Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Urban Economics 58 (2005) 367–388
JOURNAL OF Urban Economics
www.elsevier.com/locate/jue

Second-best congestion pricing schemes in the monocentric city

Erik T. Verhoef [1]

Department of Spatial Economics, Free University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

Received 12 October 2004; revised 13 June 2005

Available online 16 September 2005

Abstract

This paper considers second-best congestion pricing in the monocentric city, with endogenous residential density and endogenous labour supply. A spatial general equilibrium model is developed that allows consideration of the three-way interactions between urban density, traffic congestion and labour supply. Congestion pricing schemes are analysed that are second-best 'by design' (and not because distortions exist elsewhere in the spatial economy), like cordon charging and flat kilometre charges. Both for Cobb–Douglas utility and for CES utility, the analyses suggest that the relative welfare losses from second-best pricing, compared to first-best pricing, are surprisingly small.

JEL classification: R41; R48; D62

Keywords: Traffic congestion; Second-best pricing; Urban structure; Spatial general equilibrium

1. Introduction

Practical applications of traffic congestion pricing typically involve so-called second-best pricing regimes, which fail to charge every individual road user his or her exact marginal external congestion costs. With pay-lanes, to an increasing extent employed in

E-mail address: everhoef@econ.vu.nl.
[1] Affiliated to the Tinbergen Institute, Roetersstraat 31, 1018 WB Amsterdam.

0094-1190/$ – see front matter
doi:10.1016/j.jue.2005.06.003

the US, unpriced congestion remains existent on parallel highway lanes. In case of cordon charges, such as used in Singapore (in conjunction with tolls on major expressways and arterial roads), every road user passing the cordon pays the same charge independent of the distance travelled and route followed before and after passing the cordon, and users who remain within or outside the cordon do not pay the charge. Area charges, as recently introduced in London, impose the same charge on every vehicle observed in the area independent of the distance travelled, and leave congestion outside the area uncharged. And flat kilometre charges, as have been considered for The Netherlands, would not differentiate over time and routes.

A substantial literature has recently emerged on the economics of second-best congestion charges (e.g. Lindsey and Verhoef [6] provide an overview). Most of these studies employ partial equilibrium approaches, in which only the transport (network) market is considered explicitly. An exception is the work by authors such as Mayeres and Proost [8] and Parry and Bento [12], who study traffic congestion and road pricing for commuters in general equilibrium settings, allowing for distortions on the labour market. Their results suggest that these interactions can be of significant importance for the efficiency impacts of both congestion pricing and of the use of the associated revenues.

Another important non-transport market when evaluating congestion pricing strategies for urban areas is the (spatial) housing market. Already in the 1970s, a number of studies looked into the interactions between traffic congestion and urban land use in the context of the monocentric model (Solow and Vickrey [15]; Solow [14]; Kanemoto [5]; Arnott [3]; Sullivan [16–18]). Anas and Kim [1] and Anas and Xu [2] extended this line of research by allowing for multicentric configurations, endogenizing the emergence of centres through the explicit consideration of agglomeration forces.

The present paper considers second-best congestion pricing in the monocentric city, with endogenous residential density and endogenous labour supply, therefore focussing on the long-run effects of alternative policies. A spatial general equilibrium model is developed that allows consideration of the three-way interactions between urban density, traffic congestion and labour supply. The model would therefore, for example, allow an investigation of second-best congestion pricing with distorted spatial labour markets. This matter, however, will be addressed in a companion paper. The present paper is instead concerned with congestion pricing schemes that are second-best 'by design,'[2] like the examples mentioned above, and not because distortions exist elsewhere in the spatial economy.

Prior studies of traffic congestion in the monocentric model have typically looked at first-best congestion pricing measures, although second-best issues arising from non-optimal allocations of land to road capacity have been considered (e.g. Arnott [3]). The recent contribution by Mun et al. [10] is an exception. They focus on second-best optimal

[2] Note that these second-best schemes may eventually be more beneficial to society than a first-best scheme when also implementation and transaction costs are taken into account (such costs are ignored in this paper). Under these conditions, second-best pricing would be more efficient overall than first-best pricing, and an ambiguity in terminology arises. Milne et al. [9] for that reason distinguish between 'narrowly first-best' policies, that achieve equality between marginal costs and benefits throughout the system under consideration, and 'broadly first-best policies,' that maximize social welfare taking into account the costs of regulation and therefore may involve second-best pricing.