1974) considers local pollution and actually argues that it leads to cities being too big. We show that this argument depends on whether pollution spills over to other cities and how per capita emissions change with city size. Abdel-Rahman and Anas (2004) review this literature and also discuss the role of externalities in city system models (though not externalities arising from pollution).

On the other hand, some recent papers show that cities may be too small in equilibrium. Albouy et al. (2016) show that large cities may be too small due to the interaction of federal taxation and wedges due to land ownership. Eeckhout and Guner (2017) also show that spatially uniform taxation may lead to large cities being undersized, and that the optimum spatial tax system taxes individuals in large cities less than the current US tax system.[1]

[1] Au and Henderson (2006) show that many Chinese are too small due to the migration restrictions of

Like Albouy et al. (2016) and Eeckhout and Guner (2017), we show that cities may be too small. However, the mechanism in our paper, namely negative externalities from intercity pollution, is different.

Second, there is a small but growing literature on cities and the environment more general. Related to this paper, Gaigné *et al.* (2012) and Borck and Pflüger (2015) study the interaction of agglomeration, pollution and welfare in models with a given number (two) of cities. There are also some theoretical papers on urban structure and pollution, see Borck (2016), Borck and Brueckner (2016), Dascher (2014), Larson *et al.* (2012) and Tscharaktschiew and Hirte (2010). Finally, Glaeser and Kahn (2010) and Larson and Yezer (2015) study empirically the relation between GHG emissions or energy use and city structure. Glaeser and Kahn (2010) find that large, dense cities in the US produce fewer GHG emissions. Morikawa (2013) finds that dense cities in Japan produce lower per capita energy consumption in the service sector, and Blaudin de Thé and Lafourcade (2016) show that residents of low density suburban households use more gasoline for driving. Larson and Yezer (2015) study the effect of city size on energy use in a simulation model, finding that per capita energy use does not change with city size. A number of papers from other disciplines than economics also study the relation between city size and pollution empirically, with different results, e.g. Fragkias et al. (2013) and Sarzynski (2012). Our paper is also concerned with the relation between pollution and city size, which is essential for the comparison of equilibrium and optimal city systems.

We proceed as follows. The next section introduces the model of a symmetric city system. Section 3 presents the modeling of pollution. In section 4, we study the equilibrium and optimum size of cities with local and global pollution. Section 6 contains a numerical simulation, to get a sense of the possible divergence of optimum and equilibrium city size. In Section 5, we extend both the analytical and simulation results to the realistic case of asymmetric cities. The last section concludes.

2 The model with symmetric cities

There are m cities in the economy, whose total population is exogenous and denoted by N.[2] For now, we assume cities to be identical. The population size in each city is endogenous and given by $n = N/m$. For simplicity, the city space is linear with unit width and the

the *hukou* system.

[2] In contrast to Albouy et al. (2016), we don't consider a rural sector in the economy.

central business district (CBD) is a spaceless point located at $x = 0$, while the endogenous city border is denoted $\bar{x}$ (we focus on the right side of the city for simplicity). All individuals commute to the CBD and have identical preference given by

$$u(s, z, E) = s^{\alpha} z^{1-\alpha} E^{-\beta}, \tag{1}$$

and the budget constraint is

$$w = z + rs + tx, \tag{2}$$

where s is housing floor space (equivalently land consumption), z is consumption of a composite non-housing good, E is pollution, w is wage income, r is the housing rent per square meter, t is the commuting cost per mile, x is distance from the CBD, and $0 < \alpha < 1$, and $\beta > 0$.

Consumers choose s and z to maximize (1) subject to (2). From this we get optimal housing consumption

$$s(w - tx, r) = \frac{\alpha(w - tx)}{r}. \tag{3}$$

Consumers are mobile within and between cities, and land is rented to the highest bidder. We can now solve for households' bid rent, i.e., the maximum amount the household would be willing to pay per unit of land. Using (3) and (2) in (1) and solving $u(z, s, E) = \overline{u}$ gives

$$r(w - tx, E, v) = (w - tx)^{1/\alpha} E^{-\beta/\alpha} v^{-1/\alpha}, \tag{4}$$

where $v \equiv \alpha^{-\alpha} (1 - \alpha)^{-(1-\alpha)} \overline{u}$.

The two equilibrium conditions in the representative city are:

$$r(w - t\bar{x}, E, v) = r_A \tag{5}$$

$$\int_0^{\bar{x}} \frac{1}{s(w - tx, E, v)} dx = n, \tag{6}$$

where r_A is the agricultural land rent. Eq. (5) states that at the city border, land rent just equals the agricultural land rent. Eq. (6) says that the population n fits into the city between 0 and $\bar{x}$.

Suppose that there are external economies of scale at the city level, for instance because of gains from individual specialization. Total city production is assumed to be $Y = n^{1+\gamma}$, with $0 < \gamma < \alpha$ and the individual wage is $w = n^{\gamma}$.[3] Substituting (3) and (4) into (5) and

[3]Duranton and Puga (2004) show that several different mechanisms lead to the same functional form,

(6) and solving gives the city border and indirect utility

$$\bar{x} = \frac{n^\gamma \left[1 - r_A^\alpha (r_A + tn)^{-\alpha}\right]}{t} \tag{7}$$

$$v = n^\gamma (r_A + tn)^{-\alpha} E^{-\beta}. \tag{8}$$

Eq. (7) shows that the city expands as population grows. Note that because of the separability of utility, $\bar{x}$ is not affected by pollution. Eq. (8) shows the standard tradeoff induced by an increasing city population: on the one hand, utility increases with n due to agglomeration forces, on the other hand, it decreases because of longer commutes and competition for land, which results in higher land rents. In the next section, we model pollution in order to study how it affects this fundamental tradeoff. In particular, the pertinent question is how reallocating population among cities affects the disutility from pollution.

Note that we have assumed that land is owned by absentee landowners. As is well known, efficiency analysis requires returning differential land rents to city residents. We show in Appendix C, however, that our results hold qualitatively if land is owned by city residents.[4]

3 Pollution

Pollution in city i is given by

$$E_i = e(n_i) + \delta \sum_{j=1}^{m-1} e(n_j),$$

where $e(n_i)$ are local emissions and $0 \leq \delta \leq 1$ measures the degree of pollution spillovers. When $\delta = 0$, pollution is purely local (for instance, some forms of particulate pollution which do not diffuse over long distances). Conversely, when $\delta = 1$, pollution is purely global from the view of our city system, as is the case, for instance, for GHG emissions. Importantly, in the latter case, the environmental externality is independent of the individual's location.

An important issue in the coming analysis will be the relationship between emissions

such as gains from specialization, matching, sharing intermediate inputs, or learning.

[4] Albouy et al. (2016) study cross-city externalities arising from landownership requirements on migrants.

and city population, as captured by the function $e(n_i)$. We assume that city population affects local emissions through residents' economic activities, such as commuting, housing, and consumption of other goods whose production causes emissions. What do we know about this relation? In Section 6, we will try to estimate the population elasticity of pollution empirically, but here we briefly discuss theoretical an empirical studies that address this issue.

Borck and Pflüger (2015) present a theoretical model in which urban pollution is driven by commuting, residential energy use, industrial and agricultural production, and goods transport. They show that per capita pollution from industrial production and residential energy use decreases with city size, while pollution from commuting and goods transport increases. The total effect of city population on urban pollution is ambiguous and depends on parameters.

Some authors have estimated the relation between pollution and city population (or population density) empirically. Most of these papers estimate an equation of the form $e = Bn^{\theta}$, which we will also do in Section 6.[5] Lamsal *et al.* (2013) use cross-sectional cross-country data on NO_2 and NO_X pollution and find that the elasticity of pollution with respect to population density lies between 0.4 and 0.67. Gudipudi *et al.* (2016) study the effect of population density on CO_2 emissions and find an elasticity around 0.6, so doubling population density would reduce per-capita emissions by 24%. Fragkias et al. (2013) also estimate the effect of population on CO_2 emissions, using panel data from US cities. They find an elasticity of emissions with respect to population of 0.93. Rybski *et al.* (2016) conduct a meta-analysis of published articles that study CO_2 emissions and city size, and find that in developed countries per capita emissions decrease with city size while in developing countries per capita emissions increase with city population. However, most of these estimates seem problematic. For instance, Lamsal *et al.* (2013) use cross-sectional OLS regressions to estimate the population elasticity of pollution. But this ignores potential confounders that are correlated with population density and pollution. Fragkias et al. (2013) use panel data, but they estimate the model with random effects, which assumes that any unobserved time-invariant heterogeneity between cities is not correlated with pollution. In Section 6, we present an alternative estimate of the population elasticity of CO_2 pollution, using the same dataset as Fragkias et al. (2013).

[5] Some papers not reviewed here estimate other functional forms, where, for instance, pollution is assumed to be a quadratic function of population.

4 Equilibrium and optimum number and size of cities

The equilibrium city size in the city system is defined by the solution of $v_i = v^*$ for all i. We focus on symmetric cities. Further, we require the equilibrium to be stable, which implies $\partial v(n)/\partial n < 0$. The optimal city size is found by maximizing $mnv(n)$ with respect to n and m. Using $mn = N$, this is equivalent to maximizing $v(n)$ with respect to n. Note that, from (8) follows $v(0) = 0$ so no one would ever want not to live in a city.

4.1 Local pollution

Suppose first that pollution is entirely local, i.e. $\delta = 0$. Then migration is governed by the following utility differential

$$v(n_i) - v(n_j) = \hat{v}(n_i)e(n_i)^{-\beta} - \hat{v}(n_j)e(n_j)^{-\beta}, \qquad (9)$$
$$\text{where } \hat{v}(n) \equiv n^{\gamma}(r_A + tn)^{-\alpha},$$

and optimum city size maximizes $v(n_i) = \hat{v}(n_i)e(n_i)^{-\beta}$. We will assume that both $v(n)$ and $\hat{v}(n)$ are quasi-concave, which holds (in the neighborhood of the equilibrium and social optimum) for the parameter values used in our numerical simulations. Moreover, we assume that locally produced pollution $e(n_i)$ satisfies $e(0) = 0$ and $de/dn_i > 0$.

Since $v(n_i)$ can be shown to be inverted U-shaped, we get the standard result that equilibrium cities are too large, as in Henderson (1974). This can be seen by looking at Fig. 2. The figure shows the optimal city size $\hat{n}$ and two potential equilibrium city sizes $\tilde{n}$ and n^e.[6] Any equilibrium with city size $\tilde{n} < \hat{n}$ is unstable: if the city population were to deviate slightly from $\tilde{n}$, migration in or out of the city would occur, as indicated by the arrows. Conversely, any equilibrium with $n^e > \hat{n}$ is stable: as indicated by the arrows, a deviation from n^e would induce migration flows which restore the equilibrium. Therefore, there is a continuum of equilibria with $n^e > \hat{n}$ where $\hat{n} = n^*$ maximizes $v(n_i)$. We summarize this as:

Proposition 1 *If pollution is purely local, $\delta = 0$, cities are too large in equilibrium.*

[6] Note that there is a continuum of equilibria, so all that can be said in general is that $n^e > \hat{n}$, but the exact location of the equilibrium is indeterminate.

4.2 Global pollution

Now, let $\delta = 1$ so that pollution is global from the viewpoint of the economy. Since pollution is global, we can drop the index i from pollution E_i and write the utility difference of living in city i versus j as

$$v(n_i) - v(n_j) = E^{-\beta} \left(\hat{v}(n_i) - \hat{v}(n_j) \right). \tag{10}$$

For $E > 0$, the individual migration decision is determined by the difference $\hat{v}(n_i) - \hat{v}(n_j)$, so global pollution does not affect migration decisions. Let $\hat{n}$ denote the city size which solves $\max_n \hat{v}(n)$. Setting $\hat{v}'(n) = 0$ and solving gives

$$\hat{n} = \frac{\gamma r_A}{(\alpha - \gamma)t}. \tag{11}$$

Then, by the same argument as in Henderson (1974), there is a continuum of stable equilibria with city sizes $n^e > \hat{n}$. Fig. 3 shows possible equilibrium city sizes. As before, any equilibrium with $n^e > \hat{n}$ is stable.

The optimum city size n^* is found by maximizing $v(n) = \hat{v}(n)E(n)^{-\beta}$. The first order condition can be written

$$\hat{v}'(n) - \beta \hat{v}(n) \frac{E'(n)}{E(n)} = 0. \tag{12}$$

We know that $n^e \geq \hat{n}$ and that $\hat{n}$ maximizes $\hat{v}(n)$. Since $\beta > 0, E(n) > 0$ and $\hat{v}(n)$ is quasi-concave, evaluating (12) at $\hat{n}$ implies that $n^* < \hat{n}$ if $E'(\hat{n}) > 0$. Since $E(n) = m \cdot e(n) = \frac{N}{n} e(n)$, we find cities are definitely too large if per capita pollution is increasing in city size. Intuitively, in this case making cities larger increases pollution, which increases the disutility from pollution. This reinforces the argument in Henderson-style models which make cities too large.

However, if per capita emissions are decreasing in city size, we find $n^* > \hat{n}$. This opens up the possibility that in equilibrium, cities may be too small. However, since there is a continuum of equilibria with $n^e > \hat{n}$, cities may also be too large. Summarizing this discussion, we have:

Proposition 2 *Suppose that pollution is global, i.e. $\delta = 1$. If per capita emissions increase with n, cities are too large in equilibrium. However, if per capita emissions decrease with n, cities may be either too small or too large in equilibrium.*

Fig. 3 illustrates the case where pollution is global and per capita emissions are decreasing with city size. The blue curve depicts the function $\hat{v}(n)$ and the equilibrium city size is some $n^e > \hat{n}$. The orange curve shows the curve $v(n)$ and the optimum city size is n^*.[7] The thick red part of the $v(n)$ curve shows the part where the possible equilibrium city size (with $n^e > \hat{n}$) is smaller than the optimum size. However, the equilibrium city size may also be larger than n^*.

As Prop. 2 makes clear, in the case of global emissions whether cities are over- or undersized depends on how per-capita emissions change with city population. However, as already stated in Section 3, not much is known about this relationship. Therefore, we estimate this relationship in Section , where we use numerical simulation to gauge whether cities will be over- or undersized in equilibrium.

5 Asymmetric cities

5.1 Equilibrium and social optimum with asymmetric city sizes

We now introduce asymmetric cities into the model. To do so, we assume that an individual living in city i obtains utility

$$v_i(n_i) = A_i n_i^{\gamma} (r_A + t n_i)^{-\alpha} E_i^{-\beta}. \tag{13}$$

The variable A_i is a city level amenity, which could be a consumption amenity such as good weather or a production amenity such as good infrastructure or a favourable geographic location. Without loss of generality, we assume $A_1 = 1$ and $A_i > A_{i+1}$ for $i = 1, 2, \ldots, m-1$.

As before, pollution is given by

$$E_i(\mathbf{n}) = e(n_i) + \delta \sum_{j \neq i} e(n_j),$$

with $e = n^{\theta}$, $\theta > 0$.

We assume the number of cities m is fixed and then ask how the optimum allocation of population among these cities differs from the equilibrium one.[8] Let $\hat{v}$ be the equilibrium utility level that is attained under free migration and let the equilibrium population vector

[7] The functions have been rescaled so that $v(n^*) = \hat{v}(n^*)$ for better visibility.

[8] This differs from, e.g., Albouy et al. (2016) who study a city system with an endogenous number of asymmetric cities. Their model, however, is not taken to data from a real city system as ours. Moreover, looking at a varying number of cities in our context may not be reasonable. When we find that big cities are too small, we might want to close some small cities. Conversely, when big cities are too large, we might want to add more small cities. However, both exercises are not possible to implement in our numerical model, since we do not have data on the universe of all cities, so there is no ‘smallest’ city.

be $\hat{\mathbf{n}} = \{\hat{n_1}, \ldots, \hat{n_m}\}$. The equilibrium city size distribution satisfies $v_i(\hat{\mathbf{n}}) = v_1(\hat{\mathbf{n}}) = v$ for all $i = 1, \ldots, m$. Using (13) and setting $A_1 = 1$, we can then solve for the amenity levels that are compatible with a free migration equilibrium:

$$A_i = \left(\frac{\hat{n}_1}{\hat{n}_i}\right)^{\gamma} \left(\frac{r_A + t\hat{n}_i}{r_A + t\hat{n}_1}\right)^{\alpha} \left(\frac{E_i(\hat{\mathbf{n}})}{E_1(\hat{\mathbf{n}})}\right)^{\beta}. \tag{14}$$

Note that our formulation implies that the amenity levels are uniquely identified by the equilibrium distribution of population up to the normalization that $A_1 = 1$.

We want to compare the equilibrium city size distribution to the optimal distribution. To characterize the latter, we assume the social planner maximizes the sum of utilities

$$\max_{\mathbf{n}} \sum_{i=1}^{m} n_i v_i(\mathbf{n})$$

subject to the population constraint $\sum_{i=1}^{m} n_i = N$. Letting λ be the Lagrangean multiplier on the population constraint, the first order conditions are given by[9]

$$v_i + n_i \frac{\partial v_i}{\partial n_i} + \sum_{j \neq i} n_j \frac{\partial v_j}{\partial n_i} = \lambda, \quad i = 1, \ldots, m. \tag{15}$$

The last term on the LHS of (15) shows the pollution spillovers between cities.

The sign of $v_i - v_{i+1}$ is important in the following analysis. While $v_i = v_{i+1}$ holds at the equilibrium, suppose $v_i > v_{i+1}$ holds for all i at the social optimum. Because $\partial v_i / \partial n_i$ is negative in the neighborhood of a stable equilibrium for almost all i,[10] it must be that the optimum n_i is smaller than the equilibrium n_i in large cities, whereas the optimum n_i is larger than the equilibrium n_i in small cities. The opposite is true when $v_i < v_{i+1}$. Therefore, we can now show the following:

Proposition 3 *Assume that $\theta < 1$. Then (i) if pollution is close to local, the optimal utility is higher in larger cities. Large cities are too large and small cities are too small at the equilibrium; (ii) if pollution is close to global and the marginal damage of pollution is sufficiently large, the optimal utility is lower in larger cities. Large cities are too small and small cities are too large at the equilibrium.*

Proof. See Appendix B. ■

[9] We assume an interior solution where $0 < n_i^* < N$ for all i and that the second order conditions hold.

[10] Tabuchi and Zeng (2004) show that a stable equilibrium requires $\partial v_i / \partial n_i < 0$ for at least $m - 1$ cities.

The intuition is as follows. Suppose that pollution is local, as might be the case, say, for NO_X. Then, the indirect utility v_i is a function of its city size n_i only. As shown by Henderson (1974), the indirect utility is decreasing in n_i at a stable equilibrium.

Start from the equilibrium $v_i = v_{i+1}$ with $n_i > n_{i+1}$ and consider the effect of moving one person from the larger city i to the smaller city $i + 1$. The utility v_i rises to $v_i + \Delta_i$ whereas the utility v_{i+1} falls to $v_{i+1} - \Delta_{i+1}$ because v_i decreases with n_i. The rise Δ_i and fall Δ_{i+1} are similar in magnitude when the one person is sufficiently small relative to total city size. Since there are more people in city i, however, the sum of $n_i\Delta_i$ exceeds the sum of $n_{i+1}\Delta_{i+1}$. Therefore, it is optimal to reduce the size of larger cities and raise that of smaller cities. As a result, the utility levels in larger cities are higher than those in smaller cities at the optimum.

By contrast, if pollution is global, such as in the case of CO_2, concentrating population in bigger cities decreases total emissions if $\theta < 1$, which benefits residents in all cities. When moving one person from a smaller city $i + 1$ to a larger city i, utility of city i residents falls while that of $i + 1$ residents rises. However, due to the global externality, utility of the residents of all other cities also rises. Therefore, as long as the marginal damage of pollution is large enough, social welfare rises.

Examining (15) in Appendix B, we can further say the following. Given sufficiently large δ (i.e., close to global pollution), large cities are more likely to be too small if the housing expenditure share α, the agricultural land rent r_A, and the commuting cost t are small. In this case, the crowding effects induced by commuting and tight housing markets in larger cities are outweighed by the beneficial effect of reduced pollution for all other cities.

In order to correct the discrepancy between the equilibrium and optimal distributions of city sizes, the national government may impose location taxes and subsidies according to city size. In the case of global pollution with $\theta < 1$, in our setup, living in large cities should be subsidized to make them more attractive.[11]

[11] See also Eeckhout and Guner (2017) for an analysis of taxes related to city size.

6 Numerical simulation

6.1 Symmetric cities

We now try to assess to what extent optimum and equilibrium city size may diverge, using numerical simulation. We use the following parameter values. We set the expenditure share of housing to $\alpha = 0.24$ following Davis and Ortalo-Magné (2011), and the agglomeration elasticity to $\gamma = 0.05$ (see Combes and Gobillon, 2015, for an overview). From Borck and Brueckner (2016), we set $r_A = \$58,800$, the annual land rent of agricultural land in the US, and $t = \$503$, the annual (monetary plus time) commuting cost per mile in the US.

Since much less is known about the emissions intensity θ (see our discussion in Section 3), we estimate this parameter using US city data. Suppose that total emissions in city i in year t are $e_{it} = Bn_{it}^{\theta}$. Then, per capita emissions decrease with population size if and only if $\theta < 1$.

We can then estimate a linear regression of the form

$$\log e_{it} = c + \theta \log n_{it} + \varepsilon_{it} \tag{16}$$

where $c \equiv \log B$ is a constant, and ε is the error term.

We use data from Fragkias et al. (2013) to estimate CO_2-emissions in US core based statistical areas (metropolitan statistical areas and micropolitan areas) from 1999-2008. The dataset contains CO_2 emissions and population for 933 core based statistical areas (CBSAs). Emissions are based on data from the Vulcan Project, which quantifies U.S. fossil fuel carbon dioxide emissions at 10 km $\times$ 10 km grid cells and at the scale of individual factories, power plants, roadways and neighborhoods on an hourly basis. These are aggregated by Fragkias et al. (2013) to annual observations by CBSA.

Tab. 1 shows the summary statistics. Fig. 4 displays a binned scatter plot of log emissions against log population, where all data are pooled. Population varies from 12340 in the smallest city to 18.7 mill. in the New York metro area. Per capita emissions vary by a factor of over 200.

We start by estimating (16) by pooled OLS. Results are shown in column (1) of Tab. 2. Standard errors are clustered at the CBSA level. The coefficient on population is 0.938, and it is significantly smaller than one.[12] According to this estimate, if population doubles, per capita emissions would fall by $-(2^{\theta-1} - 1) \times 100 = 4.2\%$.

[12] Using Japanese data for 105 metropolitan employment areas in 2005, we obtain a similar result with a coefficient of 0.902, which is significantly smaller than one.

This estimate may be biased due omitted variables or reverse causality. If pollution were local, then our model would predict that individual migration decisions are based on city emissions, so population would be endogenous and OLS estimation would consequently be biased. Given that CO_2 is a global pollutant, however, this is not a concern in the present setup, since migration should be independent of local emissions. Therefore, reverse causality may not be a big concern in the current setup. However, cities may still differ in unobserved factors that affect population size and emissions. To mitigate any potential bias, we will add various fixed effects to our baseline regression.

First, in column (2), we include time fixed effects to allow for any time varying factors that are common across CBSAs and affect emissions, such as national business cycles. If these cycles were correlated with population size (say because some cities grow more than others when the economy grows) and also affect CO_2 emissions, the OLS coefficient would be biased. The coefficient in column (2), however, is the same as in the model without time effects.

In column (3) we include CBSA fixed effects. Some cities may have disproportionately many power plants that service larger geographic areas. Also, cities may differ in some unobserved dimension such as industry structure, climate, or other factors that may affect population size and emissions at the same time. As long as this heterogeneity is time invariant, we can control for it by estimating a model with CBSA fixed effects. As shown in column (3), the coefficient on population drops to 0.83 once we control for CBSA and time fixed effects. Thus, the result that $\theta < 1$ does not seem to be driven by unobserved heterogeneity among CBSAs.

Finally, to control for potential macroeconomic effects that affect regions differentially and may be correlated with city size and emissions, in column (4), we include interaction effects between year and US census divisions (there are 4 census regions and 9 divisions). As can be seen from the Table, the coefficient on population slightly drops to 0.8, and it remains significantly smaller than one.[13] Hence, doubling population would reduce per capita CO_2 emissions by $-(2^{0.8-1} - 1) \times 100 = 12.8\%$. This is our preferred estimate, since it controls extensively for time-varying regional heterogeneity, and we will use this value of θ for the numerical simulation. However, we will also use the higher value of 0.94 as a robustness check.

We now want to compare (the smallest possible) equilibrium city size, $\hat{n}$ (see (11)),

[13]The p-values for the test for $\theta < 1$ are: 0.0001 in column (1), 0.0001 in column (2), 0.045 in column (3) and .0499 in column (4). In all cases, the hypothesis that $\theta < 1$ cannot be rejected at the 5% significance level.

and optimal city size, n^*, which we now derive. Consider global emissions, $\delta = 1$. Since $e = Bn^{\theta}$, we have $E = m \cdot e = BNn^{\theta-1}$. Substituting in (8) gives

$$v(n) = (BN)^{-\beta} n^{\gamma+(1-\theta)\beta}(r_A + tn)^{-\alpha}. \tag{17}$$

Maximizing with respect to n gives

$$n^* = \frac{[\gamma + (1-\theta)\beta]\, r_A}{[\alpha - \gamma - (1-\theta)\beta]\, t}. \tag{18}$$

The maximum divergence between optimal and equilibrium city size, $n^* - \hat{n}$, is increasing in β, the index of the marginal damage of pollution. In the Appendix, we calibrate β, using central estimates of the social cost of carbon from the literature. Using the central estimate of USD 40.54 per metric ton CO_2 for 2015 (assuming 3% discounting, value updated to 2015 USD) from the recent study by Interagency Working Group on Social Cost of Carbon (2015), we find a value of $\beta = 0.022$.[14] We then get a maximum divergence of $n^*/\hat{n} = 1.1137$. Again, while cities could be oversized in equilibrium, this suggests they could be undersized by up to 11.4 percent. If $\theta = 0.94$, the optimal city size could exceed the equilibrium size by up to 3.5%.

At the high end of estimates of the social cost of carbon, we use the 95th percentile estimate of USD 118 for 2015 from Interagency Working Group on Social Cost of Carbon (2015) (again at 3% discounting, in 2015 USD), which gives a value of $\beta = 0.064$. In this case (using the original $\theta = 0.8$), we find $n^*/\hat{n} = 1.347$, so the optimum city size could exceed the equilibrium size by as much as 35 percent. Finally, estimates for the social cost of carbon increase over time. For 2050, the central and 95th percentile values from Interagency Working Group on Social Cost of Carbon (2015) are USD 77.70 and 238.74. The implied values for β are 0.042 and 0.13. Then, optimal city size could exceed equilibrium city size by up to 22% in the first and 76% in the second case. Furthermore, since climate change damages are estimated to increase over time, our model predicts that if global pollution leads to cities that are too large, the severity of the problem should increase over time.[15]

[14] The study reports averages over three different integrated assessment models. We use their average value across the three models for 2015 at 3% discounting, USD 36 as our central value.

[15] Strictly speaking, by 2050, not only climate damages but also other parameters of our model would have changed, for instance income and other parameters affecting the city system. We assume, however, that any such change is small relative to the expected increase in climate damages.

6.2 Asymmetric cities

We now simulate numerically the equilibrium and optimal city size distribution with asymmetric cities. We will pursue two different simulation exercises. In both of these, we assume a given number of cities, m, and given total population N. So we exclude the formation of new cities. In the first exercise, we assume that the city size distribution follows Zipf's law. As is well known, this is a good approximation for city systems in most countries, except at the very top and bottom of the distribution (Gabaix, 2016). We set the number of cities to $m = 180$. In the second exercise, we use the 180 largest US CBSAs. The total population is the sum of the population sizes of these 180 cities, $n = 225,678,243$ in both cases.

We then compare the equilibrium number of cities to the social optimum. The first exercise will assume that the population distribution among the cities follows Zipf's law, while the second uses a sample of the actual city size distribution among US CBSAs.

City size distribution under Zipf's law. In the first exercise, we assume that the city size distribution follows Zipf's law. As is well known, this is a good approximation for city systems in most countries, except at the very top and bottom of the distribution (Gabaix, 2016). With $n = 225,678,243$, the largest city has 39 mill. inhabitants, the second largest 19.5 mill. and the smallest city has 217,180 inhabitants. We compute the amenity levels from (14) for these given population sizes.

We first assume $\delta = 1$ so pollution is global. Fig. 5 shows the city size distribution using Zipf's law in blue and the optimal distribution in orange. For better visibility, the figure plots the equilibrium and optimal distributions assuming $\beta = 0.1$. While the two distributions with $\beta = 0.022$ are close, the biggest city is undersized by 3.5% while the smallest one is oversized by 7.7%. The welfare gain from moving to the optimal city size distribution is small, less than 0.1% of income.[16] Small welfare gains from optimal policies are also found in similar models e.g. by Eeckhout and Guner (2017) and Albouy et al. (2016).

When pollution is purely local, $\delta = 0$, we find that the divergence between optimal and equilibrium city sizes is small, for given β.[17] The largest city is oversized by about 0.2%

[16] When population is efficiently allocated, total emissions fall by 0.9%, relative to the equilibrium. Note, however, that the welfare gain from efficiently allocating population in the absence of pollution would also be small with our parameters.

[17] We note that this is partly due to the assumption that the value of β is the same for local and global pollution, but local pollution is much smaller than global pollution. Therefore, for a more realistic simu-

and the smallest city is undersized by 1.6%.

The degree of spillovers, δ, is obviously an important factor in determining whether cities are under- or oversized. Therefore, a natural question to ask is, what is the actual degree of spillovers from local emissions? To approach this question, we borrow from Borck and Brueckner (2016) who consider optimal energy taxation with local and global emissions. Their results imply that local emissions make up 60.7% of emissions from commuting and 53% of emissions from residential energy use.[18] Let us take the average of these values, 57%, so we set $\delta = 1 - 0.57 = 0.43$. We then find that the largest city is undersized by 2.5% and the smallest is oversized by 6.9%. In the past two decades local pollution has decreased relative to global pollution and this trend is likely to continue (see, e.g., Amann et al., 2013). Moreover, the damage from global warming is projected to increase over time, since GHGs accumulate in the atmosphere and warming is caused by the stock of pollution. Therefore, we tentatively conclude that using realistic parameters, the case for large cities being undersized and small ones oversized persists and will get stronger over time.

Actual US city size distribution. We now redo the exercise using cities from the actual US city size distribution. Out of all MSAs (micropolitan areas are dropped) in the year 2008, we keep the largest 180 cities (see Tab. A.1 in the Appendix). These cities comprise 90% of the total population living in MSAs (252 mill.) and 80% of the population in CBSAs (283 mill.).

We assume that the current distribution is an equilibrium and, again, compute the amenities according to (14) (the values are in Tab. A.1 in the Appendix). Fig. 5 again shows the equilibrium and optimal distribution. Zipf's law holds fairly well for the upper tail of the distribution. Note that the largest city in the sample, New York, has 18.7 mill. whereas according to Zipf's law the largest city has more than twice that many inhabitants. As shown in Fig. 5 and Tab. A.1, the three largest cities are undersized by 3.5-3.8%.[19] At the social optimum, the largest city in the sample, New York, is undersized by 696,211 of its 18.7 mill. inhabitants, while the second largest, Los Angeles, is undersized by 483,187. The smallest city in the sample is oversized by about 6.2%. Note that out of the 180 cities, 28 are undersized and the other 152 are oversized at the equilibrium. In total, moving from

lation, the marginal damage of local emissions relative to global emissions should probably be increased.

[18] It is assumed that local pollution is measured in units such that the same marginal damage value can be applied to global and local pollution.

[19] Again, the figure plots the two distributions using the largest 180 cities and $\beta = 0.1$ to better show the difference between the two distributions.

the equilibrium to the optimal allocation would require moving 5.3 mill. people or 2.4% of the total population.

Sensitivity. We now briefly describe how the results change when we vary some of the model parameters. We concentrate on the case of global pollution and the distribution obtained by Zipf's law. First, we increase θ to 0.94 to reflect a potentially higher pollution elasticity. We find that the largest city is undersized by 1.6% and the smallest oversized by 3.8%. Conversely, when θ decreases to 0.75, the biggest city is undersized by 4.4% and the smallest city is oversized by 12.5%. Since the benefit of concentration is increased the more per capita pollution decreases with city size, this finding is intuitive.

Next, suppose that in line with a high estimate of the social cost of carbon, the emissions damage, β, increases from 0.022 to 0.064 (see Section 6). Now, the largest city is undersized by 14.5% and the smallest oversized by 23%. When the agglomeration elasticity, γ, increases from 0.05 to 0.08, agglomeration becomes more efficient, and again, concentration increases at the optimum: the largest city is undersized by 4.5% at the equilibrium and the smallest oversized by 10%.[20]

On the other hand, increasing agricultural land rent r_A to \$100,000 per year increases the costs of agglomeration. The effect on the optimum size of cities is rather small, however. A very similar result obtains when the per mile commuting cost increases to \$750 per year.

7 Conclusion

The paper has analyzed the optimum size of cities in an urban model with environmental pollution. When pollution is purely local and cities are symmetric, we find that equilibrium cities are too large, mirroring the finding of Henderson (1974) and others. With asymmetric cities, this translates into the result that big cities are oversized and small cities undersized.

However, when pollution is global and per capita pollution decreases in population size, we find that in a symmetric city model, cities might be inefficiently small, contrary to the standard model. When cities are asymmetric, big cities are undersized and small cities oversized. Over the last decades, global pollution has increased relative to local pollution, and the damage from global warming increases over time. Hence, we conclude that for the future, a policy which favors big cities might actually be warranted.

[20] In fact, Tabuchi and Yoshida (2000) show that agglomeration externalities from consumption in Japanese cities are about the same size as productive externalities.

Some possibilities for future research suggest themselves. First, our analysis was based on one estimate of the population elasticity of pollution, which is a central parameter in the analysis. More robust evidence on this parameter clearly seems important. Second, we think it would be interesting to redo the quantitative analysis with data from different countries. For instance, there is a growing number of papers on Chinese cities (e.g. Au and Henderson, 2006). Since the properties of the equilibrium city system and pollution patterns in China and other developing economies is undoubtedly different from developed countries, studying equilibrium and optimum city systems in this context would seem to be relevant.

Finally, we could include city governments that maximize residents' utility. We conjecture that a city government would want a city size even smaller than the free mobility case. The reason is that in the symmetric case, the free mobility outcome is at least the n which maximizes $\hat{v}(n)$, the utility without pollution, whereas a city government would maximize roughly the utility $\tilde{v}(n)$ which would correspond to $\hat{v}(n)$ plus the disutility from local pollution, since the local government cares only for the locally produced pollution affecting its own citizens.[21] This could magnify the welfare gain from imposing optimal city sizes.

[21] This assumes global pollution. With local pollution, a city government maximizes $v(n)$ so the allocation would be efficient in the symmetric case. See, e.g. Abdel-Rahman and Anas (2004) and the literature therein.

Appendix

A Calibration of β

We now calibrate β using central estimates of the social cost of carbon from Interagency Working Group on Social Cost of Carbon (2015). The MRS between pollution and (non-housing) consumption is

$$\text{MRS} = -\frac{\partial u/\partial E}{\partial u/\partial z} = \frac{\beta z}{(1-\alpha)E}. \tag{A.1}$$

Substituting optimal consumption, $z(y) = (1-\alpha)(w-tx)$ gives $\text{MRS} = \beta(w-tx)/E$, and integrating over the city gives citywide MRS

$$\overline{\text{MRS}} = \int_0^{\bar{x}} \frac{\beta(w-tx)}{E}\frac{1}{s(x)}dx \tag{A.2}$$

$$= \frac{\beta w\left[tn + r_A - r_A^{1+\alpha}(tn+r_A)^{-\alpha}\right]}{(1+\alpha)tE}. \tag{A.3}$$

where we have substituted the optimal $s(x)$ and used (7) and (8). Finally, letting M be world population and n be city population, we get the social cost of carbon

$$\text{SCC} = \frac{\beta M w\left[tn + r_A - r_A^{1+\alpha}(tn+r_A)^{-\alpha}\right]}{(1+\alpha)tnE}. \tag{A.4}$$

We use the following parameters: world population in 2015 was $M = 7.35$ billion (source: UN World Population Prospects, `http://esa.un.org/unpd/wpp/Download/Standard/Population/`), world per capita income in 2015 was $w = \$10,743$ (source: UN National Accounts Main Aggregates Database, `http://unstats.un.org/unsd/snaama/dnllist.asp`), and total CO_2 emissions were $E = 34,649$ million metric tons CO_2 in 2011 (source: World Bank, World Development Indicators, `http://data.worldbank.org/indicator/EN.ATM.CO2E.KT/countries`). We set $n = 750,000$ and from Borck and Brueckner (2016), we use $t = \$503.53, \alpha = 0.24$, and $r_A = \$58,800$.

The target value for the social cost of carbon is USD 40.54 per ton CO_2, using the central value from Interagency Working Group on Social Cost of Carbon (2015) (converted from 2007 to 2015 USD). Using the stated parameters, setting (A.4) equal to 40.54 and solving gives $\beta = 0.022$. The other values in the text are solved likewise using different values for the social cost of carbon.

B Proof of Proposition 3

Since

$$\begin{aligned} n_i \frac{\partial v_i}{\partial n_i} &= \gamma v_i - \frac{\alpha t n_i}{r_A + t n_i} v_i - \beta \theta n_i^{\theta} v_i E_i^{-1} \\ \sum_{j \neq i} n_j \frac{\partial v_j}{\partial n_i} &= -\beta \theta \delta n_i^{\theta - 1} \sum_{j \neq i} n_j v_j E_j^{-1}, \end{aligned}$$

we have

$$\begin{aligned} & v_i + n_i \frac{\partial v_i}{\partial n_i} + \sum_{j \neq i} n_j \frac{\partial v_j}{\partial n_i} - \lambda \\ &= \left(1 + \gamma - \frac{\alpha t n_i}{r_A + t n_i} \right) v_i - \beta (1 - \delta) \theta v_i n_i^{\theta} E_i^{-1} - \beta \delta \theta Z n_i^{\theta - 1} - \lambda \\ &= X_i v_i - \beta \delta \theta Z n_i^{\theta - 1} - \lambda \qquad \text{(A.5)} \\ &= 0, \end{aligned}$$

where

$$X_i \equiv 1 + \gamma - \frac{\alpha t n_i}{r_A + t n_i} - \beta (1 - \delta) \theta n_i^{\theta} E_i^{-1}$$

and $Z \equiv \sum_j n_j v_j E_j^{-1}$ is constant across cities.

Since the expression (A.5) is the same for i and for $i+1$, we can eliminate λ as follows:

$$X_i v_i - \beta \delta \theta Z n_i^{\theta - 1} = X_{i+1} v_{i+1} - \beta \delta \theta Z n_{i+1}^{\theta - 1},$$

which can be rewritten as

$$v_i = \frac{1}{X_i} \left(X_{i+1} v_{i+1} + \beta \delta \theta Z n_i^{\theta - 1} - \beta \delta \theta Z n_{i+1}^{\theta - 1} \right).$$

Thus, the utility differential is

$$\begin{aligned} \Delta v &\equiv v_i - v_{i+1} \\ &= \frac{1}{X_i} \left[(X_{i+1} - X_i) v_{i+1} + \beta \delta \theta Z \left(n_i^{\theta - 1} - n_{i+1}^{\theta - 1} \right) \right] \\ &= \Delta V_a + \Delta V_b, \end{aligned}$$

where

$$\Delta V_a \equiv \frac{\alpha t r_A v_{i+1} (n_i - n_{i+1})}{(r_A + t n_i)(r_A + t n_{i+1})}$$

$$\Delta V_b \equiv \beta\theta \left[(1-\delta) v_{i+1} \left(n_i^\theta E_i^{-1} - n_{i+1}^\theta E_{i+1}^{-1} \right) + \delta Z \left(n_i^{\theta-1} - n_{i+1}^{\theta-1} \right) \right].$$

While $\Delta V_a > 0$, the sign of ΔV_b is indeterminate. However, the first term of ΔV_b is positive whereas the second term of ΔV_b is negative because

$$\begin{aligned} n_i^\theta E_i^{-1} - n_{i+1}^\theta E_{i+1}^{-1} &= \frac{1}{E_i E_{i+1}} \left(n_i^\theta E_{i+1} - n_{i+1}^\theta E_i \right) \\ &= \frac{\delta}{E_i E_{i+1}} \left(n_i^\theta - n_{i+1}^\theta \right) \sum_j n_j^\theta > 0 \end{aligned}$$

and

$$n_i^{\theta-1} - n_{i+1}^{\theta-1} < 0, \forall\theta \in (0,1).$$

(i) Let $\delta = 0$. Then, $\Delta V_b > 0$, and thus $\Delta v > 0$. By continuity, this also holds for δ positive but close to zero.

(ii) Let $\delta = 1$. Solving $\Delta v < 0$ for β, we have

$$\Delta v < 0 \Leftrightarrow \beta > \tilde{\beta} \equiv \frac{\alpha t r_A v_{i+1}(n_{i+1} - n_i)}{\theta Z (r_A + t n_i)(r_A + t n_{i+1})(n_i^{\theta-1} - n_{i+1}^{\theta-1})} > 0.$$

By continuity, $\Delta v < 0$ holds for sufficiently large β when δ is close to but smaller than one. ■

C Local landownership

Suppose that all land in a city is owned by residents, so the total differential land rent is distributed equally to all residents. Let income be given by $y = w + R/n$, where

$$R = \int_0^{\bar{x}} (r(x,v) - r_A) dx \tag{A.6}$$

is the total differential land rent. Rewriting (8) and (7) gives

$$\bar{x} = \frac{(w + R/n)\left[1 - r_A^\alpha (r_A + tn)^{-\alpha}\right]}{t} \tag{A.7}$$

$$v = (w + R/n)(r_A + tn)^{-\alpha} E^{-\beta}. \tag{A.8}$$

Substituting from (A.8) into $r(x, v) = (w + R/n - tx)^{1/\alpha} E^{-\beta/\alpha} v^{-1/\alpha}$ with $w = n^\gamma$ gives $r(x, v) = (r_A + tn)\,(n^\gamma + R/n)^{-1/\alpha}\,(n^\gamma + R/n - tx)^{\frac{1}{\alpha}}$. Using this in (A.6) and solving gives

$$R = \frac{n^{1+\gamma}\left[r_A^{1+\alpha} - (r_A + tn)^\alpha (r_A - \alpha tn)\right]}{(r_A + tn)^{1+\alpha} - r_A^{1+\alpha}}. \tag{A.9}$$

Finally, substituting in (A.8) gives

$$v = \frac{(1 + \alpha) t E^{-\beta} n^{1+\gamma}}{(r_A + tn)^{1+\alpha} - r_A^{1+\alpha}}. \tag{A.10}$$

which is also inverted U-shaped in n.

We then redo the simulation exercise from Section 6.2. For the city size distribution described by Zipf's law with $\delta = 1$, we find the largest city is undersized by 3.6% and the smallest is oversized by 7.3%, so results are very close to the baseline simulation. Varying δ shows that this also holds for local pollution and for the intermediate case $\delta = 0.43$.

References

Abdel-Rahman, H.M. (1988). Product differentiation, monopolistic competition and city size. *Regional Science and Urban Economics* 18, 69–86.

Abdel-Rahman, H.M. and A. Anas (2004). Theories of systems of cities. In: J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4, 2293–2339, Amsterdam: Elsevier.

Albouy, D., K. Behrens, F. Robert-Nicoud and N. Seegert (2016). The optimal distribution of population across cities. NBER Working Paper 22823.

Amann, M., Z. Klimont and F. Wagner (2013). Regional and global emissions of air pollutants: Recent trends and future scenarios. *Annual Review of Environment and Resources* 38, 31–55.

Arnott, R. Optimal city size in a spatial economy. *Journal of Urban Economics* 6, 65–89.

Au, C.C. and J.V. Henderson (2006). Are Chinese cities too small? *Review of Economic Studies* 73, 549–576.

Blaudin de Thé, C. and M. Lafourcade (2016). The carbon footprint of suburbanization: Evidence from French household data. Mimeographed.

Borck, R. (2016). Will skyscrapers save the planet? Building height limits and urban greenhouse gas emissions. *Regional Science and Urban Economics* 58, 13–25.

Borck, R. and J.K. Brueckner (2016). Optimal energy taxation in cities. CESifo working paper 5711.

Borck, R. and M. Pflüger (2015). Green cities? Urbanization, trade and the environment. IZA discussion paper 9104.

Combes, P.-P. and L. Gobillon (2015). The empirics of agglomeration economies. In G. Duranton, J. V. Henderson and W. C. Strange (eds.), *Handbook of Regional and Urban Economics* vol. 5, 247–348, Amsterdam: Elsevier.

Dascher, K. (2014). City Silhouette, World Climate. Available at `http://ssrn.com/abstract=2250673`.

Davis, M.A. and F. Ortalo-Magné (2011). Household expenditures, wages, rents. *Review of Economic Dynamics* 14, 248–261.

Duranton, G. and D. Puga (2004). Micro-foundations of urban agglomeration economies, in J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol.4, 2063–2117, Amsterdam: Elsevier.

Eeckhout, J. and N. Guner (2017). Optimal spatial taxation: Are big cities too small? CEMFI working paper no. 1705.

Fragkias, M., J. Lobo, D. Strumsky and K.C. Seto (2013). Does size matter? Scaling of CO2 emissions and U.S. urban areas. *PLoS ONE* 8(6): e64727. doi: 10.1371/journal.pone.0064727

Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives* 30, 185–206.

Gaigné, C., S. Riou and J.-F. Thisse (2012). Are compact cities environmentally friendly? *Journal of Urban Economics* 72, 123–136.

Glaeser, E. (2011). *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier*, New York: Penguin Press.

Glaeser, E.L. and M.E. Kahn (2010). The greenness of cities: Carbon dioxide emissions and urban development. *Journal of Urban Economics* 67, 404–418.

Gudipudi, R., T. Fluschnik, A. Garcìa Cantú Ros, C. Walther and J.P. Kropp (2016). City density and CO2 efficiency, *Energy Policy* 91, 352–361.

Henderson, J.V. (1974). The sizes and types of cities. *American Economic Review* 64, 640–656.

Interagency Working Group on Social Cost of Carbon, United States Government (2015). Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866. US Government Printing Office, Washington D.C.

Lamsal, L.N., R.V. Martin, D.D. Parrish and N.A. Krotkov (2013). Scaling relationship for NO2 pollution and urban population size: A satellite perspective. *Environmental Science & Technology* 47, 7855–7861.

Larson, W. and A.M. Yezer (2015). The energy implications of city size and density. *Journal of Urban Economics* 90, 35–49.

Larson, W., F. Liu and A. Yezer (2012). Energy footprint of the city: Effects of urban land use and transportation policies. *Journal of Urban Economics* 72, 147–159.

Morikawa, M. (2013). Population density and efficiency in energy consumption: An empirical analysis of service establishments. *Energy Economics* 34, 1617–1622.

Rybski, D., D.E. Reusser, A.-L. Winz, C. Fichtner, T. Sterzel and J.P. Kropp (2016). Cities as nuclei of sustainability? *Environment and Planning B*, first published on May 5, 2016, doi:10.1177/0265813516638340.

Sarzynski, A. (2012). Bigger is not always better: A comparative analysis of cities and their air pollution impact. *Urban Studies* 49, 3121–3138.

Seto, K.C., B. Güneralp and L.R. Hutyra (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences* 109, 16083–16088.

Tabuchi, T. and A. Yoshida (2000). Separating urban agglomeration economies in consumption and production. *Journal of Urban Economics* 48, 70–84.

Tabuchi, T. and D.-Z. Zeng (2004). Stability of spatial equilibrium. *Journal of Regional Science* 44, 641–660.

Tolley, G.S. (1974). The welfare economics of city bigness. *Journal of Urban Economics* 1, 324–345.

Tscharaktschiew, S. and G. Hirte (2010). The drawbacks and opportunities of carbon charges in metropolitan areas – a spatial general equilibrium approach. *Ecological Economics* 70, 339–357.

Tables

Table 1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Emissions	9330	1.542	3.924	.036	71.06
P.c. emissions	9330	8.4×10^{-6}	.000017	1.1×10^{-6}	.0002556
Population	9330	290029	994482.1	12340	1.87×10^{-7}

Table 2: CO_2-emissions and city size

	(1)	(2)	(3)	(4)
Log population	0.938***	0.938***	0.834***	0.802***
	(0.0168)	(0.0168)	(0.0978)	(0.120)
Constant	2.335***	2.343***	3.533***	3.896***
	(0.202)	(0.201)	(1.117)	(1.369)
Observations	9,330	9,330	9,330	9,330
R-squared [within]	0.681	0.682	0.128	0.147
# of CBSAs	933	933	933	933
Year fixed effects	No	Yes	Yes	Yes
CBSA fixed effects	No	No	Yes	Yes
Division×Year fixed effects	No	No	No	Yes

Standard errors are clustered at the CBSA level.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: own calculations based on data from Fragkias et al. (2013).

Table A.1: Amenity levels and population for MSAs

Rank	MSA	Amenity	Population	Optimal population	Emissions
1	New York-Northern New Jersey-Long Island, NY-NJ-PA	1	18672355	19368566	40.6523
2	Los Angeles-Long Beach-Santa Ana, CA	0.929283	12692740	13175927	24.523
3	Chicago-Joliet-Naperville, IL-IN-WI	0.877468	9384555	9712021	42.6155
4	Dallas-Fort Worth-Arlington, TX	0.809967	6158022	6309423	17.4737
5	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.803586	5906917	6044590	15.9336
6	Houston-Sugar Land-Baytown, TX	0.79822	5702270	5828906	24.0447
7	Miami-Fort Lauderdale-Pompano Beach, FL	0.791515	5454633	5568153	12.6531
8	Washington-Arlington-Alexandria, DC-VA-MD-WV	0.789769	5391607	5501839	19.3482
9	Atlanta-Sandy Springs-Marietta, GA	0.782982	5152141	5250104	25.4191
10	Boston-Cambridge-Quincy, MA-NH	0.762562	4483141	4549310	13.6544
11	Detroit-Warren-Livonia, MI	0.757858	4339504	4399464	18.7855
12	San Francisco-Oakland-Fremont, CA	0.754658	4243932	4299907	17.8689
13	Riverside-San Bernardino-Ontario, CA	0.750194	4113447	4164183	11.8552
14	Phoenix-Mesa-Glendale, AZ	0.749949	4106372	4156830	14.1292
15	Seattle-Tacoma-Bellevue, WA	0.721702	3355042	3380652	7.92053
16	Minneapolis-St. Paul-Bloomington, MN-WI	0.716596	3231982	3254448	20.0776
17	San Diego-Carlsbad-San Marcos, CA	0.707513	3022116	3039814	8.18351
18	St. Louis, MO-IL	0.696989	2792889	2806143	20.9905
19	Tampa-St. Petersburg-Clearwater, FL	0.694798	2746981	2759424	10.9384
20	Baltimore-Towson, MD	0.691538	2679819	2691115	10.236
21	Denver-Aurora-Broomfield, CO	0.680593	2463971	2471789	8.03722
22	Pittsburgh, PA	0.674867	2356802	2362930	16.1156
23	Portland-Vancouver-Hillsboro, OR-WA	0.664527	2172853	2175941	5.03929
24	Cincinnati-Middletown, OH-KY-IN	0.660888	2110942	2112924	13.3771
25	Sacramento-Arden-Arcade-Roseville, CA	0.660731	2108310	2110244	5.86701
26	Orlando-Kissimmee-Sanford, FL	0.659487	2087489	2089037	6.3697
27	Cleveland-Elyria-Mentor, OH	0.659344	2085110	2086614	12.8774
28	San Antonio-New Braunfels, TX	0.657905	2061275	2062328	9.96783
29	Kansas City, MO-KS	0.654128	1999739	1999578	12.7061
30	Las Vegas-Paradise, NV	0.648598	1912349	1910333	6.01372
31	Columbus, OH	0.641184	1800052	1795400	4.55883
32	San Jose-Sunnyvale-Santa Clara, CA	0.640857	1795231	1790460	4.66359
33	Indianapolis-Carmel, IN	0.635581	1718784	1712050	6.0172
34	Charlotte-Gastonia-Rock Hill, NC-SC	0.634421	1702338	1695167	6.24345
35	Virginia Beach-Norfolk-Newport News, VA-NC	0.631211	1657491	1649102	6.02219
36	Austin-Round Rock-San Marcos, TX	0.629492	1633870	1624828	5.41536
37	Providence-New Bedford-Fall River, RI-MA	0.6271	1601459	1591510	6.24221
38	Nashville-Davidson-Murfreesboro-Franklin, TN	0.623012	1547259	1535778	6.90834
39	Milwaukee-Waukesha-West Allis, WI	0.62232	1538232	1526495	6.60184
40	Jacksonville, FL	0.604728	1322728	1305326	7.28384
41	Memphis, TN-MS-AR	0.602921	1302060	1284213	5.0071
42	Louisville-Jefferson County, KY-IN	0.599561	1264314	1245729	9.11605
43	Richmond, VA	0.597203	1238353	1219325	6.56015
44	Oklahoma City, OK	0.5952	1216645	1197290	4.88706
45	Hartford-West Hartford-East Hartford, CT	0.59406	1204436	1184917	3.19058
46	Buffalo-Niagara Falls, NY	0.587531	1136364	1116194	5.46278
47	Birmingham-Hoover, AL	0.585626	1117101	1096833	13.5128

continued on next page

Tab. A.1 continued

Rank	MSA	Amenity	Population	Optimal population	Emissions
48	New Orleans-Metairie-Kenner, LA	0.58529	1113736	1093455	7.66965
49	Salt Lake City, UT	0.58297	1090691	1070352	2.91391
50	Raleigh-Cary, NC	0.581589	1077163	1056816	2.53995
51	Rochester, NY	0.578769	1049950	1029640	2.7185
52	Tucson, AZ	0.569877	967778	947942	2.41114
53	Tulsa, OK	0.564017	916525	897130	7.0816
54	Fresno, CA	0.563209	909630	890294	2.05444
55	Bridgeport-Stamford-Norwalk, CT	0.562524	903824	884536	3.17886
56	Albany-Schenectady-Troy, NY	0.558009	866282	847263	2.93617
57	Albuquerque, NM	0.557653	863383	844380	2.50765
58	New Haven-Milford, CT	0.556821	856622	837655	2.3126
59	Omaha-Council Bluffs, NE-IA	0.555393	845119	826202	6.44558
60	Dayton, OH	0.555115	842897	823988	2.5307
61	Bakersfield-Delano, CA	0.552004	818327	799464	4.98675
62	Allentown-Bethlehem-Easton, PA-NJ	0.551455	814050	795187	5.04325
63	Oxnard-Thousand Oaks-Ventura, CA	0.550461	806353	787482	2.29419
64	Worcester, MA	0.548434	790847	771934	2.7666
65	Baton Rouge, LA	0.548028	787767	768841	10.3903
66	Grand Rapids-Wyoming, MI	0.546107	773342	754334	2.1559
67	El Paso, TX	0.545649	769930	750898	2.64812
68	Columbia, SC	0.54235	745740	726483	3.47064
69	McAllen-Edinburg-Mission, TX	0.541094	736694	717330	3.53499
70	Greensboro-High Point, NC	0.537515	711405	691688	1.98387
71	Akron, OH	0.536347	703300	683456	1.9205
72	North Port-Bradenton-Sarasota, FL	0.535276	695944	675982	2.82686
73	Springfield, MA	0.534587	691239	671200	2.20756
74	Knoxville, TN	0.534101	687939	667846	3.10894
75	Little Rock-North Little Rock-Conway, AR	0.533206	681888	661696	2.15732
76	Stockton, CA	0.531682	671692	651334	1.37026
77	Poughkeepsie-Newburgh-Middletown, NY	0.530894	666468	646027	2.85408
78	Syracuse, NY	0.529746	658913	638356	1.90001
79	Toledo, OH	0.528919	653518	632882	3.62579
80	Charleston-North Charleston-Summerville, SC	0.527387	643613	622844	6.73385
81	Greenville-Mauldin-Easley, SC	0.524335	624245	603276	1.33841
82	Colorado Springs, CO	0.523759	620644	599649	2.63146
83	Cape Coral-Fort Myers, FL	0.5222	610984	589940	2.26724
84	Wichita, KS	0.521918	609250	588200	2.19355
85	Boise City-Nampa, ID	0.520933	603218	582158	1.66843
86	Lakeland-Winter Haven, FL	0.519544	594801	573751	4.2558
87	Youngstown-Warren-Boardman, OH-PA	0.515521	570952	550096	1.85198
88	Scranton-Wilkes-Barre, PA	0.513898	561548	540839	1.64199
89	Madison, WI	0.513254	557854	537214	3.79531
90	Des Moines-West Des Moines, IA	0.512516	553644	533089	1.66703
91	Augusta-Richmond County, GA-SC	0.510892	544471	524128	1.76162
92	Palm Bay-Melbourne-Titusville, FL	0.510518	542378	522088	1.73565
93	Harrisburg-Carlisle, PA	0.510505	542301	522013	1.60903
94	Jackson, MS	0.508897	533371	513325	1.76857
95	Ogden-Clearfield, UT	0.50803	528603	508696	1.43869

continued on next page

Tab. A.1 continued

Rank	MSA	Amenity	Population	Optimal population	Emissions
96	Chattanooga, TN-GA	0.506223	518778	499172	1.75345
97	Portland-South Portland-Biddeford, ME	0.505369	514191	494730	2.03088
98	Lancaster, PA	0.504952	511957	492567	1.29977
99	Modesto, CA	0.504402	509032	489735	1.16168
100	Provo-Orem, UT	0.502238	497639	478699	1.77947
101	Deltona-Daytona Beach-Ormond Beach, FL	0.502186	497366	478435	2.63952
102	Durham-Chapel Hill, NC	0.500749	489919	471207	6.31914
103	Santa Rosa-Petaluma, CA	0.497435	473091	454792	1.00393
104	Winston-Salem, NC	0.49695	470666	452414	11.8863
105	Lansing-East Lansing, MI	0.495532	463638	445499	2.25539
106	Spokane, WA	0.495253	462263	444141	1.01927
107	Lexington-Fayette, KY	0.494814	460112	442016	1.74198
108	Fayetteville-Springdale-Rogers, AR-MO	0.492019	446592	428569	2.30046
109	Pensacola-Ferry Pass-Brent, FL	0.491768	445392	427368	3.2278
110	Flint, MI	0.489157	433082	414988	1.24379
111	York-Hanover, PA	0.488364	429399	411263	3.27711
112	Springfield, MO	0.48834	429289	411151	1.91169
113	Visalia-Porterville, CA	0.488339	429283	411145	0.993425
114	Corpus Christi, TX	0.48701	423168	404944	2.30819
115	Reno-Sparks, NV	0.486071	418892	400598	1.44999
116	Port St. Lucie, FL	0.485769	417520	399203	3.52693
117	Asheville, NC	0.485755	417457	399139	1.62034
118	Santa Barbara-Santa Maria-Goleta, CA	0.485401	415859	397513	1.10609
119	Fort Wayne, IN	0.484556	412062	393650	1.24676
120	Mobile, AL	0.483914	409196	390734	4.38011
121	Vallejo-Fairfield, CA	0.483864	408972	390506	0.919361
122	Reading, PA	0.483586	407737	389250	1.64164
123	Canton-Massillon, OH	0.483226	406140	387627	0.815015
124	Salinas, CA	0.483199	406022	387507	6.0969
125	Huntsville, AL	0.482368	402361	383789	0.921342
126	Manchester-Nashua, NH	0.481728	399556	380946	1.02811
127	Brownsville-Harlingen, TX	0.480217	393000	374325	0.597551
128	Shreveport-Bossier City, LA	0.48007	392367	373688	4.37908
129	Killeen-Temple-Fort Hood, TX	0.480039	392237	373558	0.760663
130	Beaumont-Port Arthur, TX	0.478403	385248	366560	2.958
131	Salem, OR	0.478126	384075	365392	0.797142
132	Davenport-Moline-Rock Island, IA-IL	0.476313	376467	357875	1.57889
133	Peoria, IL	0.476196	375982	357400	5.41263
134	Montgomery, AL	0.474808	370249	351811	1.38011
135	Trenton-Ewing, NJ	0.473305	364119	345908	1.47235
136	Hickory-Lenoir-Morganton, NC	0.473276	364003	345797	3.91189
137	Tallahassee, FL	0.472592	361238	343158	1.25284
138	Fayetteville, NC	0.47121	355712	337923	0.625336
139	Evansville, IN-KY	0.47118	355591	337809	11.4147
140	Wilmington, NC	0.47015	351517	333978	2.30292
141	Rockford, IL	0.469747	349937	332497	1.19932
142	Eugene-Springfield, OR	0.469298	348176	330849	0.676484
143	Ann Arbor, MI	0.4676	341595	324696	1.09276

continued on next page

Tab. A.1 continued

Rank	MSA	Amenity	Population	Optimal population	Emissions
144	Savannah, GA	0.465921	335185	318682	2.38122
145	Ocala, FL	0.464557	330052	313817	0.756947
146	Kalamazoo-Portage, MI	0.462754	323363	307368	0.97173
147	South Bend-Mishawaka, IN-MI	0.461827	319966	304033	0.824083
148	Naples-Marco Island, FL	0.460912	316641	300723	0.876601
149	Kingsport-Bristol-Bristol, TN-VA	0.458516	308069	291989	2.22917
150	Roanoke, VA	0.457814	305596	289423	1.89997
151	Charleston, WV	0.457286	303743	287491	5.37228
152	Green Bay, WI	0.456921	302468	286158	1.79404
153	Utica-Rome, NY	0.455888	298886	282408	0.782581
154	Lincoln, NE	0.455124	296258	279661	1.54115
155	Fort Smith, AR-OK	0.454625	294551	277884	2.86569
156	Fort Collins-Loveland, CO	0.453771	291650	274888	1.28267
157	Boulder, CO	0.453537	290859	274078	1.11593
158	Columbus, GA-AL	0.452384	286985	270166	0.692666
159	Huntington-Ashland, WV-KY-OH	0.452378	286966	270147	1.67665
160	Spartanburg, SC	0.450172	279673	263115	0.778184
161	Erie, PA	0.44987	278686	262200	0.781988
162	Duluth, MN-WI	0.449832	278561	262085	2.29038
163	Lubbock, TX	0.448545	274389	258312	0.940169
164	Atlantic City-Hammonton, NJ	0.448117	273014	257093	0.864827
165	Norwich-New London, CT	0.447998	272634	256757	0.803819
166	Lafayette, LA	0.446242	267053	251818	0.942161
167	San Luis Obispo-Paso Robles, CA	0.446021	266358	251191	0.799599
168	Hagerstown-Martinsburg, MD-WV	0.445851	265823	250705	1.13058
169	Holland-Grand Haven, MI	0.444596	261906	247000	3.57025
170	Gainesville, FL	0.444525	261685	246781	1.13324
171	Clarksville, TN-KY	0.444475	261530	246628	4.90448
172	Myrtle Beach-North Myrtle Beach-Conway, SC	0.444177	260609	245701	1.01244
173	Santa Cruz-Watsonville, CA	0.442845	256520	241318	0.62579
174	Cedar Rapids, IA	0.442511	255503	240164	0.920119
175	Binghamton, NY	0.441528	252527	236721	0.868548
176	Merced, CA	0.440866	250538	234484	0.747178
177	Lynchburg, VA	0.440451	249299	233199	0.971995
178	Bremerton-Silverdale, WA	0.439647	246912	231226	0.399646
179	Amarillo, TX	0.438812	244454	230499	3.3003
180	Olympia, WA	0.438771	244332	229276	0.467105

Note: The table displays population and emissions levels by MSA for 2008 from Fragkias et al. (2013). Amenity levels are computed from (17); optimal population levels are simulated as described in the text.

Figures

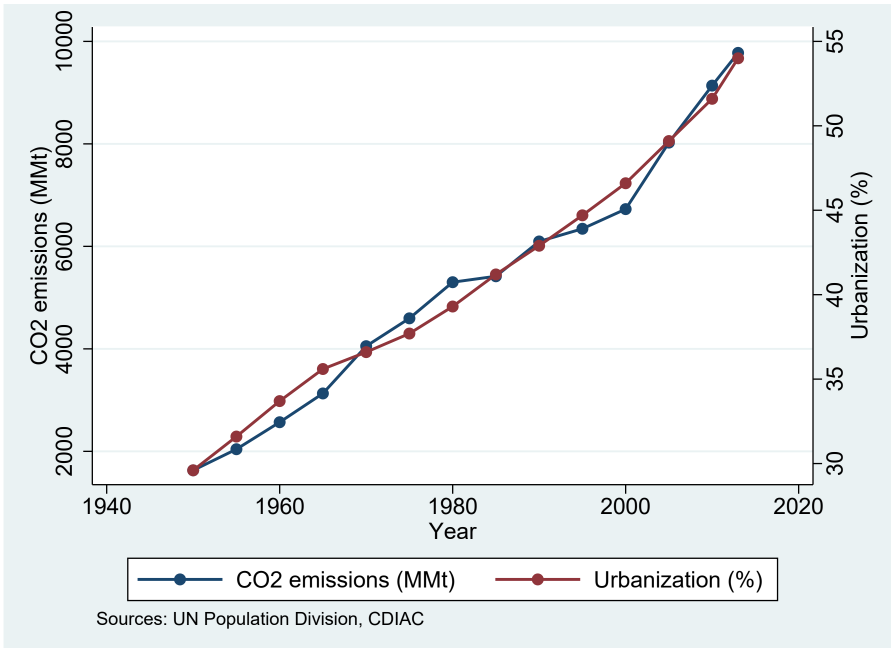


Figure 1: World urbanization and CO_2 emissions

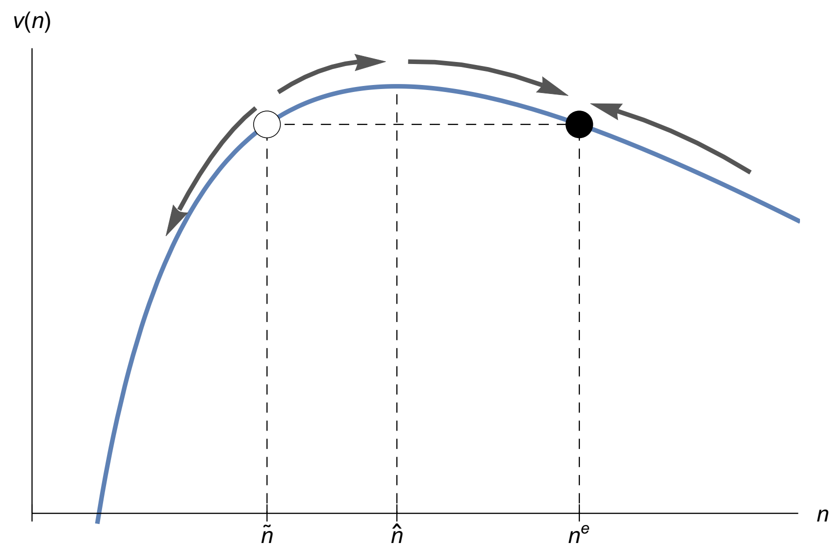


Figure 2: Equilibrium city size

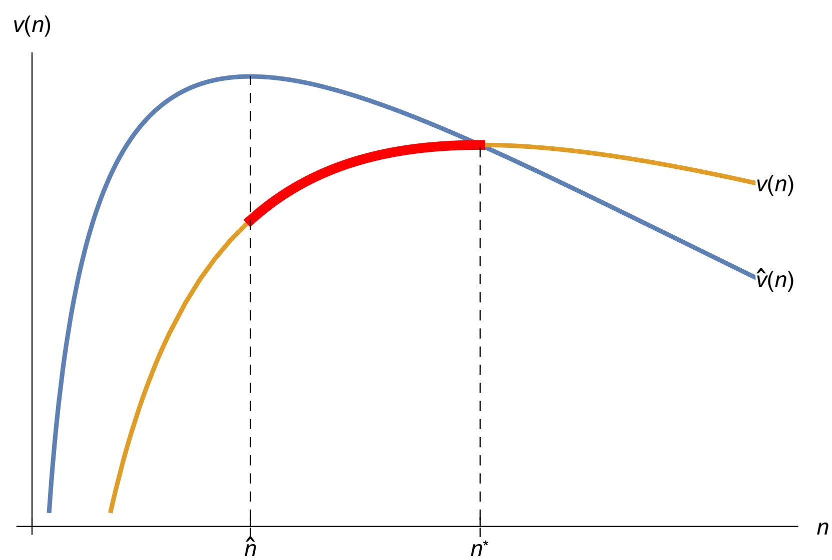


Figure 3: Equilibrium and optimum city size with global pollution

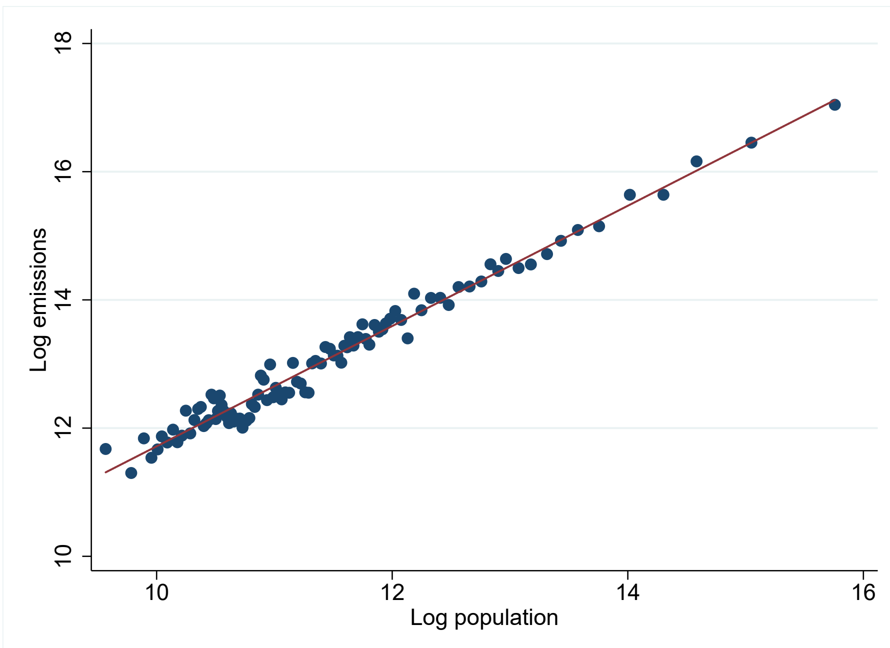
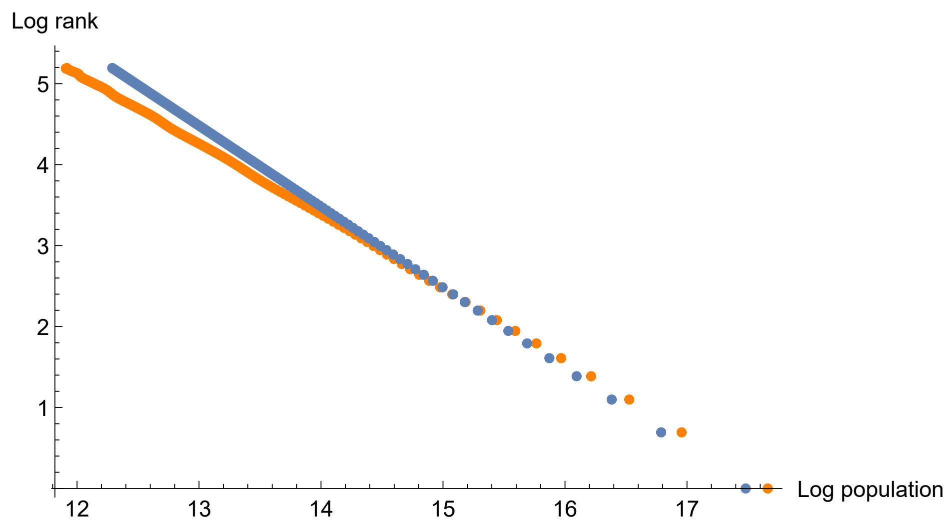
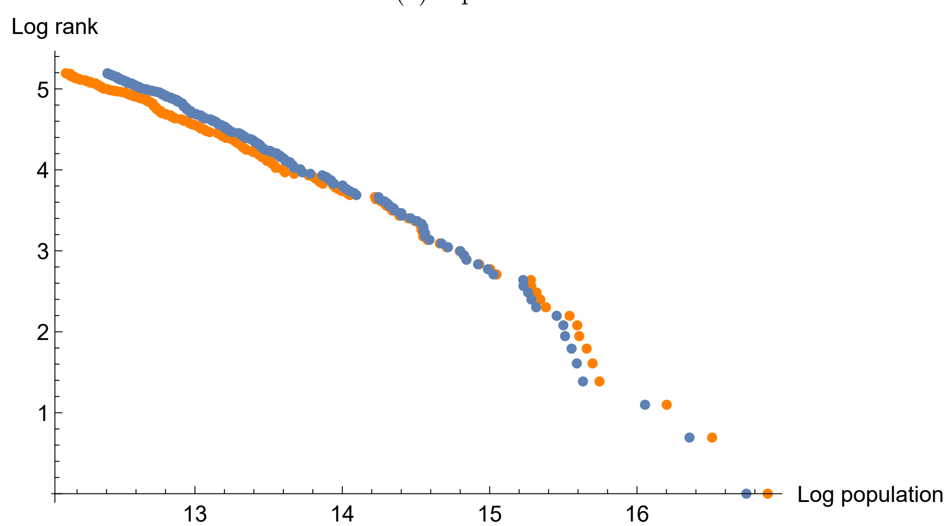


Figure 4: CO_2-emissions and city size (Source: own calculations based on data from Fragkias et al. (2013))



(a) Zipf's law



(b) Actual CBSA distribution

Figure 5: Optimal and equilibrium city size distributions (Source: own calculations based on data from Fragkias et al. (2013))