I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11354

City Size, Pollution and Emission Policies

Michael Pflüger

FEBRUARY 2018

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11354

City Size, Pollution and Emission Policies

Michael Pflüger
University of Würzburg, DIW Berlin and IZA

FEBRUARY 2018

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany
Phone: +49-228-3894-0
Email: publications@iza.org
www.iza.org

ABSTRACT

City Size, Pollution and Emission Policies*

This paper develops a micro-founded city systems model with an endogenous number of cities to explore whether local governments establish the optimal city size when production processes involve environmental pollution. Our analysis delivers two key insights. First, if an optimal scheme to regulate environmental pollution is implemented, cities chosen by local governments are never too large. They are too small if pollution is purely global, but at the optimal size, if pollution is purely local. Second, if no emission scheme is implemented or if emission policies are too lax, then cities steered by local governments, become too large, however.

JEL Classification: H73, R12, Q50

Keywords: city systems, environmental pollution, emission policies

Corresponding author:
Michael Pflüger
Faculty of Economics
University of Würzburg
Sanderring 2
97070 Würzburg
Germany

E-mail: michael.pflueger@uni-wuerzburg.de

* I would like to thank Gilles Duranton, Vernon Henderson, Yoshitsugu Kanemoto, Tomoya Mori, Yasusada Murata, Minoru Osawa, Jun Oshiro, Takatoshi Tabuchi, Kazuhiro Yamamoto, Markus Wiest as well as participants of the Tokyo Workshop on Spatial Economics 2017 (RIETI), at the UEA-NARSC-sessions in Vancouver 2017, and at research seminars at Kyoto University and Osaka University for helpful comments and discussion. I gratefully acknowledge financial support from Deutsche Forschungsgemeinschaft (DFG) through PF 360/7-1 and the hospitality of the University of Tokyo where parts of this paper were written.

1 Introduction

Are cities too small or too big? An influential recent line of research holds that, in stark contrast to public fears that the world's large cities are oversized and prone to sprawl further, these cities may actually be too small (see e.g. Albouy et al. 2016; Au and Henderson 2006; Desmet and Rossi-Hansberg 2013). A key argument in these analyses is that mobility restrictions imposed by local governments may keep city sizes too small so that, from a national perspective, the potential of these cities remains unexploited and countries thereby forego large welfare benefits.

Despite a wide recognition that environmental problems, global warming in particular, rank very high among the most important issues for today's societies, only a small literature explicitly addresses environmental concerns in the context of city-size. The purpose of this paper is to analyze the city-size issue in the face of environmental pollution within a model of city systems in the tradition of Henderson (1974) where the number of cities is endogenous. This analysis is, to the best of our knowledge, the first one to exploit this canonical model for this purpose, previous works take the number of cities as fixed.

We develop a micro-founded model with agglomeration economies due to the sharing of inputs whose production causes pollution and with a monocentric structure of cities which involves congestion due to commuting. The analysis allows for commuting costs either in terms of time ('iceberg costs') or in terms of local output ('money') and it is shown that this distinction has important consequences.

The focus of this paper is on comparing the allocation implied by the social planner (a benevolent national government) with the allocation chosen by local governments and this for the following reason. Although the evolution of cities and urban systems is shaped by various forces of self-organization and control, i.e. mobile people are attracted to locations which satisfy their pecuniary and non-pecuniary needs best, and local and national governments directly and indirectly influence city sizes, there is a growing perception that governments, notably local governments, play the crucial role in determining city sizes today.[1]

Comparing equilibrium city sizes chosen by local governments with the social optimum our analysis delivers the key novel findings when environmental pollution is global. *Specifically, if an optimal national emission policy is enacted (a permit system or, equivalently, an emission tax) but city sizes are regulated by local governments, then cities are too small and too numerous.* This novel result holds true in general, irrespective of the type of commuting costs. Intuitively, this discrepancy arises because the social planner economizes on the creation of new cities, whereas local governments do not take into account what their choice of city size implies for the number of cities. This city size bias which arises in the face of global environmental pollution in a model with an endogenous number of city locations has not been identified, yet. Our further results are intricately affected by the type of commuting costs, time versus money. In particular, we show that, in stark contrast to the above, *if commuting costs are in terms of output, cities chosen by local governments become excessively large when a national emission policy is either not in place or not stringent enough,* which may quite likely be the case in practice for political economy reasons. The distinction between time costs and monetary costs of commuting is also crucial for whether the marginal disutility of pollution is a determinant of city size and whether it matters for city size if the proceeds from an emission policy are rebated to residents or not (a positive fiscal externality in the latter case): in both cases city size is only affected when commuting costs are in terms of output.

Turning to the case of purely local pollution, a first key result is that benevolent local governments implement the first-best, irrespective of the type of commuting costs. Intuitively,

[1] Desmet and Henderson (2015) argue that "(…) while much of what we see is driven by market forces, the role of governments in economies has grown" and "Government policies and institutions strongly influence the structure of urban hierarchy." Glaeser (2013) provides a lucid review of the institutions of local governments and urban political economy. Hsieh and Moretti (2017) show that "… the constraints on housing supply in the most productive US cities effectively limit the number of workers who have access to such high productivity." Recent theoretical work highlights local governments, see e.g. Duranton and Puga (2014; 2017) and Albouy et al. (2016).

and in contrast to the case of global pollution, the social planner and local governments face an identical maximization problem under local pollution. However, similar discrepancies as under global pollution arise between the first-best and the local government solution when governments are not benevolent and commuting costs are in terms of output.

How should we assess the result that different types of commuting costs may have so different implications? There is ample evidence that both types of costs prevail in practice (Duranton and Puga 2004). This indication is important. Even though we explore the two types of commuting costs one by one and not simultaneously for reasons of tractability, we can predict from that analysis that *the implications derived for monetary commuting costs command more practical relevance when both types of commuting costs prevail*. This is so for the following reason. The stark results implied with commuting costs in terms of time (iceberg costs) derive from the fact that output and commuting costs are proportional in that case. This proportionality is broken as soon as commuting costs are (also) monetary, however, so that the results obtained for the case with commuting costs in terms of local output are a better practical guide.[2]

The policy upshot of this paper can then be summarized in two statements. First, if optimal schemes to regulate environmental pollution are implemented, cities chosen by local governments are never too large. Second, unless we can be sure that we get our environmental policies right, i.e. unless emission schemes are stringent enough, cities steered by local governments become too large.

Our analysis relates to two strands of research. First, there is the theoretical research on the optimal distribution of population across cities. Well-known insights of that research are that cities are too big under self-organization, i.e. free migration causes cities to become too large

[2] A similar argument prevails with respect to the effect of local productivity shifters (amenities, Ricardian land differences) on city size. Such productivity shifters affect city size if commuting costs are in terms of local output but not if they are time costs (e.g. Duranton and Puga 2004; 2014) which is why recent analyses of the city size distribution focus on commuting costs in terms of local output (see Behrens and Robert-Nicoud 2015 and Duranton and Puga 2017).

since migrants are not faced with the increasingly negative externalities that they impose on cities. Efficiency can be restored by competitive land-developers or by allowing for the formation of autonomous local governments which act on behalf of the atomic agents, however (Henderson 1974; Becker and Henderson 2000; Abdel-Rahman and Anas 2004). Albouy et al. (2016) provide a contrast by showing that inefficiently low city sizes may be chosen by local governments in the presence of positive fiscal externalities or when there are Ricardian differences in land. Our analysis focuses on local governments, too, but we address a novel issue, environmental pollution, and we do so within a fully micro-founded model.

There is also a growing literature that addresses the nexus between cities and the environment, much inspired by Glaeser's (2011) hypothesis that large cities make us not only richer, smarter and more productive, but also greener (Kahn and Walsh 2015; Kahn 2006). The interface between cities and environmental pollution has been addressed with a new economic geography oligopoly model by Gaigné et al. (2012). They show that Glaeser's hypothesis needs to be balanced when intra- and intercity interactions, such as longer commutes and the transport of goods are taken into account and this is reinforced by Borck and Pflüger (2017) drawing on a Krugman-type new economic geography model with endogenous lot sizes. Borck and Tabuchi (2016) set up a model with a fixed number of locations and heterogeneous amenities to highlight the difference between self-organization and the social optimum.[3] Desmet and Rossi-Hansberg (2015) analyze global warming within a flexible quantitative spatial model.

The structure of the rest of the paper is as follows. Section 2 introduces the model. Section 3 compares the social planner solution with the allocation chosen by local governments when environmental pollution is global. Section 4 addresses the case of purely local pollution. Section 5 concludes.

[3] Kyriakopolou and Xepapadeas (2016) and Schindler et al. (2017) study pollution and land use for monocentric cities, without looking at a city system, however.

2 The model

The analysis builds on the canonical model of city systems of Henderson (1974) with an endogenous number of n cities and with micro-foundations in terms of input sharing following Ethier (1982) and Abdel-Rahman and Fujita (1990), as conveniently laid out in Duranton and Puga (2014; 2004). We amend this model by assuming that intermediate inputs are produced with emissions and labor, rather than labor alone, and that emissions have a negative welfare effect on consumer-workers.

Preferences. Consumer-workers live in cities (index i), supply 1 unit of working time and consume 1 unit of housing, each. Their utility is linear in consumption X_i of a homogeneous and freely tradable final good, which is chosen to be the numéraire, and additively separable in the disutility associated with pollution Ω_i in the city, $U_i = X_i - \eta \cdot \Omega_i$, where $\eta > 0$.[4] The consumer has gross income I_i which consists of her wage w_i, a proportionate share of total land rents, TLR_i/N_i, where N_i is the (endogenous) number of city residents, and possibly also a proportionate share of the proceeds from an emission policy, TET_i/N_i. Let $R_i(0)$ denote average urban costs in spatial equilibrium in the city which comprise a consumer's expenses for housing and the commute to the CBD and back (detailed below). Her income net of urban costs (spent on the final good) is $c_i = I_i - R_i(0)$ and her indirect utility is $V_i = c_i - \eta\Omega_i$.

Pollution and spatial equilibrium across cities. Pollution in city i is the result of emissions of intermediate firms. If pollution is purely local, $\Omega_i = m_i e_i = E_i$, where m_i is the mass of intermediate firms in city i, e_i denotes the emissions of a single firm and E_i aggregate emissions in the city. If pollution is purely global, $\Omega_i = nE_i$. Consumers are mobile. Spatial equilibrium across cities commands utility equalization at a common level, $V_i = \bar{V}$.

[4] The analysis extends to more than one final output as shown in a previous version of the paper. All key results are obtained when there is one final output. I thank several discussants of this paper for urging me to simplify the analysis in this way.

Production and the wage equation. Final output in the city is produced according to the CES-production function $Y_i = B\left\{\int_0^{m_i}[y_i(h)]^{\frac{1}{1+\sigma}}\,dh\right\}^{1+\sigma}$, where $y_i(h)$ is the quantity of intermediate input h, m_i the mass of intermediates, B a productivity shifter, $0 < \sigma < 1$ and $\varepsilon \equiv (1+\sigma)/\sigma$ is the elasticity of technological substitution between any two intermediates.

Intermediates are non-tradable and produced with labor l_i and emissions e_i under increasing returns and monopolistic competition according to the cost function $C_{y_i}[y_i(h)] = {w_i}^{1-\rho}{t_i}^{\rho}\ [y_i(h)+\alpha]$, where t_i denotes the (shadow) price of emissions associated with the environmental policy, $\alpha > 0$, and $0 \leq \rho \leq 1$ is the weight of variable and fixed emissions in production. This technology extends the standard specification where labor is the only input ($\rho = 0$, see Duranton and Puga 2004) along the lines of Krugman and Venables (1995) and Tabuchi and Pflüger (2011). This cost specification can be interpreted as being supported by an explicit abatement technology as in Copeland and Taylor (1994; 2003).[5] Following the latter we impose $e_i \leq \kappa\, l_i$, where $\kappa > 0$ limits the substitution possibilities between labor and emissions to ensure that output is bounded above for a given labor input.

The quantities of intermediates are chosen to minimize the costs to produce final output Y_i. Conditional input demand is $y_i(h) = \frac{[q_i(h)]^{-(1+\sigma)/\sigma}}{\left\{\int_0^{m_i}[q_i(h')]^{-1/\sigma}dh'\right\}^{1+\sigma}}\frac{Y_i}{B}$, where $q_i(h)$ denotes the price of intermediate h. Hence, firm h faces own-price demand elasticity $-(1+\sigma)/\sigma$ and its profit-maximizing price is a constant mark-up on marginal costs, $q_i = (1+\sigma)\,{w_i}^{1-\rho}{t_i}^{\rho}$. Since all variables take on identical values for all intermediate firms due to symmetry we drop index h from now on. Free entry drives intermediates' profits to zero, $\pi_i = q_i y_i - C_{y_i} = 0$. Hence, break-even output is $y_i = \alpha/\sigma$. Aggregate labor input and emissions of intermediate firms comprise constant and variable components and are calculated as $L_i = \alpha(1-\rho)\varepsilon_i m_i (t_i/w_i)^{\rho}$

[5] More specifically, emissions are a side-product (joint output) in the process of intermediate production, but pollution can be reduced by devoting part of labor to abatement (see Copeland and Taylor 1994; 2003).

and $E_i = \alpha\rho\varepsilon m_i (t_i/w_i)^{\rho-1}$, respectively. Raising w_i/t_i lowers the demand for labor and raises the demand for emissions.

Using $m_i = [L_i(w_i/t_i)^{\rho}] \ /[\alpha(1-\rho)\varepsilon]$ implied by labor demand, as well as $y_i = \alpha/\sigma$ and the normalization $(\alpha/\sigma)^{-\sigma}(1+\sigma)^{-(1+\sigma)} \ \rho^{-\rho(1+\sigma)} \ (1-\rho)^{-(1-\rho)(1+\sigma)} = 1$ (in analogy to Duranton and Puga 2014), and applying symmetry, the aggregate production function in city i is:

$$Y_i(L_i, E_i) = B \ E_i^{\rho(1+\sigma)} \ L_i^{(1-\rho)(1+\sigma)} \tag{1}$$

Perfect competition implies that revenue equals cost in final output production, $Y_i = C_{Y_i} = w_i L_i + t_i E_i$. Employing $t_i E_i = w_i L_i \ \rho/(1-\rho)$ implied by the demand for labor and emissions at the city level, we have $w_i = (1-\rho) \ Y_i/L_i$. Using (1) the wage in the city follows as:

$$w_i = (1-\rho) \ B \ E_i^{\rho(1+\sigma)} \ L_i^{(1-\rho)(1+\sigma)-1} \tag{2}$$

Eqs. (1) and (2) deserve two comments. First, emissions have a qualitatively similar (positive) impact on aggregate output and on the wage as do productive amenities B. Of course, the difference is that pollution harms consumers. Second, when aggregate output is produced with labor and emissions, the elasticity of production with respect to labor is $(1-\rho)(1+\sigma)$, so the sharing externality is weaker due to the second factor. We impose $(1-\rho)(1+\sigma) > 1$, i.e. ρ may not be too large, to ensure that a city with positive population is viable: then, aggregate output exhibits increasing returns to labor and the wage in the city is positively related to L_i.

The urban sector. Cities are monocentric, one-sided and stretch out linearly from the CBD at $r_i = 0$ where production takes place, to the residences located at distance r_i from the CBD. The opportunity cost of land at the city border $\bar{r}_i$ is normalized to zero. Since workers consume 1 unit of floor-space, the city border is at $\bar{r}_i = N_i$. Workers commute from their residences to the CBD and back at a cost. Commuting costs comprise both the opportunity cost of time and resources, in practice. For the reason of tractability, the literature focuses on only one of these commuting cost types at a time. We follow this practice but look at both in turn since these different specifications have partly different and important implications.

(i) Iceberg commuting costs. The iceberg specification assumes that a commute to and back from the CBD reduces a consumer's unit working time by $2\,\tau\,r_i$, where $\tau > 0$ is a commuting cost parameter. This case is amply covered in the literature, so we simply state key results for further reference.[6] We use the additional sub-index *b* to indicate iceberg-costs where necessary. Total commuting cost in terms of lost labor are given by $TCCL_{bi} = \tau {N_i}^2$, bid rents (the price of land) at r_i is $R_{bi}(r_i) = 2w_i\tau(N_i - r_i)$, total land rent in the city is $TLR_{bi} = w_i\tau {N_i}^2$, average urban costs are $R_{bi}(0) = 2w_i\tau N_i$ and the labor supply in a city with N_i consumer-workers is $L_i = N_i(1 - \tau N_i)$. Inserting this into (1), a city's net aggregate output (output with respect to city population) is:

$$Y_{bi}^{Net}(N_i, E_i) = B\,{E_i}^{\rho(1+\sigma)}\,[N_i(1 - \tau N_i)]^{(1-\rho)(1+\sigma)} \tag{3}$$

(ii) Commuting costs in terms of local output. The early literature which addressed commuting costs in terms of local output took them to be linear (Becker and Henderson 2000; Duranton and Puga 2004). It is now common to generalize this specification by assuming that commuting costs are iso-elastic with respect to distance (Duranton and Puga 2014; 2017; Albouy et al. 2016; Behrens and Robert-Nicoud 2015), Combes et al. (2016) provide empirical support. Effective and notional labor supply coincide in this case, $L_i = N_i$. We apply the sub-index y for commuting costs in terms of local output where necessary. Following Duranton and Puga (2014) we assume that the commuting cost of a resident living at distance r_i from the CBD is given by $\tau\,{r_i}^{\gamma}\,(1+\gamma)/\gamma$ where $\gamma > 0$ is the mentioned elasticity, $\tau > 0$ is a commuting cost parameter, and the term $(1+\gamma)/\gamma$ is introduced to simplify expressions (recall that final output is the numéraire). As shown in Duranton and Puga (2014), total commuting costs are $TCC_{yi} =$

[6] See e.g. Abdel-Rahman and Anas (2004), Duranton and Puga (2004), Desmet and Henderson (2015).

$\tau N_i^{1+\gamma}/\gamma$, land rent at r_i is $R_{yi}(r_i) = \tau\left(L_i^{\gamma} - r_i^{\gamma}\right)(1+\gamma)/\gamma$, total land rent is $TLR_{yi} = \tau N_i^{1+\gamma}$, and average urban costs in spatial equilibrium in the city are $R_{yi}(0) = \tau N_i^{\gamma}\,(1+\gamma)/\gamma$. The city's net output Y_i^{Net} is the difference between potential output (1) and commuting costs:

$$Y_{yi}^{Net}(N_i, E_i) = B\, E_i^{\rho(1+\sigma)}\, N_i^{(1-\rho)(1+\sigma)} - \frac{\tau}{\gamma} N_i^{1+\gamma} \quad (4)$$

3 Global pollution

This section derives the social planner allocation (SP) and the allocation chosen by local governments (LG), which we also call 'market equilibrium', starting with the case of pure global pollution. Our analysis yields several new results. First, when pollution is purely global and an optimal emission scheme is implemented, cities are too small under local governments. Whereas this first result holds true in general, i.e. irrespective of the type of commuting costs, our further results are intricately affected by the type of commuting costs, time versus money. In particular, in stark contrast to the first finding, we establish that if commuting costs are in terms of output, cities chosen by local governments become excessively large when pollution is global and a national emission policy is either not in place or not stringent enough. We also show that city sizes chosen by the social planner and local governments negatively depend on the marginal disutility of pollution only if commuting costs are in terms of output. Moreover, a positive fiscal externality that arises if the proceeds from the emission policy are not rebated to city residents leaves the city size chosen by local governments unaffected if commuting costs are in terms of time, but biases cities further down if commuting terms are in terms of output.

3.1 Social planner vs. local governments.

3.1.1 General results. When pollution is purely global, as with global warming, each city resident is faced with the pollution of the total city system, $\Omega_i = n\,E_i$.

The *social planner* chooses city size, local emissions and the number of cities to maximize $U_i = X_i - \eta n E_i$, taking into account that demand in the city system equals supply, $n\,N_i X_i =$

$n\, Y_i^{Net}(N_i, E_i)$, and that the population fits into the cities, $n\, N_i = N$. Hence, $U_i = Y_i^{Net}/N_i - \eta N E_i/N_i$. Rearranging the first order conditions with respect to N_i and E_i we obtain:

$$\frac{dY_i^{Net}}{dN_i}\frac{N_i}{Y_i^{Net}} + \frac{dY_i^{Net}}{dE_i}\frac{E_i}{Y_i^{Net}} = 1 \qquad (5)$$

$$\frac{dY_i^{Net}}{dE_i} = \eta N \qquad (6)$$

Eq. (5) commands that the production elasticity of labor in the city $\varepsilon_{Y_i^{Net},N_i} \equiv \frac{dY_i^{Net}}{dN_i}\frac{N_i}{Y_i^{Net}}$ and the production elasticity of emissions in the city $\varepsilon_{Y_i^{Net},E_i} \equiv \frac{dY_i^{Net}}{dE_i}\frac{E_i}{Y_i^{Net}}$ sum up to unity. Eq. (6) requires the marginal product of emissions at the city level to be equal to the marginal damage inflicted on the total population. The number of cities follows as $n = N/N_i$.

When pollution is purely global, *local governments* take the disutility term $-\eta\Omega_i$ to be a constant.[7] Local governments thus simply choose city size to maximize per capita income net of urban costs of their residents, c_i. Let us assume that the national government implements an emission policy which fixes total emissions nE_i in the city system according to (6) through a permit system (or, equivalently, through an emission tax) and that the revenue of this policy is rebated to local governments who distribute it lump-sum to city residents. Their per capita income net of urban costs is then $c_{ji} = w_{ji} + TET_{ji}/N_i + TLR_{ji}/N_i - R_{ji}(0)$ with $j \in (b, y)$. Substituting the expressions for the wage, the proceeds from the emission scheme, total land rent and average urban costs, it follows that $c_{ji} = Y_{ji}^{Net}/N_i$ for $j \in (b, y)$. Maximizing this with respect to city size yields the condition:

$$\varepsilon_{Y_{ji}^{Net},N_i} = \frac{dY_{ji}^{Net}}{dN_i}\frac{N_i}{Y_{ji}^{Net}} = 1 \qquad (7)$$

[7] Technically, this argument commands that we think of a continuum of cities.

Local governments choose city size such that the marginal product of labor equals the average product of labor in the city, irrespective of the type of commuting costs.

A comparison with the social planner solution yields the first key novel and general result of our analysis: *when pollution is purely global cities are too small under local governments*. Intuitively, in choosing the optimal city size the social planner takes the creation of new cities into account. Condition (5) reveals that the social planner's choice of city size recognizes the productive effect of emissions at the city level. The social planner thereby economizes on the number of cities which local governments do not. This city bias, which arises in our micro-founded city systems model with an endogenous number of cities under global pollution, has not been identified in previous research.

To explore further properties of the respective allocations we need more information concerning commuting costs. The following two subsections establish the fundamental insight that city sizes chosen both by the social planner and local governments are responsive to the parameter governing the marginal disutility from pollution only if commuting costs are in terms of local output, but not if commuting costs are specified in terms of time.

3.1.2 Iceberg commuting costs. When commuting costs are of the iceberg type, socially optimal city size and emissions can be obtained recursively. This is a consequence of the proportionality of output and commuting costs when these are of the ice-berg type. The social planner chooses city size according to (5). This yields $N_{bi}^{SP} = \frac{\sigma}{\tau[(2\sigma+1)-\rho(1+\sigma)]}$ which is increasing in the agglomeration parameter σ and decreasing in the commuting cost parameter τ, a familiar result from the standard model. A novel result that obtains in our setting is that optimal city size is negatively related to ρ. This can be rationalized by noting that the sharing externality is weaker, the higher the cost share of emissions in the production of intermediates. Further, note that optimal city size is not affected by the disutility parameter η. Optimal emissions at the city level follow from (6), $E_{bi}^{SP} = a\,\eta^{-\frac{1}{1-\rho(\sigma+1)}}$ (where $a > 0$ collects constant terms). They are negatively related to the disutility parameter η and the stronger so, the higher is ρ.

As long as the optimal emission policy is in place, these emissions are also realized on the local level. City size implied by (7) is $N_{bi}^{LG} = \frac{\sigma-\rho(1+\sigma)}{\tau[(2\sigma+1)-2\rho(\sigma+1)]}$, which is positive since we imposed $(1-\rho)(1+\sigma) > 1$ in section 2, but falls short of N_{bi}^{SP}. It is easily verified that the qualitative response of N_{bi}^{LG} to changes in σ, τ and ρ is the same as for the optimal solution. Moreover, N_{bi}^{LG} is also not dependent on η.

3.1.3 Commuting costs in terms of local output. The social planner solution can no longer be obtained recursively when commuting costs are in terms of local output. Rather, from (5) and (6) we now have $N_{yi}^{SP} = \left[\frac{B\,\sigma\, E_i^{\rho(1+\sigma)}}{\tau}\right]^{\frac{1}{\gamma-\sigma+\rho(1+\sigma)}}$ and $E_{yi}^{SP} = \left[\frac{B\,\rho(1+\sigma) N_i^{(1-\rho)(1+\sigma)}}{\eta N}\right]^{\frac{1}{1-\rho(1+\sigma)}}$. Solving these two interdependent conditions yields the result that both city size and emissions are negatively related to the marginal disutility from emissions, $N_{yi}^{SP} = f_0\eta^{-f_1}$ and $E_{yi}^{SP} = f_2\eta^{-f_3}$, where $f_0, f_1, f_2, f_3 > 0$ collect constant parameters. Optimal city size has the same properties with respect to the basic parameters B, σ and τ as in the iceberg case (*3.1.2*). A key difference to that case is that the optimal city size is now negatively related to the disutility parameter η (and similarly so, to the total population N). Moreover, since the absolute value of b_1 is increasing in ρ, this impact gets magnified, the larger is ρ. Hence, optimal city size is smaller, the stronger is the weight of emissions in local production. This is crucially different from the case of iceberg commuting costs, where city size is independent from the marginal disutility of pollution.

Turning to local governments, optimal emissions E_{yi}^{SP} are supported by a national permit system (or emission tax). City size implied by (7) is $N_{yi}^{LG} = \left[\frac{B\,[\sigma-\rho(1+\sigma)] {E_{yi}^{SP}}^{\rho(1+\sigma)}}{\tau}\right]^{\frac{1}{\gamma-\sigma+\rho(1+\sigma)}}$ which falls short of the optimal one. The city size chosen by local governments is negatively affected by the marginal disutility parameter η (through E_{yi}^{SP}), just as the social planner solution. This is in stark contrast to the case of iceberg costs, where city sizes are independent of η.

3.2 Political economy considerations

The social planner allocation, i.e. the first-best optimum, could be implemented by a benevolent national government if it had control over city sizes (and, hence, the number of cities) and over emissions, where the latter could be steered through a national permit system (or an optimal emission tax). In practice, as we have stressed in the introduction, it is rather local governments (local councils, strong mayors etc., see Glaeser 2013) which control the size of cities. The national government is responsible for environmental policies, notable those that affect the total city system. If policy levers are divided this way cities become too small under global pollution, as we have seen in section 3.1.

Rather than following benevolent motives, government behavior may be driven by political economy considerations (Glaeser 2013). We consider two such instances and as we now show, it matters strongly in both cases whether commuting costs are in terms of time or money.

(i) First, national governments may implement the optimal emission policy but the proceeds of this policy may not be rebated back to city residents contrary to the assumption that we have maintained so far. Rather, these proceeds might be used for the purposes of government bureaucrats nationally or locally or in other wasteful ways. To keep the analysis simple, we do not model a political economy game, but simply assume that TET_i is not rebated to city residents. A city dweller's income net of urban costs is then $c_{ji} = w_{ji} + TLR_{ji}/N_i - R_{ji}(0)$.

With commuting costs of the iceberg type it follows that $c_{bi} = (1-\rho)\, Y_{bi}^{Net}/N_i$ which clearly falls short of Y_{bi}^{Net}/N_i. Since c_{bi} is proportional to Y_{bi}^{Net}/N_i, its maximization by local governments still yields condition (7), as in section *3.1.2*, however. Surprisingly, despite the existence of a positive fiscal externality – i.e. the central government cashes in part of the local value of production by not rebating the proceeds from the emission scheme –, city size is not scaled down under these circumstances.

Things are different when commuting costs are in terms of local output. Now $c_{yi} = (1-\rho)(Y_i/N_i) - \tau N_i{}^{\gamma}/\gamma < (1-\rho)\, Y_{bi}^{Net}/N_i$. Hence, only potential output is reduced proportionately but not commuting costs, so c_{yi} falls short of being proportionate to Y_{bi}^{Net}/N_i. This implies that the city size chosen by local governments is even lower than N_{yi}^{LG} derived in section *3.1.3*. Hence the fiscal externality matters for city size in this case.

These results provide an important background perspective on the claim that positive fiscal externalities induce local governments to choose cities that are too small (see Albouy et al. 2016). Our analysis within a micro-founded systems city model shows that this claim holds true only, when commuting costs are incurred in local output, but not when they are time costs.

(ii) A second and arguably *even more important instance of political economy* is that national governments may abstain from implementing an optimal environmental policy in the first place, say because of lobby activities on part of producers. Again we refrain from specifying a political economy game to keep things simple. Let's suppose that the firm's rents associated with emissions accrue to local residents. Then income net of urban costs is $c_{ji} = Y_{ji}^{Net}/N_i$ and this becomes the relevant maximand for local governments.

When commuting costs are in terms of time, city size is determined by (7) and given by $N_{bi}^{LG} = \frac{\sigma-\rho(1+\sigma)}{\tau[(2\sigma+1)-2\rho(\sigma+1)]}$ as in section *3.2.2*. In the absence of environmental regulation producers would drive emissions up until their marginal product is zero. In the Cobb-Douglas specification which we have chosen in accordance with Copeland and Taylor (1994; 2003) for analytical ease, this would imply infinite emissions. Rather than using a less tractable technology, we have followed these authors and introduced an exogenous upper bound. This comes into play here.[8] To arrive at emissions, note first that the labor supply in the city is $L_i = N_{bi}^{LG}(1 - \tau N_{bi}^{LG})$. Emissions at the city level then follow according to $E_i = \kappa\, L_i$.

An interesting contrast arises when commuting costs are in terms of local output. The maximization of $c_{yi} = Y_{yi}^{Net}/N_i$ by local governments is also determined by (7) and yields $N_{yi}^{LG} = \left[\frac{B\,[\sigma-\rho(1+\sigma)]E_i{}^{\rho(1+\sigma)}}{\tau}\right]^{\frac{1}{\gamma-\sigma+\rho(1+\sigma)}}$ as in section *3.2.3*. This result reveals that there are now repercussions between emissions and the city size chosen by local governments. Local producers have an incentive to increase emissions when regulation is absent and this feeds back into city size. It follows immediately that cities become too large, $N_{yi}^{LG} > N_{yi}^{SP} = \left[\frac{B\,\sigma\, E_i^{\rho(1+\sigma)}}{\tau}\right]^{\frac{1}{\gamma-\sigma+\rho(1+\sigma)}}$ iff $E_i > \{\sigma/[\sigma - \rho(1+\sigma)]\}^{1/\rho(1+\sigma)}$ which is readily possible if the upper bound on emissions is lax enough.[9] More generally, this line of reasoning clearly shows that city sizes chosen by local governments become too large not only if a national environmental policy is fully absent, but if it is not stringent enough as captured by the mentioned condition.

Hence we have another key result of our analysis: rather than being biased downwards, *cities chosen by local governments may become excessively large when pollution is global and a national emission policy is either not in place or not stringent enough*!

4 Local pollution

This section addresses the case of purely local pollution. A first result derived in the following is that the allocations chosen by the social planner and benevolent local governments coincide. This follows from the fact that, unlike the case of global pollution, the social planner and local governments face the same maximization problem here. This result holds true in general but

[8] Ossa (2011) pursues a similar strategy in his trade policy analysis.

[9] A different way to see this is by using $E_i = \kappa\, L_i$ and $L_i = N_i$ so that $N_{yi}^{LG} = \left[B[\sigma - \rho(1+\sigma)]\kappa^{\rho(1+\sigma)}/\tau\right]^{\frac{1}{\gamma-\sigma}}$.

the type of commuting costs matters in other instances, similarly as in section 3. In particular, when commuting costs are in terms of output, city sizes become excessively large if local governments do not address environmental pollution. Moreover, city size is only affected by the environmental disutility parameter when commuting costs are in terms of output. Finally, the wasteful use of the proceeds from the emission scheme biases cities downward under the latter type of commuting costs. The derivation of these results and their intuition are largely parallel to those in section 3, so we are deliberately brief in the following.

4.1 Social planner vs. benevolent local governments.

When pollution is purely local, $\Omega_i = E_i$. The *social planner* chooses N_i and E_i to maximize $U_i = Y_i^{Net}/N_i - \eta E_i$ where it is taken into account that production matches consumption and that the population fits into the cities. The first order conditions can be brought into the form:

$$\varepsilon_{Y_i^{Net},N_i} = \frac{dY_i^{Net}}{dN_i}\frac{N_i}{Y_i^{Net}} = 1 \tag{8}$$

$$\frac{dY_i^{Net}}{dE_i} = \eta N_i \tag{9}$$

Eq. (8) commands the marginal and the average product of labor in the city to be equal, and eq. (9) requires the marginal product of emissions in the city to equal marginal damage.

Benevolent *local governments* choose city size and emissions to maximize indirect utility $V_{ij} = c_{ij} - \eta\Omega_i,\ j \in (b, y)$. They use a local permit system (or a tax) which fixes $n\ E_i$ to address environmental pollution and rebate the proceeds to city residents on a per capita basis. Per capita income net of urban costs is then $c_{ji} = Y_{ji}^{Net}/N_i$. Hence, local governments face the same problem as the social planner and therefore implement the social optimum, eqs. (8) and (9).

Commuting costs: Time vs. money. With commuting costs of the iceberg-type it follows directly from (8) that $N_{bi} = \frac{(1-\rho)(1+\sigma)-1}{\tau[2(1-\rho)(1+\sigma)-1]} > 0$. This (optimal and equilibrium) city size is increasing in σ, decreasing in τ and negatively related to ρ. Moreover, city size is independent of the marginal disutility parameter η. It follows from (9) that city emissions are negatively related to the marginal disutility of pollution, $E_{bi} = g\, \eta^{\frac{-1}{1-\rho(1+\sigma)}}$, where g collects exogenous parameters. When commuting costs are monetary, city size and emissions can no longer be recursively determined, but follow simultaneously from (8) and (9) as $N_{yi} = h_0(B\, \mu/\tau)^{h_1}\eta^{-h_2}$ and $E_{yi} = h_3\eta^{-h_4}$, where $h_0, h_1, h_2, h_3, h_4 > 0$ collect constant parameters. Crucially then, city size is negatively affected by the marginal disutility in contrast to the case with iceberg costs.

4.2 Political economy considerations

Of course, local governments may also not act in the best interest of city residents. Paralleling the analysis in section 3.2 we consider two instances.

(i) First, local governments may implement the optimal local emission policy but not rebate the proceeds, so that a resident's income net of urban costs is $c_{ji} = w_{ji} + TLR_{ji}/N_i - R_{ji}(0)$. With iceberg commuting costs this implies $c_{bi} = (1-\rho)\, Y_{bi}^{Net}/N_i$, which clearly falls short of Y_{bi}^{Net}/N_i, but is proportional to Y_{bi}^{Net}/N_i, so that its maximisation implies (8), hence, the optimal city size. Things are different when commuting costs are in terms of local output. Here $c_{yi} = (1-\rho)(Y_i/N_i) - \tau N_i{}^{\gamma}/\gamma < (1-\rho)\, Y_{bi}^{Net}/N_i$, so that c_{yi} falls short of being proportionate to Y_{bi}^{Net}/N_i. Hence, city size is scaled down relative to the optimum (similar to 3.2 (*i*)).

(ii) Local governments may not implement an optimal emission policy in the first place. Suppose that firm's rents from emissions accrue to local residents. Then income net of urban costs, $c_{ji} = Y_{ji}^{Net}/N_i$, becomes the relevant maximand. With iceberg commuting costs, $N_i = \sigma/[2\sigma+1]\,\tau$, so that cities are still at their optimal size. The labor supply in the city is $L_i = N_i(1-\tau N_i)$ and emissions follow from $E_i = \kappa\, L_i$, i.e. the upper bound binds (as in section 3.2). If, alternatively, commuting costs are in terms of local output, the maximization of net urban income with respect to city size implies $N_{yi} = \left[\, [\sigma - \rho(1+\sigma)]\kappa^{\rho(1+\sigma)}/\tau\right]^{\frac{1}{\gamma-\sigma}}$, i.e. the upper bound for emissions applies.

Hence, the size of cities is driven up in this case, restricted only by the upper bound κ which may be arbitrarily large (similarly as in section 3.2).

5 Conclusion

Are cities too small or too big in the face of environmental pollution? A popular current line of thought views cities as rather being too small than too big, in particular in countries such as China, but also in the USA (e.g. Au and Henderson 2006; Desmet and Rossi-Hansberg 2011; Albouy et al. 2016).

This paper is the first one which explicitly addresses this question within a model of city systems in the tradition of Henderson (1974) where the number of cities is endogenous. For that end, a version of the canonical model of city systems is developed with micro-foundations in terms of input sharing such that the production of local intermediates goes along with emissions in addition to labor. Emissions have a negative impact on the welfare of consumers, locally and globally for the entire city system.

Our analysis delivers two key insights. First, if optimal schemes to regulate environmental pollution are implemented, cities chosen by local governments are never too large. They are too small if pollution is purely global, but at the optimal size, if pollution is purely local. Second, unless we can be sure that we get our environmental policies right, i.e. unless optimal emission schemes are implemented, the size of cities, if steered by local governments, are too large, however.

There are several avenues for future research. The framework established here could be used to address Ricardian differences in land, the heterogeneity of agents (consumer-workers) and the competitive choice of environmental policies. Moreover, the model could be extended to a dynamic setting in future research.

References

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

Abdel-Rahman, H.M. and M. Fujita (1990). Product Variety, Marshallian Externalities, and City Sizes, Journal of Regional Science 30:2, 165-183

Albouy, D., K. Behrens, Fr. Robert-Nicoud and N. Seegert (2016). The Optimal Distribution of Population Across Cities, NBER Working Paper 22823.

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

Becker, R. and J.V. Henderson (2000). Intra-Industry Specialization and Urban Development. In: Huriot, J.M. and J. Thisse (Eds.), The Economics of Cities: Theoretical Perspectives. Cambridge University Press, Cambridge, 138-166

Behrens, K. and Fr. Robert-Nicoud (2015). Agglomeration Theory with Heterogeneous Agents, in: Handbook of Regional and Urban Economics 5A, G. Duranton, J.V. Henderson and W. C. Strange (eds.), 171–245.

Borck, R. and M. Pflüger (2017). Green Cities? Urbanization, Trade and the Environment. Revised version of IZA Discussion Papers 9104, Institute for the Study of Labor (IZA).

Borck, R. and T. Tabuchi (2016). Pollution and city size: can cities be too small? CESifo Working Paper No. 6152.

Combes, P.P., G. Duranton and L. Gobillon (2016). The Costs of Agglomeration: House and Land Prices in French Cities. CEPR Discussion Paper No. 9240, published December 2012, revised June 2016, Centre for Economic Policy Research, London.

Copeland, B. R. and Taylor, M. S. (1994). North-South Trade and the Environment. Quarterly Journal of Economics 109 (3), 755-787.

Copeland, B.R. and Taylor, M.S. (2003). Trade and the Environment. Theory and Evidence. Princeton University Press.

Desmet, K. and J. V. Henderson (2015). The Geography of Development Within Countries, in: Handbook of Regional and Urban Economics 5A, G. Duranton, J.V. Henderson and W. C. Strange (eds.), 1457–1517.

Desmet, K. and E. Rossi-Hansberg (2011). Are the world's megacities too big? VOX EU, March 12.

Desmet, K. and E. Rossi-Hansberg (2013). Urban Accounting and Welfare. American Economic Review 103(6), 2296-2327.

Desmet, K. and E. Rossi-Hansberg (2015). On the Spatial Economic Impact of Global Warming. Journal of Urban Economics 88, 16-37

Duranton, G. and D. Puga (2004). Micro-foundations of Urban Agglomeration Economies, Handbook of Regional and Urban Economics 4, J.V. Henderson and J.-F. Thisse (eds.), Amsterdam: Elsevier, 2063–2117.

Duranton, G. and D. Puga (2014). The Growth of Cities, Handbook of Economic Growth 2, P. Aghion and S. Durlauf (eds.), Amsterdam: Elsevier, 468–560.

Duranton G. and D. Puga (2017). Urban Growth and its Aggregate Implications. In progress, mimeo, University of Pennsylvania and CEMFI.

Ethier, W. (1982). National and International Returns to Scale in the Modern Theory of International Trade, American Economic Review 72, 389–405.

Gaigné, C., S. Riou, and J.-F. Thisse (2012). Are Compact Cities Environmentally Friendly? Journal of Urban Economics 72(2-3): 123–136.

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

Glaeser, E. (2013). Urban Political Economy, in: A.J. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds.), Handbook of Public Economics 5, 195-256

Henderson, J.V. (1974). The Sizes and Types of Cities, American Economic Review 64, 640–656.

Hsieh, C.-T. and E. Moretti (2017). Housing Constraints and Spatial Misallocation. Mimeo.

Kahn, M. E. (2006). Green Cities: Urban Growth and the Environment. Brookings Institution Press, Washington, D.C.

Kahn, M. E. and R. Walsh (2015). Cities and the Environment, in: G. Duranton, J.V. Henderson, W.C. Strange (eds.), Handbook of Regional and Urban Economics 5, 405-465

Krugman, P. and A.J. Venables (1995). Globalization and the Inequality of Nations. Quarterly Journal of Economics 110: 857-80

Kyriakopoulou, E. and A. Xepapadeas (2016). Atmospheric Pollution in Rapidly Growing Industrial Cities: Spatial Policies and Land Use Patterns. Journal of Economic Geography, 1-28

Ossa, R. (2011). A "New Trade" Theory of GATT/WTO Negotiations. Journal of Political Economy 119:1, 122-152

Pflüger, M. and T. Tabuchi (2011). The Size of Regions with Land Use for Production. Regional Science and Urban Economics 40: 481-489.

Schindler, M., G. Caruso and P. Picard, P (2017). Equilibrium and first-best city with endogenous exposure to air pollution from traffic. Regional Science and Urban Economics 62: 12-23.