Hirte, Georg; Rhee, Hyok-Joo

Working Paper

Regulation versus Taxation

CEPIE Working Paper, No. 05/16

Provided in Cooperation with:
Technische Universität Dresden, Faculty of Business and Economics

TECHNISCHE UNIVERSITÄT DRESDEN

CEPIE Working Paper No. 05/16

Center of Public and International Economics

REGULATION VERSUS TAXATION: EFFICIENCY OF ZONING AND TAX INSTRUMENTS AS ANTI-CONGESTION POLICIES

September 2016

Georg Hirte
Hyok-Joo Rhee

Editors: Faculty of Business and Economics, Technische Universität Dresden.

This paper is published on the Open Access Repository Qucosa.
The complete Working Paper Series can be found at the CEPIE Homepage.

ISSN 2510-1196

Regulation versus Taxation: Efficiency of Zoning and Tax Instruments as Anti-Congestion Policies*

Georg Hirte†
Technische Universität Dresden

Hyok-Joo Rhee‡
Seoultech

September 2016

Abstract

We examine the working mechanisms and efficiencies of zoning (regulation of floor area ratios and land-use types) and fiscal instruments (tolls, property taxes, and income transfer), and extend the instrument choice theory to include the congestion of road and nonroad infrastructure. We show that in the spatial model with heterogeneous households the standard first-best instruments do not work because they trigger distortion of spatial allocations. In addition, because of the household heterogeneity and real estate market distortions, zoning could be less efficient than, as efficient as, or more efficient than pricing instruments. However, when the zoning enacted deviates from the optimum, zoning not only becomes inferior to congestion charges but is also likely to reduce welfare. In addition, we provide a global platform that extends the instrument choice theory of pollution control to include various types of externalities and a wide range of discrete policy deviations for any reasons beyond cost–benefit uncertainties.

JEL classification: H21; R52

Keywords: Infrastructure Congestion; Zoning; Road Tolls; Property Tax; Instrument Choice; Heterogeneity

---

*We like to thank Tony Smith for his help in elaborating the idea of the envelope methodology of land-use and transportation models, as well as Matthias Wrede and participants of various seminars and conferences for their comments. Grants of the National Research Foundation of Korea Grant (NRF-2014S1A5A2A01011212) are greatfully acknowledged. The paper is based on some ideas developed as the second author was visiting TU Dresden.

†Faculty of Traffic Sciences, Technische Universität Dresden, 01062 Dresden, Germany. E-mail: georg.hirte@tu-dresden.de

‡Corresponding author. Department of Public Administration, Seoultech, 232 Gongneung-ro, Nowon-Gu, Seoul 139-743, Republic of Korea. Phone: +822-970-6494, Fax: +822-971-4647, E-mail: rheehj@seoultech.ac.kr

1. Introduction

There has been a long-standing debate among economists on the relative efficiency of price and quantity regulation instruments to correct market failures such as externalities (e.g., Weitzman, 1974 and Kaplow and Shavell, 2002 on pollution control in environmental economics). We focus on this topic in a spatial economy and compare the efficiency of tax instruments (property taxes and congestion charges) and zoning (floor area ratio regulations and land-use type regulations) as policy instruments to control the simultaneous congestion of road and nonroad infrastructure.

This topic is important especially in city planning, simply because land use regulations (LURs) are indispensable in city planning. Consequently, it is natural that urban economists have been discussing the efficiency and equivalence of the LURs and Pigouvian tolls as instruments to mitigate congestion for several decades[1].

Research on the efficiency of LURs, henceforth also zoning, has been showing a wide variety of results depending on the specific framework used. In case of the monocentric city, it is well established that road congestion can be reduced by raising density. One form of this proposition states that in the monocentric city, lot-size zoning is equivalent to congestion charge in its effect on efficiency (Pines and Sadka, 1985; Wheaton 1998). However, this holds only in the case of fixed travel demand, in which case the lot-size zoning can modify the spatial distribution of travelers in the exactly same way as the congestion tolls (Oron et al., 1973; Wheaton, 1998). A slight generalized setting, however, immediately invalidates this proposition. Indeed, the lot-size zoning is second-best (Pines and Sadka, 1985; Kono et al., 2012) or even third-best (Pines and Kono, 2012). If agglomeration economies are present in addition to road congestion, land-use type zoning can almost be first-best if linked with a subsidy to internalize agglomeration economies in the polycentric framework (Rhee et al.,

[1] Besides, there is a sizable body of planning literature. There, the "efficiency" of development controls is not something to be disputed, and the controls are required to "grow smart" (Ewing et al., 2007).

2014). In contrast, the urban growth boundary (UGB), which is in principle second-best in analytical models (Kanemoto, 1977; Arnott, 1979; Pines and Sadka, 1985), are found to improve welfare negligibly (Brueckner, 2007) or even be absolutely harmful in the polycentric framework (Anas and Rhee, 2006). In some cases, an expansionary growth boundary is recommended instead (Anas and Rhee, 2007; Anas and Pines, 2008). The third type of land use regulation usually considered is floor area ratio (FAR) regulation. FAR regulation is found to be an effective substitute of the first-best congestion toll when a minimum FAR is applied at the city center and maximum FARs are applied in the suburb in the closed monocentric city framework (Kono et al., 2012); the efficiency requires maximum FARs everywhere in the open monocentric city framework (Kono and Joshi, 2012)[2].

There is much less literature on mitigating nonroad congestion in the theoretical spatial models. Since each new resident entails additional costs of public service provision, the (Pigouvian) property tax is an efficient congestion charge that is indispensable when non-distortionary head tax is unavailable (Hoyt, 1991; Krelove, 1993). Since zoning forces residents to consume "at least some minimum amount of housing" (Hamilton, 1975: 206) and, thus, is able to control density, the zoning works in a similar fashion to pricing instruments. In the monocentric city with negative agglomeration externalities, optimal zoning consists of a maximum FAR in the center where nonroad congestion is assumed to be stronger and minimum FARs in the suburb (Kono et al., 2009).

Until now, land use regulations have been studied only on each single type of congestion, either road or nonroad congestion, but not both. However, congestion of nonroad infrastructure such as fire protection, police services, health care, schools, sewerages, energy, and telecommunication infrastructure or waste management, is usually present together with road congestion. Tackling these two types of externalities requires employing at least two types of LURs that do not fully correlate. As this has not been studied so far, it is not obvious a priori whether there is an efficient mix of LURs or whether the LURs are equivalent to optimal pricing instruments. The first candidate from the pricing instruments is Pigouvian congestion charging and the property taxation on the differential land rent arising from congested nonroad infrastructure. This is supposed to imply self-financing of nonroads according to the Henry George theorem (Flatters et al., 1974; Stiglitz, 1977; Arnott and

[2] Optimal regulation may also depend on other existing non-optimal policies. For instance, if FAR regulation is linked with non-Pigouvian cordon tolls this implies a minimum FAR inside and maximum FARs outside the cordon ring in the monocentric city (Kono and Kawaguchi, 2015).

Stiglitz, 1979). Despite the intuition, it has not yet been studied as to whether this instrument mix is efficient in the presence of the two types of congestion. Because of the interaction between these two types of congestion, it may also happen that the optimal levels of the tax instruments deviate from pure Pigouvian levels (see Parry and Bento, 2002, on the interaction of externalities). In this case, the self-financing rule could not be sustained.

The first purpose of this study is to compare the efficiencies of pricing and regulatory instruments when both types of congestion coexist. To be able to study land-use type regulation, we need a mixed land use of residences and production. Therefore, we apply Anas-Xu (1999) type model that allows us to study these issues in a general equilibrium spatial model[3]. Our model includes features that are closely related to the real world including a general equilibrium non-monocentric model, two types of congestion, a basket of different regulation instruments, mixed land use for production and residences, and household heterogeneity. The household heterogeneity is a standard feature of the real world although it is neglected in most of the references above (exceptions are Anas and Rhee, 2006 and Rhee et al., 2014). In this case, marginal utilities of income (MUIs) may differ and the standard Pigouvian pricing is no longer first-best unless a special analytical remedy is administered, such as income transfers among households to equalize MUIs (De Palma and Lindsey, 2004).

We derive the first-order formulas of the welfare change associated with various instruments (taxes and LURs). The formulas reveal the essential differences between the model with and without household heterogeneity, between taxes and LURs, and between the spatial and non-spatial models. We, then, perform numerical simulations to compare the efficiencies. The major findings are follows.

First, we show that the free market fails to achieve the first-best efficiency in the spatial model with household heterogeneity even in the absence of market failures such as externalities. Although this result is in principle known and stems from differences in marginal utilities of income (MUIs) across heterogeneous households, the new finding is that income redistribution to equalize MUIs does not restore the first-best efficiency in the spatial model. The redistribution itself introduces a spatial distortion that nullifies the effectiveness of the conventional first-best instrument mix. The same proposition holds in the spatial economy with traffic congestion as well; Pigouvian pricing plus redistribution equalizing the

[3] As a side issue our model is able to reproduce the results from the literature mentioned above and show how different assumptions or even the limited model setup applied in those studies determine these results.

MUIs is not efficient with household heterogeneity. We can generalize this finding, that is, the standard prescription of De Palma and Lindsey (2004) holds only when the household heterogeneity is fixed, but fails if households are able to change their household type endogenously. Thus, our finding may hold in other cases too, where households change their own types by marrying, becoming parents, and switching travel modes, routes, house types, ownership types, income, or skill groups. In such cases, it turns out that only numerical simulations or empirical testing can tell which policy mix is welfare maximizing.

Second, we analytically decompose welfare changes completely and numerically approximate the changes by component for the first time. A byproduct of this exercise is the covariance term of the Feldstein (1972) type representing the impact of heterogeneity on welfare in the spatial framework. Indeed, they account for a significant proportion of the total welfare change. This makes it clear that neglecting the covariance term in the presence of household heterogeneity may induce misperceptions of other components of the effects. This result may carry over to other issues of household heterogeneity, too. In addition, referring to the welfare decomposition, we explain some puzzles raised in the literature of urban and transportation economics.

Third, the precise decomposition of the welfare change shows the different channels through which taxes and LURs affect welfare in the spatial economy with congestion[4]. While Pigouvian taxes internalize congestion externalities in the standard way, LURs have additional terms showing real estate market distortions and compliance costs of LURs. In the numerical analysis, these costs turn out to be huge and rise rapidly as the regulations get tighter.

Fourth, concerning the equivalence of LURs and Pigouvian pricing, numerical simulations confirm that LURs can indeed be as efficient as Pigouvian taxation in the second best way. This, however, is true only for a narrow range of city size, while LURs are clearly less efficient than the second-best Pigouvian charging for larger cities where infrastructure is highly congested. This is not known in the exiting literature, which claims or implicitly accepts the proposition that zoning is almost as efficient as Pigouvian tolling. In addition, we find that in the case of smaller cities, LURs can be even more efficient than Pigouvian pricing, a result not yet reported in the literature.

The first purpose of the study as stated above is about the efficiency of optimal

[4] Bertaud and Brueckner (2005) study costs of LUR in a monocentric model without externalities.

policies. An implicit assumption there is that the optimal policies are known to the planner. In contrast, the second purpose of this paper concerns the instrument choice, which is widely studied in environmental economics, when the planner cannot set the policies optimally. There are several reasons why actual policies deviate from the optimum such as conflicting aims of regulators (planners, tax authorities, or decision makers), imperfect internal decision-making and planning processes, or imperfect regulatory information (missing, wrong or uncertain). To be more specific, actual land use regulations are chosen according to the objectives that are often conflicting with one another. Compact development is an example where saving in peri-urban areas could worsen outdoor living quality in inner cities (Westerink, 2013). To our knowledge, this problem has not been studied yet in the spatial framework.

However, unlike the existing instrument choice literature, we introduce imprecision into the policies themselves rather than into the costs and benefits of pollution abatement. This analytical twist is so useful that we can evaluate *any* type of *discrete* policy deviation from optimal policies; the uncertainties of the costs and benefits of pollution abatement are a special case of our approach. In this new framework, we provide a formula, showing the welfare gain or loss at any point along the equilibrium path away from the optimum. Unlike the Weitzman's (1974) famous formula "$B'' + C''$" that holds only near the optimum, our formula in general justifies the possible superiority of externality pricing to quantity regulation for discrete deviation whatever the reason. As Pizer (1997) shows numerically, when the policy change is discrete, Weitzman's formula easily loses its predictive power. We confirm our theory by showing that indeed the welfare cost of deviation is so large for zoning failures that it is not only efficient but also safe to use tax instruments rather than the LURs. Ironically, the welfare costs arise even when policies are optimally set on average. Moreover, we find that imprecise LURs can even be worse than doing nothing at all as a means to curb congestion.

The rest of the paper is organized as follows. Section 2 presents the analytical model. Two types of infrastructure externalities are modeled: road and nonroad congestions. We characterize the first-best and second-best instruments in Section 3 and 4, respectively; we also discuss the optimal adjustment of land uses. In Section 5, we conduct numerical simulations and in Section 6 explore the consequences of deviations from optimal policies including Monte Carlo experiments. Section 7 concludes. Additionally, a glossary appears at

the end of the paper.

2. The Model

A closed metropolitan area is composed of an arbitrary number of zones. The lots inside a zone are treated identically for the purpose of residence, production, and travel. We measure distance by the distance between zone centroids. Thus, we abstract from intrazonal distance issues. Households choose where to live and work; firms choose the location of operations. Cross commuting occurs because residences and jobs are intermingled over the metropolitan area. For simplicity, we assume no trade of composite goods between zones, because the trade introduces no real difference to our analysis.

2.1. Builders and composite good producers

Builders in zone i construct office buildings using land Q_i^B and capital X_i^B according to a constant returns to scale technology $B_i = B_i(Q_i^B, X_i^B)$. The output B_i is a proxy for the structure services that users of a building enjoy and are measured by the office building's floor area. A builder has a lot in which the maximum FAR allowed is $\overline{f}_i^B$, which means $B_i \le \overline{f}_i^B Q_i^B$. In other words, imposing a maximum FAR in a zone means that the planner intends to lower the zone's market FAR below this maximum level. Let the unit land and capital rents be r_i, p_i^X, respectively. The Lagrangian of the cost minimization problem facing office builders in zone i is

$$\mathcal{L}_B = r_i Q_i^B + p_i^X X_i^B + \mu_i^B \left(B_i - B_i(Q_i^B, X_i^B) \right) + \overline{\lambda}_i^B \left(B_i - \overline{f}_i^B Q_i^B \right), \tag{1}$$

where $\mu_i^B > 0$ is the marginal cost of B_i when the FAR regulation is not binding, and $\overline{\lambda}_i^B \ge 0$ is the marginal compliance cost of the FAR regulation in zone i. When $\overline{f}_i^B$ is set sufficiently high, the FAR regulation is not binding and the problem reverts to the standard cost minimization problem. Similarly, the planner may set the minimum FAR $\underline{f}_i^B$, which requires $B_i \ge \underline{f}_i^B Q_i^B$.

There is a second type of builder, who is known as a housing builder. Denote the

total floor area of housing in zone i by H_i. Similar to the office builders, the housing builders use land Q_i^H and capital X_i^H to produce housing H_i, and the technology obeys constant returns to scale. We denote the unit rental price of housing traded in zone i by p_i^H and measure housing services by the floor area.

Firms produce composite good X_i using building services B_i and labor M_i. $X_i = S_i^X x(B_i, M_i) \equiv S_i^X x_i$, where $x(B_i, M_i)$ is homogenous of degree one in inputs and S_i^X is a multiplier, external to each firm, showing the service quality of nonroad infrastructure in zone i. We set the service level as $S_i^X \equiv S^X(B_i, H_i, K_i)$, where $\partial S_i^X / \partial B_i < 0$, $\partial S_i^X / \partial H_i < 0$ (congestible) and $\partial S_i^X / \partial K_i > 0$, where K_i is capacity of nonroad infrastructure which is under the planner's control. For simplicity, we assume that both firms and households use the same nonroad infrastructure. When land use is mixed, this approximation is not a wholly unrealistic assumption. Road congestion is an important class of infrastructure congestion, but we deal with road congestion through commuters' congested travel. Compared with smaller buildings, larger buildings are correlated with more output level, so they consume more services provided by local infrastructure and produce more network congestion from business trips and trade. The maximum FAR regulations, motivated partly by these planning concerns, target the proxy variables for their control in practice.

The city collects taxes for financing congestible infrastructure. Specifically, the *X*-good firm's profit maximization problem is

$$\max_{X_i, B_i, M_i} p_i^X X_i - (1+\tau_i^B) p_i^B B_i - w_i M_i, \tag{2}$$

where τ_i^B is the tax rate, and price terms are defined in an obvious way. In this way, we assume that one unit of the *X*-good is converted to one unit of capital input X_i^B in (1) for building construction.

At firms optimum we can derive a useful differential equation.

$$X_i dp_i^X = p_i^B B_i d\tau_i^B + (1+\tau_i^B) B_i dp_i^B + M_i dw_i - p_i^X x_i(\cdot)\left(\frac{\partial S_i^X}{\partial B_i} dB_i + \frac{\partial S_i^X}{\partial K_i} dK_i \right) \tag{3}$$

In the absence of tax and infrastructure congestion, $\tau_i^B = dS_i^X = 0$, so that we have $X_i dp_i^X = B_i dp_i^B + M_i dw_i$, a familiar general equilibrium equation relating inputs, output, and

their prices to one another.

2.2. Households

By household (i,j), we mean the representative household living in zone i and working in zone j. We differentiate types of households by commuting arrangements (i,j). For a given residence–work zone pair (i,j), the utility maximization problem of household (i,j) is

$$\max_{z_{ij},h_{ij},l_{ij}} u(z_{ij},h_{ij},l_{ij},S_i^H)+\epsilon_{ij}, \tag{4}$$

$$\text{subject to } \; p_i^X z_{ij}+(1+\tau_i^H)p_i^H h_{ij}=(8w_j-t_{ij})d_{ij}+D_{ij}, \;\; T=(8+g_{ij})d_{ij}+l_{ij},$$

where

$$D_{ij}\equiv\frac{1}{N}\Bigg[\underbrace{\sum_i r_i A_i}_{\text{land rent}}+\underbrace{\sum_i\left(t_i F_i+\tau_i^H p_i^H H_i+\tau_i^B p_i^B B_i\right)}_{\text{taxes returned}}-\underbrace{\sum_i\left(r_i R_i+p_i^X K_i\right)}_{\text{infrastructure costs}}\Bigg]+\underbrace{\left(y_{ij}-\textstyle\sum_{ij}P_{ij}y_{ij}\right)}_{\text{net income transfer}}, \tag{5}$$

$$S_i^H\equiv S^H(H_i,B_i,K_i).$$

z_{ij} denotes the composite good X consumed by household (i,j), p_i^H is the unit rental price of housing in the residence zone i, and h_{ij} is the amount of household (i,j)'s consumption of housing measured by floor area. The subscripts of the other variables are interpreted in the same way. The tax rate τ_i^H is charged on housing consumption. t_{ij} are the traffic congestion charges collected from households commuting between zone i and j. $S_i^H(.)$ is the service level of local infrastructure as rated by households and is a function of a zone's floor areas (housing and office buildings) as well as the capacity of nonroad infrastructure in that zone. Household (i,j) commutes d_{ij} days a month and work eight hours a day, while being paid w_j dollars an hour at work zone j. Each household is endowed with T hours a month, which it allocates for commuting $g_{ij}d_{ij}$, leisure l_{ij}, and working $8d_{ij}$. g_{ij} is the daily commuting time between the two zones (i,j).

Households own equal shares of the entire land in the metropolitan area, and the land rent collected is distributed equally. The metropolitan government collects taxes, uses them

for financing infrastructure, and returns what remains to households. t_i is the traffic congestion toll for cars on zone i's roads; F_i is zone i's traffic volume. Nonlabor income D_{ij} shows the fiscal arrangement to be analyzed. The planner uses the lump-sum instrument y_{ij} to equalize the MUIs of heterogeneous households in the potential first-best regime. If there is a budget deficit, the head tax is collected. P_{ij}, which we explain in the next paragraph, is the share of household type (i, j) among the fixed total population N.

The random utility term ϵ_{ij} is an i.i.d. Gumbel variate with mean zero and dispersion parameter ζ. The probability that a household most prefers zone i and j as its home–work zone pair is given by $P_{ij} = \exp \zeta V_{ij} / \sum_{mn} \exp \zeta V_{mn}$, where V_{ij} is the indirect utility of household (i, j). We measure the welfare W of residents by the expected value of the maximized utilities of the households in this metropolitan area (McFadden, 1974; Small and Rosen, 1981; Anas and Rhee, 2006):

$$W = E\left[\max_{ij}(V_{ij} + \epsilon_{ij})\right] = \zeta^{-1} \ln \sum_{ij} \exp \zeta V_{ij} \,. \tag{6}$$

Households in our model are heterogeneous because they are differentiated by their tastes for the matched pair of home-work zones. The social welfare function is a nonlinear sum of individual utilities. Residential sorting by heterogeneous households is widely observed in metropolitan areas, and the social welfare function (4) is one way of incorporating this heterogeneity. Because households differ inherently in tastes, any policy necessarily has distinct differential impacts on each type of household, and the social planner weighs the impacts using the welfare function of the heterogeneous households' differentiated evaluations.

2.3. Market equilibrium conditions

The left-hand sides of the following market clearing conditions represent demand and the right-hand sides show corresponding supply.

Building markets, Housing: $\sum_j NP_{ij} h_{ij} = H_i$ (7)

Office: input demand of X-good firms = B_i (8)

X-goods markets: $\sum_j NP_{ij} z_{ij} + X_i^B + X_i^H + K_i = X_i$ (9)

Labor markets: $M_i = \sum_j NP_{ji}(8d_{ji})$ (10)

Land markets: $Q_i^H + Q_i^B + R_i = A_i$ (11)

Further equilibrium conditions are:

Three types of zero profits, Housing and office builders and *X*-good firms (12)

In (11), we set road capacity R_i equal to the land area allocated to roads while zonal land areas are fixed at A_i.

All eight equalities exist in each zone i; the unknowns are five prices in each zone, $r_i, p_i^X, p_i^H, p_i^B, w_i$, and three outputs, H_i, B_i, X_i, in each zone. We do not list the capital inputs, X_i^H, X_i^B, of the housing and office builders as unknowns. Once the outputs B_i, X_i are known, the capital inputs are given by the input demand in the relevant markets. Because we have the same number of equations and unknowns, we should be able to solve for all the unknowns. We can express all variables of the system as functions of these eight unknowns and policy variables in each zone.

3. The Potential First-Best Regime

3.1. Theory: the first-order rate of welfare change

The planner maximizes welfare $W(\cdot)$ with respect to Pigouvian charges, infrastructural capacities $\left\{t_i, \tau_i^H, \tau_i^B, R_i, K_i\right\}_{i=1}^{\#zones}$, and income redistribution $\left\{y_{ij}\right\}_{\forall ij}$:

$$\max W(\{t_i, \tau_i^H, \tau_i^B, R_i, K_i\}_{i=1}^{\#zones}, \{y_{ij}\}_{\forall ij})$$

subject to the market equilibrium conditions and the public budget constraint.

A familiar approach is to form a Lagrangian with all the important constraints, including the market equilibrium conditions, combined by multipliers. However, this does not work because of the sheer complexity of the problem. Instead, we differentiate W with respect to a policy variable and later incorporate the equilibrium conditions.

Because the equation system (7)–(12) is composed of eight types of unknowns

specified above, we can write the first-order derivative with respect to a fiscal instrument $\phi \in \left\{ \left\{ t_k, \tau_k^H, \tau_k^B \right\}_{k=1}^{\#zones}, \left\{ y_{ij} \right\}_{\forall ij} \right\}$ as follows:

$$\begin{aligned} \frac{dW(.)}{d\phi} = & \sum_{ij} P_{ij} \sum_n \frac{\partial V_{ij}}{\partial p_n^X} \frac{dp_n^X}{d\phi} + \sum_{ij} P_{ij} \frac{\partial V_{ij}}{\partial w_j} \frac{dw_j}{d\phi} + \sum_{ij} P_{ij} \sum_{n=1,2} \frac{\partial V_{ij}}{\partial p_n^H} \frac{dp_n^H}{d\phi} + \sum_{ij} P_{ij} \sum_{n=1,2} \frac{\partial V_{ij}}{\partial p_n^B} \frac{dp_n^B}{d\phi} \\ & + \sum_{ij} P_{ij} \sum_{n=1,2} \frac{\partial V_{ij}}{\partial r_n} \frac{dr_n}{d\phi} + \sum_{ij} P_{ij} \sum_{n=1,2} \frac{\partial V_{ij}}{\partial H_n} \frac{dH_n}{d\phi} + \sum_{ij} P_{ij} \sum_{n=1,2} \frac{\partial V_{ij}}{\partial B_n} \frac{dB_n}{d\phi} + \sum_{ij} P_{ij} \frac{\partial V_{ij}}{\partial \phi}. \end{aligned} \tag{13}$$

We can derive, say, $\partial V_{ij} / \partial p_i^X$ (a derivative of indirect utility in the price of the *X*-good produced and sold in zone i) in the first term by applying the envelope theorem to the utility maximization problem.[5]. The result is $\partial V_{ij} / \partial p_i^X = -c_{ij}^M z_{ij}$. This term contains household (i, j)'s MUI, c_{ij}^M. n runs from 1 to 2, which means there are two zones in the city. This is arbitrary, and the formula holds for the city with an arbitrary finite number of zones.

On the other hand, (13) has price and quantity derivatives such as $dp_i^X / d\phi, dw_j / d\phi$, $dH_n / d\phi$. The differential equation (3) links these terms. Using the market equilibrium conditions (7)–(12), we can simplify (13) to the following intuitive form (Appendix 1):

(a) (b)

$$\begin{aligned} \frac{N}{c^M} \frac{dW}{d\phi} = & \sum_i \left(\overbrace{p_i^X x_i \frac{\partial S_i^X}{\partial B_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial B_i}}^{\text{MEC from more } B_i,\ (-)} + \tau_i^B p_i^B \right) \frac{dB_i}{d\phi} + \sum_i \left(\overbrace{p_i^X x_i \frac{\partial S_i^X}{\partial H_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial H_i}}^{\text{MEC from more } H_i,\ (-)} + \tau_i^H p_i^H \right) \frac{dH_i}{d\phi} \\ & + \sum_{i=1,2} \left(\bar{w}_i F_i g_i' - t_i \right) \left(-\frac{dF_i}{d\phi} \right) - \sum_{ij} N y_{ij} \frac{dP_{ij}}{d\phi} \quad + \text{Cov.} \end{aligned} \tag{14}$$

(c) (d) (e)

The welfare change is composed of five elements: (a) the Pigouvian term for marginal congestion costs of office services, (b) the Pigouvian term for marginal congestion costs of housing services, (c) the Pigouvian term for marginal road congestion, (d) the distortion term of spatial allocation, and (e) the covariance term due to heterogeneity of households. This is

[5] The standard approach is to apply Roy's identity, which is rather inconvenient in our general equilibrium setup.

the basic formula whose components we discuss below.

Symbols in (14) are defined as follows: $c^M \equiv \sum_{ij} P_{ij} c_{ij}^M$ is the average MUI, N_i is the number of zone i's residents, $g_i \equiv g\left(F_i, R_i\right)$ is the travel cost in zone i, $\upsilon_{ij} \equiv \left(\partial u_{ij} / \partial S_i^H\right) / c_{ij}^M$, and $\upsilon_i \equiv \sum_j \left(NP_{ij} / N_i\right) \upsilon_{ij}$ (zone i households' average marginal utility of nonroad infrastructure service S_i^H in monetary terms). The traffic in zone i, F_i, is

$$\begin{aligned} F_i = & \text{ within-zone trips + passing-through trips + incoming trips + outgoing trips} \\ = & F_{ii} + \sum_{\forall n,m \neq i} (F_{nm} + F_{mn}) + \sum_{\forall n \neq i} F_{ni} + \sum_{\forall n \neq i} F_{in}, \end{aligned}$$

where F_{ij} is the traffic volume from zone i to zone j. Cov is defined as follows:

$$\text{Cov} \equiv \frac{N}{c^M} \sum_{ij} P_{ij} \left(c_{ij}^M - c^M\right) \left(A_{ij} - \underbrace{\sum_{ij} P_{ij} A_{ij}}_{\text{Expected value of } A_{ij}} \right), \tag{15}$$

where

$$\begin{aligned} A_{ij} \equiv & (-z_{ij}) \frac{dp_i^X}{d\phi} + 8 d_{ij} \frac{dw_j}{d\phi} + (1+\tau^H)(-h_{ij}) \frac{dp_i^H}{d\phi} - \bar{w}_{ij} d_{ij} \frac{dg_{ij}}{d\phi} \\ & + \upsilon_{ij} \frac{\partial S_i^H}{\partial H_i} \frac{dH_i}{d\phi} + \upsilon_{ij} \frac{\partial S_i^H}{\partial B_i} \frac{dB_i}{d\phi} - d_{ij} \delta_k^{ij}, \end{aligned} \tag{16}$$

with $\delta_k^{ij} = 1$ when the origin–destination zone pair (i,j) contains zone k, and zero otherwise. When the policy instrument is y_{ij}, δ_{ks}^{ij} is added to (14), where $\delta_{ij}^{ks} = 0$ for $(i,j) = (k,s)$, and zero otherwise. A_{ij} is the change in consumer surplus due to the effects of changes in prices and externalities on household type (i,j) induced by fiscal instrument ϕ. The first term in (14) shows the consumer surplus change in zone i's X-good market triggered by the price change in the X-good price, p_i^X. $\bar{w}_{ij}$ is household (i,j)'s value of time. Because A_{ij} distributes with population probabilities $\left\{P_{ij}\right\}$, and the two multiplied terms in (15) are weighted by these probabilities, Cov is the covariance of individual MUIs c_{ij}^M and the rate of consumer surplus change A_{ij} that a policy change $d\phi$ engenders.

When the policy instruments are infrastructure capacities, new terms are added to (14):

$$p_k^X x_k \frac{\partial S_k^X}{\partial R_k} + N_k \upsilon_k \frac{\partial S_k^H}{\partial R_k} - r_k \quad \text{for road capacity} \quad \phi = R_k \tag{17}$$

$$p_k^X x_k \frac{\partial S_k^X}{\partial K_k} + N_k \upsilon_k \frac{\partial S_k^H}{\partial K_k} - p_k^X \quad \text{for nonroad capacity} \quad \phi = K_k . \tag{18}$$

Infrastructure capacity is expanded until the marginal expansion costs (last term in (18)) are equal to the reduced marginal congestion cost (first two terms in (18)). The same applies to road infrastructure in (17).

3.2. Discussion

3.2.1. Heterogeneity and the covariance term

Although some other studies using a spatial model present similar covariance terms (e.g., Hirte and Tscharaktschiew, 2013; Rhee et al., 2014), thus far, its consequences have not been fully explored. We can understand the role of heterogeneity using a simplified model. Suppose that there is only one good x whose unit price is p. This p is endogenous in our system, which is again a function of a policy variable ϕ. Thus, expressing the indirect utilities as $v_i\left(p(\phi)\right)$ (i is the type of households), we can write the welfare function as $W\left(v_1\left(p(\phi)\right), v_2(\cdot), \cdots, v_n(\cdot)\right)$, while omitting all the other variables for the sake of intuition (refer to the Appendix for the more detailed derivation). Then,

$$\Delta W \approx \frac{dW}{d\phi}\Delta\phi = \left(\sum_{i=1,\cdots,n} \frac{dW}{dv_i}\frac{dv_i}{dp}\frac{dp}{d\phi}\right)\Delta\phi \approx \sum_{\forall i} \frac{dW}{dv_i}\underbrace{\frac{dv_i}{dp}\Delta p}_{\approx \Delta v_i} = \sum_{\forall i} \frac{dW}{dv_i} c_i^M(-x_i \Delta p) \tag{19}$$

where dW/dv_i c_i^M is called the social MUI of household type $i \in 1,\cdots,n$ and x_i is the amount of good x consumed by type i household. The last equality follows from Roy' identity. It reveals the link between the welfare change and the change in consumer surplus

measured by the area below the inverse demand function.

If only the despot's utility matters, the covariance term does not arise from (19) simply because there is no such thing as variation in the MUIs. Next, suppose that we measure the social welfare by the area below the market demand curve. This means that we measure welfare (both individual and social) by monetary units with a uniform social MUI $dW/d\upsilon_i = c_i^M = \bar{c} = 1$. Then, (19) becomes

$$\Delta W \approx \sum_i \frac{dW}{dv_i} c_i^M (-x_i \Delta p) = \sum_i (-x_i \Delta p).$$

This expression does not have the covariance term[6].

Divide both sides of (19) by $\Delta\phi$ and take its limit. Then, expand and rearrange the terms.

$$\frac{dW}{d\phi} = c^M \sum_i \frac{dW}{dv_i}\left(-x_i \frac{dp}{d\phi}\right) + \sum_i \frac{dW}{dv_i}(c_i^M - c^M)\left(-x_i \frac{dp}{d\phi}\right), \qquad (20)$$

(a) (b)

where c^M is the expected value of $c_i^M = \sum_i P_i c_i^M$. When the function W is given by (6), we obtain $dW/dv_i = P_i$, the proportion of type i households. Then, (a) in (20) represents the terms (a)+(b)+(c) in (14).

It remains to explain how the covariance formula in (14) might be derived from (b) in (20). Since $dW/dv_i = P_i$, a simple algebra shows that the last term of (20) is nothing but the covariance.

$$\begin{aligned}\sum_i \frac{dW}{dv_i}(c_i^M - c^M)\left(-x_i \frac{dp}{d\phi}\right) &= \sum_i P_i (c_i^M - c^M)\left(-x_i \frac{dp}{d\phi}\right) - \xi \sum_i P_i (c_i^M - c^M) \\ &= \sum_i P_i (c_i^M - c^M)\left(-x_i \frac{dp}{d\phi} - \xi\right).\end{aligned},$$

The first equality holds for all real numbers ξ, because $\sum_i P_i (c_i^M - c^M) = \sum_i P_i c_i^M -$

[6] Arnott and Krauss (1998) analyze the marginal cost pricing in the presence of heterogeneous facility users, but there does not arise the covariance term. In fact, they measure the welfare by the area under the demand curve.

$(\sum_i P_i)c^M = c^M - (1)c^M = 0$. Finally, set $\xi \equiv E(-x_i[dp/d\phi]) = \sum_i P_i\,(-x_i[dp/d\phi])$ to have the covariance formula in (15).

On the basis of the discussion, we can state the following proposition:

Proposition 1: *The covariance term could arise if both of the following two conditions are met.*

Condition (a) *The welfare function contains an array of indirect utilities of heterogeneous households.*

Condition (b) *The social marginal utilities of income (social MUIs) differ between household*s.

Remark We should be careful about the precise meaning of this proposition. Let I_{ij} be the income of household (i,j). From (6), $\partial W/\partial v_{ij} = P_{ij}$, so that household (i,j)'s social MUI is $(\partial W/\partial v_{ij})(\partial v_{ij}/\partial I_{ij}) = P_{ij}c_{ij}^M$. When the social MUIs differ (i.e., there are some pairs $(i,j),(k,s)$ such that $P_{ij}c_{ij}^M \neq P_{ks}c_{ks}^M$), $c_{ij}^M = c^M$ is not guaranteed, so $P_{ij}(c_{ij}^M - \mathrm{c}^M) \neq 0$ and the covariance term does not necessarily vanish. This is why we added a qualifier "could" in the first line of the proposition. At the same time, however, social MUIs all equal does not necessarily imply zero covariance (i.e., Cov=0) as well, because $P_{ij}c_{ij}^M = P_{ks}c_{ks}^M$ for different $(i,j),(k,s)$ pairs does not necessarily imply $c_{ij}^M = c_{ks}^M$, in which case $c_{ij}^M = c_{ks}^M = \bar{c}$ with Cov=0. ■

3.2.2. Heterogeneity and the breakdown of the conventional first-best instruments

First, we show that household heterogeneity itself is a source of market failure even in the absence of infrastructure congestion. Subsequently, we introduce congestion and further show that the set of first-best instruments known in the literature fails to make the whole terms in (14)-(16) vanish. Now, suppose that the infrastructure is free of congestion, and income redistribution y_{ij} is not yet a policy variable. No congestion means a constant service quality of infrastructure, $\partial S_i^X/\partial B_i = \partial S_i^X/\partial H_i = \partial S_i^H/\partial B_i = \partial S_i^H/\partial H_i = 0$, $g_i' = 0$ at each zone i. Then, (14) reduces to

$$\frac{N}{c^M}\frac{dW}{d\phi} = \sum_i \tau_i^B p_i^B \frac{dB_i}{d\phi} + \sum_i \tau_i^H p_i^H \frac{dH_i}{d\phi} + \sum_i t_i \frac{dF_i}{d\phi} + \text{Cov.} \qquad (21)$$

If there were no covariance term, an efficient solution is achieved at $\tau_i^B = \tau_i^H = t_i = 0$ for all i in (21) (recall that infrastructure is financed by head tax). However, since there is a covariance term in (21), $\tau_i^B = \tau_i^H = t_i = 0$ for all i does not necessarily imply $dW/d\phi = 0$. This demonstrates that the heterogeneity itself is a source of market failure. Markets do not deal with household heterogeneity in an efficient way even in the absence of any other market failures.

Therefore, the presence of heterogeneous households requires an intervention to make the last term of (21) vanish. This is even true in the absence of other market failures. The standard fix of this problem is to use income redistribution y_{ij} to equalize MUIs (De Palma and Lindsey, 2004; Anas, 2012).

To see why this may not work in our spatial model, let us consider the case without any externalities and introduce transfers y_{ij} to equalize MUIs, the standard intervention to efficiently deal with heterogeneity. Head tax revenue is $N\sum_{ij} P_{ij} y_{ij}$ dollars. In this case, it turns out that we should rewrite (21) as

$$\frac{N}{c^M}\frac{dW}{d\phi} = \text{Cov} - \sum_{ij} N y_{ij} \frac{dP_{ij}}{d\phi} \tag{22}$$

Note that (22) contains $-N\sum_{ij} y_{ij}\left(dP_{ij}/d\phi\right)$, a new term which does not exist in (21). In other words, introducing transfers to equalize MUIs is not a device to ensure efficiency because it forces households to relocate, thereby altering the type composition of households.

Now, let us introduce infrastructure congestion with Pigouvian pricing and equalize MUIs through transfers. The first-order rate of welfare change is given by (14), where the second last term, (d), continues to survive under the conventional rule (i.e., marginal cost pricing coupled with redistribution):

$$\frac{N}{c^M}\frac{dW}{d\phi} = -\sum_{ij} N y_{ij} \frac{dP_{ij}}{d\phi}. \tag{23}$$

This raises the question of why the conventional intervention works in De Palma and Lindsey (2004) but fails to work in our model. De Palma and Lindsey (2004) restore market efficiency by using congestion tolls and income redistribution in the transportation market

where travelers "differ with respect to wages, values of travel time, and the congestion characteristics of their vehicles." Despite these differences in heterogeneity, in their setting, redistribution did not alter the composition of heterogeneous travelers, that is, $dP_{ij} \equiv 0$ in our terms. In contrast, the policy intervention in our setting perturbs the commuting arrangements (location pattern), so the type composition of households P_{ij} is altered implying that $dP_{ij} \neq 0$ for all (i, j) pairs. Therefore, Pigouvian congestion pricing and income redistribution only make (a), (b), (c), and (e) in (14) vanish, while leaving the spatial re-sorting term (d) intact. The result is (23). Note here that although $\sum_{ij} dP_{ij}/d\phi \equiv 0$, $\sum_{ij} Ny_{ij}\left(dP_{ij}/d\phi\right) = 0$ does not necessarily follow.

Proposition 2 (Efficiency, heterogeneity and spatial re-sorting)

1) *As a result of household heterogeneity, the free market equilibrium fails to achieve maximum welfare even in the absence of market failures.*
2) *As a result of spatial re-sorting, the conventional rule (i.e., Pigouvian tolls plus redistribution to equalize MUIs) alters the type composition of households and is not guaranteed to be first-best. Consequently, any policy mix is a candidate for the first-best, and only numerical simulations or empirical testing can tell which policy mix is welfare maximizing.*

We can interpret our result in a more general way. In any model where households are heterogeneous but can choose which type of household they belong to, the same problem may arise. There are many examples of this sort: mode, route, vehicle, and housing type choices, as well as marriage, parenting, and education. Hence, our finding should have consequences for many other issues as well.

3.2.3. Other consequences of the heterogeneity in spatial models: Pigouvian tolls

Would the presence of heterogeneity increase or decrease, say, Pigouvian tolls? It is clear from the welfare change formulas that the planner has to adjust the tolls so as to increase the covariance term as far as possible. Using transportation jargon, this clearly implies that Pigouvian tolls are likely to be suboptimal in the presence of "user"

heterogeneity. We illustrate this point using a simple example.

For the sake of exposition, suppose that the city has two zones, where zone 1 is the monocentric center with mixed land use and zone 2 has residences only. Further, assume that $A_{ij} \equiv -\bar{w}_{ij} d_{ij} \left(dg_{ij} / d\phi \right)$ in (16), $\phi = t_2$ (congestion charge at zone 2) and that there is a fixed number of commute trips in zone 1 so that the traffic at the CBD is, say, N (= population of the city). As $dg_{11}/dt_2 = dg_1/dt_2 = 0$, the welfare change associated with the covariance due to the change in t_2 is

$$\text{Cov} \times \Delta t_2 \approx \sum_{ij} P_{ij} \left(c_{ij}^M - c^M \right) \left(A_{ij} - EA_{ij} \right) \Delta t_2 = \sum_{ij} P_{ij} \left(c_{ij}^M - c^M \right) A_{ij} \Delta t_2$$

$$= -P_{21} \left(c_{21}^M - c^M \right) \bar{w}_{21} d_{21} \frac{dg_2}{dt_2} \Delta t_2 \,.$$

$$(-)$$

Workers are paid the same wage in zone 1. However, zone 2 residents pay higher travel costs which lower their value of time and, thus, reduces their labor supply; we expect that zone 2 residents' net income is lower, thus, $c_{21}^M - c^M > 0$. Therefore, a positive toll $\Delta t_2 > 0$ adds more welfare, and the household heterogeneity requires the planner to set the edge zone tolls above the Pigouvian toll. This will result in a more centralized city. However, it is not clear at all in what specific way the heterogeneity affects the policy prescription in the full model, although one finds a statement that heterogeneity tends to have a moderating effect on congestion (De Palma and Lindsey 2004: 135).

3.2.4. Other consequences of the heterogeneity in spatial models: self-financing

Spatial re-sorting $dP_{ij} \neq 0$ also makes the self-financing rule break down in our setting since Pigouvian pricing as well as standard capacity expansion is not optimal. Assume (i) service qualities, S_i^H, S_i^B, $1/g_i$, are further assumed to stay constant with regard to proportionate changes in capacities and patronage,[7] (ii) congestion is fully priced, (iii) income is suitably redistributed so as to make the MUIs all equal across heterogeneous households, and (iv) infrastructure (roads and nonroads) is expanded until the marginal

[7] When simultaneously doubling and tripling the road lanes and the number of cars on roads does not change the travel speed, this condition is met, as is true in the congestion function a la the Bureau of Public Roads.

expansion costs are equal to the reduced marginal congestion cost (Yang and Meng 2002; see (17)-(18) above). In this case self-financing holds, but efficiency is not achieved.

Proposition 3 (Self-financing)

By following the conventional rule of pricing externalities, sizing infrastructure and transferring incomes to equalize MUIs, the infrastructure budget is balanced (i.e., self-financed). However, this arrangement is no longer efficient on account of spatial re-sorting.

Later, we see that the arrangement following the conventional rule could be even disastrous.

3.2.5. Elasticity rule for Pigouvian nonroad tolls

Approximating the covariance term by zero, we can derive an elasticity rule for optimal nonroad tolls. Suppose that cross effects $\partial S_i^B / \partial H_i = \partial S_i^H / \partial B_i = 0$. Then, we can write from (21)

$$\tau_i^B p_i^B = p_i^X x_i \overbrace{\left(-\frac{\partial S_i^X}{\partial B_i} \right)}^{(+)}, \quad \tau_i^H p_i^H = N_i \upsilon_i \overbrace{\left(-\frac{\partial S_i^H}{\partial H_i} \right)}^{(+)},$$

which implies

$$\frac{\tau_i^B p_i^B}{\tau_i^H p_i^H} = \frac{p_i^X S_i^X x_i / B_i}{N_i S_i^H \upsilon_i / H_i} \times \frac{\frac{B_i}{S_i^X}\frac{\partial S_i^X}{\partial B_i}}{\frac{H_i}{S_i^H}\frac{\partial S_i^H}{\partial H_i}} \equiv \frac{\overbrace{S_i^X x_i p_i^X / B_i}^{\text{Zonal output value per floor area}}}{\underbrace{N_i S_i^H \upsilon_i / H_i}_{\text{Zonal infra benefit per floor area}}} \times \frac{\eta_i^X}{\eta_i^H}, \tag{24}$$

where $p_i^X S_i^X x_i / B_i$ is the value of zonal output per office floor area and $N_i S_i^H \upsilon_i / H_i$ is the zonal nonroad infrastructure benefits per housing floor area that N_i households of zone i experience as a whole. (24) has the following implications[8]:

[8] The covariance term as well as the spatial re-location term imply that Pigouvian tolls are not first-best, while

Summary 1 (Elasticity rule) *Assume zero covariance and zero cross effects.*

1) *Property taxes are proportional to the elasticities of the service qualities in floor areas. This proportionality is weighted by the value of affected economic welfare, which are output values in the case of offices and residential service utilities in the case of housing.*
2) *There is no a priori reason to believe that business properties should be taxed more or less than housing.*

4. Analysis of the Second-Best Regimes

4.1. Theory

In light of the first-order welfare change (14), we may refer to the congestion tolls on building structures (i.e., (a) and (b) in (14)) as differentiated property taxes levied on congestible local public services. That is, we can regard property taxes as a type of congestion charge. In the real world, property taxes are not necessarily sufficiently differentiated among property types and localities. To take an extreme example, let $\tau_i^H = \tau_i^B \equiv \tau$ (constant) for all zones i and all property types. Even in this case, however, we expect the welfare performance of this tax scheme not to be poor. The reason is that although the tax rate τ is fixed, the tax bill per unit of floor area, τp_i^H, τp_i^B, is higher in the central business district (CBD) than in suburbs. So, the congestion charge continues to be higher in the CBD than in the suburbs[9].

Concerning zoning, imagine that there is a city called *Zoning City*, where the only instruments available are zoning instruments $\{s_k, \overline{f}_k^H, \overline{f}_k^B, \underline{f}_k^H, \underline{f}_k^B\}_{k=1}^{\#zones}$, which are residential land shares s_k in nonroad land, maximum FARs $\overline{f}_k^H$, $\overline{f}_k^B$, and minimum FARs $\underline{f}_k^H$, $\underline{f}_k^B$ in each zone k. Superscripts H, B denote housing and office (business) buildings, respectively. In the Zoning City, the LURs are fully differentiated over different zones and

---

in some studies where spatial heterogeneity is present Pigouvian terms are said to be first-best (Anas and Xu, 1999; Anas and Rhee, 2006, 2007). Anas (2012) also states that transfers to equalize MUI are usually not consistent with Pigouvian tolling and that the Henry-George Theorem does not hold in this case. Our study reveals that the reason for this outcome is the spatial misallocation resulting from equalization transfers.

[9] We refer readers interested in spatial analysis of the second-best optimal property tax system to other literature (e.g., Kono and Pines, 2013).

property types, and infrastructure is financed by head tax. After modifying the utility maximization problem and market equilibrium conditions in accordance with the Zoning City's setup, the planner maximizes welfare W with respect to the zoning instruments:

$$\max W\left(\left\{s_k,\ \overline{f}_k^{\,H},\ \overline{f}_k^{\,B},\ \underline{f}_k^{\,H},\ \underline{f}_k^{\,B}\right\}_{k=1}^{\#zones}\right) \tag{25}$$

subject to (a) market equilibrium conditions,

(b) fixed infrastructural capacities.

The rates of welfare change with respect to the regulatory instruments are as follows:

$$\frac{N}{c^M}\frac{dW}{ds_k} = \sum_i \overbrace{\left(p_i^X x_i \frac{\partial S_i^X}{\partial H_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial H_i}\right)}^{\text{MEC by an extra } H_i} \frac{dH_i}{ds_k} + \sum_i \overbrace{\left(p_i^X x_i \frac{\partial S_i^X}{\partial B_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial B_i}\right)}^{\text{MEC by an extra } B_i} \frac{dB_i}{ds_k}$$

$$+\sum_i \underbrace{\overline{w}_i F_i \frac{\partial g_i}{\partial F_i}}_{\text{MEC in traffic}} \left(-\frac{dF_i}{ds_k}\right) + \underbrace{\left(r_k^H - r_k^B\right)}_{\substack{\text{Size of} \\ \text{distorted rents}}} A_k + \text{Cov}_k^{LS}, \tag{26}$$

$$\frac{N}{c^M}\frac{dW}{df_k^B} = \sum_i \left(p_i^X x_i \frac{\partial S_i^X}{\partial H_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial H_i}\right) \frac{dH_i}{df_k^B} + \sum_i \left(p_i^X x_i \frac{\partial S_i^X}{\partial B_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial B_i}\right) \frac{dB_i}{df_k^B}$$

$$+\sum_i \overline{w}_i F_i \frac{\partial g_i}{\partial F_i}\left(-\frac{dF_i}{df_k^B}\right) + \underbrace{\left(\overline{\lambda}_k^B Q_k^B \delta^{Max} - \underline{\lambda}_k^B Q_k^B \delta^{Min}\right)}_{\text{Marginal compliance cost}} + \text{Cov}_k^{FARB}, \tag{27}$$

$$\frac{N}{c^M}\frac{dW(.)}{df_k^H} = \sum_i \left(p_i^X x_i \frac{\partial S_i^X}{\partial H_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial H_i}\right) \frac{dH_i}{df_k^H} + \sum_i \left(p_i^X x_i \frac{\partial S_i^X}{\partial B_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial B_i}\right) \frac{dB_i}{df_k^H}$$

$$+\sum_i \overline{w}_i F_i g_i' \left(-\frac{dF_i}{df_k^H}\right) + \underbrace{\left(\overline{\lambda}_k^H Q_k^H \delta^{Max} - \underline{\lambda}_k^H Q_k^H \delta^{Min}\right)}_{\text{Marginal compliance cost}} + \text{Cov}_k^{FARH}. \tag{28}$$

f_k^H could be either the maximum FAR $\overline{f}_k^{\,H}$ or the minimum FAR $\underline{f}_k^{\,H}$ for housing, whichever is applicable. When the maximum FAR regulation is binding, $\delta^{Max} = 1,\ \delta^{Min} = 0$ in the second lines of (27)–(28); when the minimum FAR regulation is binding, $\delta^{Max} = 0,\ \delta^{Min} = 1$. The first three terms in each formula of (26)–(28) are the first three terms in (14) with the Pigouvian taxes set to zero. r_k^H, r_k^B are unit rents of residential and business

land in zone *k*, respectively. The terms before the covariance terms are new terms appearing in the Zoning City. They show the regulatory costs arising from distorted real estate markets. Marginal compliance costs, $\bar{\lambda}_k, \underline{\lambda}_k$, are the Lagrangian multipliers associated with the maximum and minimum FAR regulations, respectively, that appear in (1). We, further, see that the zoning instruments interact with each other via its effects on services (structures) and congestion.

We can discuss the regulatory terms in (26)–(28) a little more. Suppose that starting at a situation with a positive regulative allocation of land to residential use, $s_1 > 0$, more land is allocated to residential use in zone 1 (that is, higher s_1 with $\Delta s_1 > 0$). Then, the residential land rent there will fall c.p. and the business land rent will rise with $r_1^H - r_1^B < 0$. In this case, $(r_1^H - r_1^B)\Delta s_1 = (-)\times(+) < 0$, so this land share adjustment is accompanied by the welfare loss from zone 1's land market. When, starting at $s_1 < 0$ the residential share in zone 1 is adjusted downward instead, we continue to have $(r_1^H - r_1^B)\Delta s_1 = (+)\times(-) < 0$. Similarly, the last terms in (27)–(28) are easily shown to be always negative for a small adjustment of the LURs. The intuition is that every binding regulation causes compliance costs because it forces the producers to adjust their operation in a direction they do not want.

Summary 2 *(Analytical structure)*

1) *Pricing and quantity instruments work on the same variables to reduce congestion: volume of structures and traffic.*
2) *However, the LURs have additional terms associated with the distortionary cost of real estate markets (structures and land).*

4.2. Discussion

4.2.1. Implications for applied models

Before proceeding, we note that (14) and (26)–(28) were derived by adding individual market effects over different markets using the differential equation (3). Finally, all the market effects cancel out, and only those representing market imperfections show up. In this sense, the first-order welfare changes as formulated in (14) and (26)–(28) are an *envelope* result of the general equilibrium land-use-transportation model à la Anas and Kim

(1996). Because the result is of the envelope type, we do not need all the market information to calculate the welfare impact of a policy; we need only the impacts on market imperfections, already present or newly introduced.

Summary 3 (*Envelope result)*

Calculating the general equilibrium welfare change requires only information of market imperfections, existing or newly introduced.

The envelope result has one important implication on the applied research using the land-use and transportation models. One common platform used in this class of models is the spatial input–output framework, in which a regional economy is composed of a closely knit network with interregional and interindustry linkages.[10] One problem with this ambitious modeling effort is the huge and demanding data set required to run the models, which is mostly unavailable. Now, note that the first-order derivatives (14), (26)–(28) have no terms, at least ostensibly, related with interregional and interindustry technical linkages. Reformulating the original production function $X_i = S_i^X x(B_i, M_i)$ so as to consider these linkages, we continue to have the identical derivatives (14) and (26)–(28) due to the zero profits. Therefore, there is some room to compromise the burdensome data requirement with the theoretical rigor of the applied models.

Therefore, we could measure the net welfare gain or loss differently from Echenique et al. (2012) and Jun (2012). For example, Jun (2012) evaluates the welfare impact of the restricted provision of land for factories in the capitol region of Korea. To calculate the production loss, he required regional input–output data. The formulas in (26)–(28) state instead that it is enough to measure the distortionary cost of the real estate markets affected. In the case of land supply restriction, Jun (2012) already has zonal areas A_k (second last term in (26)). Because the only imperfection in his model is traffic congestion, which can be calculated easily in our approach, the only challenging task is to measure the multiplicative term $(r_k^H - r_k^B)$ in (26).

[10] There is voluminous literature on the benefits and costs of smart growth and compact development in the planning field (Ewing and Cervero, 2010; OECD, 2012). However, Echenique et al. (2014) and Jun (2012) are the only studies that, to the authors' knowledge, analyze together the economic costs and benefits of LURs within one modeling framework of regional economies.

4.2.2. Implications in the monocentric city: links to the literature

The existing literature on the FAR regulations is predominantly monocentric with no business land use. To link our discussion to the literature and to gain more insight, imagine a city composed of two zones, where zone 1 (CBD) accommodates both office buildings and housing and zone 2 is completely residential[11]. Furthermore, suppose that the number of workdays d_{ij} is fixed at $\bar{d}$ to be tractable and to fit standard monocentric approaches. Because of the fixed number of working days, zone 1's traffic volume is fixed at $F_1 \equiv N$ at each working day. This is equivalent to the standard monocentric model's setup, in which traffic volume is usually fixed at N at the CBD point.

Assume that LURs $(s_1; \underline{f}_1^B, \underline{f}_1^H; \bar{f}_2^H)$ are the only policies available and that the covariance terms are negligible. Now, imagine that the city planner slightly adjusts land shares and FARs prevailing in the free markets with the aim of improving welfare. Since F_1 is fixed, $dF_1/\mathrm{d}\phi = 0$, $\phi \in \left\{s_1, \underline{f}_1^B, \underline{f}_1^H, \bar{f}_2^H\right\}$ and the associated welfare changes are shown as

$$\Delta W \approx \text{nonroad ext.} \times \Delta s_1 + \bar{w}_1 F_2 g_2' \underbrace{\left(-\frac{dF_2}{ds_1}\right)}_{(+)} \Delta s_1 + \underbrace{\left(r_1^H - r_1^B\right) A_1 \Delta s_1}_{\leq 0}, \tag{27}$$

$$\Delta W \approx \text{nonroad ext.} \times \Delta f_1^H + \bar{w}_1 F_2 g_2' \underbrace{\left(-\frac{dF_2}{df_1^H}\right)}_{(+)} \Delta f_1^H + \underbrace{\left(\bar{\lambda}_1^H Q_1^H \delta^{Max} - \underline{\lambda}_1^H Q_1^H \delta^{Min}\right) \Delta f_1^H}_{\leq 0}, \tag{28}$$

$$\Delta W \approx \text{nonroad ext.} \times \Delta f_2^H + \bar{w}_1 F_2 g_2' \underbrace{\left(-\frac{dF_2}{df_2^H}\right)}_{(-)} \Delta f_2^H + \underbrace{\left(\bar{\lambda}_2^H Q_2^H \delta^{Max} - \underline{\lambda}_2^H Q_2^H \delta^{Min}\right) \Delta f_2^H}_{\leq 0}, \tag{29}$$

$$\Delta W \approx \text{nonroad ext.} \times \Delta f_1^B + \bar{w}_1 F_2 g_2' \left(-\frac{dF_2}{df_1^B}\right) \Delta f_1^B + \underbrace{\left(\bar{\lambda}_1^B Q_1^B \delta^{Max} - \underline{\lambda}_1^B Q_1^B \delta^{Min}\right) \Delta f_1^B}_{\leq 0}. \tag{30}$$

Δs_1 is the discrete change in zone 1's residential land share. The other notations are interpreted in the same way. (27)–(30) reveals the interplay between different types of congestion and different zoning policies. We can use these equations to reproduce different

[11] This setup is more general than most existing models, in which the CBD is described as a point, meaning that the whole metropolitan land is used exclusively for residences.

results from the literature by adjusting assumptions.

For instance, in case of traffic congestion only, the first terms in (27)–(30) vanish. Because the last terms are all negative in (27)–(30), the only way the LURs can improve welfare is to make the second terms in (27)–(30) positive. This immediately implies $\Delta f_1^H > 0$ from (28) (a minimum FAR regulation in the center) and $\Delta f_2^H < 0$ from (29) (a maximum FAR regulation at suburbs). Our assumptions concerning F_2 are $\partial F_2 / \partial f_1^H < 0$ and $\partial F_2 / \partial f_2^H > 0$, i.e. higher buildings in the CBD and lower buildings in suburbs pushes people into the CBD, thus, lowering traffic in suburbs. This reproduces the prescription to metropolitan-wide road congestion in the monocentric literature (e.g., Pines and Kono, 2012; Kono et al, 2012). Our theory shows that this outcome is valid only when the increase in density in the CBD does not cause the service quality of nonroads at the city center to deteriorate (see the first terms in (26)–(28) and (27)–(30)). When such adjustment congests the CBD too much, it could be even better to decentralize activities from the CBD to other zones, as in Anas and Rhee's (2007) case of growth boundaries.

If there is no road congestion, the second terms vanish in (27)–(30). Lowering density by maximum FAR is then the prescription where maximum FAR is tighter in the CBD (see Kono et al. (2010) for "population" externalities). In this way, our model allows evaluating zoning in the presence of two types of congestion. Optimal zoning policy then depends on the relative strength of congestion types as well as on the instrument mix available, as we show next.

4.2.3. Optimal adjustment of land uses

Convert the first-order derivative, say, (26) into a new formula composed of elasticities and shares of external costs as follows:

$$\frac{dW}{ds_k} = \sum_i EC_i^H \varepsilon_{H_i s_k} + \sum_i EC_i^B \varepsilon_{B_i s_k} - \sum_i \underbrace{\bar{w}_i F_i g_i}_{\substack{\text{Total transp.} \\ \text{cost of zone } i}} \varepsilon_{g_i s_k} + \underbrace{ALR_k}_{\substack{\text{Agg. land} \\ \text{rent of zone } k}} \varepsilon_{A_k s_k}, \qquad (31)$$

where $EC_i^H \equiv \left(p_i^X x_i \partial S_i^X / \partial H_i + N_i \upsilon_i \partial S_i^H / \partial H_i \right) H_i$ is the external cost present in zone i's nonroad infrastructure due to the marginal adjustment in housing stock H_i; EC_i^B is the external cost present in zone i's nonroad infrastructure due to the marginal office stock

change B_i; and $ALR_k \equiv [r_k^H s_k + r_k^B(1-s_k)]A_k$ is the aggregate land rent of zone *k*, while ε_{Hs}, ε_{Bs}, ε_{gs}, ε_{As} are the elasticities of housing stock, office stock, congestion and aggregate land rents w.r.t. the share of land usable for housing s_k (land-use type regulation) The coefficient terms before the elasticities in (31) play the role of relative weights when the elasticities are added over all markets and externalities.

The FAR regulations have the similar expression. Denote by ε_{Hf}^H, ε_{Bf}^H, ε_{gf}^H, ε_{Qf}^H the elasticities of the housing stock, office stock, congestion and the office floor area with respect to to the FAR regulation in zone k on housing, respectively. Then, we have

$$\frac{dW(.)}{df_k^H} = \sum_i EC_i^H \varepsilon_{H_i f_k^H}^H + \sum_i EC_i^B \varepsilon_{B_i f_k^H}^H - \sum_i \underbrace{\bar{w}_i F_i g_i}_{\substack{\text{Total transp.}\\ \text{cost in zone } i}} \varepsilon_{g_i f_k^H}^H + \underbrace{\lambda_k^H \overbrace{f_k^H Q_k^H}^{\text{Floor area}}}_{\substack{\text{Total compliance cost of}\\ \text{zone } k\text{'s housing builders}}} \varepsilon_{f_k^H Q_k^H{}_i f_k^H}^H . \quad (32)$$

We can derive a similar formula for the FAR regulation of office buildings, f_k^B.

Summary 4 *(Optimal adjustment of LURs)*

The planner should adjust the LURs, f_k^H, f_k^B, s_k, *so as to make* (31)–(32) *positive, while assigning more weight to the terms with higher cost shares and higher elasticities.*

In summary, Pigouvian tolls as well as property taxes mimicking Pigouvian tolls are only second best because they cannot consider differences in MUIs, and because property taxes fail to account fully for traffic congestion. In addition, land-use regulations entail compliance costs and, thus, cannot be first best even if they were to account for household heterogeneity. Considering the sheer complexity of the model, we cannot decide of whether zoning is as efficient as Pigouvian tolls. To explore this issue, we now turn to numerical exercises.

5. Numerical Examination

5.1. Calibration

We examine a hypothetical metropolitan area, linear in shape, which accommodates a population of 1.2 million in a fully circular nonmonocentric metropolitan area. The

population density is 14 persons/hectare. The population is smaller than mid-sized American metropolitan areas and density is set accordingly. We use a Cobb–Douglas function for the *X*-good producers, $X_i = S_i^X M_i^{\mu} B_i^{1-\mu}$. Housing builders produce housing according to the CES-technology

$$H_k = \left[\alpha_H \left(Q_k^H \right)^{\rho_H} + \left(1 - \alpha_H\right) \left(X_k^H \right)^{\rho_H} \right]^{1/\rho_H} .$$

Office buildings B_k are produced similarly to housing. We use the utility function

$$u_{ij} = \alpha \ln \left[(1-\alpha_U) z_{ij}^{\rho_U} + \alpha_U h_{ij}^{\rho_U} \right]^{1/\rho_U} + \beta \ln l_{ij} + \ln S_i^H .$$

With no harm to the major point of the study, we set the number of workdays d_{ij} at 20.8 days a month.

Table 1 Reference parameters

Geography and Population

Zone 1 & 5: 7 km, Zone 2 & 4: 5 km, Zone 3: 4 km

$N = 1.2$ million persons (2 dependents/household)

Population density: 14.0 persons/hectare on average (endogenous in each zone)

Production

X-good producers: $\mu = 0.8$ (labor cost share),

$1 - \mu = 0.2$ (land cost share)

Builders of housing and office buildings: Land cost share = 30%

$\rho_H = \rho_B =$ -0.923 (elasticity of factor substitution = 0.52)

$\alpha_H = 0.875$, $\alpha_B = 0.915$

Household-workers

Household income = $50,000/year

Housing expenditure = 30% of the household income

Utility function: $\alpha = 0.4$, $\beta = 0.6$, $\rho_U =$ -0.786, $\alpha_U = 0.475$

Time endowment $T =$ 500 hours/month

Number of workdays: $\bar{d} = 20.8$ days/month

$\zeta = 6$ (dispersion parameter)

We adjust the cost shares and elasticities of substitution according to empirical studies and consumer expenditure surveys (Koenker, 1972; Shoven and Whalley, 1977; Polinsky and Ellwood, 1979; McDonald, 1981; Thorsnes, 1997). We specify the service qualities of

infrastructure as $S_i^X = a_X \left[K_i / (B_i + H_i) \right]^{\delta_X}$ and $S_i^H = a_H \left[K_i / (B_i + H_i) \right]^{\delta_H}$. In line with Yeoh and Stansel (2013), we choose $\delta_X \in [0, 0.11]$. More problematic is to set the coefficients of the function S_i^H. We resolve the difficulties in setting the coefficients $a_X, \delta_X, a_H, \delta_H$ in such a way that the uniform property tax rate of 0.95%, as applied to the stock value of properties, covers the cost of infrastructure (roads and nonroads together). We use a discount rate of 5% to convert tax rates into flow rates. We use the Bureau of Public Roads function for the congested travel time, $g(F_i, R_i)$. Table 1 displays the parameters used for the simulations.

Table 2. Technical details of the simulations

(a) Section 5.2

City type	Road budget	Nonroad budget	Road tolls	Prop. tax rates
Base City	Roads and nonroads all financed by head tax		No tolls	No prop. tax
Tolled City	Road tolls exactly cover road budget in each zone.	Prop. taxes exactly cover nonroad budget in each zone.	Endogenous[1]	
First-best City	Road and nonroad tolls charged in the same way as the Tolled City.		Endogenous (income redistribution used too)	
PT City	Prop. tax rev. = sum of metro-wide road and nonroad budgets		No road tolls	Endogenous
Zoning City, LS only	Financed with head tax		Neither road tolls nor prop. taxes	

Note: Capacities of nonroads in the Tolled City are determined by (17)-(18). Capacities of road and nonroads of the other cities follow those of the Tolled City.

(b) Section 5.3

City type	Road budget	Nonroad budget	Road tolls	Prop. tax rates
Base City	Roads and nonroads all financed by head tax		No tolls	No prop. tax
Tolled City	Not necessarily balanced[1]	Prop. taxes exactly cover nonroad budget in each zone.	Policy variable	Endogenous[2]
PT City	Financed with head and prop. taxes; prop. taxes ≠ sum of road and nonroad budgets.		No road tolls	Policy Variable
Zoning City, LS only	Financed with head tax		Neither road tolls nor prop. taxes	

[1] Any budget deficit or surplus of the city is equally shared among the residents in the form of head tax (deficit) or government transfer (surplus).

[2] Although nonroad capacities are fixed here, land rents and the prices of capital inputs are not fixed in our general equilibrium model. So, property tax rates are "endogenous" to cover variable input costs.

5.2. Efficiencies when optimal policies are known

5.2.1. Basic Simulations

We perform simulations to compare the efficiencies of LURs, congestion charges, and uniform property taxes which are often used in reality. We define five types of city: Zoning City, PT (property tax) City, Tolled City, Base City and First-best City.

The *Tolled City* sets Pigouvian congestion charges for roads and nonroads while balancing the road and nonroad budgets in each zone separately and adjusting capacities, too ((17)-(18) hold). In this city, all the terms except for the last two terms in (14) are set equal to zero. In the model with no user heterogeneity, this scheme is not only self-sufficient but also first best. All other types of cities use the same level of road and nonroad capacities as the Tolled City in order to avoid experimental noise that might be introduced by differing infrastructural capacities.

In the *PT City*, traffic congestion is not priced and a single uniform property tax rate is applied to both residential and business properties irrespective of the location of a structure. This rate is set so as to precisely cover the combined expenditure for roads and nonroads in the metropolitan area. In the *Zoning City*, the planner knows the optimal mix of LURs and adjusts residential land shares s_i and FARs f_i^H, f_i^B in each zone to maximize welfare (6). The expenditure for roads and nonroads is financed with a head tax. In the *First-best City*, congestion externalities are fully priced as in the Tolled City. In addition, the last two terms are set to zero as well by income redistribution y_{ij}, meaning that the sum of all the terms in (14) vanishes in the First-best City. The *Base City* is the laissez-faire city where none of the policies is available and infrastructure is financed by head taxes. Table 2(a) summarizes the major features of the cities.

Table 3. Land use patterns of the Base and Zoning Cities under reference population

(a) Land use

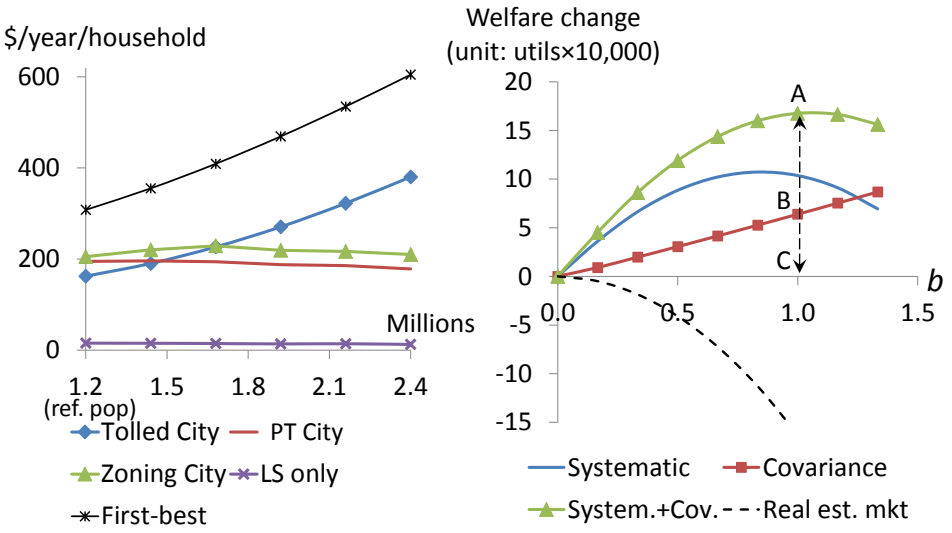
Zone		3 (CBD)	4	5
Share of business land	Base City	0.15	0.32	0.49
	Zoning City	0.14	0.30	0.46
	Tolled City	0.15	0.32	0.49
Share of residential Land	Base City	0.39	0.54	0.50
	Zoning City	0.40	0.56	0.53
	Tolled City	0.39	0.54	0.51

Note: In the Base City, roads' land share at the CBD = 1− (business land 0.15 + residential land 0.39) = 0.46 (=46%). The land for roads in zone 1 and 5 is of pretty small size less than 1% of the total zonal land area.

(b) Floor area ratios

Zone		3 (CBD)	4	5
Business buildings	Base City	0.69	0.49	0.17
	Zoning City	0.70	0.47	0.15
	Tolled City	0.69	0.48	0.17
Residential buildings	Base City	0.66	0.45	0.15
	Zoning City	0.68	0.43	0.13
	Tolled City	0.66	0.44	0.14

Note: The entries for the Zoning City are optimal LURs, strictly binding to every builder and landowner. The FAR numbers here are just indexes and the absolute size carries no meaning.



(a) Welfare gain (b) Welfare loss of imprecise LURs

Figure 1 Welfare performance

Table 3 shows that the Zoning City only minimally adjusts land uses and floor area ratios compared to the Base City. Figure 1(a) shows the welfare gains that various types of city attain over and above the Base City's welfare. Zoning and property taxes are shown to move from being more efficient than to being as efficient as and, eventually, to being less efficient than Pigouvian pricing (i.e., the Tolled City) as metropolitan population increases. This means two things. First, zoning is almost as efficient as Pigouvian tolls. This coincides with the existing literature. For example, Kono et al. (2012) show that FAR regulations in the monocentric city achieve 80% of the welfare gain obtained by the first-best city. According to Rhee et al. (2014), land share adjustment together with production subsidies achieve 99% of the first-best instruments' welfare gain. However, the proposition that zoning is almost as efficient as Pigouvian tolls is valid over only a narrow range of metropolitan population. Second, it is surprising that Pigouvian congestion charges (i.e., the Tolled City) could be less efficient than zoning, which has not been reported in the literature. This cannot occur in the model without heterogeneity.

Summary 5 *(Equivalence of LURs and Pigouvian tolls)*

When optimal policies are known, zoning could be more efficient than, as efficient as, or less efficient than Pigouvian tolls. That is, in contrast to the finding in the theoretic literature that zoning is very efficient (Pines and Sadka, 1985; Wheaton, 1998; Kono et al., 2012; Rhee et al., 2014), zoning could be far less efficient than pricing instruments for large cities. Conversely, zoning could be more efficient than congestion charges for small cities.

5.2.2. Discussion: covariance term and real estate market distortions

The covariance term explains the apparent anomalies in Figure 1(a). To see this more clearly, we create a straight path of the LUR policy changes, parameterized by b, from the Base City to the Zoning City, and path integrate the welfare function along this straight line using (26)–(28) (Yu and Rhee, 2013). In Figure 1(b), b=0 is the Base City, and $b = 1$ is the Zoning City. The curve "Systematic" is welfare explained by the first four terms in concert in (26)–(28); the curve "Covariance" is welfare explained by the covariance terms in (26)–(28). The covariance part explains 38% of the welfare gain of the Zoning City over the Base City (i.e., BC/AC=0.38). This share is astonishingly large and is shown to increase even further as

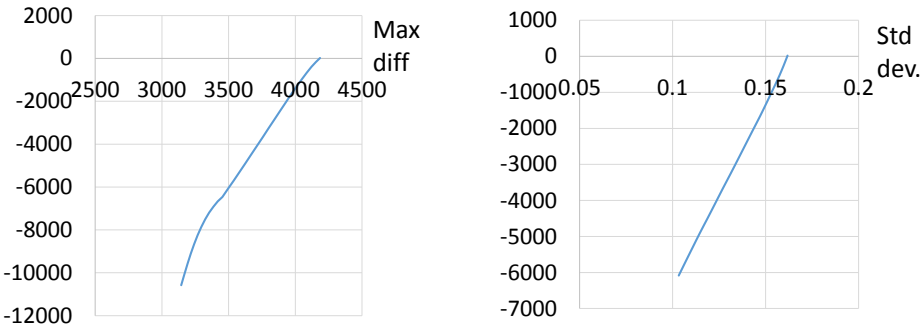
population increases.

How do the anomalies in the equivalence results arise? In smaller cities with lower externalities, zoning improves on redistribution effects in the presence of household heterogeneity while market distortions are relatively small. With increasing city size externalities increase and zoning becomes less and less efficient compared to Pigouvian tolls because redistribution effects become relatively less important and land market distortions or compliance costs increase fast (the dotted curve in Figure 1(b)). The same applies to property taxes. Because the zoning city applies a policy mix that is better than any single zoning instrument our finding holds also for any policy applying only a single zoning instrument as considered in the literature on zoning.

As a reference, we also report a zoning city in which only land shares s_i are adjusted. The "LS only" curve represents that city in Figure 1(a). Comparing the Zoning City curve with the LS Only curve suggests that the efficiency gain of the Zoning City arises mostly from FAR regulations. Figure 1(a) simply shows that any policy not considering household heterogeneity would fail to improve welfare much.

5.2.3. Discussion: breakdown of the conventional rule (spatial distortions)

Household heterogeneity renders the conventional rule for correcting the congestion externalities disastrous. In the nonspatial model where the market intervention does not alter the type composition of user heterogeneity, the combination of congestion charges and income redistribution, i.e., the conventional rule, restores efficient market outcomes. In the Tolled City, we fully priced congestion externalities, so the first three summations in (14), (a)+(b)+(c), vanished. In the First-best City, we also made the *sum* of the remaining two terms in (14), (d)+(e), zero using income transfer y_{ij}. In contrast, the conventional rule makes only (a)+(b)+(c)+(e) vanish in our spatial model, while ignoring the spatial re-sorting term, $\text{(d)} = -\sum_{ij} N y_{ij} \left(dP_{ij} / d\phi \right)$.



Note: Max diff= max income − min income of the households
Std dev.=standard deviation of MUIs × 10,000
Y-axis: welfare loss ($/year/household)

Figure 2 Conventional rule breaks down.

Let us examine how problematic the conventional rule is in our spatial model. As the utility function is Cobb-Douglas, the MUI is the inverse of household income. Building upon this observation, we vary differences in MUIs through redistribution to lower the difference between richer and poorer households in addition to Pigouvian tolling. Figure 2 shows the results when we successively add or subtract a fixed amount of income, say $100 dollars. When there is no re-sorting, the covariance term vanishes and welfare reaches its maximum if income is equal, that is, the maximum income difference is zero. However, as Figure 2 shows, welfare does not increase but declines the smaller this difference. The reason is the relocation of households as a response to redistribution. The highest point in Figure 2 marks the First-best City where the last two terms (e)+(d) vanish on account of redistribution[12]. Figure 2 demonstrates that setting (a)+(b)+(c)+(e) to zero according to the conventional rule, is disastrous when the spatial re-sorting term is ignored in the model where type composition itself varies in response to a policy intervention.

[12] The highest point in the figure does not look like $308, which should be the intercept of the curve First-best in Figure 1(a). The reason is that we did not normalize the household income to $50,000/year for easier computing.

6. Instrument Choice: Efficiencies When Optimal Policies Are Not Known

6.1. Setup

In Subsection 5.2, the planner could pinpoint an optimal policy (first-best or second-best policy). In reality, however, the instrument choice and actual implementation are not free of hindrance and complications due to various reasons: imperfect property rights, multiple externalities, market power, unobservable behavior, imperfect information, administrative capacity, and politics (Fischel, 1985: ch.10; Bennear and Stavins, 2007). Further issues are discussed in the political economy of instrument choice (e.g., influence of lobbying groups by Grossman and Helpman, 1994; overview of transport literature by Hepburn, 2006). Consequently, the instrument level chosen is likely to deviate from the theoretic optimum. In this subsection, we consider this issue and explore its consequences for the optimal instrument design.

To evaluate the welfare impact of deviations from optimal policies, we choose the following experimental setting. Imagine three cities each with 1.2 million inhabitants and the same road and nonroad capacity as the Tolled City in Section 5.2. The first city, *Base City*, uses spatially differentiated property taxes to finance local road and nonroad infrastructure. Neither other taxes nor transfers are available.

In the second city, *Tolled City*, the planner sets road tolls to maximize welfare W in (6), while considering the covariance term and collecting spatially differentiated property taxes to balance nonroad budgets in each zone. As these road tolls take into account the covariance term, the road tolls deviate from Pigouvian tolls; any surplus or deficit from road tolls is redistributed. In this Tolled City, we assume that income transfer is not available. Because the planner sets the road tolls considering the covariance term and ignoring the road budget balance, the welfare is higher than that of Figure 1(a). In Figure 1(a), road and nonroad budgets were all constrained to be exactly balanced in each zone. Next, we vary road tolls in each zone around these second-best tolls that we just obtained, and calculate changes in welfare caused by deviations from the second-best road tolls. Here, the tolls set differ from the standard Pigouvian tolls equaling marginal external cost of congestion[13]

[13] Of course, this higher welfare cannot happen in the model with homogeneous households where balanced budgets and full pricing are two salient features of the first-best city when the service quality stays the same with the proportionate change in infrastructure capacity and patronage.

The third city, *PT City*, levies spatially uniform property taxes but does not charge road tolls. This type of city is observed more commonly in the real world. It varies the tax rates for housing τ^H and business buildings τ^B and uses head tax recycling to maximize welfare. The intention is to check the sensitivity of the welfare change in the setting where the city is minimally constrained in using the property taxes. Table 2(b) summarizes the major features of the city types examined in this subsection.

6.2. Theory

To measure the welfare loss induced by deviations from optimal policies, we parameterize the congestion charges by $b \geq 0$, where $b = 0$ corresponds to the Base City and $b = 1$ to the Tolled City. Similarly, we parameterize the LURs by $\phi > 0$, where $\phi = 1$ corresponds to the Zoning City and $\phi \in (0,1)$ corresponds to the Base City.[14] We perturb b, ϕ around 1 and denote the welfare change by ΔW^{Tolled} for the Tolled City and by ΔW^{Zoning} for the Zoning City. By construction, $\Delta W^{Tolled} < 0$ and $\Delta W^{Zoning} < 0$. We express these differentials up to the second order of the parameters to obtain

$$\Delta W^{Tolled} \approx \frac{1}{2}\frac{d^2 W^{Tolled}}{db^2}\bigg|_{b=1} (\Delta b)^2, \quad \Delta W^{Zoning} \approx \frac{1}{2}\frac{d^2 W^{Zoning}}{d\phi^2}\bigg|_{\phi=1} (\Delta \phi)^2, \tag{33}$$

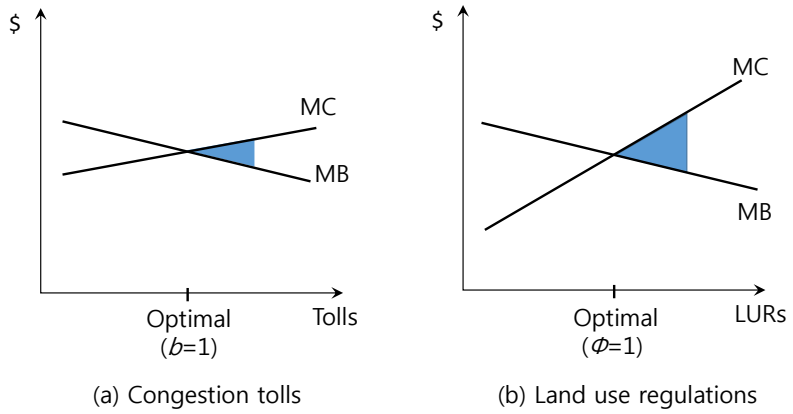
where W^{Tolled}, W^{Zoning} are welfare of the Tolled and Zoning Cities, respectively. We collect the terms from (33) under the rubric of B'' for benefits and C'' for costs to obtain the formulas containing second-order terms similar to Weitzman (1974) and Laffont (1977). Implicit in our exercise is that uncertainties have been introduced into the policy instruments and that welfare is given for certainty. This means that welfare is ex post welfare. Unlike Weitzman (1974), this twist greatly facilitates the exposition, as we shall see.

There are two reasons why (33) is not of much help. First, the general equilibrium nature of our framework is too complex to provide an unambiguous sign from analytics. Second, the policy intervention is discrete, so the local approximation (33) working only near the optimums could be completely wrong for the discrete changes introduced by the Zoning

[14] The usage of ϕ here slightly differs from its usage in Figure 1(b). In Figure 1(b), the Base City corresponds to $\phi = 0$.

and Tolled Cities. Therefore, Weitzman and Laffont's second-order formulas, which hold only around the optimums, do not allow obtaining the welfare differential with discrete changes. We have to rely on alternative ways to evaluate discrete derivations from optimal policies (see Pizer, 2002, for mitigating climate change).

Once again, the first-order welfare changes (14), (26)–(28) show a way to get around this difficulty. One critical difference between (14) and (26)–(28) is that (14) has two terms, one positive and the other negative, mutually offset inside parentheses, but (26)–(28) do not. Because the terms inside parentheses of (14) cancel out at the optimum, the welfare change will be small in association with a given change in b, Δb, away from the optimum ($b=1$). Note that (14) holds at any equilibrium near or far from the optimum. Figure 3 is a graphical illustration consistent with this observation in which the marginal benefit (MB) curve is located close to the marginal cost (MC) curve in panel (a) and far apart from the marginal cost curve in panel (b).

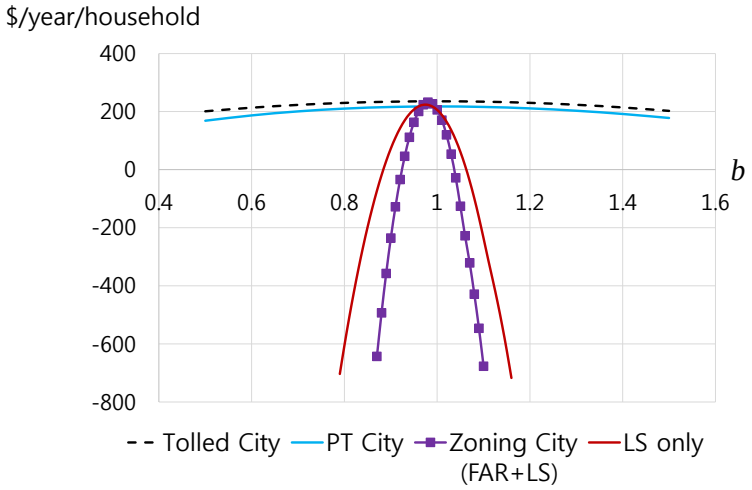


(a) Congestion tolls (b) Land use regulations

Figure 3 Deviation from optimal policies

From previous simulations we can get some intuition. Table 3 shows that the Zoning City only slightly perturbs the Base City, and the dotted curve in Figure 1(b) indicates that when the deviation is large, the welfare cost by the second last terms in (26)–(28) will soon overwhelm all the other effects. Figure 3(b) is a graph consistent with this interpretation. To remind the readers, the two curves are located far apart compared to Figure 3(a). Again, note

that (26)–(28) hold at any point near or far from the optimum. Once we have a figure resembling Figure 3, we are likely to obtain a larger welfare loss in the Zoning City than in the Tolled City for the same given deviation $\Delta b = \Delta\phi$. In the end, we need simulations to see whether we can corroborate this intuition.



Note: The graphs were drawn under the reference population of 1.2 million.

Figure 4 Sensitivity of welfare under imperfect information

6.3. Simulation

We perform simulations to validate the insight provided by (14) and (26)–(28), or equivalently Figure 3. Indeed, Figure 4 confirms our conjecture. It shows new profiles of welfare gain over the Base City under the reference population of 1.2 million. On the horizontal axis, 1.2 means that policies were set 20% higher than the optimal values. The curve labeled "LS only" is the Zoning City, in which only land shares are perturbed, while FARs are fixed at the levels of Subsection 5.2. Conversely, if we perturb the FARs only and hold land shares fixed at the levels of Subsection 5.2, we obtain a shape closely following the "LS only" curve. So, we omit the "FAR only" curve in Figure 4. The city types relying on the tax instruments rarely experience large welfare losses compared to the second-best optimal welfare (approximately $220). It does not matter much whether the planner fails to pinpoint the optimal tax policies. In contrast, when the planner is wrong about the zoning policy, the cost is huge. Because a small deviation could bring about a strictly negative net benefit,

zoning could be even worse than doing nothing.

One problem with Figure 4 is that all the policy instruments are inflated or deflated by the same ratio to obtain the graphs. To explore whether the outcome is robust to other pattern of deviation form optimal policy, we conduct another experiment. This time, we allow actual policies to deviate from the optimal policies by setting differing ratios among changes in the instruments. As pricing instruments are extremely insensitive to deviation (e.g., policy failure or information errors), it is sufficient to check the sensitivity of the LURs against the deviation.

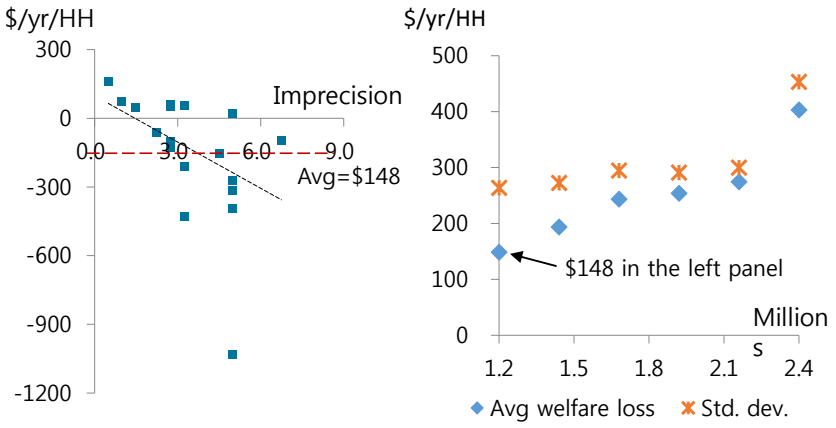
Now, we introduce discrete random variables $\theta_i, i=1,2,3$ that could be either positive or negative. This random variables can take values $\{-0.15;-0.07;0;0.07;0.15\}$ with equal probability of 1/5.[15] We denote the optimal LURs (Zoning City) by s_i^* (residential land share), f_i^{H*} (FAR of housing) and f_i^{B*} (FAR of office buildings). Lacking the precise information of the second-best LURs, the planner adopts actual land-use controls as follows:

$$s_i = (1+\theta_1)s_i^*, \quad f_i^H = (1+\theta_2) f_i^{H*}, \quad f_i^B = (1+\theta_3) f_i^{B*},$$

where θ_i denotes the percentage deviation from the optimal level of the instrument in the Zoning City. We randomly set all θ_i and calculate the associated welfare change from the Base City. We repeat this independently 20 times. At each trial, we measure the imprecision of zoning by $100\sum_{i=1,2,3}\left(\theta_i - \hat{\theta}_i\right)^2$, where $\hat{\theta}_i$ is the average of θ_i's.

In Figure 5(a), we plot the 20 welfare changes resulting from this experiment against the degree of regulatory imprecision. Welfare is higher than the Base City in 7 trials (dots above the *x*-axis) and lower than the Base City in 13 trials (dots below the *x*-axis). Although the planner's policy is correct on average, due to imprecision he loses on average \$148 (Figure 5(a)) instead of winning more than \$200 (Figure 4).

[15] This uniform distribution is not new at all. Refer to Pizer (2002: 415, 417) for a similar practice in estimating the costs and benefits of climate change mitigation.



(a) Welfare gain and loss (b) Varying population

Figure 5 Welfare loss of imprecise zoning

Next, we vary the metropolitan population N, perform the same experiments leading to Figure 5(a), and tabulate the average welfare losses. Figure 5(b) is the result. We calculate the standard deviation of each experiment composed of twenty independent trials and superimpose the standard deviations on the average welfare losses. As in Figure 5(a), the planner always loses on average. The average loss and the standard deviation increase as the metropolitan area becomes more congested. We may narrow the range of the values that θ_i takes on or we may use different (symmetric) distributions for θ_i. Nonetheless, unless the range is much smaller than ± 0.15, we obtain essentially the same results of welfare loss.

The experimental results of the LURs coincide with what has been reported in the literature on metropolitan-wide development controls. For example, building height restriction in Bangalore, India has resulted in a loss of 1.5% to 4.5% of household income (Bertaud and Brueckner, 2005). Similar welfare costs are reported for the greenbelts around metropolitan areas in Seoul, Korea by Lee (1999) and in the UK by Cheshire and Sheppard (2002). One persistent difficulty is that it is not easy to set land use regulations optimally in the real world.

While these studies calculate the costs of one single instrument of zoning, our study shows that a policy mix of zoning also bears costs. We further find that these costs can be more than compensated by benefits near the optimum zoning policy. However, net costs arise

if zoning is set away from the optimum. They increase more than proportionally to the deviation from optimal zoning. These findings are not yet reported in the zoning literature

Summary 6

1) *Depending on policy accuracy, the efficiency ranking of the LURs and tax instruments varies greatly.*
2) *Policy decision-making and the imperfectness of regulatory information needed for LURs or taxes might cause nonoptimal choice of zoning and taxes. The welfare costs of suboptimal zoning are so huge that zoning is strongly adverse compared to Pigouvian tolls and property taxes. It is even possible that zoning is worse than doing nothing at all.*

6.4. Instrument choice in pollution control and city planning

The formulas (14) and (26)–(28) should not come as a surprise, because they are conceivable from Harberger (1971) and are related to the traditional view of instrument choice in environmental economics. Nevertheless, these formulas considerably extend the discussion of existing instrument choice literature; they provide a global platform that extends the instrument choice theory of pollution control to various types of externalities and to a wide range of policy deviations beyond cost-benefit uncertainties.

Strictly speaking, Weitzman's (1974) famous formula $B''+C''$ holds only near the optimums. Many authors adopt this local formula to differentiate the global superiority of different policies (Hoel and Karp, 2002) or resort to graphs to draw global implications and conclusions (Stavins, 1996; Kaplow and Shavell, 2002). Pizer (2002) emphasizes the danger of this practice. He calculates that the welfare differential between prices and quantities is about five times larger with the linearization of Weitzman in contrast to the application of a full welfare analysis.

This practice simply testifies to the difficulty in deriving formal expressions applicable to a discrete policy change. In the process, uncertainties of benefits and/or costs are introduced, only to have the same local characterization of the welfare differential between prices and permits (Laffont, 1977). In contrast, we examine deviations from optimal policies that might be caused by quite a number of different reasons found in the planning and decision process, such as lobbying or information failures. The trick was to incorporate

the uncertainties or imprecision into the policies rather than into the costs and benefits of pollution control.

Let us illustrate how nicely (14) and (26)–(28) explain the superiority of prices to quantities in the instrument choice literature whose setting is much simpler than ours is. Again, Pizer (1997: 2) reports essentially the same simulations as ours in controlling greenhouse gases and summarizes that "slightly more stringent targets lead to dramatic welfare losses." This is not a coincidence. Pizer sets the tax price equal to the almost constant marginal benefit of pollution abatement (Figure 5 in his report, a special case of our nonlinear marginal harm schedules). In addition, permits are bound to entail distortionary costs under uncertainties (the second-terms of (26)–(28) in our setting). Thus, the welfare calculus in Pizer's study has essentially the same mathematical structure as our study of the cost–benefits associated with the two competing instruments, although the mathematical structure is not explicit in his study.

Now, since the instruments solving the equation (14) = 0 are first best and his pricing always equals the true marginal damages due to the flat marginal harm schedule, the flat pricing should work better than permits in Pizer, even under (cost) uncertainties. Because quantities entail not a small cost of market distortions, as the second-last terms in (26)–(28) show, the welfare differential should be large between pricing and permits. In other words, the same cost–benefit calculus applies to Pizer's and our studies; it is natural that pricing works much better than permits in his study and climate control literature in general with stock pollutants, when either one of the instruments is to be used.

Consequently, although (14) and (26)–(28) are intended to be discerning for nonstock externalities, they turn out to apply equally well to stock pollutants. Our theory and simulations show that these findings from environmental policies in principle carry over to congestion policies in a very different spatial framework that considers various nonstock externalities, land use, and heterogeneous households. It is in this sense that (14) and (26)–(28) and Figure 3 provide a global platform that extends the instrument choice theory of pollution control to various types of externalities and to a wide range of policy deviations beyond cost-benefit uncertainties.

Summary 7

1) *The responsibility to control externalities differs between owners of individual facilities in the case of pollution control and public authorities managing infrastructure in our case. However, similar economics is at work, so that* (14), (26)–(28) *are equally applicable to the instrument choice in both environmental economics and city planning.*
2) *The superiority of prices to quantities in controlling externalities stems from the cost–benefit structure of welfare changes, which is uniquely associated with each type of policy.*

7. Conclusion

We found that some of the well-known results on congestion policies no longer hold if there are heterogeneous households, multiple land uses, and different types of congestion. It turns out that redistribution effects can be so large that any policy not considering redistribution is far from being first-best. This result demands further research to explore whether household heterogeneity has similarly strong implications on optimal policy design for other regulation issues. Furthermore, the result that even in a first-best framework Pigouvian taxes and redistribution to equalize marginal utilities of income do not provide a first-best solution, might be relevant for other cases in which re-sorting across heterogeneous household types is feasible (e.g., heterogeneity according to education status, family status, and income type).

The findings concerning zoning are surprising. The outcome from the literature that zoning is almost as efficient as Pigouvian tolls is valid only for a narrow range of a city's population. Furthermore, zoning could be even better than congestion charges. Such findings are not yet known in the literature. The findings point to the significance of redistribution issues when household heterogeneity is present. Consequently, some of the generally accepted findings are challenged. Whether this holds true for other regulation policies is an open issue.

Policies might not be optimally set for various reasons, such as imperfect information, we provide an evaluation of the efficiency losses of deviations from optimal policy. We

conclude from this exercise that enacting zoning is a very risky undertaking because this regulation policy could worsen welfare below the no-policy case with congestion. This issue is serious, considering that zoning is a standard instrument in city and regional planning.

Of course, land-use regulations usually have other objectives too, such as separation of incompatible uses. Our study is limited in this respect. However, even then, the efficiency problem remains and might offset all or part of the benefits that land-use planning intends to reap. The strong distortion that zoning imposes on real estate markets is the reason for the high welfare costs of small deviations from optimal regulation. The same distortionary cost will continue to work in models with other types of externalities, notably, production and consumption externalities, and in models of government competition among municipalities.

We expect that this vulnerability of land-use regulations to non-optimal choice is relevant to policy of other types of regulation applied to various markets. This is already known with respect to misperception (e.g., Kaplow and Shavell, 2002). However, in our study, we show that any deviation for any reason from the optimal regulation causes this robustness problem of the instrument. Accordingly, further research is needed on the local robustness of other regulatory instruments around the optimum. We are confident that this issue is replicated in other markets and regulatory policies. Therefore, it could be a very common issue and should be considered when exploring optimal instrument design.

Acknowledgement

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014S1A5A2A01011212). Authors thank Professor Tony Smith at the University of Pennsylvania for his help in elaborating the idea of the envelope methodology of land-use and transportation models. We thank Professor Matthias Wrede's comments at the GfR Winter Seminar. We also appreciate the comments made by the participants at the ITEA-conference at Oslo, the Real Estate Seminar at the University of Regensburg, seminars at TU Dresden, the winter meeting of the German speaking section of ERSA/RSAI, and the workshops of the Korean Society for Public Administration Theory and the Korea Planning Association.

References

Anas, A., 2012. The optimal pricing, finance and supply of urban transportation in general

equilibrium: a theoretical exposition. Economics of Transportation 1, 64-76.
Anas, A. and Kim, I., 1996. General equilibrium models of polycentric urban land-use with endogenous congestion and job agglomeration. Journal of Urban Economics 40, 232-256.
Anas, A. and Pines, D., 2008. Anti-sprawl policies in a system of cities. Regional Science and Urban Economics 38, 408-423.
Anas, A. and Rhee, H.-J., 2007. When are urban growth boundaries not second-best policies to congestion tolls? Journal of Urban Economics 61, 263-286.
Anas, A. and Rhee, H.-J., 2006. Curbing excess sprawl with congestion toll and urban growth boundaries. Regional Science and Urban Economics 36, 510-541.
Anas, A. and Xu, R., 1999. Congestion, land-use, and job dispersion: a general equilibrium model. Journal of Urban Economics 45, 451-473.
Arnott, R., 2007. Congestion tolling with agglomeration externalities. Journal of Urban Economics 62, 187-203.
Arnott, R., 1979. Unpriced transport congestion. Journal of Economic Theory 21, 294-316.
Arnott, R. and Krauss, M., 1998. When are anonymous congestion charges consistent with marginal cost pricing? Journal of Public Economics 67, 45-64.
Arnott, R. and Stiglitz, J., 1979. Aggregate land rents, expenditure on public goods, and optimal city size. Quarterly Journal of Economics 93, 471–500.
Bennear, L. and Stavins, R., 2007. Second-best theory and the use of multiple policy instruments. Environmental and Resource Economics 37, 111-129.
Bertaud, A. and Brueckner, J., 2005. Analyzing building height restrictions: predicted impacts and welfare costs. Journal of Urban Economics 35, 109-125.
Brueckner, J., 2007. Urban growth boundaries: an effective second-best remedy for un-priced traffic congestion? Journal of Housing Economics 16, 263–273.
Cheshire, P. and Sheppard, S., 2002. The welfare economics of land-use planning. Journal of Urban Economics 52, 242-269.
De Palma, A. and Lindsey, R., 2004. Congestion pricing with heterogeneous travelers: a general equilibrium welfare analysis. Networks and Spatial Economics 4, 135-160.
Echenique, M. et al., Hargreaves, Anthony J., Mitchell, Gordon, and Namdeo, Anil, 2012. Growing cities sustainably. Journal of the American Planning Association 78, 121-137.
Ewing, Reid . et al., 2007. Growing Cooler: The Evidence on Urban Development and Climate Change. Washing-ton, D.C. USA: Urban Land Institute.
Ewing, R. and Cervero, R., 2010. Travel and built environment: a meta-analysis. Journal of the American Planning Association 76(3), 265-294.
Feldstein, M., 1972. Equity and efficiency in public sector pricing: the optimal two-part tariff. Quarterly Journal of Economics 86, 175-185.
Fischel, W., 1985. The Economics of Zoning Laws. Baltimore, U.S.A.: Johns Hopkins University Press.
Flatters, F. et al., 1974. Public goods, efficiency, and regional fiscal equalization. Journal of Public Economics 3, 99-112.
Grossman, G. and Helpman, E., 1994. Protection for sale. American Economic Re-view 84, 833-850.

Hamilton, B., 1975. Zoning and property taxation in a system of local governments. Urban Studies 12, 205-211.
Harberger, A., 1971. Three postulates for applied welfare economics: an interpretive essay. Journal of Economic Literature 9, 785-797.
Hepburn, C., 2006. Regulation by prices, quantities, or both: A review of instrument choice. Oxford Review of Economic Policy 22, 226-247.
Hirte, G. and Tscharaktschiew, S., 2013. The optimal subsidy on electric vehicles in German metropolitan areas: A spatial general equilibrium analysis. Energy Economics 40, 515-528.
Hoel, M. and Karp, L., 2002. Taxes versus quotas for a stock pollutant. Resource and Energy Economics 24, 367-384.
Hoyt, W., 1991. Competitive jurisdictions, congestion, and the Henry George Theo-rem. When should property be taxed instead of land? Regional Science and Urban Economics 21, 351-370.
Joshi, K. and Kono, T., 2009. Optimization of floor area ratio regulation in a growing city. Regional Science and Urban Economics 38, 502-511.
Jun, M.-J., 2012. The economic costs and transportation benefits of Seoul's industrial land use controls, Urban Studies 49, 1791-1810.
Kanemoto, Y., 1977. Cost-benefit analysis and the second best land use for transportation. Journal of Urban Economics 4, 483-503.
Kaplow, L. and Shavell, S., 2002. On the superiority of corrective taxes to quantity regulation. American Law and Economics Review 4, 1-17.
Koenker, R., 1972. An empirical note on the elasticity of substitution between land and capital in a monocentric housing market. Journal of Regional Science 12, 299–305.
Kono, T. and Joshi, K., 2012. A new interpretation of the optimal density regulations: closed and open city. Journal of Housing Economics 21, 223-234.
Kono, T., 2010. Necessity of minimum floor area ratio regulation: a second-best policy. Annals of Regional Science 44, 523-539.
Kono, T. et al., 2012. Optimal regulation on building size and city boundary: an effective second-best remedy for traffic congestion externality. Regional Science and Urban Economics 42, 619-630.
Kono, T. and Pines, D., 2013. Optimal spatial structure of property tax in congested and non-congested monocentric cities. Manuscript available in the web.
Krelove, R., 1993. The persistence and inefficiency of property tax finance of local public expenditures. Journal of Public Economics 51, 416-435.
Laffont, J., 1977. More on prices vs. quantities. Review of Economic Studies 44, 177-182.
Lee, C.-M., 1999. An intertemporal efficiency test of a greenbelt: assessing the economic impacts of Seoul's greenbelt. Journal of Planning Education and Research 19, 41-52.
McDonald, John F., 1981. Capital-land substitution in urban housing: A survey of empirical estimates. Journal of Urban Economics 9, 190–211.
McFadden, D., 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics, edited by Paul Zarembka. New York: Academic Press, 105-142.
OECD, 2012. Compact City Policies: A Comparative Assessment. OECD Green Growth

Studies, OECD Publishing. http://dx.doi.org/10.1787/9789264167865-en.
Oron, Y. et al., 1973. Optimum vs. equilibrium land use pattern and congestion toll. RAND Journal of Economics 4, 619-636.
Parry, I. and Bento, A., 2002. Estimating the welfare effect of congestion taxes: the critical importance of other distortions within the transport system. Journal of Urban Economics 51, 339–365.
Pines, D., 2005. Containing the city boundary as a second-best policy. Unpublished manuscript. Tel Aviv University.
Pines, D. and Sadka, E., 1985. Zoning, first-best, second-best, and third-best criteria for allocating land to roads. Journal of Urban Economics 17, 167-183.
Pines, D. and Kono, T., 2012. FAR regulations and unpriced traffic congestion. Regional Science and Urban Economics 42, 931-937.
Pizer, W., 1997. Prices vs. quantities revisited: the case of climate change. Discus-sion paper 98-02, Resources for the Future.
Pizer, W., 2002. Combining price and quantity controls to mitigate global climate change. Journal of Public Economics 85, 409-434.
Polinsky, A. and Ellwood, D., 1979. "An empirical reconciliation of micro and grouped estimates of the demand for housing." Review of Economics and Statistics 61, 199–205.
Rhee, H.-J. et al., 2014. Zoning in cities with traffic congestion and agglomeration economies. Regional Science and Urban Economics 44, 82-93.
Shoven, J. and Whalley, J., 1977. Equal yield tax alternatives: general equilibrium computational techniques. Journal of Public Economics 8, 211–224.
Small, K. and Rosen, H., 1981. Applied welfare economics with discrete choice models. Econometrica 49, 105-130.
Stavins, R., 1996. Correlated uncertainty and policy instrument choice. Journal of Environmental Economics and Management 30, 218-232.
Stiglitz, J., 1977. The theory of local public goods. In: M. Feldstein, R.P. Inman (eds.), The Economics of Public Services, London: MacMillan, 274-333.
Thorsnes, P., 1997. Consistent estimates of the elasticity of substitution between land and non-land inputs in the production of housing. Journal of Urban Economics 42, 98–108.
Weitzman, M., 1974. Prices vs. quantities. Review of Economic Studies 41, 477-491.
Wheaton, W., 1998. Land-use and density in cities with congestion. Journal of Urban Economics 43, 258-272.
Westerink, J. et al., 2013. Dealing with sustainability trade-offs of the compact city in peri-urban planning across European city regions. European Planning Studies 21(4), 473-497.
Yang, H. and Meng, Q., 2002. A note on "highway pricing and capacity choice in a road network under a build-operate-transfer scheme." Transportation Research Part A 36, 659-663.
Yeoh, M. and Stansel, D., 2013. Is public expenditure productive? Evidence from the manufacturing sector in U.S. cities, 1880-1920. Cato Journal 33(1), 1-28.
Yu, S. and Rhee, H.-J., 2013. A numerical analysis of land-use-transportation model as a form of analytical tool. Journal of Korean Society of Transportation 31(2), 33-44.

Appendix 1. Derivation of the first-order welfare change

Form the Lagrangian of the household's utility maximization problem.

$$L = u_{ij}(.) + c_{ij}^{M}\left((8w_j - t_{ij})d_{ij} + D_{ij} - p_i^X x_{ij} - (1+\tau_i^H)p_i^H h_{ij}\right) + c_{ij}^{T}\left(H - (8+g_{ij})d_{ij} - l_{ij}\right)$$

We calculate each term of (11) one by one. An example follows.

$$\frac{\partial V_{ij}}{\partial p_n^X} = \frac{\partial}{\partial p_n^X}\left[u_{ij}(.) + c_{ij}^{M}\left((8w_j - t_{ij})d_{ij} + D_{ij} - p_i^X x_{ij} - (1+\tau_i^H)p_i^H h_{ij}\right) + c_{ij}^{T}\left(H - (8+g_{ij})d_{ij} - l_{ij}\right)\right]$$

$$= c_{ij}^{M}\left(\frac{\partial D_{ij}}{\partial p_n^X} - \delta_{in}x_{ij}\right) = c_{ij}^{M}\left(\frac{-K_n}{N} - \delta_{in}x_{ij}\right)$$

We have applied the envelope theorem to have the formula. After calculating this type of derivatives, we substitute them into (11).

The first seven terms of (11)

$$= \sum_{ij} P_{ij}\frac{\partial V_{ij}}{\partial w_j}\frac{dw_j}{dt_k} + \sum_{ij} P_{ij}\sum_{n}\frac{\partial V_{ij}}{\partial p_n^X}\frac{dp_n^X}{dt_k} + \sum_{ij} P_{ij}\sum_{n=1,2}\frac{\partial V_{ij}}{\partial p_n^H}\frac{dp_n^H}{dt_k} + \sum_{ij} P_{ij}\sum_{n=1,2}\frac{\partial V_{ij}}{\partial p_n^B}\frac{dp_n^B}{dt_k}$$

$$+ \sum_{ij} P_{ij}\sum_{n=1,2}\frac{\partial V_{ij}}{\partial r_n}\frac{dr_n}{dt_k} + \sum_{ij} P_{ij}\sum_{n=1,2}\frac{\partial V_{ij}}{\partial H_n}\frac{dH_n}{dt_k} + \sum_{ij} P_{ij}\sum_{n=1,2}\frac{\partial V_{ij}}{\partial B_n}\frac{dB_n}{dt_k}$$

$$= \sum_{ij} P_{ij}c_{ij}^{M}(8d_{ij})\frac{dw_j}{dt_k} + \sum_{ij} P_{ij}\sum_{n} c_{ij}^{M}\left(\frac{-K_n}{N} - \delta_{in}x_{ij}\right)\frac{dp_n^X}{dt_k}$$

$$+ \sum_{ij} P_{ij}\sum_{n=1,2}\frac{c_{ij}^{M}}{N}\left(\tau_n^H H_n - N(1+\tau_i^H)\delta_{in}h_{ij}\right)\frac{dp_n^H}{dt_k} + \sum_{ij} P_{ij}\sum_{n=1,2}\frac{c_{ij}^{M}}{N}\tau_n^B B_n\frac{dp_n^B}{dt_k}$$

$$+ \sum_{ij} P_{ij}\sum_{n=1,2}\frac{c_{ij}^{M}}{N}\left(A_n - R_n\right)\frac{dr_n}{dt_k}$$

$$+ \sum_{ij} P_{ij}\sum_{n=1,2}\left(\frac{\partial u_{ij}}{\partial S_i^H}\frac{\partial S_i^H}{\partial H_i}\delta_{in} + \frac{c_{ij}^{M}}{N}\tau_n^H p_n^H\right)\frac{dH_n}{dt_k} + \sum_{ij} P_{ij}\sum_{n=1,2}\left(\frac{\partial u_{ij}}{\partial S_i^H}\frac{\partial S_i^H}{\partial B_i}\delta_{in} + \frac{c_{ij}^{M}}{N}\tau_n^B p_n^B\right)\frac{dB_n}{dt_k}$$

$$= \sum_{ij} P_{ij}\left(c_{ij}^{M} - c^{M} + c^{M}\right)(8d_{ij})\frac{dw_j}{dt_k} - c^{M}\sum_{n}\frac{K_n}{N}\frac{dp_n^X}{dt_k} - \sum_{ij} P_{ij}\sum_{n} c_{ij}^{M}\delta_{in}x_{ij}\frac{dp_n^X}{dt_k}$$

$$+ \sum_{ij} P_{ij}\frac{c_{ij}^{M}}{N}\left(\sum_{n}\tau_n^H H_n\frac{dp_n^H}{dt_k} - \sum_{n} N(1+\tau_i^H)\delta_{in}h_{ij}\frac{dp_n^H}{dt_k}\right) + \frac{c^{M}}{N}\sum_{n=1,2}\tau_n^B B_n\frac{dp_n^B}{dt_k}$$

$$+ \frac{c^{M}}{N}\sum_{n=1,2}\left(A_n - R_n\right)\frac{dr_n}{dt_k} + \sum_{ij} P_{ij}\left(\sum_{n}\frac{\partial u_{ij}}{\partial S_i^H}\frac{\partial S_i^H}{\partial H_i}\delta_{in}\frac{dH_n}{dt_k} + \frac{c_{ij}^{M}}{N}\sum_{n}\tau_n^H p_n^H\frac{dH_n}{dt_k}\right)$$

$$+ \sum_{ij} P_{ij}\left(\sum_{n}\frac{\partial u_{ij}}{\partial S_i^H}\frac{\partial S_i^H}{\partial B_i}\delta_{in}\frac{dB_n}{dt_k} + \frac{c_{ij}^{M}}{N}\sum_{n}\tau_n^B p_n^B\frac{dB_n}{dt_k}\right)$$

Using (3) and the market equilibrium conditions, we can simplify the above as follows:

The first seven terms of (11)

$$= \frac{c^M}{N}\sum_i \left[\left(p_i^X x_i \frac{\partial S_i^X}{\partial B_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial B_i}\right) + \tau_i^B p_i^B\right]\frac{dB_i}{dt_k} + \frac{c^M}{N}\sum_i \left[\left(p_i^X x_i \frac{\partial S_i^X}{\partial H_i} + N_i \upsilon_i \frac{\partial S_i^H}{\partial H_i}\right) + \tau_i^H p_i^H\right]\frac{dH_i}{dt_k}$$

$$+ \sum_{ij} P_{ij}\left(c_{ij}^M - c^M\right)(8d_{ij})\frac{dw_j}{dt_k} - \sum_{ij} P_{ij}\left(c_{ij}^M - c^M\right)x_{ij}\frac{dp_i^X}{dt_k} - \sum_{ij} P_{ij}\left(c_{ij}^M - c^M\right)(1+\tau_i^H)h_{ij}\frac{dp_i^H}{dt_k}$$

$$+ \sum_{ij} P_{ij}\left(c_{ij}^M - c^M\right)\upsilon_{ij}\frac{\partial S_i^H}{\partial H_i}\frac{dH_i}{dt_k} + \sum_{ij} P_{ij}\left(c_{ij}^M - c^M\right)\upsilon_{ij}\frac{\partial S_i^H}{\partial B_i}\frac{dB_i}{dt_k} \qquad (34)$$

In the similar fashion, we can calculate the last term of (11).

The last term of (11) $= \sum_{ij} P_{ij}\frac{\partial V_{ij}}{\partial t_k}$

$$= \sum_{ij} P_{ij}\left[\frac{c_{ij}^M}{N}\left(-Nd_{ij}\delta_k^{ij} + F_k + \sum_{i=1,2} t_i \frac{\partial F_i}{\partial t_k}\right) - c_{ij}^M \sum_{ml}\frac{\partial P_{ml}}{\partial t_k} y_{ml} - c_{ij}^M \bar{w}_{ij} d_{ij}\frac{\partial g_{ij}}{\partial t_k}\right]$$

$$= -\sum_{ij} P_{ij}\frac{c_{ij}^M}{N} Nd_{ij}\delta_k^{ij} + \sum_{ij} P_{ij}\frac{c_{ij}^M}{N}\left(F_k + \sum_{i=1,2} t_i \frac{\partial F_i}{\partial t_k}\right) - c^M \sum_{ml}\frac{\partial P_{ml}}{\partial t_k} y_{ml} - \sum_{ij} P_{ij} c_{ij}^M \bar{w}_{ij} d_{ij}\frac{\partial g_{ij}}{\partial t_k}$$

$$= -\frac{1}{N}\sum_{ij} P_{ij}\left(c_{ij}^M - c^M + c^M\right)Nd_{ij}\delta_k^{ij} + \frac{c^M}{N}\left(F_k + \sum_{i=1,2} t_i \frac{\partial F_i}{\partial t_k}\right)$$

$$- c^M \sum_{ml}\frac{\partial P_{ml}}{\partial t_k} y_{ml} - \sum_{ij} P_{ij}\left(c_{ij}^M - c^M + c^M\right)\bar{w}_{ij} d_{ij}\frac{\partial g_{ij}}{\partial t_k}$$

$$= -\frac{c^M}{N}\sum_{ij} P_{ij} Nd_{ij}\delta_k^{ij} + \frac{c^M}{N}\left(F_k + \sum_{i=1,2} t_i \frac{\partial F_i}{\partial t_k}\right) - c^M \sum_{ml}\frac{\partial P_{ml}}{\partial t_k} y_{ml} - \frac{c^M}{N}\sum_{ij} NP_{ij}\bar{w}_{ij} d_{ij}\frac{\partial g_{ij}}{\partial t_k}$$

$$- \frac{1}{N}\sum_{ij} P_{ij}\left(c_{ij}^M - c^M\right)Nd_{ij}\delta_k^{ij} - \sum_{ij} NP_{ij}\left(c_{ij}^M - c^M\right)\bar{w}_{ij} d_{ij}\frac{\partial g_{ij}}{\partial t_k}$$

$$= \frac{c^M}{N}\sum_{i=1,2}\left(\bar{w}_i F_i g_i' - t_i\right)\left(-\frac{\partial F_i}{\partial t_k}\right) - c^M \sum_{ij} y_{ij}\frac{\partial P_{ij}}{\partial t_k} \qquad (35)$$

$$- \sum_{ij} P_{ij}\left(c_{ij}^M - c^M\right)d_{ij}\delta_k^{ij} - \sum_{ij} NP_{ij}\left(c_{ij}^M - c^M\right)\bar{w}_{ij} d_{ij}\frac{\partial g_{ij}}{\partial t_k}$$

Combine (34) and (35) to have the desired formula. We can derive the derivatives of the other policy variables in the similar fashion.

Appendix 2. Glossary

Abbreviations LUR land-use regulations; FAR floor area ratio; MUI marginal utility of income

Zone index i, j, k, s

Transport $g_i \equiv g_i(F_i, R_i)$ travel time in zone i (hour/km), g_{ij} travel time between zone i and j; t_i, t_{ij} are tolls per kilometer similarly defined; F_i, F_{ij} traffic volumes similarly defined; R_i road capacity in zone i; K_i capacity of nonroad infrastructure such as water, sewage, power, gas, and telecommunications that do not consume land per se

Producers X_i composite goods = *X*-goods; M_i labor input to X-good production; B_i building input to X-good production (unit: floor area); Q_i^H, Q_i^B land inputs for the production of residential and commercial buildings, respectively; X_i^H, X_i^B capital inputs (=capital converted from X-goods by one-to-one) for the production of residential and commercial buildings, respectively; H_i housing units measured in floor area; B_i: business buildings measured in floor area

Household *N* metro population; N_i zone i's residents; household (i, j): the representative household living in zone i and working in zone j; z_{ij} composite consumed by household (i, j); q_{ij} lot size consumed by household (i, j); *T* time endowment; d_{ij} number of commuting days of household (i, j); V_{ij} indirect utility function of the household (i, j); W welfare function = expected maximum utility; ε_{ij} idiosyncratic taste term for home-work zone pairs; λ dispersion parameter of the idiosyncratic terms; u_{ij}, V_{ij} ordinary and indirect utility functions, respectively; c_{ij}^M, c_{ij}^T Lagrangian multipliers of income and time, respectively; c^M weighted average of c_{ij}^M's; $\bar{w}_{ij} \equiv c_{ij}^T / c_{ij}^M$; $\bar{w}_i$ value of time of the travelers in zone i

Prices r_i unit land rent in zone i; w_i hourly wage offered in zone i; p_i^X price of composite good = price of capital inputs to housing and business building producers; p_i^B unit rental price of office floor area; p_i^H unit rental price of housing floor area; $\bar{w}_i$ value of time of zone i's travelers; $\bar{w}_{ij}$ value of time of household (i, j)

Externalities υ_i zone i households' marginal utility of nonroad infrastructure; $S_i^X(B_i, H_i, K_i)$ quality of nonroad infrastructure in zone i as perceived by zone i's X-good producers; $S_i^H(B_i, H_i, K_i)$ quality of nonroad infrastructure in zone i as perceived by households living in zone i

Taxes τ_i^B, τ_i^H property tax rates imposed on one dollar rental price of office buildings and housing units, respectively; t_i, t_{ij} traffic congestion tolls in zone i and for the trips whose O-D is (i, j), respectively; y_{ij} income transfer for household (*i*,*j*)

Regulations s_i residential land share of zone i; $\overline{f}_i^H, \underline{f}_i^H$ maximum and minimum floor area ratios, respectively, imposed on housing units in zone i; $\overline{f}_i^B, \underline{f}_i^B$ maximum and minimum floor area ratio imposed on business buildings in zone i, respectively; $\overline{\lambda}_i^H, \underline{\lambda}_i^H$ marginal compliance cost when maximum and minimum FARs, respectively, are imposed on housing units in zone i; $\overline{\lambda}_i^B, \underline{\lambda}_i^B$ marginal compliance cost when maximum and minimum FARs, respectively, are imposed on office buildings in zone i