Transportation Research Part B 108 (2018) 106–126

Contents lists available at ScienceDirect

Transportation Research Part B

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

Demand imbalances and multi-period public transport supply

Daniel Hörcher*, Daniel J. Graham

Railway and Transport Strategy Centre, Department of Civil Engineering, Imperial College London, United Kingdom

ARTICLE INFO

Article history:
Received 30 May 2017
Revised 15 December 2017
Accepted 16 December 2017

Keywords:
Public transport
Transport supply
Crowding
Capacity optimisation
Demand imbalances

ABSTRACT

This paper investigates multi-period public transport supply, i.e. networks in which capacity cannot be differentiated between links and time periods facing independent but nonidentical demand conditions. This setting is particularly relevant in public transport, as earlier findings on multi-period road supply cannot be applied when the user cost function, defined as the sum of waiting time and crowding costs, is nonhomogeneous. The *presence* of temporal, spatial and directional demand imbalances is unavoidable in a public transport network. It is not obvious, however, how the *magnitude* of demand imbalances may affect its economic and financial performance. We show in a simple back-haul setting with elastic demand, controlling for total willingness to pay in the network, that asymmetries in market size reduce the attainable social surplus of a service, while variety in maximum willingness to pay leads to higher aggregate social surplus and lower subsidy under efficient pricing. The analysis of multi-period supply sheds light on the relationship between urban structure, daily activity patterns, and public transport performance.

1. Introduction

Public transport is supplied by multi-product firms in the sense that services are normally provided along predetermined lines with several stops in both directions. From an economic point of view, each direction of each inter-station section in various time periods can be considered as an individual market with or without demand interactions between them. Due to operational constraints, these markets are often served with the same capacity. As demand for public transport is hardly identical in separate markets, first-best capacity provision can never be feasible in reality. From the operator's point of view, this constraint translates into the fact that public transport services are subject to demand imbalances, and a second-best capacity has to be determined in a multi-period framework.

The presence of demand fluctuations is hardly questionable. However, their *magnitude* may differ across a wide range. In this paper, we focus on the simplest case of transport supply under demand imbalances: the back-haul problem, with independent demand curves. We show using a supply optimisation model that the magnitude of demand imbalances can have a crucial impact on the average crowding experience of passengers. Moreover, beyond the optimal capacity, the economic and financial performance of the service is also affected by the differences in ridership in jointly served markets, controlling for the aggregate scale of operations. In Section 3.2 we show that the bigger the deviation in market size, the lower the amount of social surplus that a public operator can achieve, and the more subsidies it will need to cover its losses. By contrast, imbalances in willingness to pay between joint markets reduce the optimal subsidy and leave more aggregate benefits

* Corresponding author.
E-mail address: d.horcher@imperial.ac.uk (D. Hörcher).

https://doi.org/10.1016/j.trb.2017.12.009
0191-2615/

for society. The core message of the paper is that the level of demand asymmetry as an *external factor* has significant impact on the economic and financial performance of public transport provision.

The magnitude of demand imbalances in the context of road provision is out of the main focus of transport research. Small and Verhoef (2007, Section 5.1.1) derive that the standard self-financing result of Mohring and Harwitz (1962) survives in the multi-period setting as well, assuming (i) constant returns to scale in congestion technology, (ii) neutral scale economies in capacity provision, (iii) perfectly divisible capacity, and (iv) time-varying first-best static congestion pricing where the toll equals to the marginal external congestion cost. The first assumption implies that the user cost function is homogeneous of degree zero, so that Euler's theorem can be applied to relate the impact of marginal demand and capacity deviations on the user cost of travelling. The empirical literature confirms that the three assumptions on cost functions are not far from reality in road transport, and therefore the transport economics community did not see much potential in further investigating the optimisation of road supply in a multi-period context specifically.[1] In public transport, however, user costs are far from constant returns to scale due to the well-known Mohring effect (Mohring, 1972). This paper fills in an important gap in the literature with the analysis of multi-period supply optimisation in public transport.

The paper presents two main lines of research. It contributes to the literature of public transport economics with a number of additional *theoretical* insights. We discuss

- **T1** the cohabitation of Mohring-type waiting time benefits and negative crowding externalities in a public transport model,
- **T2** the application of the Cost Recovery Theorem in the presence of waiting time as well as crowding externalities, and
- **T3** the interplay between frequency and vehicle size provision when demand is unevenly distributed between jointly served markets.

The more *policy oriented* branch of the paper investigates the effect of the magnitude of demand imbalances on

- **P1** second-best choice of frequency and vehicle size,
- **P2** the resulting peak and off-peak occupancy rates,
- **P3** maximum social surplus that can be reached with second-best supply, considering constant total willingness to pay for the service, and
- **P4** the amount of subsidy which is required to cover the financial deficit under efficient pricing.

The upcoming sections are structured as follows. Section 2 sets the field for subsequent analyses with a baseline supply optimisation model and a discussion of theoretical research questions T1 and T2 in the list above. Then, Section 3, the backbone of the paper, deals with the investigation of second-best supply in the back-haul problem. In particular, Section 3.1 begins with a simple inelastic demand setting which enables us to uncover the mechanics behind theoretical topic T3 above, while Section 3.2 presents core insights on major research objectives P1–P4 in connection with demand imbalances. The most relevant research outcomes are summarised in Table 3. Finally, Section 4 outlines an agenda for future research and Section 5 concludes.

2. Fundamentals of public transport supply

In transport economics theory, the main topics of interest in supply optimisation include (i) decision rules for optimal capacity setting, (ii) the determinants of short-run marginal social costs of service usage that form the basis for efficient pricing, and (iii) the degree of self-financing under socially optimal pricing. This section follows the same steps of analysis for the specific case of public transport.

2.1. Earlier literature on public transport capacity

Jara-Díaz and Gschwender (2003) provide a comprehensive review of the evolution of early capacity models. Most of these contributions kept the methodological framework of assuming inelastic demand, constructing a social cost function, and minimising it with respect to the optimal frequency and other supply-side variables.

Waiting time: The most common elements of public transport models since Mohring (1972) consider waiting time as a user cost and frequency as a decision variable. These imply scale economies in user costs, as high demand leads to high frequency, low headways, and lower expected waiting time for all users. We further investigate this mechanism in Section 2.3.

Cycle time: Several authors model that cycle time (i.e. the running time of vehicles) may be a function of the number of boarding and alighting passengers at intermediate stops through dwell times. This makes the case for a negative consumption externality, because boarding imposes additional travel time cost on passengers already on board. This feature is an important component of capacity models focusing primarily on bus operations,[2] e.g. Jansson (1980), Jara-Díaz and Gschwender (2003), Jara-Díaz and Gschwender (2009), Tirachini et al. (2010) and Tirachini (2014).

[1] Exceptions including Bichsel (2001) and Lindsey (2009) focused on second-best scenarios with pricing restrictions and uncertainty.

[2] Note, however, that travel times of rail services are generally much less sensitive to the number of boarding and alighting passengers than buses with a front door boarding policy. Assuming endogenous train length, dwell times are definitely not linear in the number of boardings, because the optimal number of doors may increase with demand.