Externalities in the Urban Economy

Erik T. Verhoef
Peter Nijkamp

*Department of Spatial Economics, Faculty of Economics and Business Administration,
Vrije Universiteit Amsterdam, and Tinbergen Institute.*

EXTERNALITIES IN THE URBAN ECONOMY

Erik T. Verhoef and Peter Nijkamp[*]
Department of Spatial Economics
Free University Amsterdam
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
Phone: +31-20-4446094
Fax: +31-20-4446004
Email: everhoef@econ.vu.nl

This version: 25/09/03

Abstract

This paper is concerned with the economics of urban externalities. We start by reviewing the literature on urban externalities, and observe that although many interesting contributions have been made, there seems to be sufficient scope and need for further research, both theoretically and empirically. We identify what we believe to be important advances to be pursued in future research on urban externalities. These include (1) the explicit consideration of mutual interactions between externalities; (2) a thorough analysis of the relationship between these externalities and urban form; and (3) a clear focus on (realistic) second-best policies. The importance of these issues is illustrated by developing a simple urban general equilibrium model in which we study the interactions between agglomeration externalities and pollution from commuting. Our results show that seemingly impossible findings from a non-spatial perspective, namely a simultaneous stimulation of agglomeration externalities and a reduction of environmental externalities, is in fact the outcome of first-best policies in our spatial model. Moreover, while the incentives from road pricing and labour subsidies would seem to be perfectly opposite in a non-spatial setting, leaving one of the two instruments redundant, our results show that their welfare effects may, in contrast, turn out to be strongly super-additive when a spatial perspective is taken.

[*]Both authors are affiliated to the Tinbergen Institute, Roetersstraat 31, 1018 WB Amsterdam.

1. Cities in perspective

Over the past hundred years cities have developed into engines for the economic development of our globe and its regions. At present, the city mirrors part of a global network society by acting as a nodal point in an interlinked information and communication configuration (see Castells 1996). But whatever appearance a city may have had in the history of mankind, it has always formed the cradle of civilization. The key role of the city in ancient times is eloquently presented in a fascinating study of Tulleken (1988), when he argues:

"*Yet by 3000 BC, an astonishingly different panorama was unfolding. Along the length of the valley, magnificent cities sprawled on the riverbanks. Around them, fields of grain spread like a tide of fecundity across the once desolated flatlands. Groves of date palms swayed in the wind, offering fruit and strade. Within the massive walls that ringed the cities, temples towered over both streetscape and plain. There were brick places and mansions and street after street of comfortable bouses. People thronged the avenues and marketplaces; in hundreds of workshops artisans turned out all manner of goods, from pottery to sparkling jewelry. On holy days, processions of the worshipful wound through the streets to the temples. What had happened in this land the Greeks later called Mesopotamia, 'between the rivers', was the most crucial event in human history: the birth of civilization*" (p. 1).

Cities are a centre of socio-economic interplay, human confrontation, political dialectics, birthplaces of civilization, centres of science and art, and a melting pot of cultures. According to Jane Jacobs (1969), cities generate economic growth *inter alia* from the disordered order of human interaction. In the urban economics literature, we find the concept of agglomeration advantages, which means that a spatial clustering of economic activities (industries, households, public services) leads to various types of economies of scale, which cannot be generated elsewhere. Sometimes a distinction is made into localization advantages, urbanization advantages, scale advantages, urban externalities and the like. They all point at the fact that a geographic iuxtaposition may lead to win-win situations for all actors involved.

It has to be admitted that city life does not only have positive benefits, but also several disadvantages. A discussion on city life is often witnessing uneasy feelings. O'Sullivan (2000) quotes two authors who express contrasting views on the merits of the city: *"Cities have always been the fireplace of civilization, whence light and heat radiated out into the dark"* (Theodore Parker) and *"I'd rather wake up in the middle of nowhere than in any city on earth"* (Steve McQueen) (p.1). Despite the existence of mixed feelings about the city, the idea that the city is a 'blessing in disguise' is still prevalent. This has also to do with the great variety of roles cities are able to play. We will mention here a few of such important roles of cities, without striving for an exhaustive list.

Shelter role
The city is offering settlement facilities for numerous people, based on its scale advantages in housing many citizens. Shelter has even become a human right, and cities are able to care for the housing needs of people. From this perspective, cities offer a significant contribution to a sustainable human habitat.

Religious role
In the early biblical history the city was often regarded as the source of evil (Babylon, Nineveh). But in the later history we observe a more positive appreciation of the city. Jerusalem was the seat of King David and the New Jerusalem became even a metaphor for a total re-birth of mankind.

Cultural role
Historically, the city was the place where arts and sciences were flourishing. Venice, Bologna, Padua, Paris, Augsburg, Amsterdam and many other places offer an overwhelming evidence of the favourable seedbed conditions of an urban way of life for the advancement of culture.

Political role
Democracy was a new type of governance which found its seedbed in the city. And still nowadays political power is largely concentrated in cities and governments have established their premises in cities. Deconcentration of physical government facilities (*e.g.*, premises) has never become very successful. Administrative functions are usually executed in capital cities of countries, or at least in cities with a critical political mass.

Economic role
The city is the market place for economic activity. It is also the place where usually products are designed and often manufactured. Furthermore, it is a marketplace where capital is supplied and advisory services are offered. In addition, the city is – as a result of various types of agglomeration advantages – a very efficient way of organizing production and consumption.

Social role
Cities house thousands of people who are through the associative nature of city life able to communicate with a great number of others, intensively or less intensively. But they have a social contact and communication spectrum which far exceeds that of a uniform distribution of people.

Engineering role
The city is also the cradle of technological inventions and innovations. It brings together craftsmanship, technical expertise, hardware, software and orgware. As a result, cities are still the breeding places for the genesis of new products and services.

Network role

In an emerging network society cities become more and more the virtual centres of global network forces. The city brings together a *triple-C* potential: communication, competence and creativeness. Despite doomsday scenarios on the 'death of distance' and on the threats to city life, it is more plausible that cities continue to reinforce their role in local, regional, national and international networks.

The manifold strategic functions of the city have also attracted many negative forces which might erode city life. Congestion, pollution, poor health conditions and criminality are examples of phenomena which exert a threat for survival of the modern city. The present paper will, within limitations, focus attention in particular on the non-market dimensions of the urban economy. In other words, we will address in particular the externalities involved with a modern urban system as a nucleus of a network economy. In this context, not only the transport sector, but also urban land use (including residential sites, job places and public or private facilities) deserve a prominent place in our discussion. In evaluating urban polices for such a complex urban system, we have to recognize the feasibility of first-best policy instruments vis-à-vis second-best policy instruments in order to care for some realism in our analysis. This sense of political realities is also prompted by the need to consider explicitly socio-economic and spatial equity impacts of externalities in an urban setting, particularly as unregulated urban markets may create many impediments to a balanced urban economic development.

2.   Urban externalities

The spatial-economic architecture of the city can still powerfully be described by means of Von Thünen's land rent theory. In a monocentric world, distance friction together with competition among land users will lead to concentric rings of economic activity reflecting the users' bidding power. Clearly, this model has many limitations: it is static, it is based on uniform land use, it draws on equal spatial accessibly and it neglects (positive and negative) externalities. Several of these limitations have thoroughly been analysed in the urban economics literature (see *e.g.* Miyao 1981 and Fujita 1989), while more recently also insights from the new economic geography have been added to the body of knowledge on urban development (see Fujita *et al.*, 1999).

Agglomeration economies are still at the heart of modern theories on urban growth. Geographical clustering offers a great variety of economies of scale and scope (including transaction benefits, contact opportunities and search advantages), so that we observe at a global scale a continual urbanization process. The agglomeration forces are apparently so strong that the shadow sides of urban areas are far exceeded by positive density economies. The urban economy is full of externalities, both positive and negative, and the question is how the existence of such market failures affects urban development. Examination of such externalities (ranging from social costs like waste or criminality to social benefits like

increased contact potentials) requires a thoughtful economic framework that addresses, for instance, the urban activity market, the property conditions of real estate and land, and the supply of public and private goods. The role of governments in case of market failures has extensively been discussed in the history of economic thinking by Pigou (1920) and Coase (1960). Market organization, as well as (the distribution of) property rights, are the main socio-economic arrangements determining the effectiveness of government interventions (Webster and Wai-Chung Lai, 2003).

Thus, the modern urban scene offers many perspectives, of both a policy and a theoretical economic nature. It is widely recognized that the emerging network economy is prominently shaped by an urban force field. Cities exhibit a wide array of attractive (e.g. the proximity to a wide variety of goods, services and jobs) and unattractive (e.g. expensive housing, lack of open space and environmental quality) features, and are hence governed by centripetal and centrifugal forces, often leading to complex spatial dynamic processes. Although in the long run market forces may lead to some equilibrium between positive and negative aspects of city life, it is doubtful whether the result would represent an efficient spatial-economic configuration. Many positive and negative features of city life involve what economists call externalities: unpriced effects that economic agents impose upon one another. Important examples from the urban arena include traffic congestion, noise and smell, pollution, agglomeration advantages, and ethnic segregation and/or concentration. Free markets involving externalities typically do not yield the most efficient outcome. And this may have serious implications for the city, and for urban policies.

Given the high degree of urbanization in most contemporary societies, and given the importance of the aforementioned externalities in urban firms' and citizens' daily lives, and given most citizens' concerns with the quality of urban life as well as the economic health and general well-being of the cities they live in, the economic analysis of these phenomena and the evaluation of policy options to deal with them is of great social importance, in addition to the academic interest that this type of analysis may provoke.

If the urban economy is fraught with externalities, a closer look at this phenomenon is no doubt warranted. The definition of externalities – unpriced effects that actors impose upon other actors – naturally implies that externalities are more important in urban areas than elsewhere, both absolutely and relatively. Urban areas are denser, and proximity makes the occurrence of unpriced spill-overs more likely. Nevertheless, most economic studies of externalities take a non-urban perspective, and the spatial dimension is typically lacking from the analysis. But this dimension may often be crucial in assessing the implications of externalities and policies, as both receptors and creators of externalities may often respond to these primarily in terms of spatial behaviour. For instance, spatial segregation often results from the desire to live close to some groups or activities and less close to others; potential visitors may avoid central city smell and noise by shopping elsewhere; and firms may seek a location near other firms in the hope of benefiting from knowledge spill-overs or labour market pooling. In aggregate, such behaviour may have important consequences for urban densities and urban form. And given the discrepancies that often exist between the private and

social cost and benefits of individuals' location decisions, the resulting aggregate urban configuration may be far from optimal from the social perspective, indicating a clear potential and need for welfare enhancing policies.

Urban externalities have been studied from various perspectives in the prior literature. This broad literature encompasses contributions from various disciplines, including urban economics, geography, planning and sociology. It is therefore relevant to identify what we believe to be important advances to be pursued in future economic research on urban externalities. These include (1) the explicit consideration of mutual interactions between externalities, and (2) between these externalities and urban form; and (3) the focus on (realistic) second-best policies. It is important to motivate these priorities.

Interactions between externalities are important because these may strongly affect economic policy guidelines for (optimally) coping with external effects. To give a simple example, the transport economics literature has provided substantial underpinning for the use of pricing instruments in coping with traffic congestion externalities ('road pricing'). By setting charges equal to the travel time delay costs that road users impose on others (the 'marginal external costs'), unregulated traffic volumes will be reduced to optimal levels, and substantial social cost savings can be realized, outweighing the social benefits foregone for those priced off the road. Society as a whole would thus benefit. True as this conclusion may be in the context of a transport model that ignores relations between the transport system and the urban economy, it may require reconsideration when those travelling during peak hours are mainly workers who contribute to positive agglomeration economies during their workdays. These externalities may be reduced, or even eliminated, when the charge induces these workers to change their job location, or to work at home. The interaction between externalities would, in this example, call for a downward adjustment in congestion charges. In contrast, an upward adjustment might be in order if negative environmental externalities (pollution) result either from the commute itself, or from production activities following the commute.

Evident as these qualitative observations may seem, and despite the potentially far-reaching policy implications in an urban context where multiple externalities typically prevail simultaneously, such issues have only since recently been investigated in the economics literature, usually applying non-spatial modelling approaches (see, for instance, Parry and Bento, 1999, on the interactions between inefficient labour and transport markets in the context of congestion pricing; and Verhoef, Button and Van den Bergh, 1997, on the policy implications arising from interactions between pollution from transport and from production, from a spatial perspective). There is certainly a need to make significant contributions to this important and growing literature, in particular by exploring such interactions while explicitly accounting for the spatial urban dimension.

Indeed, it may often be important not only to consider interactions between externalities as such, but also interactions between externalities and urban form. These interactions may run in both directions: the existence of externalities may affect urban form (*e.g.*, when citizens choose to live further from the centre to avoid noise or crowding), and

urban form may affect the severity of externalities (*e.g.*, traffic congestion will depend on the spatial arrangement of residential and job locations). Such interactions may affect the optimal design of policies directly, may also affect the consequences of these policies, and may in the third place have an indirect impact on the above mentioned interactions between externalities (*e.g.*, if congestion pricing would lead to a relocation of firms to less dense areas, the above mentioned negative effect of congestion pricing on agglomeration externalities may become even larger insofar as these agglomeration externalities decrease with the distance between firms).

Thirdly, it is also necessary to explicitly consider what economists call 'second-best' policies. These are policies that are imperfect from the strict economic perspective, but that are often much more realistic from a practical viewpoint. An example, again from the transport economics literature, is when a road operator can only put charges on a limited number of links in a network, instead of all links as is required for an optimal, 'first-best', control (Verhoef, 2002). Economic policy rules under such second-best circumstances become much more complicated than under hypothetical first-best conditions. The reason is that economic distortions caused by the instrument itself should be taken into account when using that instrument. In the network example, prices equal to marginal external costs on those links that can be charged are not (second-best) optimal. It is even not inconceivable that such 'quasi first-best pricing' would lead to welfare losses, instead of the gains that are hoped for. Second-best policies are therefore certainly not an academic *curiosum*, but should instead be seen as of much greater relevance than first-best pricing that has traditionally been considered in the literature. Given the complex nature of cities, where multiple externalities, and other market failures, exist simultaneously, and given the practical unavailability of first-best instruments, the development of analytical frameworks that allow consideration of second-best policies in urban economic models deserves high priority.

Thus, the field of urban externalities offers a rich research spectrum which ought to be explored more thoroughly. In this endeavour, several important distinct fields may be distinguished, such as traffic congestion in or around the city, environmental externalities and urban environmental quality, agglomeration externalities and industrial and services clustering, and neighbourhood externalities in relation to social and ethnic segregation and urban housing markets. It is noteworthy that all these externalities do not only refer to inefficiently operating urban markets, but involve also various types of distributive effects. In the practice of policy making and analysis, equity aspects tend to play as important a role as efficiency aspects, and this is certainly not different for the urban policy arena.

The above-mentioned range of urban economic and urban policy issues have intensively been discussed in the literature. In the next section we will offer a concise and selected overview of various types of research, distinguished into theoretical and applied studies.

3. An overview of urban externalities studies

As mentioned, the economic literature on (urban) externalities is vast and covers many aspects of the urban economy. A very selective presentation of such studies with a major emphasis on policy is given here, while making a distinction between theoretical and empirical contributions.

Theoretical studies

Recent years have demonstrated an avalanche of theoretical urban policy evaluation studies. It is generally believed that the relevant long-run impacts of policies – on externalities as well as on urban form – can be identified solidly only by using spatial general equilibrium approaches. Partial equilibrium approaches would often miss exactly those (spatial) interactions that were identified above as important in an urban context. For example, an analysis that would take into account the relocation of households due to the existence of environmental externalities in a city centre, or a policy designed to reduce these, but would fail to account for the implied change in labour supply in the city centre and the associated change in demand for goods supplied in this centre, would give a biased picture of the effects of the policy, and of how it should be designed in a (second-best) optimal way. The evaluation would namely be based on the characteristics of an urban system out-of-equilibrium, which therefore cannot be a long-run stable configuration.

No single analytical and integrated framework is currently available that can readily be used to analyse all key research questions of urban economic policy. The development of proper analytical modelling tools is still a fertile field, and is among the greatest challenges to be met in urban economics.

The theoretical work along these lines should address fundamental questions surrounding the conceptually sound economic modelling of (multiple) externalities in a spatial equilibrium setting. It may build upon the contemporary urban and spatial economic tradition (for overviews, see for instance Fujita, 1989; Fujita, Krugman and Venables, 1999; and Anas, Arnott and Small, 1998), and then seek to extend this tradition where appropriate. It need not be wise to focus on one single stream of modelling approaches, or indeed to pursue one single general analytical methodology. In contrast, it seems of great importance to design and test different analytical model specifications in studying urban externalities in the context of policy analysis, as these may yield qualitatively different and sometimes contrasting insights.

One example concerns the modelling of agglomeration externalities. Various approaches have been proposed in the literature (Duranton and Puga, 2004, provide an excellent overview), varying from spatial modifications of the Dixit-Stiglitz (1976) model of monopolistic competition where agglomeration effects arise through consumers' preferences for heterogeneity, as in Fujita, Krugman and Venables (1999), to reduced-form formulations where production efficiency increases in the aggregate urban production level (e.g. Sullivan, 1986) or labour supply (*e.g.*, Arnott, 1979).

Another example concerns alternative ways of modelling traffic congestion, which may take the form of static stationary-state congestion (*e.g.*, as in the spatial urban model by

Anas and Kim, 1996), dynamic flow congestion (as in Chu, 1995), or dynamic bottleneck congestion (as in Vickrey, 1969).

For both examples, it is not unlikely that conclusions on the relation between the externality and urban form, as well as on the spatial impacts of externality regulation, may vary across different model specifications (as in fact demonstrated elegantly in Arnott, 1998). Such differences are of course important to identify – and conversely, if similarities are to be identified, this would often be a welcome establishment of robustness of results over alternative analytical formulations, too. Not surprisingly, similar differences in modelling approaches can be distinguished for other urban externalities of interest.

Likewise, diverging insights may be derived for models in which, for example, the city is 'open' or 'closed' (*i.e.*, whether or not in- and outmigration is considered), a single city versus a system of cities is considered, labour supply is fixed versus endogenous, a single good or multiple goods are produced in the city, *etc*. Given the potentially decisive effects of such model characteristics on the insights generated, it is a research challenge to explicitly consider different, 'competing' formulations in investigating the issues at hand.

New analytical economic studies on urban externalities can nevertheless be expected to have a number of fundamental characteristics in common. They will typically be based on principles of utility maximization by consumers and profit maximization by producers. The models will have to describe spatial equilibria, where possible in continuous space and otherwise in discrete space, so that urban land use patterns are endogenous. Analyses will typically first have to be performed in the context of linear-space, static, often monocentric urban models in the tradition of Alonso (1964) and Muth (1969), before possible explorations are made into, for instance, two-dimensional (ground) space, dynamics, or multi-centric configurations.

Relatively simple urban externalities models – like the one we will present in Section 4 below – will clearly suffer from a higher level of abstraction, and hence a more limited degree of realism. The reason that such models nevertheless have been – and still are – extensively used in the urban economics literature is that they allow the identification of fundamental economic insights in a manageable setting. The next step then typically concerns the investigation of whether these insights carry over to more complex, more realistic, but less transparent settings. In other words, the insights from the simpler models help explain the results of the more elaborate models, which otherwise will be harder or even impossible to understand due to the complex interactions occurring.

For the same reason, new types of desired extensions will typically not be implemented simultaneously, but rather one-by-one, and depending on whether the type of problem studied would call for that extension. For example, two-dimensional ground space is relevant when the objective of study concerns traffic congestion on a network. Dynamics become relevant when path-dependency issues may play an important role, for instance when studying segregation issues. And polycentric configurations may be relevant when studying, again, transport networks, or agglomeration externalities in the context of sub-centres.

As is customary in the contemporary urban economics literature, it is to be expected that closed-form analytical solutions will typically not exist – or if they do, they will be too complex to offer any clear insights – so that numerical simulation modelling will form an important tool in identifying the comparative static or dynamic properties of various possible equilibria (some examples of how such approaches can be used are, among many others, Anas and Kim, 1996; Tabuchi, 1998; Fujita, Krugman and Venables, 1999; and Verhoef and Nijkamp, 2002). And finally, full-fledged urban externalities models will be used to assess the theoretical mutual interactions between urban form and the spatial incidence of the externalities studied; to assess the implied efficiency and welfare impacts; and to derive policy implications for first-best and second-best policy instruments.

Towards empirical studies

Next to theoretical model development, there have also been many advances in empirical studies. Nevertheless, substantial further research will be needed to better understand the causes, nature and consequences of urban externalities, and to identify policy implications. This subsection will again outline some general principles of such research.

A first goal of empirical research studies on the urban economic system will, for some externalities, be to identify the very nature of these externalities. This is in particular relevant for the 'container-notion' of agglomeration externalities. Practical definitions of agglomeration externalities may encompass a plethora of market-internal and external relations. A market-internal relation would, for instance, be a firm's desire to economize on transport costs for its inputs and outputs, from and to other firms, and thus to locate near these other firms. Insofar as these costs are market internal (and congestion would be priced optimally), it is to be expected that firms would face economically optimal incentives to agglomerate, and no market intervention would be called for. On the other hand, agglomeration effects such as caused by knowledge spill-overs would reflect genuine externalities, that are not optimally reflected in market prices; at least not if the land market is competitive (following Henderson, 1985, one might argue that a city developer would have the incentive to optimally internalize such external benefits). These effects would therefore potentially ask for government intervention. It is therefore important to have a clear picture not only of the overall importance of agglomeration advantages in the urban economy, but also of which part of these reflect market failures that should be accounted for in (second-best) policies, and which part should be left to the market.

A second goal is to assess the empirical relevance of the various externalities studied, and thus to provide inputs for the calibration of the theoretical models developed. The very nature of externalities – implying that these are unpriced in free markets – often makes their empirical measurement and valuation a difficult task. Nevertheless, various attempts to measure the empirical relevance of the external effects considered in this study have been made. For example, the relevance of traffic congestion externalities in urban areas can be assessed by using estimates of total (and marginal) travel time losses and uncertainty, in combination with estimates of what is known as the 'value of time' and 'value of

unreliability' (see for instance Small, 1992). Agglomeration externalities have, for instance, been investigated empirically by analysing the impacts of the size of a sector (for 'localization economies'), or of the urban economy as a whole (for 'urbanization economies'), upon a sector's productivity (e.g. Henderson, 1986; Mun and Hutchinson, 1995). Ciccone and Hall (1996) consider the relation between urban density and productivity. Environmental externalities in urban areas have, for example, been monetized using hedonic price methods (see Freeman, 1993, for an overview). And the empirical economic effects of segregation have for instance been assessed by looking at the relation between the degree of segregation and indicators such as high school drop-outs, employment and single motherhood (Cutler and Glaeser, 1997).

It is noteworthy that also outside the field of (urban) economics, a wealth of literature is available that deals with similar phenomena. One can think of traffic engineering in the context of congestion, geography and planning for agglomeration, environmental studies for environmental externalities, and sociology and cultural anthropology for segregation. Such studies will guide us in further understanding the necessary inputs for the modelling exercises, but also more general empirical patterns of economic development in the city. One of the main challenges in an applied context appears to be the question of how people would respond to urban economic policy instruments.

An important methodological choice in empirical valuation studies is between revealed preference studies, studying actual market behaviour of the relevant agents, and stated preference studies, using questionnaire type of approaches. Both techniques have been used widely in valuation studies, and both have their respective advantages and disadvantages (see for instance Perman, McGilvray and Common, 1999, for a discussion in the context of environmental externalities). There have been many recent advances, especially in the field of contingent valuation studies, and such studies may assume a more prominent methodological place in the study or urban externalities.

4. A modelling framework for urban externalities: analyzing first-best and second-best policies for multiple externalities

It is instructive, at this point, to present an illustrative modelling framework for analyzing first-best and second-best regulation when multiple externalities exist simultaneously in an urban economy. For reasons of simplicity, we have chosen the example of a traditional monocentric city, and assume that only one single product is produced in its (assumed spaceless) CBD. Two externalities are present in our city. First, agglomeration effects that are external to the individual firm explain why all production is concentrated in the CBD – and perhaps why the city exists in the first place. Agglomeration externalities are modelled in reduced form (*i.e.* no micro-foundation is spelled out explicitly), and cause marginal productivity to increase with total labour supplied, for example, because of knowledge spill-overs. This aggregate positive relation between labour supply and productivity is consistent with most structural models of agglomeration benefits (Duranton and Puga, 2004). The

second externality is caused by the fact that commuters travelling to the CBD pollute the urban environment.

A spaceless analysis of the associated policy dilemma might suggest a simple trade-off, stipulating that depending on whether the marginal agglomeration externality or the marginal environmental externality dominates, a first-best policy mix would either mean that aggregate labour supply and hence commuting should be increased, or that both should be decreased. Under fixed vehicle technologies, a simultaneous stimulation of agglomeration externalities and reduction of pollution from commuting would seem impossible. We have chosen this example to highlight the importance of taking a spatial perspective on analyzing urban externalities, and will present the result that when doing so, first-best policies indeed can result in such an attractive – but counterintuitive – 'win-win' outcome. Moreover, while the incentives from road pricing and labour subsidies would seem to be perfectly opposite in a non-spatial setting, leaving one of the two instruments redundant, our results will show that their welfare effects may, in contrast, turn out to be strongly super-additive when a spatial perspective is taken.

The model deployed here is inspired by Verhoef and Nijkamp (2002), who consider the trade-offs between agglomeration externalities and pollution from production in a mono-centric city; the presentation of the model in Section 4.1 below draws heavily from theirs. The main features of our model are the following. We consider a general spatial economic equilibrium: the land market, the labour market and the market for the industrial product in the city are simultaneously in equilibrium. Although we have a 'closed city', with a given population, labour supply is flexible and endogenous because a household can vary the hours of labour supplied. The firms are assumed to cluster in a spaceless CBD, and only use labour in their production process. The perfectly competitive firms have identical linear production functions. Agglomeration economies in the city are represented by 'Marshallian' externalities: marginal and average productivity increases in the aggregate labour supply. Pollution from commuting arises in a fixed proportion of total kilometres driven, and disperses over the entire city and hence affects the local environmental quality in a spatially non-differentiated manner. Therefore, proximity to the CBD yields a benefit in terms of lower commuting costs, but no relative disbenefit due to a poorer environmental quality. We aim to investigate the market equilibrium that results in such a configuration, and compare it with equilibria that would arise under various policies: first-best policies, and two second-best policies that only internalize the congestion externality and the agglomeration externality, respectively.

4.1.   *The analytical model[1]*

In this section, we present the details of the analytical model. Before turning to a detailed description of consumers' behaviour, firms, and a characterization of general equilibrium, some introductory remarks are in order. First, z will be used to denote a one-dimensional continuous urban space. The location of the spaceless industrial area is at $z=0$, and the

---

[1] This section draws heavily from Verhoef and Nijkamp (2002).

residential area stretches from $z=0$ to $z=z^*$, with z^* being the endogenous city boundary. Commuting time increases linearly with z at a rate t: there is no traffic congestion. At the boundary of the city, the equilibrium residential bid-rent must be equalized to the exogenous and constant agricultural bid-rent r_A, because land should go to the highest bidder.

It is assumed that all excess land rents above r_A are redistributed among the city's population. Alternatively, we could have used the 'absentee land-lord assumption', which seems less realistic, as it assumes that none of the land rents generated in the city would be used for consumption in the city. At the same time, it is less plausible to assume that *all* land rents generated in the city are redistributed among the population, as this would imply that the endogenous city size could be expanded costlessly. The present representation would correspond to the situation where the public authority of the city buys the urban land against the relatively low rural land price, implying an equivalent per-unit-of time price of r_A, and redistributes all excess rents generated in the city among its population. It is a convenient assumption in the sense that it easily allows us to consider households with similar initial endowments. If residential land were privately owned, someone paying a high rent would at the same time receive that high rent, leaving him no worse off in budgetary terms than someone paying a low rent. As a result of these assumptions, some share of the urban production will not be consumed in the urban area, but will be exported in exchange for the purchase of land against the agricultural rent.

Some final assumptions and remarks are to be made. All consumers and producers are assumed to be price-takers. Households are identical, and so are firms. The industrial product can be transported costlessly, and the world-market price of the industrial good p is used as the *numéraire*. We now turn to the various actors in the city and the resulting equilibrium issues.

Consumers

The closed city has N households, which we treat as a continuum of single economic entities. A household's utility depends on the consumption of the industrial good y, on the consumption of space or the size of the residence s, on the consumption of free time or leisure T_f, and on the environmental quality, Eq. A household's financial budget then consists of the wage rate w times the amount of hours worked T_w, plus the redistributed excess land rents (R in total, R/N per household), plus the share in the government's surplus or deficit as it arises from taxing road use and/or subsidizing labour supply (B in total, B/N per household). In equilibrium, the household's budget is fully spent on the consumption of y and s, and – if levied – on a road toll for commuters (τ_R per unit of distance). A household's given time budget is denoted by T, and can be spent on leisure (T_f), work (T_w) and commuting (T_c). All prices are treated parametrically by the (price-taking) households; w denotes the wage rate, p the price of the industrial good, and r the rent. Commuting does not require financial outlays other than possibly a total toll (over the full trip) $\tau_R \cdot z$, but costs time at a given rate t: $T_c = t \cdot z$. The number of commuting trips made by an individual is assumed to be proportional with the amount of effective working time supplied (T_w).

A household's simultaneous labour supply and consumption decisions can be modelled by using the 'gross budget', that would be available under the maximum possible amount of time worked, and to let the household 'buy back' leisure time against the prevailing wage rate w. Observing that the household's optimization problem is dependent on the residential location z, it can then be written as:

$$\underset{y(z),s(z),T_f(z)}{Max} U\big(y(z),s(z),T_f(z),Eq(z)\big)$$

$$s.t. \frac{R+B}{N} + (w-\tau_R \cdot z) \cdot \big(T-t\cdot z - T_f(z)\big) - p\cdot y(z) - r(z)\cdot s(z) = 0 \tag{1}$$

with:

$$R = \int_0^{z^*} r(z) - r_A \, dz \tag{2a}$$

and:

$$B = \int_0^{z^*} n(z)\cdot \big(T-t\cdot z - T_f(z)\big)\cdot \tau_R \cdot z \, dz - \int_0^{z^*} n(z)\cdot \big(T-t\cdot z - T_f(z)\big)\cdot \sigma_L \, dz \tag{2b}$$

where $n(z)$ gives the density of households at z, and σ_L is the second possible policy instrument of labour subsidies. The gross budget available at location z is thus defined as:

$$M(z) = \frac{R+B}{N} + (w-\tau_R \cdot z)\cdot (T-t\cdot z) \tag{3}$$

A spatial equilibrium requires that utility $U(z)$ be constant over z for all $0<z\leq z^*$ (and exceeds $U(z)$ for $z>z^*$): if a higher utility could be reached at any z, the rent $r(z)$ would be bidden up, as households would like to move from their original location to z in order to enjoy the higher utility level prevailing there. This implies a particular equilibrium pattern of land-rents. We can be more explicit about this when postulating a specific form for the utility function. We will be using a simple Cobb-Douglas structure here:

$$U(z) = y(z)^{\alpha_y} \cdot s(z)^{\alpha_s} \cdot T_f(z)^{\alpha_f} \cdot Eq^{\alpha_e}$$

with : $\quad \alpha_y + \alpha_s + \alpha_f = 1$ $\tag{4}$

Because utility is ordinal and any monotonic transformation of a given utility function represents the same preferences, the constraint on the parameters can be added without loss of generality. This utility function has the specific property of a unitary elasticity of substitution, implying that the gross income shares spent on y, s and T_f will be constant and given by the relative sizes of the parameters α. Specifically, the conditional demands for y, s and T_f are:

$$y(z) = \frac{\alpha_y \cdot M(z)}{p} \tag{5a}$$

$$s(z) = \frac{\alpha_s \cdot M(z)}{r(z)} \tag{5b}$$

$$T_f(z) = \frac{\alpha_f \cdot M(z)}{w - \tau_R \cdot z} \tag{5c}$$

and the indirect utility – for analytical convenience defined as the logarithm of the maximum utility achievable under given prices and wage – can be written as:

$$V(z) = \alpha_y \cdot \ln \alpha_y + \alpha_s \cdot \ln \alpha_s + \alpha_f \cdot \ln \alpha_f + \ln\left(\frac{R + B}{N} + (w - \tau_R \cdot z) \cdot (T - t \cdot z)\right)$$
$$- \alpha_y \cdot \ln p - \alpha_s \cdot \ln r(z) - \alpha_f \cdot \ln(w - \tau_R \cdot z) + \alpha_e \cdot \ln Eq \tag{6}$$

The condition that *V* in (6) be constant over space implies:

$$V'(z) = \frac{-w \cdot t - \tau_R \cdot T + 2 \cdot z \cdot \tau_R \cdot t}{\dfrac{R + B}{N} + (w - \tau_R \cdot z) \cdot (T - t \cdot z)} - \alpha_s \cdot \frac{r'(z)}{r(z)} + \alpha_f \cdot \frac{\tau_R}{w - \tau_R \cdot z} = 0 \tag{7}$$

where a prime denotes a 'space derivative' (with respect to location). Equation (7) gives a first-order differential equation for *r(z)*, which can be solved to yield:

$$r(z) = K \cdot (w - \tau_R \cdot z)^{-\frac{\alpha_f}{\alpha_s}} \cdot (B + R + N \cdot (T - t \cdot z) \cdot (w - \tau_R \cdot z))^{\frac{1}{\alpha_s}} \tag{8}$$

where K is a constant of integration. Invoking the equilibrium condition that $r(z^*)=r_A$, we can solve for K:

$$K = \frac{r_A}{(w - \tau_R \cdot z^*)^{-\frac{\alpha_f}{\alpha_s}} \cdot (B + R + N \cdot (T - t \cdot z^*) \cdot (w - \tau_R \cdot z^*))^{\frac{1}{\alpha_s}}} \tag{9}$$

To derive the equilibrium level of utility, we must be explicit about the production structure. This issue will be addressed below. We first conclude this analysis with a few identities. We can find the population density *n(z)* as the inverse of the 'lot-size' *s(z)*:

$$n(z) = \frac{1}{s(z)} \tag{10}$$

The total population is given, so that:

$$\int_0^{z^*} n(z)\,dz = \int_0^{z^*} \frac{1}{s(z)}\,dz = N \tag{11}$$

Total labour supplied equals:

$$L = \int_0^{z^*} n(z) \cdot (T - t \cdot z - T_f(z))\,dz \tag{12}$$

Total local consumption of the city's product equals:

$$Y = \int_0^{z^*} n(z) \cdot y(z) \, dz \tag{13}$$

The total amount of land consumed in the city must be equal to z^*, which is by definition true:

$$z^* = \int_0^{z^*} n(z) \cdot s(z) \, dz = \int_0^{z^*} 1 \, dz = z^* \tag{14}$$

Finally, we define environmental quality such that a virgin state corresponds to $Eq=1$, while the worst possible state, where an increase of consumption of other goods cannot even increase utility, occurs when $Eq=0$. Eq is assumed to decrease in a simple linear fashion in total emissions from commuting E, which in turn depends in a linear fashion on the total amount of commuting (weighted by distance), Km, in a fixed proportion ε (*i.e.*, ε gives the emissions per unit of distance driven):

$$Eq = 1 - E = 1 - Km \cdot \varepsilon = 1 - \int_0^{z^*} n(z) \cdot \left(T - t \cdot z - T_f(z)\right) \cdot \varepsilon \cdot z \tag{15}$$

Producers

There is a continuum of firms, each of which is infinitesimally small relative to the market and takes all prices as given. The industrial output is homogeneous, and agglomeration externalities in our model thus arise from a more efficient production when aggregate labour supply L increases. These agglomeration effects are summarized in an efficiency measure A, which individual firms take as given, but that is endogenous on the city level to represent agglomeration economics. Firms have a simple linear production technology with one input (labour). A firm's production function thus exhibits constant returns to scale, and therefore qualifies for application of Euler's theorem. Therefore, also when the urban aggregate production function exhibits increasing returns to scale due to agglomeration externalities, we can model the firms' aggregate behaviour using a 'derived aggregate production function', in which the efficiency parameter A is treated parametrically. The following derived aggregate production function applies:

$$Q = A \cdot L \tag{16}$$

We will use the following functional form for A in the numerical model:

$$A = 1 + \alpha_L \cdot L^{aggl_L} \tag{17}$$

Therefore, we use a simple exponential relation, where $aggl_L$ in reality is likely to be (and will restricted to be) smaller than one. The fixed unitary factor can be interpreted as the efficiency level for the first firm to locate in the CBD, and might reflect the presence of local infrastructure *etc.* A more than sufficient, but drastic condition for maintaining a monocentric configuration would be to set $A=0$ for locations outside the CBD. As we wish to focus attention on monocentric configurations, we will make this assumption.

Perfect competition drives profits to zero, with the result that the following equality holds:

$$p \cdot A = w - \sigma_L \tag{18}$$

General spatial equilibrium

The model described above has 19 unkowns, some of which are functions of z. These unknowns are $V(z)$, w, $M(z)$, $r(z)$, R, B, K, $y(z)$, Y, $s(z)$, $n(z)$, $T_f(z)$, L, z^*, Eq, E, Km, Q and A (recall that r_A, p and N are given; all other scalars are clearly parameters). The 19 equations needed to solve this system are (2ab), (3), (5a-c), (6), (8)-(13) and (15)-(18); note that (15) represents three equations. For other types of utility and production functions, as long as they imply unique conditional (factor) demands, a similar equality of numbers of equations and unkowns should in principle hold. We refrain from a formal analysis of existence, uniqueness and stability of equilibria and optima in our model.

In our list of equations, we did not include the 'aggregate demand equals aggregate supply' relation, which in our partly open system reads:

$$p \cdot (Q - Y) = r_A \cdot z^* \tag{19}$$

Equation (19) states that the value of the city's production in excess of its local consumption should be just sufficient to pay for the purchase of land against the exogenous terms of trade r_A/p. The share of local production not exported is consumed locally. The reason for not including this equilibrium condition explicitly is that it will be automatically satisfied under the zero profit condition and exhaustion of consumers' financial budgets – as in fact dictated by Walras' Law. To see why, first observe that zero profits imply that:

$$p \cdot Q = (w - \sigma_L) \cdot L \tag{20}$$

The exhaustion of consumers' total financial income implies (in aggregate terms) that the sum of redistributed land rents, the redistributed government surplus and wage income should be equal to the sum of expenses on the local product, rents and the road toll:

$$R + (\tau_R \cdot Km - \sigma_L \cdot L) + w \cdot L = p \cdot Y + (r_A \cdot z^* + R) + \tau_R \cdot Km$$
$$\Rightarrow (w - \sigma_L) \cdot L = p \cdot Y + r_A \cdot z^* \tag{21}$$

Substitution of (21) into (20) immediately yields (19).

It is not possible to obtain any further analytical (equilibrium) results for our model. We therefore now move to the results of a numerical illustration, to study the comparative static properties of the free-market and some second-best and first-best equilibria.

4.2.   A numerical example: base-case equilibrium

The numerical model is fully consistent with the analytical model just presented. Given that the model was built solely for the purpose of this paper, no particular attention was paid to its computational speed, and an algorithm was used that can most briefly be described as one that

successively equilibrates the various markets, repeating the procedure until convergence (according to the criteria set) is reached.[2] Given the nature of the problem, a mathematical software package was used that is capable of numerical solution of differential equations.

Parameters		Base-case equilibrium values of some key variables	
α_L=0.03 $aggl_L$=0.5 α_s=0.15 α_y=0.2 α_f=0.65 α_e=0.1 ε=1·10^{-7}	N=1000 T=1 t=1·10^{-5} p=1 r_A=0.005 Taxes and subsidies: τ_R=0 σ_L=0	Y=298 Q=399 A=1.49 w=1.49 L=268 Km=2.09·10^6 E=0.209 Eq=0.791 z^*=20 256.5	$p·Q$=399 $w·L$=399 $p·(Q–Y)$=101 $r_A·z^*$=101 R=122 B=0

Table 1. Parameters and equilibrium values of some key variables in the base-case scenario

The left two columns of Table 1 show the parameter values we assumed for the base-case of our simulation model. The city has 1000 inhabitants, with a time endowment that is normalized to unity. Some 30% of the time available after commuting will be spent working (which appears reasonable when excluding 8 hours sleeping time from the daily time budget: it would then mean an average of 33 hours per week over the whole year), and 42% (100·0.15/(0.15+0.2)) of the monetary budget will be allocated to rent, the remaining share being used for the consumption of the locally produced good. The price p is the *numéraire* in the price dimension. The right two columns show that in equilibrium, the average productivity is 49% higher due to agglomeration benefits than would be the case without them. Around 25% of the urban product is exported for the purchase of land. The maximum commuting time at z^* is 0.2; which implies an average of 3.2 hours (for a round trip) per day (of 16 hours) over a full year. The environmental quality has reduced to 79% of the virgin state. Because α_e is set at 0.1, this means that an inhabitant would be indifferent between obtaining a virgin state environmental quality or a 2.4% increase in the availability of all other goods (including leisure time) in the utility function, keeping everything else constant. These figures should give an adequate overall impression of the assumed base-case.

Figure 1 shows the spatial patterns of some variables of interest in the base equilibrium. The upper-left panel shows the equilibrium land rents, which have the expected convex shape. Land rents near the CBD are around four times as high as the agricultural rent applying at the fringe. As a result, households near the fringe will occupy nearly four times as much space as households near the CBD, as shown in the upper-right panel. Total expenses on space are somewhat lower near the fringe, because the monetary income is lower than in the centre due to a smaller number of hours worked (as shown in the lower-left panel). As a result, households near the fringe also consume less of the local product. The smaller

---

[2] Once the starting values are set appropriately, the model reaches convergence within a minute on a modern PC.

consumption of leisure time near the fringe, despite the smaller amount of hours worked, is explained by the extra time spent commuting (not shown in Figure 1). Finally, the lower-right panel shows that indeed a spatial equilibrium is obtained: utility – calculated directly according to (4), with the endogenously determined equilibrium values of $y(z)$, $s(z)$, $T_f(z)$ and Eq substituted in – is constant over space, and the fluctuations shown are evidently due to numerical imprecision only.
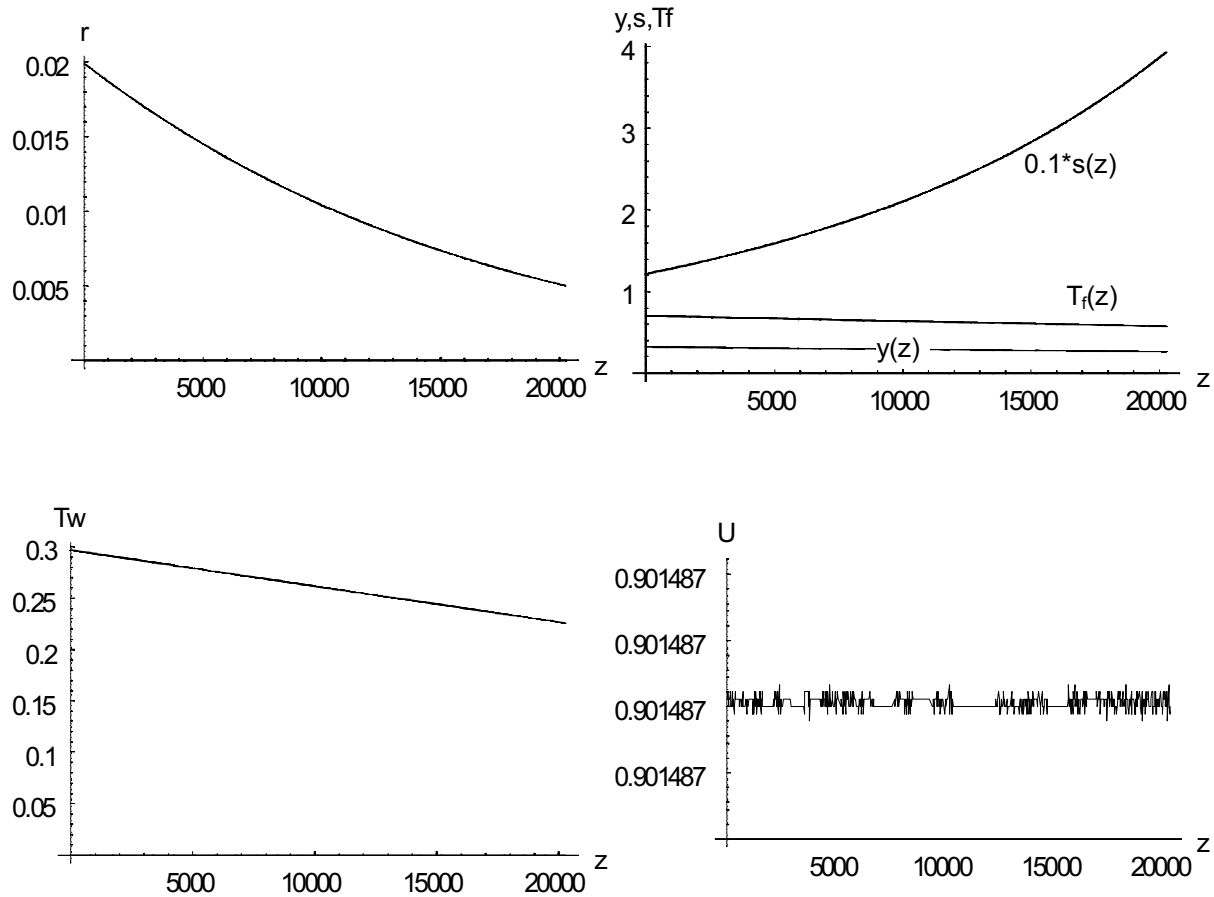

Figure 1. Equilibrium land rents (upper-left panel), consumption (upper-right panel), labour supply (lower-left panel) and utility (lower-right panel)

4.3.   First-best regulation

Given the simultaneous existence of two externalities in our model, it should be no surprise that the free-market equilibrium ('base equilibrium' in the sequel) is not efficient. Given that environmental externalities arise in a fixed proportion to kilometres driven, and agglomeration externalities depend exclusively on the amount of labour supplied, first-best policy instruments are easily identified. These consist of the combination of a road tax τ_R for every kilometre driven, and a subsidy σ_L on every unit of labour supplied. In the optimum, the levels of these instruments should be equal to the marginal externality. These levels were determined numerically, and the first column of Table 2 shows the corresponding values, as well as the

resulting values for some key (non-spatial) endogenous variables as a proportion of their values in the base equilibrium.

	First-best	*Second-best road pricing*	*Second-best labour subsidy*
τ_R	$2 \cdot 10^{-5}$	$7.5 \cdot 10^{-6}$	0
σ_L	0.25	0	0.05
B	−30.9	14.3	−13.7
Y	107.97 %	97.55 %	102.76 %
Q	103.53 %	96.72 %	102.31 %
A	100.49 %	99.53 %	100.32 %
w	117.26 %	99.53 %	103.68 %
L	103.02 %	97.17 %	101.98 %
Km	91.31 %	91.45 %	102.58 %
E	91.31 %	91.45 %	102.58 %
Eq	102.29 %	102.25 %	99.32 %
z^*	90.48 %	94.26 %	100.98 %
R	122.44 %	100.28 %	104.24 %

Table 2. Key characteristics of first-best and second-best optima (% relative to base equilibrium)

The results clearly reject the hypotheses that one might infer from a non-spatial analysis. These would be that there is a simple trade-off between the two externalities, stipulating that depending on whether the marginal agglomeration externality or the marginal environmental externality dominates, a first-best policy mix would either mean that aggregate labour supply and hence commuting should be increased, or that both should be decreased; and that one of the two instruments would be redundant as they both affect the same margin of behaviour (labour supply, and in a fixed proportion to this, commuting). Under the first-best policy mix, labour supply increases by 3.0% and total commuting decreases by 9.7%. As a result, environmental quality improves by 2.3%, and marginal productivity increases by 0.49%.[3] This favourable combination can be realized because average commuting distances decrease: z^* falls by 9.5% compared to the base equilibrium. Instead of giving conflicting incentives, the road tax in an indirect way in fact contributes to the goal of stimulating labour supply. It gives a direct incentive to shorten commuting distances, which in turn gives an incentive to increase labour supply T_w, which – as was shown in Figure 1 – increases with proximity to the CBD.

Figure 2 compares the spatial patterns of two endogenous variables under first-best regulation to the base equilibrium. The left panel shows that land rents rise over most of the

---

[3] The implied elasticity of average productivity (A) with respect to aggregate production (Q) of 0.14 (= 0.49/3.53) 0.14 is well in line with the empirical estimates reviewed by O'Sullivan (2002) for localization economies, which range between 0.02 and 0.27 (p. 50).

city, and more strongly so for locations closer to the CBD (near the new fringe, land rents of course must have fallen as $r(z^*)=r_A$ continues to hold). This is a direct consequence of the imposition of the road tax – increasing the per-unit-of-distance total commuting costs and hence making the equilibrium bid-rent steeper – and is indirectly aggravated by the substitution away from land consumption. Recall that excess land rents are redistributed over the city's population, so that the city's net land costs ($r_A \cdot z^*$) nevertheless decrease. The right panel shows that labour supply increases over most of the city, but again more strongly so for locations closer to the CBD. These spatial differences are again explained by the increase in per-unit-of-distance total commuting costs, which decreases the shadow price of leisure time more strongly for more distant locations; compare (5c).
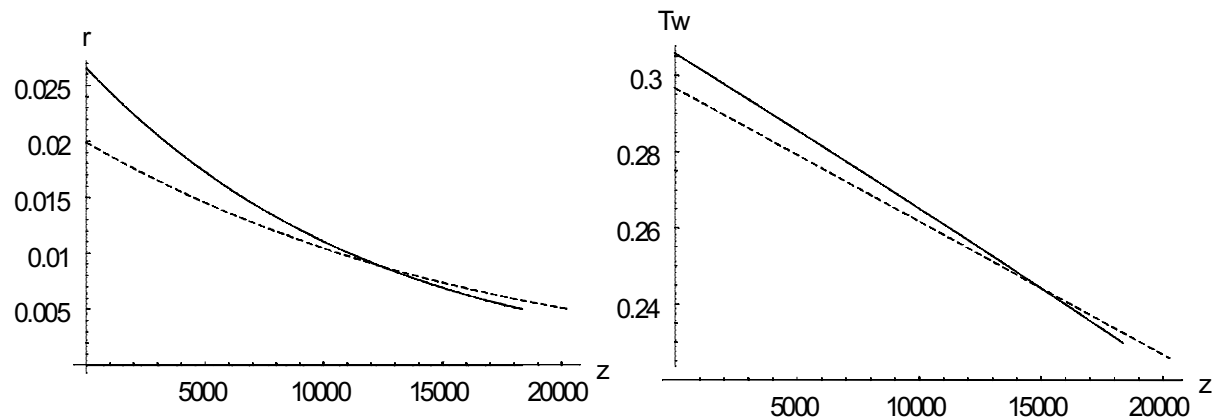

Figure 2. First-best land rents (left panel) and labour supply (right panel); base equilibrium dashed

4.4. Second-best regulation

First-best policies are, tautologically, the best choice when the city authority seeks to maximize the citizens' welfare. At the same time, the practical relevance of such idealized may be smaller than most economists would hope, for a variety of reasons. First-best policies may be considered too complex, may require 'too much' cooperation between different governmental authorities (in our example the city's department of transport and the department of economic affairs), *etc.* Second-best policies then come into the picture, and deserve – we believe – much more attention from the urban economist. We will therefore consider two possible second-best policies, namely the use of a road tax τ_R in isolation, and the use of a labour subsidy σ_L in isolation. Again we have refrained from the (notoriously difficult) task of deriving analytical expressions for the second-best optimal levels of these instruments, and have instead derived these numerically for our model. The second and third column in Table 2 provide the detailed impacts of these second-best policies for the non-spatial endogenous variables of interest, while Figures 4 and 5 below show the resulting impacts on the spatial patterns of land rents and labour supply.

First, however, Figure 3 shows how the relative welfare gain, expressed as the proportion of the first-best efficiency gain (pragmatically expressed as the Equivalent Variation for the median-location household in the base equilibrium) that can be achieved

with these second-best instruments, as a function of their levels (also expressed as the proportion of the first-best levels). Figure 3 conveys a number of messages. The first is that second-best policies are considerable less efficient than first-best policies, with the relative efficiency gain peaking just below 0.3 for the road tax, and just above 0.05 for the labour subsidy. Because the sum of these relative gains is around 0.35, it appears that the joint use of these instruments yields a gain nearly three times as high as the sum of the gains when the instruments are used in isolation. In other words, the efficiency gains from the two instruments are (strongly) super-additive. The results in the second and third column in Table 2 show why: when used in isolation, the road tax reduces agglomeration benefits, while the labour subsidy increases pollution. The explanation for these negative side-effects is simple, and similar to what one would infer from a non-spatial analysis: a road tax discourages commuting and hence labour supply (as shown also in the right panel of Figure 4), and a labour subsidy stimulates labour supply and hence commuting (as shown also in the right panel of Figure 5). As we discussed above, these unfavourable side-effects of the policies when used in isolation can be avoided when using them in combination.
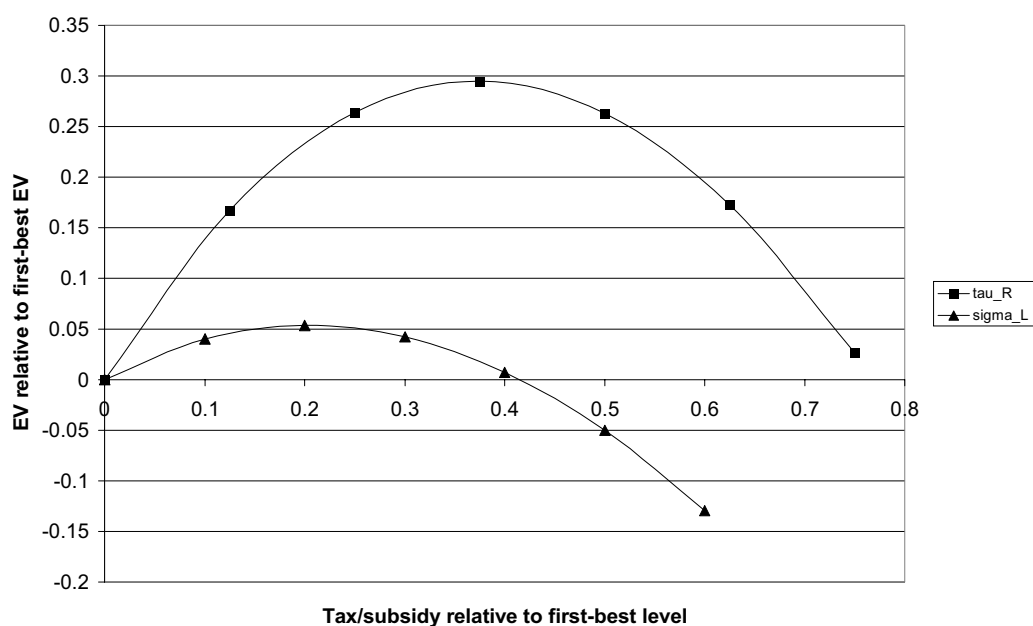

Figure 3. Second-best regulation: equivalent variation as a proportion of first-best equivalent variation, as a function of tax/subsidy levels

But there is another lesson to be drawn from Figure 3, and that is that the second-best optimal use of the two instruments is at a level significantly below their first-best levels (at around 20% for the labour subsidy and at around 40% for the road tax). The explanation is that the second-best optimal levels trade off the 'good news' from the two instruments (enhancing agglomeration benefits for the labour subsidy and reducing pollution for the road tax) against the bad news (reducing agglomeration benefits for the road tax and enhancing pollution for the labour subsidy). The consequence is that a naïve use of the instruments, set at their first-

best levels despite the fact that the other instrument is not used, would lead to efficiency gains below the second-best gains. In fact, in our numerical example, for both instruments an efficiency loss would result from such naïve policies.

This is of course a nice illustration of our earlier claim that for an urban economy, expectedly fraught with many positive and negative externalities, it seems of great importance to develop modelling tools allowing for a simultaneous consideration of multiple externalities. The risk of performing 'less than second-best' is otherwise not inconceivable. Likewise, our simple model has illustrated why it is important to analyze urban externalities in a general spatial equilibrium setting: the optimal use of the two instruments in combination and the super-additivity of their efficiency gains could otherwise not have been identified. And the third priority for research that we mentioned – the importance of considering second-best issues when analyzing inherently second-best instruments – is illustrated by the large discrepancy between 'truly' second-best and naïve first-best levels for the use of these instruments.
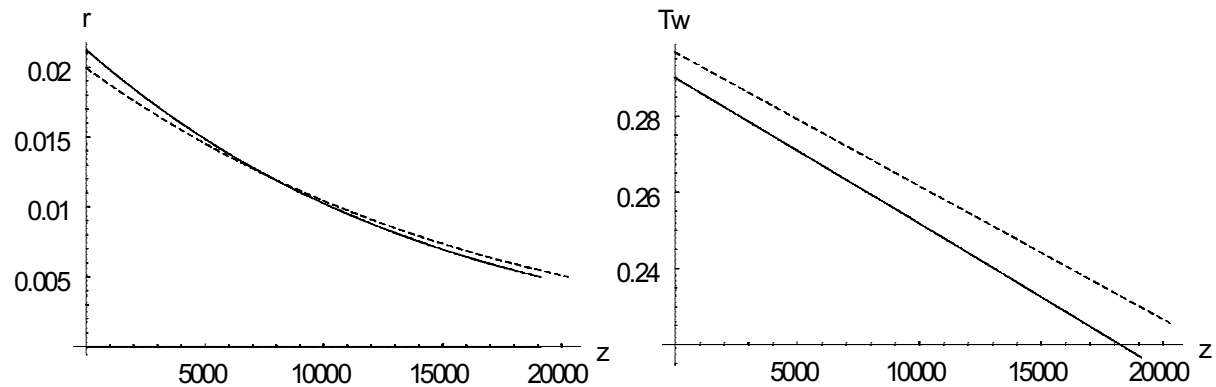

Figure 4. Second-best road pricing: land rents (left panel) and labour supply (right panel); base equilibrium dashed
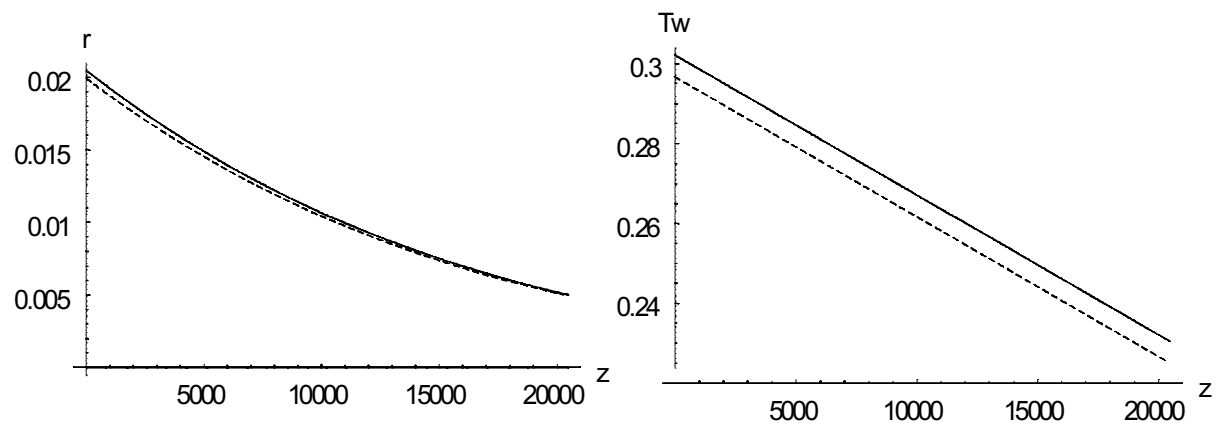

Figure 5. Second-best labour subsidies: land rents (left panel) and labour supply (right panel); base equilibrium dashed

5.   Conclusion

This paper has addressed the economics of urban externalities. We started by reviewing the literature on urban externalities, and observed that although many important contributions of course have been made, there seems to be sufficient scope and need for further research, both theoretically and empirically. We identified what we believe to be important advances to be pursued in future research on urban externalities. These include (1) the explicit consideration of mutual interactions between externalities; (2) a thorough analysis of the relationship between these externalities and urban form; and (3) a clear focus on (realistic) second-best policies.

The importance of these issues was illustrated by developing a simple urban general equilibrium model, in which we studied the interactions between agglomeration externalities and pollution from commuting. Our results show that what seems impossible from a non-spatial perspective, namely a simultaneous stimulation of agglomeration externalities and reduction of environmental externalities, is in fact the result from first-best policies. The reason lies in spatial adaptations, in particular a shortening of commuting distances which directly reduces pollution, and indirectly – via the increase in labour supply that follows from a greater proximity to the CBD – stimulates agglomeration benefits. Moreover, while the incentives from road pricing and labour subsidies would seem to be perfectly opposite in a non-spatial setting, leaving one of the two instruments redundant, our results show that their welfare effects may, in contrast, turn out to be strongly super-additive when a spatial perspective is taken. The model thus underlines the importance of using spatial general equilibrium frameworks when analyzing urban externalities.

Our results furthermore provided a nice illustration of our claim that for an urban economy, expectedly fraught with many positive and negative externalities, it seems of great importance to develop modelling tools allowing for a simultaneous consideration of multiple externalities. The risk of performing 'less than second-best' is otherwise not inconceivable. And finally, the large discrepancy between 'truly' second-best and naïve first-best levels for these instruments underlines the importance of seriously considering second-best issues when analyzing inherently second-best instruments.

Evidently, the model we presented falls in the category of conceptual 'simple static monocentric models' as we called them in our review. It remains to be seen which additional insights can be obtained on the issues that we have studied from more elaborate modelling settings. But, we hope to have convinced the reader that this would be just one among many challenges facing the urban economist interested in the economics of urban externalities.

References

Alonso, W., **Location and Land Use Harvard**, Cambridge MA, 1964.

Anas, A., R. Arnott and K. Small, Urban Spatial Structure, **Journal of Economic Literature**, vol. 36, 1998, pp. 1426-1464.

Anas, A. and I. Kim, General equilibrium models of polycentric land use with endogenous congestion and job agglomeration, **Journal of Urban Economics**, vol. 40, 1996, pp. 232-256.

Arnott, R., Optimal city size in a spatial economy, **Journal of Urban Economics**, vol. 6, 1979, pp. 65-89.

Arnott, R., Congestion tolling and urban spatial structure, Journal of Regional Science, vol. 38, 1998, pp. 495-504.

Ciccone, A. and R.E. Hall, Productivity and the density of economic activity, **American Economic Review,** vol. 86, 1996, pp. 54-70.

Coase, R.H., The Problem of Social Cost, **Journal of Law and Economics,** vol. 3, 1960, pp. 1-44

Cutler, D.M. and E.L. Glaeser, Are ghettos good or bad?, **Quarterly Journal of Economics**, vol. 112, 1997, pp. 827-872.

Chu, X., Endogenous trip scheduling: the Henderson approach reformulated and compared with the Vickrey approach, **Journal of Urban Economics**, vol. 37, 1995, pp. 324-343.

Dixit, A.K. and J.E. Stiglitz, Monopolistic competition and optimum product diversity, **American Economic Review**, vol. 67, 1976, pp. 297-308.

Duranton, G. and D. Puga, Micro-foundations of Urban Agglomeration Externalities. In J.V. Henderson and J.-F. Thisse **Handbook of Regional and Urban Economics** (Vol. 4), North-Holland, Amsterdam (2004 – forthcoming).

Freeman, A.M. III, Property value models, in: **The Measurement of Environmental and Resource Values: Theory and Methods Resources for the Future**, Washington D.C., 1993, pp. 367-420.

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

Fujita, M., P. Krugman and A. Venables, **The Spatial Economy**, MIT Press, Cambridge MA, 1999.

Henderson, J.V., Efficiency of resource usage and city size, **Journal of Urban Economics**, vol. 19, 1986, pp. 47-90.

Henderson, J.V., The Tiebout model: bring back the entrepreneurs, **Journal of Political Economy**, vol. 93(2), pp. 248-264.

Jacobs, J., **The Economy of Cities**, 1969, Random House, New York.

Miyao, T., **Dynamic Analysis of the Urban Economy**, Academic Press, New York, 1981.

Mun, S. and B.G. Hutchinson, Empirical analysis of office rent and agglomeration economies: a case study of Toronto, **Journal of Regional Science**, vol. 35, 1995, pp. 437-455.

Muth, R.F., Cities and Housing Chicago, Chicago Ill, 1969.

Parry, I.W.H. and A.M. Bento, Revenue recycling and the welfare effects of congestion pricing. Working Paper, Resources for the Future, Washington, 1999.

Perman, R., Y. Ma, J. McGilvray, and M. Common, **Natural Resource & Environmental Economics** (2nd edition), Addison Wesley, Longman Ltd., UK, 1999.

Pigou, A., **The Economics of Welfare**, MacMillan, London, 1920.

Small, K.A., Urban Transportation Economics. Fundamentals of Pure and Applied Economics Harwood Chur, 1992.

Sullivan, A.M., A general equilibrium model with agglomerative economies and decentralized employment, **Journal of Urban Economics**, vol. 20, 1986, pp. 55-74.

Tabuchi, T., Urban agglomeration and dispersion: a synthesis of Alonso and Krugman**, Journal of Urban Economics**, vol. 44, 1998, pp. 333-351.

Vickrey, W.S., Congestion theory and transport investment, **American Economic Review**, vol. 59, 1969, 251-260.

Verhoef, E.T., Second-best congestion pricing in general networks: heuristic algorithms for finding second-best optimal toll levels and toll points, **Transportation Research**, vol. 36B, 2002, pp. 707-729.

Verhoef, E.T. and P. Nijkamp, Externalities in urban sustainability: environmental versus localization-type agglomeration externalities in a general spatial equilibrium model of a single-sector monocentric industrial city, **Ecological Economics**, vol. 40, 2002, pp. 157-179.

Verhoef, E.T., J.C.J.M. van den Bergh and K.J. Button, Transport, spatial economy and the global environment, **Environment and Planning**, vol. 29A, 1997, pp. 1195-1213.

Webster, C., and L. Wai-Chung Lai, **Property Rights, Planning and Markets**, Edward Elgar, Cheltenham, 2003.