Working Papers

www.cesifo.org/wp

Optimal Energy Taxation in Cities

Rainald Borck
Jan K. Brueckner

CESIFO WORKING PAPER NO. 5711
CATEGORY 10: ENERGY AND CLIMATE ECONOMICS
JANUARY 2016

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 5711*

Optimal Energy Taxation in Cities

Abstract

This paper presents the first investigation of the effects of optimal energy taxation in an urban spatial setting. Rather than exploring the effects of a carbon tax, our approach is to derive the supplements to existing taxes that are needed to support the social optimum. We then analyze the effects of these taxes on urban spatial structure. Emissions are generated by housing consumption and commute trips, and the optimal tax structure has a tax on commuting, housing floor space, and land. These taxes reduce the extent of commuting and the level of housing consumption while increasing building heights, generating a more-compact city with a lower level of emissions per capita.

JEL-codes: Q580.

Keywords: emissions, energy, taxation, monocentric city.

Rainald Borck
University of Potsdam
Faculty of Economic and Social Sciences
August-Bebel-Str. 89
Germany – 14482 Potsdam
rainald.borck@uni-potsdam.de

Jan K. Brueckner
Department of Economics
University of California, Irvine
3151 Social Science Plaza
USA – Irvine, CA 92697
jkbrueck@uci.edu

January 2016

Optimal Energy Taxation in Cities

by

Rainald Borck and Jan K. Brueckner*

1. Introduction

In step with growing concerns about the impact of global warning, urban research has increasingly focused on the energy consumption of cities. This research reflects the recognition that residential and commercial land-uses are important generators of greenhouse gas (GHG) and particulate emissions along with the transportation and industrial sectors. Their importance is seen in Table 1, which shows energy use by sector, with emissions from electricity generation "distributed" according to the final users of the electricity. As can be seen, when their electricity use is taken into account via the distribution method, the residential and commercial sectors each account for an appreciable 16.9% of total emissions, with their 33.8% total exceeding the shares of industry and transportation. Therefore, economic analysis of policies designed to control GHG emissions should ideally include these two real-estate sectors in its focus along with other sources.

In advancing this goal, some researchers have studied the relationship between a building's energy use and its structural characteristics, with notable contributions by Costa and Kahn (2011), Chong (2012) (who also draws a link to climate), and Kahn, Kok and Quigley (2014). Using a hedonic approach, Eichholtz, Kok and Quigley (2010) ask whether the market values green buildings, finding that energy-efficient commercial structures indeed command higher rents. Glaeser and Kahn (2010) extend the focus beyond residential energy use to include emissions from driving and public transit, generating a ranking of US cities according to their overall carbon footprints. Zheng, Wang, Glaeser and Kahn (2011) extend this approach to Chinese cities.

In parallel with these empirical efforts, other researchers have imbedded energy usage into the familiar monocentric-city model of urban economics, with the ultimate goal of appraising the effect of urban policy interventions on GHG emissions. These studies, which rely on numerical simulations of realistically calibrated urban models, include Larson, Liu and Yezer's (2012) study evaluating the energy-use impacts of higher gasoline taxes, better vehicle fuel efficiency, urban greenbelts, and housing density restrictions. Larson and Yezer (2015) ask how such policy impacts vary with city size, while Borck (2015) explores the effect of building-height limits on urban GHG emissions. Tscharaktschiew and Hirte (2010) study the impact of emissions taxes and congestion on emissions from commuting.

Although this second group of studies has greatly increased our understanding of the links between urban spatial structure, energy use, and GHG emissions, an important set of questions remains unanswered. In particular, no study has analyzed *optimal* urban form when housing and commute trips generate GHG emissions. Moreover, while a carbon tax offers the simplest way to regulate GHG emissions, no study has asked whether the levels of *existing urban taxes* could be altered so as to achieve the same optimal outcome.[1] The present paper remedies both these omissions. We add energy use and GHG emissions to the housing sector of the standard urban model, doing so in a novel and realistic fashion, while also recognizing the GHG emissions from commuting. With emissions assumed to reduce consumer utilities, the analysis then develops the conditions that characterize the optimal city, which embody a trade-off between the environmental gains from lower emissions and the losses from achieving them. The form of these optimality conditions reveals how existing urban taxes (real estate taxes and gasoline taxes) can be used to generate the optimum. With this theoretical foundation, numerical simulation analysis then derives the changes in urban form that follow from imposition of the optimal taxes. The simulation results thus show how urban spatial structure responds to optimal energy taxation.

More specifically, the model relies on principles from the engineering and architecture literatures by assuming that residential energy use from heating and cooling depends on a building's exposed surface area, reflecting heat transfer through exposed surfaces. According to Ching and Shapiro (2014), a building's energy use per square foot of floor space is proportional to its surface area per square foot of floor space, with surface area including the sides of the building along with the roof. Since the roof area stays constant as the height of the building increases, surface area increases less rapidly than floor space as height grows. The result is energy economies from building height, with energy use per square foot of floor space falling as height increases, a pattern seen in the empirical results of Larson et al. (2012).

If a building's total energy use (and hence its total GHG emissions) just depended on its square footage of floor space, the appropriate residential energy tax would just be a tax per square foot of space. But with surface area mattering instead, the analysis shows that residential energy taxes should include a tax per square foot of floor space *along with a tax on the building's footprint*, which captures energy usage that depends on the roof area (equal to the footprint). When buildings completely cover the land, as in the standard urban model, the footprint tax is just a tax on the entire land input, hence a land tax. By raising the land cost to the developer, this land tax encourages the construction of energy-efficient tall buildings. Note that the land tax adds to the tax burden on land already inherent in the housing tax.[2]

In addition to these taxes on residential land-use, the model prescribes a commuting tax per mile to address GHG emissions per mile driven. This prescription emerges from a model without traffic congestion, in contrast to the work of Larson et al. (2012) and Larson and Yezer (2014), where congestion is realistically modeled.

We calibrate the model in the most realistic possible fashion and then use it to predict the impact on urban spatial structure from imposing the optimal taxes on floor space, land, and commuting. The numerical results thus allow a comparison of the optimal city, where the GHG externality is addressed, to the laissez-faire city, where no intervention is undertaken. With emissions generated by housing consumption and commuting, the expectation is that optimal energy taxation will reduce the levels of both activities, leading to a city that is more spatially compact than a city without such taxes. The simulations show whether this broad conjecture is confirmed while illustrating the details of the city's adjustment to taxation.

In addition, by relying on three separate taxes rather than a carbon tax to generate the social optimum, we are able to ask some interesting second-best questions and provide numerical answers. In particular, we set one or two of the taxes equal to zero and find the optimal value(s) of the remaining tax(es), showing the characteristics of the resulting second-best city. In one exercise, for example, the optimal value of the commuting tax is computed when the housing and land taxes are constrained to be zero, with the effects on urban structure then shown. In this case, the commuting tax has to do the work of three separate taxes in restraining emissions, with suboptimal results.

As is well known, welfare analysis in urban economics is best carried out by focusing on a fully-closed city, where resources leakages are absent (see Pines and Sadka (1986)). The rental income from land in such a city accrues to the residents rather than leaking to absentee owners, and tax revenue is also redistributed to the residents in lump sum fashion, eliminating another potential leakage. The city simulated in the paper has both these features.

One of the paper's innovations is its modeling of energy economies from building height, and this feature's connection to previous work should be noted. The models of Larson et al. (2012) and Larson and Yezer (2014) include a similar feature, although in a discrete fashion. In particular, energy use per square foot is assumed to decrease discontinuously as building height passes through several discrete critical points, in contrast to the present continuous formulation. The model of Borck (2015), by contrast, includes no energy benefits from tall buildings. His exercise of imposing building-height limits therefore generates no sacrifice on this dimension, but the resulting supply restriction, by raising the price of floor space throughout the city, reduces residential emissions by shrinking individual dwelling sizes. The urban sprawl created by height limits, however, has an offsetting effect on emissions from commuting. The present paper borrows from Borck's (2015) approach while incorporating height economies.

The plan of the paper is as follows. Section 2 presents the theoretical analysis. Section 3 explains the setup of the simulation model, and Section 4 presents the simulation results. Section 5 offers conclusions.

2. Model

2.1. The setup

The model is based on the standard model of a monocentric city, adapted to include energy use. In addition, the surface area of buildings, previously not an issue in urban modeling, plays a prominent role, as explained above. To incorporate surface area, suppose that buildings are square, occupying a land area of ℓ and completely covering that land, as in the standard urban model.[3] Structural density (capital per unit of land) is S and floor space per unit of land is given by $h(S)$. The h function is the intensive form of a constant-returns floor-space production function, and it satisfies $h' > 0$ and $h'' < 0$. Since floor space per unit of land is the most natural index of building height, $h(S)$ can also be viewed as the height of the building. Therefore, each of the four sides of the building has area $h(S)\sqrt{\ell}$ (height $\times$ width), and the area of the roof is ℓ. Surface area is then

$$4h(S)\sqrt{\ell} \ + \ \ell. \tag{1}$$

Letting e denote energy use per unit of surface area, the building's energy use is e times (1). Energy use per unit of land is then given by

$$\frac{(4h(S)\sqrt{\ell} \ + \ \ell)e}{\ell} \quad = \quad \frac{4h(S)e}{\sqrt{\ell}} \ + \ e. \tag{2}$$

The second term is energy use per unit of land due to heat transfer through the building's roof, while the first term captures heat transfer through the sides. It is clear from (2) that a building occupying more land has greater energy efficiency per unit of land, which would prompt the developer to increase ℓ, an incentive that is absent in the standard urban model (where ℓ is matter of indifference).[4] To abstract from this issue, we fix ℓ, and for convenience set the value at 16 by choice of units of measurement, so that energy use per unit of land becomes

$$h(S)e \ + \ e. \tag{3}$$

Again, the last term captures energy use due to heat transfer through the roof (whose area matches the lot size), while the first term captures energy use from heat transfer through the building's sides, a transfer that is proportional to the floor space it contains. Note that the presence of the additive e term in (3) means that energy use increases less rapidly than floor space, implying energy economies from building height. Equivalently, dividing (3) by square footage ($h(S)$) shows that energy use per square foot is $e \ + \ e/h(S)$, an expression that is smaller in a taller building. Each unit of residential energy use generates ψ units of emissions, so that (using (3)) residential emissions per unit of land are given by $\psi(h(S)e \ + \ e)$.

Let the cost per unit of energy be normalized to unity, and assume that the developer bears the building's energy cost. In addition, let p denote the price per square foot of housing, r denote rent per unit of land, and i denote the price per unit of capital. Then, using (3), the developer's profit per unit of land is

$$ph(S) \ - \ iS \ - \ r \ - \ \textit{energy cost per unit of land} \ = \ (p-e)h(S) \ - \ iS \ - \ e \ - \ r. \quad (4)$$

In the absence of taxes, the developer would choose S to satisfy

$$(p-e)h'(S) \ = \ i, \quad (5)$$

and land rent r would be determined by the zero-profit condition:

$$r \ = \ (p-e)h(S) \ - \ iS \ - \ e. \quad (6)$$

The form of both conditions is familiar from the standard urban model (see Brueckner (1987)).

Energy is also used as workers commute to the CBD. Let the cost per mile of commuting (on a round-trip basis) be denoted t, so that commuting from a residence x miles from the CBD costs tx per period. The parameter t includes private energy costs, as reflected in the cost of fuel. Suppose that emissions per round-trip mile of commuting are given by γ, so that the energy used in commuting from a distance x generates γx worth of emissions.

Several other sources of residential energy use have been omitted from the model: kitchen appliances, such as refridgerators and stoves, and hot-water heaters. These sources can be viewed as generating a fixed amount of energy use that does not increase proportionally with the physical size of the dwelling.[5] This fixed usage presumably accounts for empirical findings showing that residential energy use per square foot of floor space falls as dwelling size rises (see, for example, Larson, Liu and Yezer (2012)). Since a city's aggregate energy use from household appliances will thus be roughly proportional to the number of dwellings but unaffected by urban form, we omit it from the analysis.

2.2. Emissions and energy taxes

Let μ denote the social damage from each unit of emissions. Then, the taxes needed to support the social optimum can be derived from the model. These taxes are as follows:

- A tax of $\tau_q = \mu\psi e$ per square foot of floor space, addressing emissions from energy use due to heat transfer through the sides of a building
- A tax of $\tau_\ell = \mu\psi e = \tau_q$ per unit of land, addressing emissions from energy use due to heat transfer through a building's roof
- A tax of $\tau_t = \mu\gamma$ per mile of commuting, addressing emissions due to energy use in commuting

To demonstrate the need for these taxes analytically, let utility be given by $v(c, q, G)$, where c is consumption of a nonhousing good, q is consumption of housing floor space, and G gives the level of emissions affecting the city's residents. Two equivalent approaches to the planning problem are possible, following the past literature. Under the first approach, the planner minimizes the city's resource consumption, subject to several constraints: achievement of a fixed utility level u for its residents, the requirement that the city fits its population, and a condition giving total emissions G.[6] Under the second approach, which is dual to the first, the planner maximizes the common utility level of urban residents subject to a resource constraint, the population constraint, and the G condition. Since the first approach is somewhat simpler, the present analysis follows it.

To start, observe that the fixed-utility constraint, which can be written $v(c, q, G) = u$ for some constant u, implies $c = c(q, G)$, with the derivatives of this function equaling minus the marginal rates of substitution: $c_q = -v_q/v_c < 0$ and $c_G = -v_G/v_c > 0$ given $v_G < 0$ (subscripts denote partial derivatives). Using the $c(q, G)$ function and letting $\overline{x}$ denote the distance to the city's edge and r_a denote the opportunity cost of land (agricultural rent), the city's resource consumption is given by

$$\int_0^{\overline{x}} 2\pi x \left[iS \; + \; \frac{h(S)}{q}(c(q, G) + tx) \; + \; h(S)e \; + \; e \; + \; r_a \right] dx. \tag{7}$$

In (7), the choice variables S and q are implicitly functions of x. The first term in the integrand captures capital usage, the second term equals c consumption plus commuting cost per person at distance x multiplied by the population at x. That population equals the area $2\pi x dx$ of the ring at x times h/q, where h/q gives population density (housing square footage per unit of land divided by square feet per dwelling). The remaining terms capture building energy use and the opportunity cost of land.

Letting L denote the city's fixed population, the population constraint is written

$$\int_0^{\overline{x}} 2\pi x \frac{h(S)}{q} dx \quad = \quad L, \tag{8}$$

and the multiplier associated with this constraint is λ. Total emissions G satisfy

$$\int_0^{\overline{x}} 2\pi x \left(\psi[h(S)e + e] \ + \ \gamma \frac{h(S)}{q} x \right) dx \quad = \quad G, \tag{9}$$

where $2\pi x(xh(S)/q)dx$ gives total commute miles for consumers living in the ring at x. The multiplier associated with this constraint is $\mu > 0$, the social damage from an extra unit of emissions. While the formulation based on eq. (9) is most easily interpreted under the assumption that emissions affecting the city are all locally generated, the model can also incorporate the intercity pollution externalities underlying the global-warming process with some simple redefinitions.[7]

The planner chooses values of G and $\overline{x}$ and values of S and q at each distance to minimize (7) subject to (8) and (9). After forming a Lagrangean expression using (7)–(9), the optimality conditions for S and q are generated by differentiating inside the integrals, while the condition for $\overline{x}$ comes from differentiating with respect to the limits of integration. After a modest amount of manipulation (see the Appendix), these conditions reduce to equations that identify the taxes required to support the optimum. The first equation is

$$c(q, G) \ + \ q\frac{v_q}{v_c} \ + \ (t + \mu\gamma)x \quad = \quad -\lambda, \tag{10}$$

Recognizing that the consumer will set v_q/v_c equal to the price p per square foot of housing, the first two terms correspond to total consumption expenditure in a decentralized equilibrium. The tx term is the money cost of commuting, but (10) shows that this cost must be supplemented by a tax of $\mu\gamma \equiv \tau_t$ per mile traveled, as in the third bullet point above. The term $-\lambda$ is constant over x and corresponds to the common income of consumers in a decentralized equilibrium. Note that, ignoring differences in automobile fuel efficiency, the commuting tax has the same form as the gasoline taxes levied throughout the world.

The next condition is

$$\left(\frac{v_q}{v_c} - e - \mu\psi e\right) h'(S) \quad = \quad i. \tag{11}$$

Comparing to the profit-maximization condition (5) and recognizing $v_q/v_c = p$, (11) implies that the net price received by the developer per unit of floor space should be reduced below $p - e$ by the amount $\mu\psi e \equiv \tau_q$, an emissions tax per square foot of floor space (as in the first bullet point above).

The laissez-faire equilibrium condition determining the distance $\overline{x}$ to the edge of the city would set r evaluated at $\overline{x}$ equal to r_a, and using (6), this condition is written

$$(\overline{p} - e)h(\overline{S}) \ - \ i\overline{S} \ - \ e \quad = \quad r_a, \tag{12}$$

where $\overline{p}$ and $\overline{S}$ are the p and S values at $\overline{x}$. By constrast, the optimality condition for $\overline{x}$ reduces to

$$\left(\frac{\overline{v_q}}{\overline{v_c}} - e - \mu\psi e\right) h(\overline{S}) \ - \ i\overline{S} \ - \ e \ - \ \mu\psi e \quad = \quad r_a, \tag{13}$$

where v_q/v_c is also evaluated at $\overline{x}$. Comparing (12) and (13) indicates that, in addition to the tax of $\tau_q = \mu\psi e$ per square foot of housing floor space, a tax per unit of land equal to $\mu\psi e \equiv \tau_\ell = \tau_q$ is also needed, which reduces land rent by that amount (as in the second bullet point above). With these two taxes subtracted in the equilibrium condition, it then corresponds to the optimality condition.

Note that the housing tax corresponds to a standard property tax (levied, however, as an excise tax instead of an ad-valorem tax), while the land tax matches taxes of this type levied in some cities (in excise not ad-valorem form, however). Observe also that the property tax, if levied in ad-valorem fashion, is equivalent to separate ad-valorem taxes levied at a common rate on land and housing capital (see Brueckner and Kim (2003)). An additional ad-valorem land tax on land would add to the tax burden, with the combined taxes equivalent to a split-rate tax structure that taxes land and capital at different rates (see Oates and Schwab (1997)). While the excise form of the current housing and land taxes disrupts this simplicity, it remains true that the land tax adds to the tax burden on land already present in the housing tax.

Recall that the multiplier μ appearing in the tax terms equals the marginal social damage from emissions. From the first-order condition for G, the multiplier equals

$$\mu \quad = \quad -\int_0^{\overline{x}} 2\pi x \frac{h(S)}{q}\frac{v_G}{v_c}dx \quad > \quad 0. \tag{14}$$

The integral weights the MRS between G and c by population and sums across distance to yield the social damage from an extra unit of G.[8]

It is important to note that, because the planning problem portrays a city where the cost of land is the agricultural opportunity cost and where taxes are absent from the objective function, the corresponding decentralized city must have several features. First, the rental income generated in the city must accrue to its residents. In particular, the city must be "fully closed" in the sense of Pines and Sadka (1986), with differential land rent (the amount in excess of r_a) earned as income by the residents. The residents are thus viewed as owning a corporation that acquires the city's entire land area from its outside owners at a rental price r_a, with the land then rented to the residents themselves in a competitive market. The residents thus share the aggregate rental income net of r_a generated by the city, in effect paying rent to themselves. Second, since tax revenue is absent from the planning problem, the revenue from the energy taxes must be redistributed to the residents on an equal per capita basis. With these two requirements, the differential rent and tax revenue generated in the city stays within it, as envisioned in the planning problem. The ensuing numerical analysis imposes both requirements.

2.3. Predicting the impact of energy taxation

A main goal of the numerical analysis presented in section 3 is to illustrate the impact on the spatial structure of the city from levying optimal energy taxes. In principle, these effects might be predictable in advance through an appropriate comparative-static analysis, relying on Pines and Sadka's (1986) extension of Wheaton's (1974) comparative-static analysis to the present context of a fully closed city.

Unfortunately, however, the required comparative statics cannot be inferred from Pines and Sadka's results. Imposition of the land tax, for example, can be viewed as equivalent to an increase in the agricultural rent r_a, which Pines and Sadka (1986) analyze. However, the present tax change corresponds to an increase in r_a combined with an increase in income equal to the rebated per capita land-tax revenue, whose impact cannot be inferred from the results they present. A similar point applies to the effects of the commuting tax. Moreover, as mentioned above, the tax on housing square footage is similar to a standard property tax, whose effects are analyzed by Brueckner and Kim (2003). While they show that the property tax causes the city to shrink spatially when the elasticity of substitution between housing and c does not exceed unity (as under the Cobb-Douglas preferences imposed below), Brueckner and Kim's model is not fully closed, nor does it incorporate redistribution of tax revenues.

The previous literature thus cannot be used directly to predict the separate effects of the three taxes in the current model, and the need to predict their *combined effects* makes the prediction task even more daunting. Hence, in the next section, we present results from numerical simulations.

3. Simulation Setup

3.1. Preliminaries

To evaluate the effect of imposing the optimal energy taxes, we numerically compare the urban equilibrium without any taxes to the equilibrium where the taxes are imposed, relying on the optimal tax formulas. The approach thus shifts from the orientation of the planner, whose goal was to minimize the city's resource consumption, to analysis of equilibria, knowing that an equilibrium where taxes are imposed according to the optimal formulas is efficient.

We impose specific functional forms in order to simulate the model numerically. Parameters are taken partly from published sources and partly calibrated to replicate key features of American cities. The utility function is assumed to take the following form:

$$v(c, q, G) \;\equiv\; (1-\alpha)\log c \;+\; \alpha \log q \;-\; \nu G, \tag{15}$$

where $0 < \alpha < 1$ and $\nu > 0$ is the marginal damage from emissions. While the first two terms correspond to standard Cobb-Douglas preferences over c and q, total emissions G appears in the third linear term, which is not logged. Using data from US metropolitan areas (MSAs), Davis and Ortalo-Magné (2011) show that the expenditure share of housing is remarkably constant across MSAs and over time, so that the Cobb-Douglas assumption seems justified. They estimate an average expenditure share of housing of $\alpha = 0.24$, which is the value we use. The parameter ν is set at 0.05. This value is chosen to generate a plausible reduction in total emissions in moving to the first best (about 1/3).

In the consumer budget constraint, income is set equal to the 2011 US value of median household income, equal to \$51,324. Commuting costs per mile are made up of monetary and time costs of commuting and are set at $t = \$521.77$ per mile per year (see the Appendix for details). City population is set at $L = 750,000$ households.[9]

As in Bertaud and Brueckner (2005), housing production is assumed to be Cobb-Douglas, which yields the intensive production function $h(S) = \rho S^{\beta}$, where $\beta < 1$. Ahlfeldt and McMillen (2014) use data from several cities to estimate the elasticity of substitution between land and capital and find that it is close to one, which supports the Cobb-Douglas assumption. In the simulation, we set $\rho = 0.00005$ and $\beta = 0.745$. Agricultural land rent r_a is set at \$58,800 per square mile (see the Appendix).

Based on calculations in the Appendix, we set γ, annual CO_2 emissions per mile of commuting, at 147.521 kg. To get this value, total commuting distance is multiplied by 500 (to get round-trips on 250 workdays) and then by a conversion factor that converts person-miles of commuting (assuming an average mix of commuting modes) into CO_2 equivalents.

As explained in the appendix, we set emissions ψ per kilowatt-hour of residential energy use at 0.2964. This ψ value is then multipled by energy use e per square foot of building surface area to get emissions per square foot of surface area. The computation of e is unfamiliar and closely tied to the model, and it proceeds as follows. For a square k-story building with floor space of q, surface area is $A_k = 4kH\sqrt{q/k} + q/k$, where H is the height of one story (the first term is the area of the sides and the second the area of the roof, assumed flat). The Residential Energy Consumption Survey (RECS) provides q values for detached single-family homes of 1, 2, and 3 stories, and assuming $H = 12$ feet, the previous formula can be used to compute $A_k,\ k = 1, 2, 3$. The resulting median surface-area value across the three types of houses is given by $A = 8,458.82$. Next, we use the RECS to get median energy use for space heating and air conditioning across all detached single-family houses (with different numbers of stories), which equals 60,444 thousand BTUs or 17,714.38 kwh. We associate this median value with the median single-family surface area A, which allows us to divide 17,714.38 kwh by $A = 8,458,82$ to get a value for energy use per square foot of surface area. This value equals $e = 1.985$ kwh, and it can then be applied to buildings of all heights.[10]

Although the exposition of the model assumes for simplicity that all land is used for housing, we assume in the simulations that a fraction 0.75 of each annulus is available for residential use. Finally, it should be noted that, while the chosen value of t reflects the existing level of gasoline taxes, the model and its calibration does not reflect existing real estate taxes, which are in effect assumed to be zero. For the property tax, this assumption is appropriate given that the distortionary nature of tax means that it would not be present at the social optimum. By contrast, since the gasoline tax component of t can be viewed as a nondistortionary road user fee, its presence in the calibrated model is not inconsistent with efficiency.

3.2. Solution procedure

In analysis of the standard urban model, the consumer maximizes utility, and the requirement that the maximized value equals a spatially invariant utility level u then determines the housing price p as a function of u and the other variables of the model (see Brueckner (1987)). This function is found by substituting the budget constraint $c + pq = y - tx$ (y is income) into $v(c, q)$ from (15) with G suppressed, yielding $v(y - tx - pq, q)$. The first-order condition $v_q/v_c = p$ for choice of q is then solved simultaneously with the condition $v(c, q) = u$

to determine p and q as functions of the remaining variables. The resulting p solution, denoted $p(x, y, t, u)$, gives the price per square foot of housing that allows a utility-maximizing consumer living at distance x and earning income y to reach utility u. Under the Cobb-Douglas preferences in (15) (with $\nu = 0$), $p(x, y, t, u) \equiv B(y - tx)^{1/\alpha} u^{-1/\alpha}$, where B is a constant. The function $q(x, y, t, u)$ gives the solution for housing consumption as a function of these same variables.

With p determined, the housing developer's profit maximization problem then yields the conditions (5) and (6), with $e = 0$ holding in the standard model. These conditions give solutions for S and r as functions of the same variables that determine p: $S(x, y, t, u)$ and $r(x, y, t, u)$. Finally, population density, given by $D = h(S)/q$, also can be written in the same fashion making use of the solutions for S and q: $D(x, y, t, u)$.

The arguments of these functions are modified in the current framework. Since u in the standard model comes from consumer maximization over just c and q, the utility argument must be replaced by $u + \nu G$ in the current framework (see (15)). In addition, commuting cost per mile t is replaced by $t + \tau_t$ to capture the tax on commuting. Finally, letting R denote total differential land rent and T denote total tax revenue, the income y must be replaced by $y + (R + T)/L$ to capture redistribution of equal per capita shares of total differential land rent and taxes.[11] Therefore, p is now written as $p(x,\ y + (R + T)/L,\ t + \tau_t,\ u + \nu G)$.

In addition, the S and r functions now depend on this same new list of arguments along with e and τ_q. These dependencies of S can be seen in (11), where $\mu\psi e = \tau_q$ and the MRS expression is replaced by the modified p function. The S that satisfies the equation then depends on the arguments of p and on e and τ_q. The r dependencies can be seen from the LHS of (13). Land rent r is given by the LHS expression in (13) with p in place of the MRS and the bars removed, so that r depends on the arguments of p along with e and τ_q.

To solve the model, the first step is to set land rent at $\overline{x}$ equal to agricultural rent r_a plus the land tax τ_ℓ, with the condition written as

$$r(\overline{x},\ y + (R + T)/L,\ t + \tau_t,\ u + \nu G,\ e,\ \tau_q) \quad = \quad r_a + \tau_\ell. \tag{16}$$

This condition is used to solve for utility u as a function of the remaining variables (which include $\overline{x}$ and r_a). The u solution is then substituted back into the r function and into the S function and the D function, which has the same arguments as r and S. When this substitution is made, G drops out as a determinant of r, S and D, but all three variables now depend on $\overline{x}$ and $r_a + \tau_\ell$. Letting the new functions giving this dependence be denoted $\widehat{r}$, $\widehat{S}$ and $\widehat{D}$, the set of equilibrium conditions that need to be solved can be written down.

The condition analogous to (8) stating that the city fits its population is written

$$\int_0^{\overline{x}} 2\pi x \ \widehat{D}(x,\ y + (R+T)/L,\ t+\tau_t,\ e,\ \tau_q,\ \overline{x},\ r_a + \tau_\ell)dx \quad = \quad L. \tag{17}$$

The condition stating that differential land rent integrates to R is

$$R \quad = \quad \int_0^{\overline{x}} 2\pi x \ [\widehat{r}(\cdot) - r_a]dx. \tag{18}$$

where the arguments of $\widehat{r}$ are suppressed. Note that R appears on both sides of this condition. The condition giving total tax revenue is

$$T \quad = \quad \tau_t \int_0^{\overline{x}} 2\pi x \widehat{D}(\cdot) x dx \ + \ \tau_q \int_0^{\overline{x}} 2\pi x h(\widehat{S}(\cdot))dx \ + \ \tau_\ell \int_0^{\overline{x}} 2\pi x dx, \tag{19}$$

where the arguments of $\widehat{S}$ and $\widehat{D}$ are suppressed. Note that, since T appears in these arguments, it is present on both sides of this condition.

Finally, using (14), the condition giving μ, the marginal social damage from emissions, is given by

$$\mu \quad = \quad -\int_0^{\overline{x}} 2\pi x \widehat{D}(\cdot)\widehat{MRS}(\cdot)dx, \tag{20}$$

where $\widehat{MRS}(\cdot)$ is the function corresponding to $v_G/v_c = -\nu/v_c$, which has the same list of arguments as $\widehat{D}$.

Since all the taxes depend on μ (recall $\tau_t = \mu\gamma$ and $\tau_q = \tau_\ell = \mu\psi e$), the taxes that appear in (17)–(19) are endogenous, depending on the endogenous μ from (20). But since (20) also involves the endogenous variables R, T, $\overline{x}$ (with μ also appearing on the RHS via the taxes), the entire set of conditions (17)–(20) constitutes a simultaneous equation system that determines solutions for the four endogenous variables R, T, $\overline{x}$, and μ, with the μ solution then yielding the optimal taxes.

The solution for G is given by a modified version of (9), where S in the first term is replaced by $\widehat{S}(\cdot)$ and h/q is replaced by $\widehat{D}(\cdot)$. Since G does not appear in the arguments of $\widehat{S}$ and $\widehat{D}$, the modified (9) thus gives G in terms of the other endogenous variables, whose values are determined by (17)–(20).

It should be noted that endogeneity of the taxes would be eliminated if the MRS expression in (20) (and in the original equation (14)) were a constant. This case emerges, for example, if preferences over c and q in (15) take the Leontief form, making v_c a constant, denoted ϕ. With v_G equal to the constant ν, the MRS is then μ/ϕ and (20) gives $\mu = L\mu/\phi$, yielding exogenous taxes via the tax formulas.

To solve equations (17)–(20), we use an iterative procedure. It starts with guesses for initial values of R, T and μ. Given these values, the population condition (17) is solved for $\overline{x}$. With the solution in hand, the integrals in (18)-(20) then are computed, using the initial guesses of R, T and μ in evaluating the integrands. The integrals then give updated values of the variables R, T and μ, which are substituted in (17), yielding a new solution for $\overline{x}$. The process continues until convergence is achieved, which occurs after relatively few iterations. The equilibrium value of G is then computed from the modified (9).

4. Simulation Results

4.1. No-tax equilibrium

We first solve for the no-tax equilibrium. The procedure is to set $\tau_t = \tau_q = \tau_\ell = 0$ and then to solve (17) and (18) for $\overline{x}$ and R. Figures 1–5 show the spatial contours of p, q, r, $h(S)$, and D in the no-tax city, represented by the gray curves, and Table 2 gives the central (CBD) values of these variables. The solution gives $\overline{x} = 29.59$, which implies an average commuting distance of 14.35 miles, slightly longer than the average commute for workers in MSAs of 1–3 million inhabitants (13.74 miles, from National Household Travel Survey (NHTS)). Units of q are chosen such that the average dwelling size is 2,196 square feet, with q rising from 1,498.85

at the CBD to 4,340.35 at the city border $\overline{x}$.[12] The housing price p falls from \$8.67 per square foot at the CBD to \$2.14 at $\overline{x}$, while land rent r falls from \$14.2 million per square mile to $\$58,880 = r_a$. Building height $h(S)$ falls from 23.69 at the CBD to 0.397 at $\overline{x}$, and population (dwelling) density falls from 4290.27 dwellings per square mile to 24.87 (average density is 272.68).[13] Despite the presence of rent redistribution and the emissions externality, these spatial patterns are familiar from the standard urban model. Total CO_2 emissions G in the city are 2.489×10^9 kg (2.49 million metric tons), and per-capita emissions equal 3318.72 kg. Residential energy use is responsible for 54% of total emissions, with commuting responsible for the balance of 46%.

4.2. The first-best equilibrium

We now turn to the model solution when emissions taxes are levied. The optimal taxes are given by $\tau_q^* = \tau_\ell^* = \$0.72$ per square foot and $\tau_c^* = \$179.90$ per mile (recall that these are annualized values). On average, the housing tax corresponds to an ad valorem tax of 10.1% on housing rent, the land tax amounts to 7.5% of land rent, and the commuting tax to 34.5% of commuting costs. Note that the average rates of the housing and land taxes are given by τ_q/p and τ_l/r averaged across the city's x values, while the rate of the commuting tax, which is just τ_t/t, is spatially invariant.

Table 2 gives the central values of $p,\ q,\ r,\ h(S)$, and D in the taxed city, and Figures 1–5 show the spatial contours of these variables, which are represented by the black curves. The figures show that, relative to the no-tax city, the p, r, $h(S)$, and D contours rotate clockwise, while the q contour rotates counterclockwise.

In response to the optimal taxes, the city shrinks spatially, with the urban boundary lying at $\overline{x} = 18.03$. Compared to the no-tax case, the spatial extent of the city thus shrinks by 30 percent. This finding confirms the expectation that energy taxation makes cities more compact by discouraging long commutes, reducing housing consumption, and increasing building heights. In the taxed city, dwelling size q rises from 1241.59 square feet at the CBD to 2763.33 at the new $\overline{x}$. Average dwelling size is 1717.18 square feet, 22% lower than in the no-tax equilibrium. The housing price p falls from \$10.95 per square foot at the CBD to 3.82 at $\overline{x}$, while land rent r falls from \$27.1 million per square mile to $r_a = 58,880$. Building height $h(S)$

falls from 38.50 at the CBD to 1.18 at $\overline{x}$, while population density falls from 8,415 dwellings per square mile to 115.9 (average density is 735).

Total emissions G in the taxed city are 1.69×10^9 kg and per-capita emissions equal 2257.78 kg, with both values naturally smaller than in the untaxed city (the reduction is 32%). Residential energy use is now responsible for 52% of total emissions, with commuting responsible for the balance of 48%.

The compensating variation associated with the welfare gain in moving to the first best can be computed. It equals the reduction in income needed to restore the no-tax utility level when all the endogenous variables are held at the levels prevailing in the first-best city. The compensating variation is close to 1.5% of income, a relative modest level similar to the gains from correcting other externalities in a monocentric city (Brueckner's (2007) computed gain from correcting unpriced congestion is 0.7% of income).

The contour rotations in Figures 1–5 are similar to the effects of an increase in the commuting-cost parameter t in the closed-city version of the standard model. While a higher commuting cost per mile is a consequence of the present model's commuting tax, partly accounting for this similarity of effects, many additional forces are at work in generating them. These forces include responses to the land tax τ_ℓ, which tends to raise the cost of land and thus encourages developers to economize on land in production of housing, tending to raise S and building height $h(S)$. But since the housing tax τ_q (which is analogous to a property tax) is a tax on output of housing floor space, it tends to depress S and $h(S)$, offsetting the effect of the land tax. The housing tax also tends to reduce the dwelling size q as consumers substitute toward nonhousing consumption. The tax's effects on $h(S)$ and q, both being negative, have an ambiguous effect on population density (h/q), as discussed in detail by Brueckner and Kim (2003). These varied tax effects are mediated by the impacts of redistribution of differential land rent and tax revenue, adding to the complex interplay of forces affecting urban form in the taxed city. Interestingly, though, this interplay yields qualitative impacts similar to the effects of increase in commuting cost in the standard model.

To get a sense of the magnitudes of the optimal taxes, consider first a comparison of the average housing and land-tax rates to actual US property-tax rates, expressed as a percentage of rent rather than value. Recall that a standard ad-valorem property tax is absent from the model, with the rate set to zero. Letting κ denote the property-tax rate on value and θ denote the discount rate, the property-tax rate expressed as a percentage of rent is given by $\kappa/(\kappa+\theta)$.[14] Assuming $\theta = 0.04$ and using a representative 1.5% property-tax rate,[15] so that $\kappa = 0.015$, this expression reduces to 0.27, indicating that the existing property tax claims about 25% of rent. The average housing and land-tax rates of 10.1% and 7.5% from the model are well below this value.

The model's value of the commuting cost parameter t includes the gasoline tax, and the optimal commuting tax can be generated by increasing this tax. To judge the magnitude of the required increase, the 34.5% rate of the commuting tax can be converted into a required increase in the tax per gallon of gasoline, as follows. The annual t value of \$521.77 per mile is based on an overall money and time cost of \$0.83 per mile, with a money cost of \$0.55 per mile (see the Appendix). The 34.5% commuting tax rate (equal to τ_t/t) implies a \$0.29 increase in this \$0.83 overall cost per mile. With an average US gasoline tax of \$0.487 per gallon[16] and an average light-vehicle fuel economy of about 20 miles per gallon,[17] the gasoline tax per mile is $0.487/20 = 0.024$/mile. Achieving the required \$0.29 increase in commuting cost per mile thus requires a \$0.29 increase in the gasoline tax per mile, or a $20 \times 0.29 = \$5.80$ increase in the gasoline tax per gallon, to a value of \$6.287. Imposing the optimal commuting tax would therefore require a roughly 12-fold increase in the gasoline tax. The resulting tax per gallon is about 50% higher than the highest European gasoline taxes, which are around \$4.00 per gallon.[18]

This gasoline tax is much higher than the optimal value computed by Parry and Small (2005), who derive an optimal \$1.01 US tax per gallon, taking into account emissions along with accident and congestion externalities. Our much higher value depends on the structure and calibration of the model, particularly the assumed value of the emissions parameter ν in the utility function, and it would be reduced with a smaller ν, as seen in the sensitivity analysis below.

4.3. Second-best optima

Instead of solving the equation system (17)–(20), the equilibrium in a city with optimal emissions taxes can be found in a different, equivalent manner. Under this approach, the taxes are treated as parameters, with (17)–(19) solved for $\overline{x}$, R, T conditional on τ_t, τ_q and τ_ℓ. Then, the value of u conditional on the taxes is determined by (16). The optimal values of the taxes (the ones that maximize u) can then be determined by a search procedure. Note that this procedure makes no use of the optimal tax formulas, and thus does not require computation of a value for μ, the marginal social damage from emissions.

This approach gives the same numerical answers as the original approach and thus need not be used in finding the first-best equilibrium. But use of the approach is necessary in investigating the properties of second-best optima, which are utility-maximizing equilibria with one or two of the three taxes constrained to equal zero. With some taxes set at zero, the utility-maximizing value(s) of the remaining tax(es) can be found using a search procedure.[19]

First, we constrain the commuting tax to be zero. The resulting optimal housing and land taxes are much higher than in the first-best case, while no longer being equal. The second-best optimal taxes are $\tau_\ell = 1.972$ and $\tau_q = 1.046$, with the housing tax now amounting to 14.5% of housing rents and the land tax to 22.9% of land rents, on average. The second-best city's $\overline{x}$ value, equal to 17.22, is about 5% smaller than the first-best value of 18.03. When commuting is not taxed, setting the housing and land tax at their first-best levels would lead to a city that is too spread out, since commuting costs are below social costs. Hence, both the housing tax and land tax must be raised, making the city more compact, even more so than the first best city. Interestingly, this second-best city has a density contour that lies between the relatively flat one of the no-tax city and the steep contour of the first-best city, as seen in Figure 6 (the second-best contour is dashed). The same observation applies to the building-height contour, while the p contour rotates counterclockwise relative to the first best contour, showing the reduced value of access to the CBD in the absence of the commuting tax (the q contour rotates clockwise). Table 2 gives the central values D, $h(S)$, q, p and r in the second-best city along with emissions per capita.

When the housing tax is set to zero, the second-best optimal land and commuting taxes are $\tau_\ell = 0.996$ and $\tau_t = 233.873$ (the average rates are 9.7% and 44.8%). Again, both taxes are set higher than the first-best rates. These taxes lead to a spatial city structure similar to that of the first-best city. The boundary distance $\overline{x}$ is 17.83, 98.8% of the first-best value. The absence of a housing tax leads to larger dwellings (and a lower p) at all distances relative to the first-best city, tending to increase the city's spatial area, but this effect is partly countered by the higher taxes on land and commuting. The increased land tax in particular leads to taller buildings at all distances, so that the pattern of population density is very similar to that in the first-best city. Table 2 again provides central data for this second-best city.

Next, we set the land tax to zero. The optimal second-best housing and commuting taxes are $\tau_q = 0.879$ and $\tau_t = 223.027$, implying average tax rates of 10.7% and 42.7%. In this case, the urban boundary is $\overline{x} = 20.42$, which is 13% above the first-best level. The absence of the land tax thus causes the city to expand far beyond the efficient size. The increases in the housing and commuting taxes partially compensate for the absent land tax. But dwellings are larger at all distances than in the first-best city (p is lower), and buildings are shorter everywhere. As a result, population density is lower at all distances out to the first-best boundary, accounting for the larger $\overline{x}$. See Table 2 for further information.

Finally, we set both housing and land taxes at zero, so that the commuting tax is the only second-best tax. Its optimal value is then $\tau_t = 316.819$ or 60.7% of average commuting costs. The urban boundary is $\overline{x} = 21.76$, which is 21% above the first-best level. As in the previous exercise, the absence of taxes on housing and land causes dwellings to be too large and buildings to be too short. While the commuting tax increases strongly to counteract this tendency, the city is much larger than optimal and population density is inefficiently low in the center. See Table 2 for more information. As would be expected, the compensation variations associated with all of these second-best tax schemes are smaller than the first-best value (the values are well below 1% of income).

If policy makers were to consider use of existing urban taxes to counteract a city's GHG emissions, it would be natural for them to focus on the gasoline tax, not heeding this paper's prescription for additional taxes on land and housing. The second-best optimal commuting-tax rate of 60.7% requires a \$0.50 increase in the baselineline \$0.83 commuting cost per mile, which translates (using the previous procedure) into a \$10.00 increase in the gasoline tax, to a value of \$10.487 (more than a 20-fold increase).

4.4. Sensitivity analysis

This subsection provides sensitivity analyses. One by one, we vary some of the more interesting parameters, increasing (or in one case, decreasing) each of the parameters by 50% of the benchmark value. The results of these exercises are shown in Table 3.

First, we vary the marginal disutility of emissions, ν. In addition to exploring the issue just discussed, this exercise is important because, in the economics of climate change, there is considerable controversy regarding discount rates, uncertainty, tipping points and so on. These factors would affect the marginal social damage from emissions and hence, optimal taxes. Therefore, we decrease ν by half to $\nu = 0.025$. While the no-tax equilibrium is obviously unaffected, the optimal taxes fall by 48% each, to $\tau_q = \tau_\ell = 0.347$ and $\tau_t = 87.133$. As a result, $\overline{x}$ is 23% larger and emissions per capita are 18% higher than in the previous first-best city. Emissions are reduced by 20% compared to the no-tax city, rather than by 32%, as in the previous first-best case.[20] The new τ_t value leads to a commuting tax rate of 17%, requiring a \$0.14 increase in the baseline \$0.83 cost per mile. This increase requires a \$2.82 increase in the gasoline tax per gallon to a level of \$3.31, less than the maximum European level.

Conversely, when ν increases by 50% to $\nu = 0.075$, emissions are reduced by 14% and $\overline{x}$ shrinks by 17% relative to the previous first-best city. The optimal taxes rise by 54%. Compared to the no-tax city, emissions are reduced by 41%.

Second, we increase income by 50% from the benchmark value of \$51,324, to \$76,986. This value corresponds to the household income in very rich metro areas such as San Francisco or Boston. In the standard urban model, such an increase leads to higher average housing consumption, longer commutes, and urban sprawl. Obviously, these effects increase CO_2 emissions. In the first-best city, the income increase leads to a 39% increase in $\overline{x}$ and a 57% increase in emissions per capita relative to the benchmark first-best city. Interestingly, the optimal taxes each fall by about 2%. The apparent intuition is that the emissions externality affects utility linearly, so that when income rises, the 'distortionary' effects of taxation in the housing market weigh more strongly than the increased emission damage.

Next, we increase population from $L = 750,000$ to 1,125,000. Comparing first-best cities, the population increase leads to a 10% percent decrease in $\overline{x}$. Emissions increase by 25%

while emissions per capita decrease by 17%, matching a pattern observed by Larson and Yezer (2015).[21] Each of the taxes increases by 57% percent. The result that optimal city size with larger population is *smaller* in the presence of optimal emissions taxes is striking. This outcome is due to the fact that, in the absence of taxes, the city expands spatially by a modest 6% while emissions increase strongly, by 41%. As long as the emissions damage is large enough, the increase in optimal taxes is so large that the optimal city shrinks, relative to the benchmark. If we reduce the value of ν sufficiently, this result is reversed, with the optimal city expanding spatially when population increases.

A further exercise is to increase annual commuting cost t to \$782.65 per mile. Comparing first-best cities, the commuting-cost increase leads to a 20% decrease in $\overline{x}$ and a 21% decrease in emissions per capita, with both changes partly reflecting the higher private cost of commuting. Despite this improvement in incentives, each tax increases by 3%.

Finally, we vary the energy efficiency of buildings, e, and the emissions intensity γ of commuting. Comparing first-best cities, an increase in e (which could also be caused by a change in the local temperature) reduces $\overline{x}$ by 10% and raises emissions per capita by 17%. The $\overline{x}$ change appears to partly reflect greater private incentives to reduce housing consumption in the face of lower residential energy efficiency. In response to the higher e, τ_q and τ_ℓ rise by 53%, while τ_t rises by 2%.[22] Again comparing first-best cities, an increase in γ reduces $\overline{x}$ by 7% and raises emissions per capita by 14%. Since the γ change generates no private commuting-cost increase, the $\overline{x}$ change mostly reflects the 51% increase in τ_t. The housing and land taxes rise by only 1%.

5. Conclusion

This paper has presented the first investigation of the effects of optimal energy taxation in an urban spatial setting. Rather than exploring the effects of a carbon tax, our approach is to derive the supplements to existing real estate and gasoline taxes needed to support the social optimum, analyzing their effects on urban spatial structure. This exercise is carried out using a model that incorporates emissions economies from tall buildings. Since emissions are generated by housing consumption and commuting, optimal taxation reduces the levels of both activities, generating a more-compact city with a lower level of emissions per capita. In response to optimal taxation, the distance to the urban boundary shrinks by 30%, the city's central building height rises by 62% (helping to raise central population density by 96%), and emissions per capita fall by 32%.

The paper also carries out second-best exercises, with the most instructive being an exercise where the housing and land taxes are set at zero, so that the commuting tax must do all the work in limiting emissions from both residences and commute trips. In this case, the second-best optimal commuting tax would correspond to a gasoline tax of more than $10 per gallon, a twenty-fold increase over the current US average tax. While this increase partly reflects the assumed value of the individual utility loss from emissions, it also reflects the expanded role of the commuting tax.

Future research could add detail to the model, especially on the commuting side, following the lead of Larson et al. (2012). Their model includes traffic congestion that in turn affects travel speed, along with a realistic relationship between speed and emissions. Adding such features could improve the accuracy of the commuting tax levied in the current model while incorporating the overall connection between traffic congestion and urban spatial structure. In conjunction with the incorporation of congestion, a further extension could add public transit along with household heterogeneity in the valuation of commute time.

Appendix

A1. Planning-problem derivations

The Lagrangean expression for the planning problem is generated by subtracting the RHS expressions in (8) and (9) from the left-hand sides, multiplying the resulting expressions by the multipliers λ and μ, and adding (7). The first-order conditions for S, q, G and $\overline{x}$ are

$$S: \qquad i \;+\; \frac{h'(S)}{q}[c(q,G)+tx] \;+\; h'(S)e \;+\; \lambda\frac{h'(S)}{q} \;-\; \mu\psi h'(S)e \;-\; \mu\gamma\frac{h'(S)}{q}x \;=\; 0 \tag{$a1$}$$

$$q: \qquad -\frac{h(S)}{q^2}[c(q,G)+tx] \;-\; \frac{h(S)}{q}\frac{v_q}{v_c} \;-\; \lambda\frac{h(S)}{q^2} + \mu\gamma\frac{h(S)}{q^2}x \;=\; 0 \tag{$a2$}$$

$$G: \qquad -\int_0^{\overline{x}} 2\pi x\frac{h(S)}{q}\frac{v_G}{v_c}dx \;+\; \mu \;=\; 0 \tag{$a3$}$$

$$\overline{x}: \qquad i\overline{S} \;+\; \frac{h(\overline{S})}{\overline{q}}[c(\overline{q},G)+t\overline{x}] \;+\; h(\overline{S})e \;+\; e \;+\; r_a \;+\; \lambda\frac{h(\overline{S})}{\overline{q}} - \mu\psi[h(\overline{S})e+e]$$

$$-\; \mu\gamma\frac{h(\overline{S})}{\overline{q}}\overline{x} \;=\; 0. \tag{$a4$}$$

Rearranging ($a2$) yields (10), and ($a3$) is the same as (16). Solving ($a2$) for λ and substituting in ($a1$) yields (11) after rearrangement, and substituting in ($a4$) yields (13) after rearrangement.

A2. Data sources and calibration calculations

Income y is set at the 2011 value of median household income in the US, which is $51,324. The source is *Household Income: 2012, American Community Survey Briefs*, By Amanda Noss, U.S. Department of Commerce Economics and Statistics Administration, U.S. Census Bureau, September 2013 (`https://www.census.gov/prod/2013pubs/acsbr12-02.pdf`).

To compute commuting cost per mile, t, we follow Bertaud and Brueckner (2005). We use the median hourly wage of $17.09 (from Bureau of Labor Statistics, `http://www.bls.gov/oes/current/oes_nat.htm`) and value it at 50% (Small (2012)) to get an hourly time cost of commuting \$8.545. Assuming that rush hour traffic moves at 30 miles per hour, the implied time cost per mile of commuting is \$0.28. Adding a money cost of automobile operation equal to \$0.55/mile (the current Federal allowance), total commuting cost per mile is then \$0.83. Multiplying by 1.25 workers/household, by 250 work days/year and again by 2 to convert to a round-trip basis, annual commuting cost per mile is \$521.771/year.

The computation of agricultural rent r_a again follows Bertaud and Brueckner (2005). We take the average value of farm real estate per acre in 2011, \$2300 (the source is United States Department of Agriculture (2015), Land Values: 2015 Summary, `http://www.usda.gov/nass/PUBS/TODAYRPT/land0815.pdf`). To convert this number to annual rent, we use a discount rate of 4% to get a rent per acre of $\$2,300/0.04 = \$92$, yielding a land rent per square mile of $r_a = \$58,880$.

To get a value for γ, CO_2 emissions per mile of commuting, we use data from the Carbon Trust to find following conversion factors for commuting (see Carbon Trust, Conversion factors: Energy and carbon conversions 2011 update `http://www.carbontrust.com/media/18223/ctl153_conversion_factors.pdf`). For cars, we use the value for average petrol for cars (0.3358 kg CO_2/mile); for public transport, we use $1/2 \times$ value for bus [0.1488 kg per passenger-km] + $1/2 \times$ value for subway [0.0736 kg per passenger-km], which gives 0.1112 kg per passenger-km or 0.1789 kg per passenger-mile; for cycling/walking we use a value of 0. These values are then weighted by modal shares (82%/11%/7%) to get the value of 0.279 kg CO_2/mile. We then multiply by 250 workdays per year and by 2 to get an annual round-trip value of 139.347 kg CO_2/mile. Note that, although our t value pertains to automobile commuting because it is then easily computed, the γ value embraces all commuting modes.

For emissions from residential energy use, we again use data from Carbon Trust. The conversion factors for heating and cooling are 0.5246 kg CO_2/kwh for electricity, 0.1836 for natural gas, and 0.2674 for fuel oil. We weigh these values by the percentage shares of households using the three energy sources (35, 52, 7, from RECS, `http://www.eia.gov/consumption/residential/`) to get a conversion factor of $\psi = 0.2978$ kg CO_2/kwh.

Table 1: 2013 Emissions by Sector
(millions of metric tons CO_2 equivalent)

Electricity-generation emissions
are distributed to final user

implied sector	*volume*	*percentage*
Industry	1922.6	30.0%
Transportation	1810.3	27.1%
Residential	1129.1	16.9%
Commercial	1126.7	16.9%
Agriculture	646.4	9.7%
Total	6673.0	100%

Columns do not sum since emissions from
US Territories are excluded. Source is
Environmental Protection Agency (2015, Table ES-7)

Table 2: City Characteristics

	No-tax	*First best*	*Second best* $\tau_t = 0$	*Second best* $\tau_q = 0$	*Second best* $\tau_\ell = 0$	*Second best* $\tau_\ell = \tau_q = 0$
city border $\overline{x}$	29.59	18.03	17.22	17.83	20.42	20.76
emissions per capita	3318.72	2257.78	2362.05	2316.77	2238.44	2382.87
central density D	4290.27	8415.71	5742.55	8549.1	7847.33	9780.53
central bldg. height $h(S)$	23.69	38.50	27.76	40.65	36.40	45.27
central dwelling size q	1498.85	1241.59	1312.2	1290.54	1259.03	1256.18
central housing price p	8.67	10.95	10.20	10.43	10.92	10.82
central land rent r	14.2 m	21.1 m	17.0 m	29.1 m	25.3 m	33.9 m
commuting tax τ_t	0	179.896 (34.5%)	0	233.873 (44.8%)	223.027 (42.7%)	316.820 (60.7%)
housing tax τ_q	0	0.717 (10.1%)	1.046 (14.5%)	0	0.879 (10.7%)	0
land tax τ_ℓ	0	0.717 (7.5%)	1.972 (22.9%)	0.996 (9.7%)	0	0

Table 3: Sensitivity Analysis

(*percentage change in first-best value relative to benchmark first-best value*)

	ν falls to 0.025	ν rises to 0.075	y rises to \$76,986	L rises to 1,250,000	t rises to \$782.65/mile	e rises by 50%	ψ rises by 50%
city border $\overline{x}$	+23%	−17%	+39%	−10%	−20%	−10%	−7%
emissions per capita	+18%	−14%	+57%	−17%	−21%	+17%	+14%
commuting tax τ_t	−48%	+54%	−2%	+57%	−3%	+2%	+51%
housing tax τ_q	−48%	+54%	−2%	+57%	−3%	+53%	+1%
land tax τ_ℓ	−48%	+54%	−2%	+57%	−3%	+53%	+1%

Figures

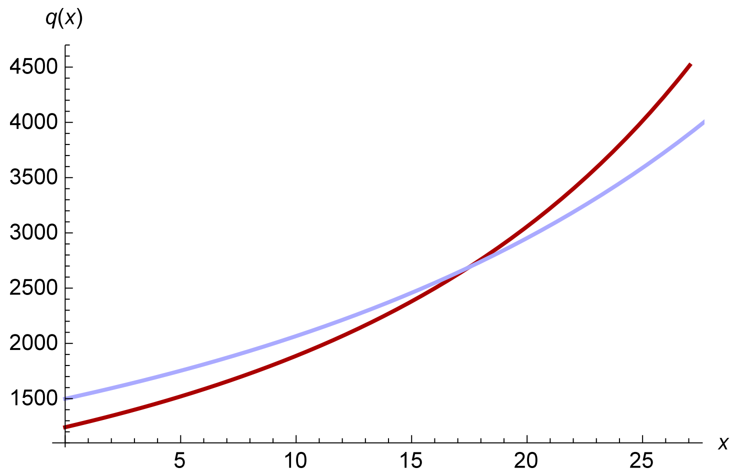

Figure 1: Dwelling size in the first best (dark red/black) and no-tax city (light blue/gray)

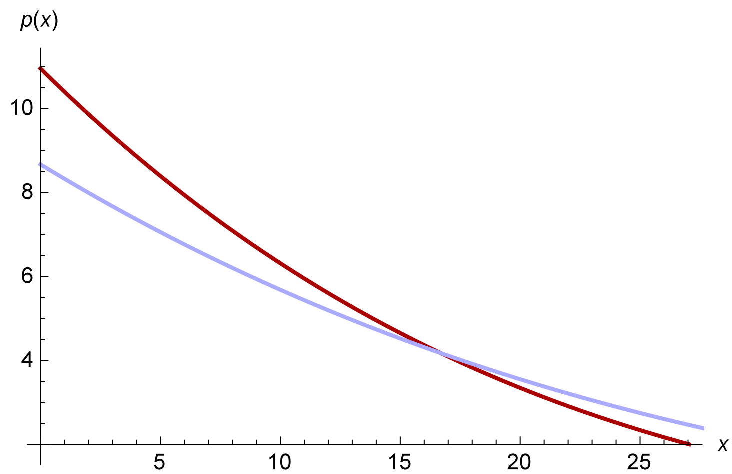

Figure 2: Housing price in the first best (dark red/black) and no-tax city (light blue/gray)

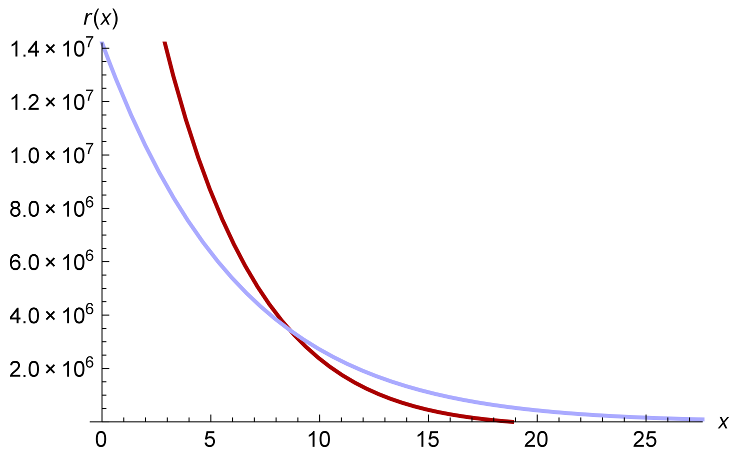

Figure 3: Land rent in the first best (dark red/black) and no-tax city (light blue/gray)

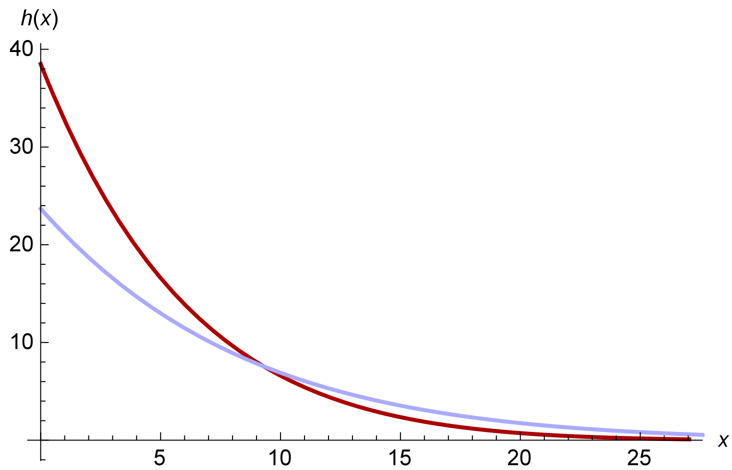

Figure 4: Building height in the first best (dark red/black) and no-tax city (light blue/gray)

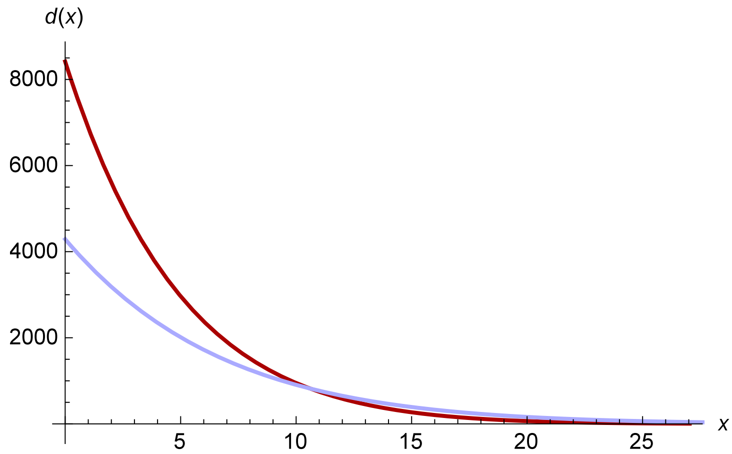

Figure 5: Population density in the first best (dark red/black) and no-tax city (light blue/gray)

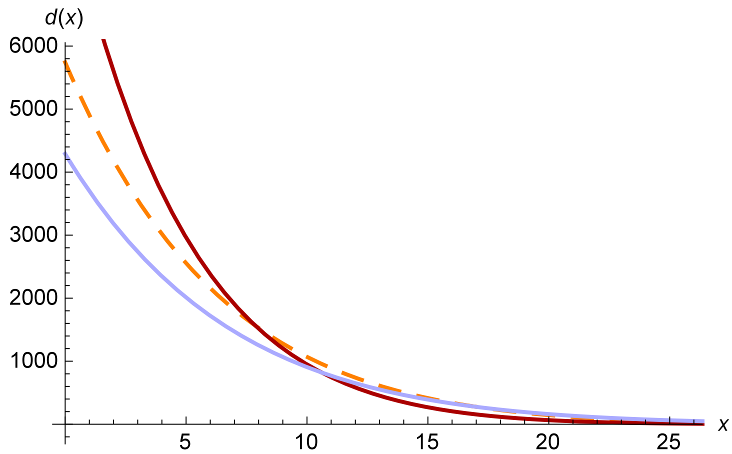

Figure 6: Population density in the first best (dark red/black), second best ($\tau_t = 0$, dashed) and no-tax city (light blue/gray)

References

AHLFELDT, G.M., MCMILLEN, D.P., 2014. New estimates of the elasticity of substitution between land and capital. Working paper, Lincoln Institute of Land Policy.

BERTAUD, A., BRUECKNER, J.K., 2005. Analyzing building-height restrictions: predicted impacts and welfare costs. *Regional Science and Urban Economics* 35, 109-125.

BORCK, R., 2015. Will skyscrapers save the planet? Building height limits and greenhouse gas emissions, Working paper, University of Potsdam.

BORCK, R., PFLÜGER, M., 2015. Green Cities? Urbanization, trade and the environment. IZA discussion paper 9104.

BRUECKNER, J.K., 1987. The structure of urban equilibria: A unified treatment of the Muth-Mills model. In: Mills, E.S. (Ed.), *Handbook of Regional and Urban Economics*, Vol. 2, North Holland, Amsterdam, pp. 821-845.

BRUECKNER, J.K., 2007. Urban growth boundaries: An effect second-best remedy for unpriced traffic congestion? *Journal of Housing Economics* 16, 263-273.

CHING, F.D.K., SHAPIRO, I., 2014. *Green Building Illustrated*. John Wiley & Sons, Hoboken, N.J.

CHONG, H., 2012. Building vintage and electricity use: Old homes use less electricity in hot weather. *European Economic Review* 56, 906-930.

COSTA, D.L., KAHN, M.E., 2011. Electricity consumption and durable housing: Understanding cohort effects. *American Economic Review* 101, 88-92.

DAVIS, M.A., ORTALO-MAGNÉ, F., 2011. Household expenditures, wages, rents. *Review of Economic Dynamics* 14, 248-261.

EICHHOLTZ, P.M.A., KOK, N., QUIGLEY, J.M., 2010. Doing well by doing good: Green office buildings. *American Economic Review* 100, 2494-2511.

US ENVIRONMENTAL PROTECTION AGENCY, *Inventory of US Greenhouse Gas Emissions and Sinks, 1990-2013.* Washington, D.C. (`http://www.epa.gov/climatechange/emissions/usinventoryreport.html`).

FULLERTON, D, WEST, S.E., 2002. Can taxes on cars and on gasoline mimic an unavailable tax on emissions? *Journal of Environmental Economics and Management* 43, 135-157.

Fujita, M., 1989. *Urban Economic Theory.* Cambridge University Press, Cambridge.

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

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

Joshi, K.K., Kono, T., 2009. Optimization of floor area ratio regulation in a growing city. *Regional Science and Urban Economics* 39, 502-511.

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

Larson, W., Yezer, A., 2014. The energy implications of city size and density. *Journal of Urban Economics* 90, 35-49.

Nordhaus, W., 2013. DICE-2013 model, beta version of April 22, 2013, Yale University.

Oates, W.E., Schwab, R.M., 1997. The impact of urban land taxation: The Pittsburgh experience. *National Tax Journal* 50, 1-21.

Parry, I.W.H., Small, K.A., 2005. Does Britain or the United States have the right gasoline tax? *American Economic Review* 95, 1276-1289.

Pines, D., Sadka, E., 1985. Zoning, first-best, second-best and third-best criteria for allocating land for roads. *Journal of Urban Economics* 17, 167-183.

Pines, D., Sadka, E., 1986. Comparative statics analysis of a fully closed city. *Journal of Urban Economics* 20, 1-20.

Small, K.A., 2012. Valuation of travel time. *Economics of Transportation* 1, 2-14.

Song, Y., Zenou, Y., 2006. Property tax and urban sprawl: Theory and implications for US cities. *Journal of Urban Economics* 60, 519-534.

Tscharaktschiew, S., Hirte, G., 2010. The drawbacks and opportunities of carbon charges in metropolitan areas: A spatial general equilibrium approach, *Ecological Economics* 70, 339-357.

Wheaton, W.C., 1974. A comparative static analysis of urban spatial structure. *Journal of Economic Theory* 9, 223-237.

White, M.J., 2004. The ‘arms race’ on American roads: The effect of sport utility vehicles and pickup trucks on traffic safety. *Journal of Law and Economics* 47, 333-356.

Zheng, S., Wang, R., Glaeser, E.L., Kahn, M.E., 2011. The greenness of China: Household carbon dioxide emissions and urban development. *Journal of Economic Geography* 11, 761-792.

Footnotes

*We thank Howard Chong for steering us toward the engineering/architecture literature on residential energy use, and we are grateful to Sofia Franco, Georg Hirte Tatsuhito Kono, Will Larson, and Tony Yezer for detailed comments and to Ben Leard and other conference participants for additional comments. Any shortcomings in the paper, however, are our responsibility.

[1]In similar fashion, Fullerton and West (2002) show that, in treating automobile emissions, a carbon tax can be replaced by taxes with other features that acheive the same outcome (i.e., a gas tax that depends on fuel type, engine size, and installed pollution control equipment, or a vehicle tax that depends on mileage).

[2]For additional analysis where adjustment of building heights serves to amerliorate an externality, see Joshi and Kono (2009). With unpriced traffic congestion, population in a monocentric city is insufficiently concentrated, and this paper shows that a second-best best remedy is building-height regulations that impose a minimum near the center and a maximum in the suburbs.

[3]Among buildings with a given footprint area, square buildings have the smallest surface area (see below).

[4]Taken literally, the model implies that developers should construct buildings with the biggest possible footprint, limited only by the city's street grid.

[5]Energy use from appliances may, of course, show a modest increase with dwelling size (from larger refridgerators and hot-water heaters or additional televisions), but omission of this effect is acceptable as an approximation.

[6]See Fujita (1989) for another use of this approach.

[7]In particular, suppose that a fraction ω of total emissions G is particulate matter that creates local pollution while the remainder is GHG. The GHG component affects residents of other cities, while the GHG created in those cities affects residents of the given city. Let the economy contain n identical cities, and let η and ξ denote the disutilities per unit of particulates and GHG, respectively. Then, the emissions term in the utility function is $\eta\omega G + \xi(1-\omega)G + (n-1)\xi(1-\omega)G = [\eta\omega + n\xi(1-\omega)]G$, so that G remains the appropriate utility-function argument. In the simulation model below, where G is assumed to enter linearly in preferences, its coefficient ν can be viewed as equal to $\eta\omega + n\xi(1-\omega)$. Under this

formulation, the planner would face n cities with identical populations, and would choose common values of G and the other variables to minimize global resource consumption (n times (7)) subject to the constraints (8) and (9). See Borck and Pflger (2015) for a related analysis of global emissions in a framework with two cities, as well as Gaigné, Riou and Thisse (2012).

[8]The dual version of the planning problem starts by deriving income-compensated demand functions for c and q conditional on G, denoted by $c(p, G, u)$ and $q(p, G, u)$. In (7)–(9), the first function is substituted in place of $c(G, u)$ and the second is substituted in place of q. Then, (7) is set equal to I, which gives the economy's total endowment of c. Finally, u is maximized subject to the three modified constraints, with p, G, and $\overline{x}$ treated as choice variables along with u. The optimality conditions in (10), (11) and (13) again emerge. See Pines and Sadka (1986) for another use of this approach.

[9]With an average household size of 2.6, the city would then have 1.95 million inhabitants.

[10]By ignoring the possible irregular shapes of single-family houses, this calculation may lead to a biased value of e, but the result is acceptable as an approximation.

[11]In reality, tax revenues might be used to subsidize energy-efficient public transit or building modifications designed to reduce energy use. Analysis of these options would require a more detailed model.

[12]We solve the model, and then rescale the resulting q values by multiplying by a factor ξ that makes the average value in the city equal to 2,196 square feet. This average value is given by $1/N$ times the integral of q, weighted by population density, over x. Then, the p solution at each x is divided by ξ, as is the floor space tax. This procedure follows Bertaud and Brueckner (2005).

[13]This population density is similar to that of Buffalo-Niagara Falls, NY, or Dallas-Fort Worth-Arlington, TX, according to the 2010 Census [www.census.gov]. In all MSAs with population between 1.5 and 2.5 million, the average population density is 540 people, or 208 households, per square mile.

[14]To derive this expression, note that property value P is determined by the relationship $P = (p - \kappa P)/\theta$, with P equaling the discounted value of the flow of rent minus taxes. Solving yields $P = p/(\kappa + \theta)$, so that the tax liability as a percentage of rent is given by $[\kappa p/(\kappa + \theta)]/p = \kappa/(\kappa + \theta)$.

[15]See, for example, Song and Zenou (2006).

[16]See the following webpage from the American Petroleum Institute: http://www.api.org/oil-and-natural-gas-overview/industry-economics/fuel-taxes/gasoline-tax.

[17]US Department of Transportation data at the following link show miles per gallon for the US fleet of cars and light trucks of 23 and 17, respectively. With light trucks constituting about 40% of the overall light vehicle fleet (White (2004)), average miles per gallon is around 20. (http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_statistics/html/table_04_23.html)

[18]See the following webpage from US Department of Energy: http://www.afdc.energy.gov/data/10327.

[19]As mentioned earlier, Borck (2015) studies building-height limits as a different second-best tool for combating global warming. The intuition is that, by tightening housing supply, lower building heights may depress housing consumption, thus reducing emissions. However, the population-density contour in a city with building-height limits is too flat, compared to a city with first-best taxation.

[20]Using results from Nordhaus' (2013) DICE model, the reduction of CO_2 emissions in 2020 is 19% in his optimal taxation case, 34% if the goal is to limit global warming to 2 degrees centigrade, and 57% with low discounting, as in the Stern Review.

[21]This result emerges when the population increase is caused by an exogenous increase in amenities in an open-city context.

[22]This analysis also applies to the effects of a change in ψ, emissions per unit of residential energy usage.