Economics of Transportation 1 (2012) 64–76

Contents lists available at SciVerse ScienceDirect

Economics of Transportation

journal homepage: www.elsevier.com/locate/ecotra

The optimal pricing, finance and supply of urban transportation in general equilibrium: A theoretical exposition☆

Alex Anas *

State University of New York at Buffalo, 405 Fronczak Hall, Amherst, NY, USA

ARTICLE INFO

Article history:
Received 6 March 2012
Received in revised form
16 May 2012
Accepted 21 May 2012

Keywords:
Optimal urban transportation
Pricing
Finance
Supply

ABSTRACT

We present a general equilibrium framework of optimal allocation treating the pricing, finance and supply of urban transportation. Uncongested public transportation technology with economies of scale supports the city's existence; and a congested road system subject to constant returns limits urban size. Optimal investment in public transit and in roads follow Samuelson's rule. With optimally determined urban population, roads are fully financed by Pigouvian congestion tolls while aggregate differential land rents fully finance public transit and any other activity with internal or external economies of scale (Henry George Theorem). Marshallian agglomeration from production in the core and the suburbanization of jobs to avoid congestion are treated. We also see how the optimal rules and the Henry George Theorem are modified when the demand for location is determined by a random utility model.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

It is hard to imagine any model of the urban economy with any level of spatial detail that can ignore personal transportation. Conversely, ignoring the urban economy in transportation studies leads to wrong predictions. The generation of travel is strongly linked to land use, to housing and job location decisions and to the pricing, finance, modal mix and supply of transportation. This interdependence and feedback between transportation and the urban economy has been center stage in urban economic theory ever since theoretical models of urban land use have been formulated in the1960s.

The purpose of this article is to present in a nutshell and in a form that is stripped from unnecessary details, the most fundamental results on the interrelationships between urban economics and transportation that have emerged out of five decades of theory. The extant theoretical literature contains only some of these results. The unified extended treatment exposited here has not appeared anywhere and the simplicity, generality and power of the basic results has become largely obscured by the demanding analytical requirements of the "workhorse of urban economics": the monocentric model.

The monocentric model assumes that all jobs are pinned at the central business district (CBD) and treats the urban space around the CBD as a continuous dimension.[1] The analytical difficulty from treating space as continuous has limited the scope of many contributions and the eagerness of scholars to engage in extending broadly this basic workhorse model. In particular, extensions of partial equilibrium models to general equilibrium ones have lagged and several issues have been considered relatively intractable. This has been especially true for analytical models with decentralized employment and job sub-centers that conform to realistic patterns and go beyond straightforward extensions of the monocentric framework[2]; and for models with more than one mode of urban travel (e.g. roads and public transportation).[3]

To make it possible to obtain fundamental results with little if any loss of generality, this article embraces an approach that holds the promise of making urban economics at once more digestible, more communicable and more analytically pleasant, and quite easily extendable compared to the continuous workhorse model.

☆ "In writing this article I benefitted from my teaching at the Kuhmo-Nectar Summer School on Transportation Economics in 2007 through 2010 at Urbino, Amsterdam, Copenhagen and Valencia and also from some of my lectures in Ph.D. level courses in urban and public economics."

* Tel.: +1 716 645 8663; fax: +1 801 749 7805.
E-mail addresses: alex.anas@verizon.net, alexanas@buffalo.edu

[1] Mohring and Harwitz (1962), Alonso (1964), Mills (1967) and Muth (1969) employed the continuous model. But Strotz (1965), whose contribution got less attention, utilized a model of discrete annular rings.

[2] Brueckner (1979) made a simple monocentric model with jobs serving local residences at each location; White (1988) relaxed the monocentricity assumption to include suburban employment. Sullivan (1986) formulated a general equilibrium model of a one-dimensional city with a suburban job center and a central business district.

[3] Two modes were treated by Capozza (1973, 1976) in a one dimensional monocentric model and by Anas and Moses (1979) in a two-dimensional model with discrete transit lines. LeRoy and Sonstelie (1983) examined the location patterns of low and high income groups in a one dimensional monocentric city with both roads and public transit. Kim (1979) used a linear programming model of Edwin Mills to demonstrate by means of non-empirical simulations that an urban area of two million population can support a public transit system.

2212-0122/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ecotra.2012.05.001

The approach treats space as discrete. Once this is accepted, then a natural starting point is that of a core–periphery setup of just two discrete urban land areas. The core is an area within a metropolis that represents the CBD and the inner or central city; the periphery is a surrounding area that represents the suburbs.[4] This undeniably fits urban reality all around the world, quite well. In this article, we generalize the core–periphery setup and show how to exploit its potential. Given the limited space available to this article, we focus on the most basic issues.

First, in Section 2, we use the core–periphery setup to model monocentric urban spatial equilibrium with public transit in the core and roads in the periphery. We do a comparative static analysis of this core–periphery model in general equilibrium, showing that its qualitative properties are consistent with the corresponding analysis of the continuous monocentric model obtained by Wheaton (1974) who first did the comparative statics analysis of the continuous-space Alonso model with a homogeneous urban population.

Second, in Section 3, we deal with optimal allocation when there are two modes of transportation. We do this by extending the general equilibrium model of Section 2. The obvious choices are public transportation in the form of subways and private transportation which requires roads. Although transportation economists have long recognized the consumer's choice between these two modes especially after the seminal work of McFadden (1974), what is really important here but neglected in the theoretical literature is a crucial difference between the two modes in their cost-structure on the supply side. Public transport infrastructure is subject to great economies of scale arising from the high fixed costs of tunneling or of surface high capacity corridors.[5] The higher the investment in such infrastructure, the lower the private cost of individual users of public transportation becomes. The reason for this reduction in private cost is not only the higher speeds induced by the higher investment but also the shorter access when the number of tunnels and lines increases (Anas and Moses, 1979), each serving a narrower urban area. It is reasonable, in a theoretical setting such as ours, to follow the dominant convention treating public transportation as not congestible. Our second mode, roads and bridges accessed by private cars, is assumed – again following the dominant convention – to be constant returns to scale, requiring an extensive land input. With congestion, the private cost of road transportation increases with its use and at an increasing rate and its marginal social cost is above the average private cost.

In Section 3, armed with the above two-mode context we exploit the question of the optimal pricing, finance and supply of public transit and roads. To make sure there is no confusion, throughout this paper *pricing* refers to road charges. It would refer more generally to gasoline taxes, public transit fares, or any instruments that prices travel per trip, per mile, per passenger or per gallon of fuel. *Finance* refers to the source of the funds that cover the public provision cost of a road or public transit system. In general, such funds may come from the pricing or tax instruments, but also from sources such as land taxes, income taxes or head taxes. *Supply* refers to the quantity, or quality, extent and location of a transport system that is provided. In a general equilibrium context, *pricing–finance–supply* are treated in a unified manner and are jointly determined.

Since both forms of transport infrastructure are public goods and externalities are also present, markets cannot optimally resolve pricing, finance and supply. A variety of planning regimes are then potentially available, and they vary from minimalist public intervention to first-best optimality with many lower-best regimes in the middle. These regimes are distinguished from one another according to the optimal provision rules and the pricing/tax instruments available and actually used to finance the supply of transportation that is provided.

Our main focus in this article is on first-best optimality.[6] We formulate the social optimum as a mixed regime, in which the planner controls the fiscal instruments and the markets allocate resources and determine prices subject to the values of the fiscal instruments chosen by the planner. The first-best allocation is found by the planner optimizing the choice of the fiscal instruments, taking into account how the markets will respond. A central result original to this paper is that the economy of scale in public transportation is one of the unsung reasons that support a larger optimal urban population size. That is, as people concentrate in the same area where they can share a public transit infrastructure, the cost per trip of that infrastructure decreases with the number of trips made on it. Therefore, a public transport system provides a second-nature cause for a city to be formed at a larger size.[7] An associated central result, but this one over sung in the existing literature, is that congestion on urban roads increases with urban population. Since, in road transport, there are no or only small economies of scale but rising congestion, this diseconomy offsets the benefits from further gains on the public transit side.

At an optimal allocation, the level of investment in public transit and in roads are each determined according to Samuelson's rule for the supply of a public good (Samuelson, 1954), that is expansion of such investments continues until the sum of the private marginal benefits over all users of each infrastructure just equals the marginal cost of supplying more of that infrastructure. An optimal city size is reached where the marginal benefits of public transit and the marginal costs of road transportation are balanced. At an optimally sized population for a region or an urban area the Henry George Theorem holds (Flatters et al., 1974; Arnott and Stiglitz, 1979).[8] We will see that in our basic model this means that the aggregate rent surplus generated by the urban concentration completely pays for the total social cost of the public transport infrastructure, while the aggregate revenue from Pigouvian tolling of road congestion completely pays for the cost of roads.[9]

The basic model of Section 3 which includes within a planner's social optimization problem the spatial equilibrium model of Section 2 yields the results described in the above paragraph and is offered as a framework for examining a number of extensions. In Section 4 we consider what happens if there are Marshallian externalities in the core's production process, often cited as the reason for the existence of cities. The results of Section 3 extend nicely to this case and we show how the Henry George Theorem generalizes: the rent surplus from the core now finances not only the social cost of the public transportation

[4] The two-zone approach owes its beginnings to Pines and Sadka (1981). It has been refined and extended in Anas and Pines (2008, in press) some results from which are discussed in Section 7 of this article.

[5] See the transportation chapters in the editions of the urban economics text by Mills and Hamilton (1989).

[6] This is deliberate due to space limitations. Some discussion of lower-best regimes is in a later section focusing on the use of urban growth boundaries as policy instruments.

[7] First-nature reasons for the existence of cities are natural amenities at the site of the city or local public goods as in Arnott and Stiglitz (1979), scale economies internal to the firm as in Dixit (1973), or Marshallian agglomeration economies external to the firm (Marshall, 1920). Abdel-Rahman and Anas (2004) show how the various forms of scale economy are reduced-form-identical.

[8] See Arnott (2004) for a recent comprehensive discussion of the testing of the Henry George Theorem.

[9] Vickrey (1963, 1969) proposed that road congestion be priced according to Pigou's (1920) corrective taxation of a negative externality. See also Baumol (1972).

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات