Pricing public transport services

Jan Owen Jansson, Johan Holmgren and Anders Ljungberg

Linköping University Post Print

N.B.: When citing this work, cite the original article.

Original Publication:
Jan Owen Jansson, Johan Holmgren and Anders Ljungberg, Pricing public transport services, 2015, Handbook of research methods and applications in transport economics and policy, 260-308.

http://dx.doi.org/10.4337/9780857937933.00022

Copyright: Edward Elgar Publishing.
http://www.elgaronline.com/

Postprint available at: Linköping University Electronic Press
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-125188
13. **Optimal Public Transport Pricing**

Jan Owen Jansson, Linköping University, Sweden  
Johan Holmgren, Molde University College - Specialized University in Logistics, Norway  
Anders Ljungberg, Transport Analysis, Sweden

13.1 Introduction

This chapter aims at outlining pricing policy for public transport that maximizes the social surplus, that is, the sum of the producer surplus and the consumer surplus, while internalising possible system-external costs. It starts by presenting the door-to-door transport cost as a key concept in price theory for public transport, and then first principles of optimal pricing valid for all modes of public transport are laid down. These principles are applied to urban (short-distance) public transport in sections 2-5 and to interurban (long-distance) public transport in section 6. Section 7 summarises the methodological conclusions.

13.2 Methodological focus and first principles of optimal pricing

On the cost-side the point of departure for the analysis is the fact that the total costs of each transport system are borne by three groups of individuals: the service producers, the service users, and some of the system-external outsiders.

\[ TC = TC_{prod} + TC_{user} + TC_{ext} \]

This total cost concept is usually summarized as the total social costs. For simplicity it is left without a suffix, which also goes for the average social cost (AC) and marginal social cost (MC).

At the level of the individual traveller, it is important to note that, as distinct from most individual transport, a journey involving public transport consists of more than one stage: the concept of the door-to-door trip cost is meant to pay attention to this fact. A close relative to the door-to-door cost concept is the generalized cost (GC) which in many cases should replace ordinary price in the demand function. GC should, in optimum, include charges made
on the travellers as proxies for the producer and external cost components in the door-to-door trip cost. \( GC = P + AC_{user} \), where \( P \) stands for the charge(s) on the traveller and \( AC_{user} \) for the non-monetary sacrifices of the transport service users.

### 13.2.1 Door-to-door trip costs

In analogy with standard supply system cost theory, the door-to-door trip cost by public transport can be divided into the *distribution* cost, including the cost of the first and last stages, and the *production cost* of the main stage, that is the public transport itself.

The significance of this division of the trip cost for modal choice is schematically illustrated in Figure 13.1 where the door-to-door trip cost is given for three main modes of public transport as functions of trip length.

![Figure 13.1. Illustration of the emergence of modal niches in public transport: door-to-door generalized costs against trip length for public transport by road, rail and air](image)

On the vertical axis of Figure 13.1 the relative distribution costs typical of road transport by bus, rail transport, and air transport are shown. The underway cost is given as proportional to
trip length for simplicity, as a compromise between its two main components, the progressively increasing passenger time costs and the degressively increasing costs of the transport producer with respect to trip length. The infrastructure costs are represented by track and terminal charges on the operators passed on to the travellers through the fares.

Each mode of public transport has a niche because a faster mode has substantially higher distribution costs than a slower mode, as a rule. High speed is obtained at the sacrifice of environmental disturbances that require a location of airports far out of cities. Large railway stations and overground rail track that still exist in densely populated urban areas were built long ago when, with the values of today, encroachment and disturbance costs were generally underestimated.¹

The traditional niche for rail and middle-distance bus transport has diminished concurrently with the private car diffusion. High-speed rail transport has in recent times made discernible inroads into middle-distance air transport markets. The present modal split in Sweden for non-urban travel is illustrated in Figure 13.2.

Source: Swedish National Travel Survey 2011-2012

Figure 13.2. Modal split of non-urban travel of different trip length

¹ A variation of this kind of comparative cost illustration takes demand density as the main cost determinant along the horizontal axis, giving rise to somewhat different modal niches, where rail and air change positions, but bus remains the least-cost mode in the initial range (compare to Allport, 1981).
As seen in the diagram, individual car travel is dominant in the non-urban travel market up to trip lengths in the region of 400-800 km.

What about urban transport? Short-distance trips which are made in urban areas constitute the largest trip category in terms of number of trips. They cannot be represented in Figure 13.2 for expository reasons. The thin, shaded column in Figure 13.1 is the relevant range, but the cost functions in Figure 13.1, illustrating middle- and long-distance door-to-door trip costs, should not be extrapolated backwards. The urban modal split is a more complex question where town planning and urban transport policy are intertwined. The simple model of Figure 13.1 has to be modified quite considerably. To begin with, let us exemplify the structure of the door-to-door trip costs in urban areas by the figures of Table 13.1. The underlying data relate to work trips by bus compiled in comprehensive surveys carried out by the Transport and Road Research Laboratory in England in the 1970s and 1980s.

The salient feature of the cost structure is the dominance of the user costs, comprising both the distribution cost and the travel time cost. It can be assumed that the producer costs per passenger were reasonably close to the level of fares at that time, which was in the range of 20 – 40 per cent of the total social cost as seen in Table 13.1.

Table 13.1. Door-to-door cost structure of travel to/from work by bus in UK in 1976
(index numbers; total GC for a 2 km trip = 100)

<table>
<thead>
<tr>
<th>Cost component</th>
<th>Trip length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 km</td>
</tr>
<tr>
<td>Walking time</td>
<td>28</td>
</tr>
<tr>
<td>Waiting time at bus stop</td>
<td>24</td>
</tr>
<tr>
<td>Riding time</td>
<td>25</td>
</tr>
<tr>
<td>Change of bus²</td>
<td>2</td>
</tr>
<tr>
<td>Fare</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
</tr>
</tbody>
</table>

Source: Adaptation of some results in Webster (1977), Webster et al. (1985)

The non-monetary user costs are highly relevant both for investment CBA and price theory. The density of the network of lines and the frequency of public transport services on each particular line are just as important for the transport system optimization as the characteristics

---

² For the individual commuter who has to change buses or trains, this inconvenience carries a relatively high cost (Currie et.al. 2005) and therefore such trips by public transport are quite rare, and the average cost of vehicle transfer for all commuters is low.
of the transport vehicles themselves. The role played by the user costs for working out the optimal level and structure of fares is a main concern in what follows.

**13.2.2 Standard price theory extended by taking due account of user costs**

In the ideal world we are presuming, \( MC = MC_{\text{prod}} + MC_{\text{user}} \), because the transport system external costs, both those caused by the vehicles in the public transport system concerned and the competing modes of transport, are assumed to be internalized by, for example, appropriate fuel taxation.

The following extension of standard price theory implies that ordinary price (\( P \)) is replaced by the generalized cost (\( GC \)) that is made up of the sum of the average non-monetary user cost and the price, or fare in the present case. Where the former component is something like two thirds of \( GC \), as indicated in the aforementioned example (Table 13.1), this approach makes a great deal of methodological difference when the production factor inputs have a significant influence on the user costs besides contributing to output in line with the production function. This is demonstrated in an algebraic model where supply and demand are laid down as follows:

- **Total output** = \( Q = f(X_i) \)
- **Production factors** = \( X_1 ... X_i ... X_n \)
- **Total producer costs**, \( TC_{\text{prod}} = \sum_i p_i X_i \)
- **Total non-monetary user costs**, \( TC_{\text{user}} = Q \cdot h(X_i) \)
- **Generalized cost per trip**, \( GC = P + h(X_i) \)

Trip demand is a function of the generalized cost. In its reversed form the demand function is represented by \( GC \) as a function of \( Q \), written \( g(Q) \).

For the social surplus maximization, form the Lagrangian,

\[
\pi = \int_0^Q g(Q) dQ - \sum_i p_i X_i - Q \cdot h(X_i) - \lambda [Q - f(X_i)] ,
\]

(13.1)

and take the partial derivatives of \( \pi \) with respect to \( X_i \) and \( Q \) and set them equal to zero:
\[
\frac{\partial \Pi}{\partial X_i} = -p_i - Q \frac{\partial h}{\partial X_i} + \lambda \frac{\partial f}{\partial X_i} = 0 
\]  
(13.2)

\[
\frac{\partial \Pi}{\partial Q} = g(Q) - h(X_i) - \lambda = 0 
\]  
(13.3)

\[
\frac{\partial \Pi}{\partial \lambda} = Q - f(X_i) = 0 
\]  
(13.4)

Separating \( g(Q) \) in (13.3) into \( P \) and \( h(X_i) \) yields:

\[
\lambda = P 
\]  
(13.5)

Inserting \( P \) for \( \lambda \) in (13.2) finally gives:

\[
\frac{p_i + Q \frac{\partial h}{\partial X_i}}{\frac{\partial f}{\partial X_i}} = P 
\]  
(13.6)

In the exceptional case where the production factors have no influence on the user costs, the ratio of the factor price \( (p_i) \) to the marginal product \( \frac{\partial f}{\partial X_i} \) of every production factor is equal to the product price \( (P) \) in optimum. In the normal case where some user costs are affected by the production factor inputs, this should be taken into account by an additional term in the numerator such as \( Q \frac{\partial h}{\partial X_i} \) in (13.6) above. As a foretaste of what is discussed at greater length in what follows, it can be mentioned that when \( X_i \) stands for plant size, the user cost component can act as a relatively small addition to the optimal product price. On the other hand, when \( X_i \) stands for the number of plants in the production system considered, the user cost component in (13.6) is negative that could substantially reduce the optimal product price. In public transport systems this is known as the “Mohring effect”\(^3\).

It is true that the generalized cost concept is not problem-free, but we argue that at least in urban transport systems attempts to keep to conventional microeconomic supply and

demand analysis meet with greater problems, because it requires that the quality of service is kept constant along the expansion path.

13.2.3 A fundamental theorem

It is easily demonstrated that irrespective of which, or how many of the production factors are marginally increased, the resultant optimal product price \( P \) will be the same, provided that the initial position is a least-cost factor combination.

Taking the differentials of the production function and the total costs, \( TC = TC_{\text{prod}} + TC_{\text{user}} \), we have:

\[
dQ = \sum_i dX_i \frac{\partial f}{\partial X_i}
\]

\[
dTC = dTC_{\text{prod}} + dTC_{\text{user}} = \sum_i dX_i \left( p_i + q \frac{\partial h}{\partial X_i} \right) + dQ \cdot h(X_i)
\]

The ratio of \( dTC \) to \( dQ \) constitutes the system marginal cost, \( MC \). Making use of (13.6) above, the expression written within brackets in (8) can be replaced by \( P \frac{\partial f}{\partial X_i} \), and \( MC \) can be written:

\[
MC = \frac{dTC}{dQ} = \frac{P \sum_i dX_i \frac{\partial f}{\partial X_i} + h(X_i) = P + h(X_i) = GC}{\sum_i dX_i \frac{\partial f}{\partial X_i}}
\]

The differentials of those factors assumed to be fixed in the price-relevant “run” obviously take a value = 0. The number of factors of production assumed to be fixed – none, or any combination of factors – would not affect the ratio of \( dTC \) to \( dQ \), which in all cases is equal to \( P + h(X_i) \), because the summation terms in the numerator and the denominator of the first term will always be the same, and \( h \) is the average user cost in the initial situation. The general optimality condition is \( MC = GC \), just as was stated at the outset.\(^4\)

\(^4\) Compare also Jara-Diaz (2007).
13.2.4 The short- or long-run marginal cost controversy

A corollary of the fundamental theorem is that for substantially larger factor increments than the next to infinitesimal ones that are considered in the formulation of the general marginal cost expression (9) above, the marginal cost curves defined by different number of variable factors will be more and more divergent. There is a systematic pattern of the increasing divergence. The more factors that are variable, the flatter the slope of the MC-curve. This pattern is illustrated in Figure 13.3. The MC-curve (1) – the steepest one – applies where only one factor of production is variable, and in the other extreme case, where all factors are variable, which is true in the planning stage before the plant or system of plants in question are built, the slope may even be negative as illustrated in the diagram.

![Figure 13.3](image_url)

*Figure 13.3. dTC/dQ in optimum and out of optimum for different subsets of production factors assumed to be variable*

At first this may seem counterintuitive. Long-run marginal cost pricing is often believed to result in a higher price than short-run marginal cost pricing. The truth is that in optimum the price will be the same, but out of optimum the two pricing principles give different results. GC = LRMC would be lower than GC = SRMC when considering increases in the production from the existing optimal level, and the other way round. This agrees with common sense: when factor fixity is pronounced, it is difficult and costly to increase
production on short notice to meet a surge in demand, and cost savings are correspondingly
difficult to make in the short run when demand suddenly decreases.

When dealing with public transport, in particular buses on a common road network, the
question of the relevant “run” should not be an issue. The road network can be regarded as
given, while the vehicles – that is the “plants” of the system – are mobile and possible to
vary in number, size and type on relatively short notice.

13.2.5 First-best versus second-best
Four of the following five sections deal with urban public transport. Why is methodology for
optimal pricing and investment in urban public transport particularly important? As opposed
to practically all other activities in modern society, including interurban transport, there is a
blatant lack of progress in this field which is a main reason for the degeneration of urban life
in many large cities. The former EU commissioner for transport has described the urge for
“improving the quality of life in our cities” in the following way:

*Those of us who live and work in cities, and that is the majority of the population of Europe,
are increasingly aware of the deteriorating quality of life there. The pressure on space, the
growing problem of urban pollution, the widening variations in living standards are all
making our cities increasingly difficult places to run and certainly to enjoy.* (Neil Kinnock,
1998)

Improved public transport has a key role in meeting these challenges. To make the most of it,
the goal should be maximization of the social surplus which requires both a widening of
perspective and a deepening of the knowledge of central aspects of the supply and demand in
this field.

Public transport policy cannot alone solve the urban transport problem. Town planning and
transport infrastructure policy also play important roles. A big problem for economic analysis
is that a “market solution” involving marginal cost pricing of each competing modes of
transport is insufficient for obtaining a first-best optimum. Corner solutions requiring
regulations, including for example, complete prohibition of motor traffic in some sections of
the road network, may be a necessary part of the first-best solution. To illustrate one central
aspect of this complex of problems, Figure 13.4 suggests that in a typical European city with a
population of 1 million a network of city-bus lines, unobstructed by other traffic, would be an
unbeatable mode of public transport in the central city, as is also demonstrated by the
following model analysis. Where these preconditions do not exist – unfortunately, the normal
case in reality – the consequent relatively slow buses are challenged by underground rail transport even on short distances, in spite of the fact that the distribution cost of metro services is twice or more higher than that of surface transport by bus. For longer commuter trips, over ground rail and Bus Rapid Transit (BRT) on separate lanes could also get the better of the private car, provided that cars are charged optimal prices for road use and parking.

Figure 13.4. Supplementary illustration to Figure 13.1 above: door-to-door generalized costs against trip length for urban public transport

The price theory with applications in this chapter is normative in the sense that it aims at the first-best optimum. Also for practical analytical reasons second-best solutions are left out of consideration, because these would involve widely different public transport subsidization depending on the car traffic conditions (see Parry and Small 2009, Proost et.al. 2002, Proost and Dender 2008).
When it comes to non-urban public transport, the question of first-best vs. second-best is not as important, because the preconditions for the main competitor – the private car – are not as widely different as in an urban context, and transport infrastructure policy could not make such a great difference.

13.3 Urban public transport 1: central city bus line network

It is helpful to divide the total urban transport market into two parts, at least in the case of large cities: the market for intra-central-city travel, and the market for commuter travel between the suburbs and the central city. In the latter case each particular bus line or commuter train service can be regarded as the pricing-relevant transport system, while the whole central city bus line network has to be dealt with as a whole.

The system approach is quite demanding in this case in view of all the natural differences in the urban landscape, the structure of settlements and the transport infrastructure of different central cities. The detailed design of urban bus transport systems is a matter of thorough operations analysis in each particular case. However, here we look for the general features and seek to apply the preceding modified microeconomic supply and demand paradigm to public transport services. It is helpful to start by a “base case” where the complications of the typical peakiness of the demand for commuter train and bus line services can be ignored. Short business trips during office hours, shopping trips when shops are open, and leisure trips in the evenings, together make for a reasonably uniform expected travel demand in time and space the whole day and early evening as regards intra-central city travel in large cities. Fluctuations in actual demand are mainly random, which are met by providing reserve capacity; literally full capacity is attained only in extreme situations. The buses concerned have ample room for standing passengers that constitute the reserve capacity. For quality reasons it is desirable that all passengers are seated. Let us define full practical capacity as a state where the expected seat occupancy rate is unity. In a second model in the next section the focus is on work and school commuting between the suburban residential areas and the central city where workplaces and schools for higher education are concentrated. Systematic peaks in time and space is the salient feature of this demand.

To design an optimal network of bus lines in the central city is nevertheless a very complex matter when it comes to details such as the exact drawing of each particular line and the
appropriate location of bus stops. For the present purpose a highly stylized design is sufficient. We can resort to the classic justification of “urban transport parables” by R. Strotz.

...due to the immense complexities of the problem a pre-scientific approach has to be adopted ... telling simple little stories, each of which highlights a particular though ubiquitous problem. From each of these we wish to draw a moral, a principle that ought not to be overlooked when a more complex situation is to be faced. (Strotz 1965 p.128)

We are assuming ideal conditions for the central city bus line network in question, corresponding to the illustration of the relative door-to-door trip cost by unobstructed bus transport in Figure 4 above. Our intention is not to picture how it normally works, but how it could work under first-best circumstances.

The following symbols are used in the analysis:

A = central city area
B = number of bus trips per hour in this area
Y = B/A = trips generated per hour and km² (density of demand)
L = average trip length
N = number of busses in the system
S = number of seats per bus (=bus size)
V = cruising speed (bus speed excluding dwell time at stops)
R(S, V) = overall bus speed
H = bus service hours per day
t = boarding/alighting time per passenger
C(S) = day cost of bus with driver(s)
r = riding time cost per hour
W1 = walking time cost per hour
W2 = waiting time cost per hour
α = walking time per trip in proportion to the distance between parallel bus lines
β = waiting time per trip in proportion to the headway
k = composite constant in the distribution cost formula: k = 2 α·β·w₁·w₂
13.3.1 The producer costs

On the present assumption of homogeneous demand, and provided that the buses are fully occupied – on average, all seats are occupied – total passenger-kilometres per hour \((BL)\) equal total bus-seat-kilometres per hour produced in the system

\[ BL = SNR(S,V) \]  \hspace{1cm} (13.10)

The traffic operation costs per day of the service producer can be assumed to be equal to the product of the number of buses in operation \((N)\) and the “day cost” \((C)\) of the buses, consisting of bus capital and running costs including garaging, maintenance and repair costs, and the wage cost of the driver(s). The day cost is assumed to be a function of the main design characteristics, bus size \((S)\) and cruising speed \((V)\), while the wage cost is a dominant fixed component, independent of bus size. Expressed per day of operation, the total producer cost of traffic operations is written:

\[ N \cdot C(S,V) = \text{producer costs per day} \]  \hspace{1cm} (13.11)

Depending on their exact definition, the constitute an appreciable part of the total costs of public transport enterprises. It is customary to exclude the overhead costs from the total traffic operation costs in marginal cost analysis. We disagree with this custom, but unfortunately there is no unobjectionable way of including the corresponding production factors in the production function. For the following theoretical discussion, this does not matter because we just consider increases in \(N\) and \(S\) in the formulation of the price-relevant costs. In section 3.9 the overhead cost problem is taken up for discussion.

13.3.2 The user costs

The user costs are divided into the “distribution cost” and the riding time cost. The former includes the costs of catching the bus involving both walking and waiting and another walk to the final destination.

In the present model the distribution costs are fully taken into account by considering a network of bus lines. Thereby both the walking and waiting efforts of bus riders are explicit objects of adjustment in the striving for social surplus maximization.
It is assumed that the density of demand for travel in the central city within office and shopping hours is reasonably uniform, that is, the points of origin and destination are evenly spread in the central city. Travel is made by bus on 2n lines forming a grid network. The time and effort of a representative traveller per trip can be obtained as follows:

The average walking time to/from bus stops is proportional to the distance between parallel bus lines in the direction heading for – the proportionality constant is denoted $\alpha$ – and the average waiting time at the bus stops is proportional to the time distance between the buses (headway) – the proportionality constant is denoted $\beta$. An efficiency condition for the network design is that the average walking time cost and the average waiting time cost per trip should be the same, irrespective of the relative costs of walking and waiting time. This may seem counter-intuitive, but is obtained when the density of lines and the frequency on each line are adjusted in response to the relative walking and waiting time costs, $w_1/w_2$. This can be shown as follows:

The average walking time per trip, involving two walks, comes to $\alpha \sqrt{A} / 2n$, where $A$ is the central city area, $n$ is the number of northbound and southbound as well as eastbound and westbound lines in the grid and walking speed is assumed to be 4 km/hour.

The headway is obtained by first taking the distance in kilometres between the buses on each line, which is the ratio of the double route length, $2\sqrt{A}$ and the number of deployed buses per route $N/2n$. Dividing this by the overall speed, $R$, gives the headway. Applying the proportionality constants $\alpha$ and $\beta$, the total walking and waiting cost per trip can consequently be written:

$$AC_{distr} = \frac{\alpha w_1 \sqrt{A}}{2n} + \frac{4\beta w_2 n \sqrt{A}}{NR}$$

(13.12)

The optimal grid density can be separately determined by minimizing $AC_{distr}$ with respect to $n$, given the central city area, the total number of buses and the overall bus speed. Increasing $n$ means that the walking distance per bus trip is reduced, while the headway is increased.

$$\frac{\partial AC_{distr}}{\partial n} = -\frac{\alpha w_1 \sqrt{A}}{2n^2} + \frac{4\beta w_2 \sqrt{A}}{NR} = 0$$
If the walking cost per hour, $w_1$, is high relative to the waiting cost per hour, $w_2$, it means that the density of the network should be high and the frequency of service on each particular line relatively low, and the other way round.

Inserting the value of $n^*$ in the distribution cost expressions above, it is found that the efficient walking and waiting costs per trip are equal.\(^5\) The final distribution cost as:

$$AC_{\text{distr}} = 2\sqrt{\frac{2w_1w_2}{NR} \frac{\alpha \mu \lambda}{A}}$$  

(13.14)

The parameter values in the numerator of the square root expression (13.14) can finally be merged into a single constant, $k$.

$$AC_{\text{distr}} = 2\sqrt{\frac{kA}{NR}}$$  

(13.14a)

The riding time cost is easily laid down. The average riding time cost per passenger trip is the product of the riding time cost per hour ($r$) and the average trip length ($L$) divided by the overall speed ($R$).

$$AC_{\text{ride}} = \frac{rL}{R}$$  

(13.15)

### 13.3.3 Alternative formulations of the price-relevant cost

We are now going to apply the optimal price formula (6) above in the model of a central city grid network. For the social surplus maximization the applicable lagrangian expression, corresponding to (13.1) above from which the optimal price formula (13.6) is derived, is now written:

$$\pi = H\int_{0}^{B} g(B) dB - TC_{\text{prod}} - TC_{\text{user}} - \lambda \left[ B - \frac{NSR(S,V)}{L} \right]$$  

(13.16)

\(^5\) Compare to Chang and Schonfeld (1991) where further references are given
where $TC_{prod} = N \cdot C(S,V)$, and $TC_{user} = HB \left[ 2 \left( \frac{kA}{NR(S,V)} \right) + \frac{rL}{R(S,V)} \right]$

The system optimization involves adjustments of both the density of bus lines and the frequency of service on each line in accordance with the aforementioned principle, and the determination of size and number of buses. Assuming full capacity utilization, the latter two factors determine the former two qualities of service.

There are thus two main production factors or “design variables” involved, the number of buses ($N$) and the bus size ($S$), and consequently two ways of formulating the price-relevant cost, designated $P_N$ and $P_S$. The third main design variable that could come into question, the cruising speed ($V$), we choose to keep constant. Urban traffic conditions are the principal determinant of $V$, which is assumed to be out of control of the bus system planner.

The former alternative corresponds to the usual approach to optimal public transport pricing, which has been different variations of the “average cost of a marginal bus” approximation, that is, the production cost per passenger of an additional bus minus the benefits to the existing users of an additional bus per additional passenger. The deduction is called “the Mohring effect”. Its relevance for pricing has been questioned, and there are admittedly some remaining empirical problems. Both “half the headway” assumption and the value of the waiting time at the bus stops in the distribution cost formula, let alone the evaluation of the disguised waiting time, are marked by uncertainty. It is then reassuring, and useful for long-distance public transport pricing to be able to calculate the optimal price in an alternative way that results in a value that is less dependent on the user cost component. Considering an increase in bus size ($S$) has this advantage. However, we start with the usual approach.

13.3.4 Optimal bus fare 1: the price-relevant cost of increasing the number of buses

When considering $N$ as the factor of change in the derivation of the price-relevant cost, the result corresponding to (13.6) above is:
\[ \lambda = HP_N = \frac{\partial TC_{prod}}{\partial N} + HB \left( \frac{\partial AC_{distr}}{\partial N} + \frac{\partial AC_{ride}}{\partial N} \right) \]

The riding time cost according to (13.15) is independent of \( N \), so besides the producer cost, just the user distribution cost plays a role for the price-relevant cost. The derivative of \( AC_{distr} \) with respect to \( N \) is negative, so a deductible user cost component appears in the expression for \( P_N \).

\[ P_N = \frac{C}{H} + B \left( -\frac{1}{N} \sqrt{\frac{kA}{NR}} + 0 \right) \]

Observing that \( \frac{CL}{HSR} = AC_{prod} \) and \( BL = NSR \), we finally have

\[ P_N = AC_{prod} - \frac{1}{2} AC_{distr} \]

A seemingly very simple result is obtained. Since both \( AC_{prod} \) and \( AC_{distr} \) depend on \( N \) and \( S \), this expression does not say exactly what the optimal price is along all the expansion path, only that the optimal price is below \( AC_{prod} \).

**13.3.5 Optimal bus fare 2: the price-relevant cost of increasing bus size**

In particular the waiting time component of \( AC_{distr} \) in the expression for \( P_N \) can be difficult to get straight. When considering \( S \) as the factor of change this is a smaller problem. Applying formula (6) above now gives the following basis for the price-relevant cost derivation.

\[ \lambda = HP_S = \frac{N \frac{\partial C}{\partial S} + B \left( \frac{\partial AC_{distr}}{\partial R} + \frac{\partial AC_{ride}}{\partial R} \right) \frac{\partial R}{\partial S}}{\frac{\partial B}{\partial S}} \]

Since the overall speed (\( R \)) to some degree depends on \( S \) (whereas \( R \) is independent of \( N \), given \( S \)), the derivatives of the two user costs components with respect to \( S \) had better be
calculated in two steps. Each one can be calculated as the product of the derivative with respect to \( R \) and the derivative of \( R \) with respect to \( S \).

Observing that \( \frac{\partial B}{\partial S} \) is equal to \( NR(1+\frac{\partial R}{\partial S})/L \), all derivatives in (13.18) can be transformed into more comprehensible elasticities, and we have:

\[
P_S = \frac{E_{CS}AC_{prod} + E_{RS}\left(\frac{1}{2}AC_{dist} + AC_{ride}\right)}{1 + E_{RS}} \tag{13.18a}
\]

\( E_{CS} \) stands for the elasticity of the day cost of buses with respect to bus size, and \( E_{RS} \) for the elasticity of overall speed with respect to bus size. \( E_{RS} \) is negative but likely to be small, so the denominator of (13.18a) is minutely smaller than unity.

The product of \( E_{CS} \) and \( AC_{prod} \), after dividing by \( 1+E_{RS} \), stands for the (quality-unadjusted) marginal producer cost \( MC_{prod}^S \), given full capacity. \( AC_{prod} \) is substantially greater than \( MC_{prod}^S \) owing to considerable economies of bus size (which will be confirmed by the following empirical exploration). The total user cost contribution to \( P_S \) takes a low value in this case. What is notable, however, is that the user cost component now constitutes an addition to the producer cost component. When the chosen factor combination minimizes the social costs for a given level of output, \( P_N = P_S \).

### 13.3.6 Empirical exploration of two key relationships, \( C(S) \) and \( R(S) \), and a closer look at the overhead costs

In order to apply the formulas for \( P_N \) and \( P_S \), there are two key relationships that should be estimated: the relationship between the bus capital and running costs and bus size, \( C(S) \), and the relationship between overall speed and bus size, \( R(S) \). As it is assumed that the cruising speed (\( V \)) is independent of bus size, the latter task boils down to finding out how the time of boarding and alighting depends on bus size.

### 13.3.7 Bus costs and bus size

The main item of the total traffic operation costs is the cost of drivers, which can be assumed to be independent of bus size. Next comes the costs of the buses themselves. The
bus running cost mainly consists of the cost of fuel and wear and tear (repair and service). The bus standing costs mainly consists of capital cost and the cost of garaging.

Jansson (1984) shows linear relationships between the bus costs and bus size based on data from *The Commercial Motor Journal* for both bus running cost and bus standing cost. The data concern British urban traffic conditions in the 1970s. The economies of bus size are as seen particularly pronounced in the running costs.

Regression obtained:

\[
\text{bus running cost per mile (pence)} = 11 + 0.14S \quad (R^2=0.94) \tag{13.19a}
\]

\[
\text{bus standing cost per week (£)} = 6.5 + 0.72S \quad (R^2=0.98) \tag{13.19b}
\]

Apart from the large fixed cost of the driver, the appreciable ordinates of the linear relationships imply that the size-elasticities are increasing with bus size from 0.1 to 0.6 and from 0.4 to 0.9, respectively, in the range of 10-100 seats.

It is interesting that recent data from Sweden indicate that similar linear relationships still apply, as appears from figure 13.6.
Source: Own estimation based on data collected from operators and the Regional Public Transport Authorities in 2012.

Figure 13.6a. Bus running cost against bus size

Figure 13.6b. Bus standing cost against bus size

Regression obtained:

\[
\text{bus running cost per km (Eurocent) } = 22 + 1,03S \quad (R^2 = 0,98) \tag{13.20}
\]

\[
\text{bus standing cost per day (EUR) } = 33 + 2,38S \quad (R^2 = 0,97) \tag{13.21}
\]

The span of the size-elasticities coming out for the running cost is somewhat higher than in the earlier study (0.3–0.8 versus 0.1-0.6), whereas the span of the size-elasticities for the standing cost is about the same.

13.3.8 Dwell time and bus size

Bus productivity is not simply proportional to the product of bus size (holding capacity) and running speed. The time the vehicle has to spend at stops matters, too. The question is, how much does the time requirement of boarding/alighting hamper output?

The dependence of overall speed \((R)\) on bus size is in the first place due to the fact that the time of standing still at stops (the dwell time) depends on \(S\), given the rate of capacity utilization. The total of dwell time in the system per hour has two components. One is proportional to the number of stops, provided that all buses call at each stop, and can be included in the cruising time per km defining the cruising speed, \(V\). The other is assumed to be proportional to the number of boarding passengers (first “proportionality hypothesis”). The boarding/alighting time proper per hour in the system is then equal to \(tB\), which spread over \(N\) buses comes to \(tB/N\) per bus. From the production function (13.10) the ratio \(B/N\) can be replaced by \(SR/L\), and overall speed \((R)\) can be written:
The importance of $L$ and $t$ for urban bus transport is quite striking. For example, given the bus size, an average trip length of 10 km and a low value of $t$, for example 1 second, the overall speed reduction is about 2 per cent compared to driving without passengers (but calling at every stop), while in a case where $L$ is just 1 km and $t$ is 5 seconds, the boarding and alighting delays would reduce overall speed by half.

There have been several studies over the years of dwell time, and how it is affected by different factors. Usually regression analysis is applied in order to explain variation in data collected by manual counts and in some cases (e.g. Fernández et.al. 2009) by video surveillance. It is obvious that the b/a-times vary with contextual factors such as the demographic profile of the users and the type of vehicle used. The preparation time varies between 2,38 seconds (York, 1993) up to 12 seconds (Aashtiani and Iravani, 2002) while boarding time varies between 0,5 seconds when prepaid cards were used (TRB, 2000) up to 16,22 seconds in cases where the passenger pays in cash and the driver has to return change (Tirachini, 2013). Alighting times vary between 0,4 seconds when using six door buses (TRB, 2000) up to 4,9 seconds (Li et.al., 2006). A concise overview of previous results is provided by Tirachini (2013).

Only a few of these studies relate dwell time to number of passengers boarding and alighting that makes it possible to test the proportionality hypothesis. Mention can be made of Guenthner and Hamat (1988), Dueker et.al. (2004) and Rajbhandari et. al. (2003). The latter tried several different functional forms in estimating dwell times ($d$) against the total number of passengers ($Z$) boarding and alighting at a stop. The best model was found to be:

$$d = 7.31 \cdot Z^{0.76}$$

(13.23)

Thus they found that the total dwell time is increasing with the number of passengers but at a declining rate. Such a conclusion must be due to the fact, first that stopping and starting time could be less dependent on bus size than actually boarding/alighting time and secondly that the observations of boarding and alighting were made on different kinds of vehicles and/or in situations where different kinds of payment systems were in use. When the number of passengers in the system increases, larger vehicles with more doors will be used and it will be
efficient to invest in faster payment systems. However, given the vehicle type and payment system it can be assumed that the b/a time per passenger is constant as long as the bus is not crowded (i.e. filled beyond practical capacity). A linear approximation of (13.23) has been obtained by regression analysis that makes it possible to divide total dwell time \(d\) into fixed preparation and stopping and starting time \(c\) and time for actual boarding and alighting. The result is:

\[ d = c + tZ = 31.4 + 2.09Z \]  

(13.24)

In the calculations in following sections b/a time per passenger is therefore set to 2.09 seconds.\(^6\)

### 13.3.9 The overhead cost problem

The overhead costs pose different problems for the practical cost accountant and the theoretical economist. The former struggles with the allocation of the overheads between different lines of the business and further on between different products, while the economist is mainly concerned with the question of to what extent the overheads are to be viewed as variable or fixed costs, that is: how do they depend on the total output volume? In older American literature on transport economics, in the times of strong anti-trust policy initiatives and deregulation, a topical question was if that could go too far. Would not disintegration of large concerns increase total overhead costs in the industry in question? However, at least so far as the transport industries are concerned it was found that overhead costs develop in proportion to the size of the company. To illustrate this, a chart for a number of bus companies in Britain is presented in Figure 13.7 where overhead costs are plotted against fleet size. The proportionality hypothesis is also borne out in linear shipping. A cross-section analysis of a large number of shipping lines made by Ferguson et al. (1960) showed that administrative costs amounted to 10 per cent of gross revenue, regardless of the size of the fleet. A similar result was reported in an investigation of American shipping lines ten years later (Devanny et al. 1972). Cross-section studies of the total costs of (American) railroad companies and bus transport companies do by and large support the proportionality hypothesis.

\(^6\) It can be noted that if a second order polynomial function is used in the regression, \(t\) is found to range from 2.75 seconds (when having only one passenger boarding and alighting) to 2.09 seconds if there are 140 passengers boarding and alighting (assuming full capacity utilization with a double articulated bus and that 70 passengers alight and 70 board at a stop).
Current Swedish data point in the same direction. In figure 13.8 the relationship between the costs of traffic operations and the reported overhead costs⁷ in the 21 Swedish counties are given. The figures are county averages of the years 2006-2011 and are shown in logarithms. There is an apparent linear relationship between the costs of traffic operations and the overhead costs implying that there are no economies of scale in the overheads.

⁷ These costs include administrative costs (including planning), marketing costs and cost for controlling and enforcing payments. (From the Swedish governmental agency called Transport analysis)
Figure 13.8. The relationship between traffic operations costs and overhead costs, county averages 2006 – 2011.

Using all annual figures (126 observations) from the Swedish counties, the average ratio of the overheads to the traffic operation costs is found to be 0.1 with only small variations between counties.

Based on this empirical evidence the fixity misconception as regards overheads is rejected. The remaining question is how this finding should be used in the derivation of the price-relevant marginal cost?

Theoretically, the correct procedure would be first to include salaried staff \( (M) \) in the production function, \( Q = f(N,S,V,M) \). Secondly, the cost of salaried staff \( (sM) \) should be added to \( TC_{\text{prod}} \), and thirdly, it should be examined to what extent, if at all, the user cost are affected – presumably reduced – by increasing \( M \). That factor of production should, in principle, be treated in the same way as the other factors and design variables in the formulation of the lagrangian expression used for the social surplus maximization,

\[
\lambda = HP_M = \frac{\partial TC_{\text{prod}}}{\partial M} + HB \frac{\partial AC_{\text{user}}}{\partial M}
\]

\( (13.25) \)
The problem in practice is obviously that there is insufficient knowledge of the appearance of a production function that includes $M$, from which the derivative $\partial B/\partial M$ could be calculated, as well as of how variations in $M$ would affect the user costs. In this predicament it is tempting to resort to the common, yet misconceived idea that although the overheads may be proportional to output in the long run, they are basically fixed in the short run, and therefore could be ignored when calculating the short-run marginal cost, the only one supposed to be price-relevant. The fallacy in this reasoning is that the equality of the price-relevant short-run and long-run marginal costs in optimum is overlooked, and as an unfortunate consequence the average variable cost ($AVC$) is used as a proxy for the short-run marginal cost. This is not just a rough approximation, but a systematic underrating of price-relevant cost.\(^8\)

The correct approximation in the present model is to add a 10 per cent mark-up to the price-relevant marginal cost corresponding to the share of the overheads in the total producer costs.

### 13.3.10 Costs and prices in conclusion

When we now have surveyed the two key relationships identified in the theoretical discussion, and got a clue of how to tackle the overhead cost problem, the social marginal cost ($MC$) pointed out in the theoretical analysis above as the critical function in the search for the ideal output and optimal fare, can be established and given a numerical representation.

A slight adjustment of the symbols used so far is appropriate for the base case conclusions. The central city size ($A$) is in fact inconsequential for the result in the model as long as the density of demand ($Y$), that is, bus trips generated per km\(^2\), remains the same. Then the distribution cost expression can be reformulated by substituting $BL/S$ from the production function (13.10) for $NR$ in (13.14a), and we have:

$$AC_{distr} = 2 \frac{kS}{YL}$$ (13.26)

### 13.3.11 Diagrammatical illustration

The expansion path is defined by the least-cost conditions at each level of demand ($Y$). This first of all means that both $N$ and $S$ must take particular optimal values for each value of $Y$. Diagrammatically the result is summarized in Figure 13.9. The ideal output is found where the inverse demand – $GC$ as a function of the density of demand (not shown in Figure 13.9) –

---

\(^8\) See further in Jansson (1984), chapter 4
and the system marginal cost \((MC)\) intersect. The optimal fare equals the difference between \(MC\) and \(AC_{\text{user}}\), shaded in the diagram. As seen, with the parameter values given and the key relationships found in the empirical work discussed in the preceding section, the optimal fare takes values from one fourth to two thirds of \(AC_{\text{prod}}\) in a wide range of the demand density.

\[\text{Figure 13.9. Final result in the central city network model}\]

A more detailed picture of the generalized cost and its components including the optimal fare in a low-density case, a medium-case, and a high-density case are presented in Table 13.2.
Table 13.2. Numerical illustration of the result of urban bus transport system optimization assuming uniform demand in time and space\(^9\)

<table>
<thead>
<tr>
<th>System characteristics</th>
<th>Low density of demand X=150</th>
<th>Medium density of demand X=890</th>
<th>High density of demand X=9650</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal bus size (S^*)</td>
<td>12</td>
<td>20</td>
<td>33</td>
</tr>
<tr>
<td>Generalized cost, GC</td>
<td>2.36</td>
<td>1.86</td>
<td>1.56</td>
</tr>
<tr>
<td>(AC_{\text{walk}})</td>
<td>0.57</td>
<td>0.30</td>
<td>0.12</td>
</tr>
<tr>
<td>(AC_{\text{wait}})</td>
<td>0.57</td>
<td>0.30</td>
<td>0.12</td>
</tr>
<tr>
<td>(AC_{\text{ride}})</td>
<td>1.03</td>
<td>1.06</td>
<td>1.10</td>
</tr>
<tr>
<td>(AC_{\text{prod}})</td>
<td>0.76</td>
<td>0.50</td>
<td>0.34</td>
</tr>
<tr>
<td>Optimal fare</td>
<td>0.20</td>
<td>0.20</td>
<td>0.23</td>
</tr>
</tbody>
</table>

It is striking that within the wide span of the density of trip demand considered, the optimal bus size does not exceed 33 seats, and that the optimal fares are not higher than 0.23 Euro. Both features are mainly explained by the assumed absence of systematic variations in demand in the central city during the day. The inevitable random fluctuations of demand is assumed to be accommodated by sufficient space for standing passengers, which means that, on average, bus occupancy can be equal to the number of seats. These very favourable circumstances are meant to represent an extreme that certainly is worth aiming at, but is very difficult to attain due e.g. to traffic congestion and the fact that in reality complete absence of peaks and troughs in the demand in time and space, even within office hours is rarely in existence. In the following section the overriding importance of the demand peakiness in a different segment of the urban travel market is elucidated. First, a word of caution should be given.

13.3.12 A word of caution: using the formulas out of optimum

In public transport pricing policy discussions where the Mohring effect is prominent, the importance of bus size is often overlooked, which can lead astray. A main point of our approach is that system optimization is a necessary condition for laying down the optimal fare in the whole range of demand. It can be instructive to show what the result can be if this condition is not wholly fulfilled.

In section 13.3.3 above two formulas for the optimal fare are derived, one where the number of buses is the factor that is changed, and another where the bus size is incrementally \(^9\) To get whole number values of the optimal bus size in the table, the three \(X\)-values could not be chosen to equal 100, 1000 and 10 000 as was first intended.
changed. In Figure 13.10 the result of applying the formulas for $P_N$ and $P_S$ as given in (13.17a) and (13.18a) is illustrated in a wide range of the density of demand ($Y$). However, values of $Y$ less than 3000 trips per km$^2$ and hour is the normal range in large cities. In a case where the bus size is assumed to be the same in the whole range, the two formulas give quite different results except at one level of $Y$, for which the assumed common bus size is optimal.

The middle curve in Figure 13.10 represents $P_{opt}$, that is the optimal fare along the expansion path, where both the number of buses and the bus size are continuously varied$^{10}$. Just as the width of the shadowed band in Figure 13.9, $P_{opt}$ is nearly constant except in the very initial range.

![Figure 13.10. $P_N$ and $P_S$ out of optimum](image)

As seen, neither $P_N$ nor $P_S$ is constant with respect to the density of demand. Given the bus size ($S=25$), $P_N$ is steadily rising from being negative for low volumes of travel to an upper limit for very high volumes where the negative user cost component goes toward zero. On the other hand, $P_S$ is slightly falling with the volume of travel, given the bus size. Only one particular bus size is, of course, optimal in a certain situation, so $P_{opt}$ is intersecting the rising $P_N$ function and the falling $P_S$ function at their point of intersection.

A practical conclusion of this exercise is that since $P_S$ is reasonably close to $P_{opt}$ in the whole range of $Y$ (except in the initial low density interval), $P_S$ is a safer bet when it comes

$^{10}$ Setting $P_N$ according to (13.17) equal to $P_S$ according to (13.18), and assuming full capacity utilization give us two equations from which $N$ and $S$ can be expressed as functions of $Y$.  

28
to applications of the price theory. When it is possible to approximately predetermined the price by the Ps approach, the transport system optimization is greatly facilitated.
13.4 Urban public transport 2: Commuter traffic

Now the focus is on markets where the peakiness of demand in time and space is the salient feature. In addition to random demand fluctuations there are large systematic differences in the demand between the peak hours (typically two hours in the morning and two hours in the late afternoon) and the rest of the service hours, the off peak period. There are also spatial peaks and troughs, the prime example being the large difference in patronage between the main haul and the back haul in the peak period. The primary mode of transport for meeting the demand of suburban commuters in larger cities is rail-borne trains (over- and/or underground), so the pricing principles outlined in this section concern both buses and trains. In some basic respects the principles are the same for both modes, and there is no need for an explicit distinction in the analysis. Where there are relevant, important differences – for example, train capacity can be increased by adding carriages – this should, of course, be recognized.

An analytically important aspect of both modes in commuter traffic is that each line can be considered in isolation. 11 The individual commuter train line or bus line constitute the "system" in the cost model irrespective of ownership. In case a single body for coordinating the public transport exists, the total overhead costs constitute a “common cost” for all lines and has to be allocated between the individual lines in order to get a complete system cost model. This could be done in the same way as argued in the previous section concerning a coordinated central city bus line network, by just adding a fixed percentage of the traffic operation costs representing the overheads to get the total producer costs.

13.4.1 Preliminaries of peak-load pricing

In view of the marked peakiness of the demand, a modified methodology has to be applied. As distinct from the preceding model of a central-city network of bus lines, a multi-product case is at hand, and the first task is to define the relevant products. Total travel demand is separated into the following six subcategories:

\[ Q_1 = \text{Peak passenger flow per hour on the main haul in the critical section} \]

\[ B_1 = \text{Total peak hour trips} \]

\[ 11 \text{In very compact built-up areas without "green fingers" between the suburbs “the circle town” model is applicable (Jansson 1984, 1997), where the number of lines are variable, and the distribution cost has two main determinants – the density of bus lines and the frequency of service on each particular line – just as the previous model of a central city grid network.} \]
$B_{11} = \text{Peak hour trips on the main haul}$

$B_{12} = \text{Peak hour trips on the backhaul}$

$Q_2 = \text{Off-peak passenger flow per hour}$

$B_2 = \text{Total off-peak trips per hour in both directions}$

Both the (hourly) intensity of these demand categories and their duration matter for the analysis, so the following additional symbols are introduced:

$H_1 = \text{Peak period duration}$

$H_2 = \text{Off-peak period duration}$

The capacity constraint is assumed to be binding in optimum at least for the first demand category. To assume full capacity utilisation all along the line, as was done in the previous model, would be unrealistic on a line where commuters are dominant. It is necessary to distinguish passenger kilometres, that is, the product of passenger trips ($B_1$) and average trip length ($L$), and passenger flow per unit of time ($Q_1$) in the peak period. The ratio $Q_1/B_1$ is designated $\varphi$.

The “peak vehicle requirement” that determines the number of vehicles is normally limited to a relatively small portion of the line, the so called critical section. The maximum passenger flow in the critical section ($Q_1$) is decisive, and the trip length ($L$) is of secondary importance. Were the passenger flow constant all along the line, the product of the passenger flow and the circuit distance would equal total passenger-kilometres per unit of time. There is no way of obtaining this ideal state in commuter traffic in the peak period because of the large systematic differences in trip generation along the line, in the first place between the main haul and the backhaul. It can be safely taken for granted that even a zero fare on the backhaul would not fill up the vehicles required in the critical section of the main haul. A very high fare on the main haul would make no difference because of nearly zero cross-elasticity (save the very long-run where the location of workplaces and places to live might substantially change).

The equality of the passenger flow in the critical section, and the consequent capacity requirement which replaces the production function in the previous model is written:
\[ Q_i = \frac{NSR}{D} \quad (13.27) \]

However, the number of trips \((B_i)\) plays a role also in the present model – a double role, in fact: on the cost side it matters that the boarding/alighting \((b/a)\) puts some claim on the capacity by slightly increasing the dwell time at stops and stations, and on the user benefit side it matters that the larger the number of trips \((B_i)\), the higher the total benefit of increases in the frequency of service and changes in other service qualities.

Since \(Q_i\) and \(B_i\) are different and not necessarily connected entities in the present case as opposed to the previous model, the full capacity expression 13.21) for the overall speed \((R)\), is now inapplicable. As before

\[ R = V \left( I - \frac{tB_i}{N} \right), \quad (13.28) \]

but since full capacity all along the circuit cannot be assumed, we have to leave it at that; \(B/N\) cannot be replaced by \(SR/L\), that facilitated the application of the previous model, but should on its own be allowed to affect overall speed. Combining (13.27) and (13.28) gives the capacity constraint that also takes account of the role played by \(B_i\) on the cost-side:

\[ Q_i = \frac{SV(N - tB_i)}{D}, \quad (13.29) \]

In this case it is both analytically and practically convenient to distinguish the two demands on capacity made by travellers: occupying a seat in the critical section, and the time required for boarding and alighting. Introducing occupancy charges as well as \(b/a\)-charges, the former can be determined as the marginal cost of increasing \(Q\), and the latter by separately calculating the marginal cost, of increases in \(B\) leaving the consequent demand for a seat on the bus out of consideration.\(^{12}\)

A complication that has to be tackled in this connection is that due to a time table restriction, there is an additional, acquired difference between the costs of main haul and backhaul trips. For main haul trips, irrespectively of whether or not a passenger occupies space in the critical section, a boarding/alighting charge is theoretically justified in addition to the occupancy charge.\(^{12}\)

\(^{12}\) The division between occupancy charges and \(b/a\)-charges is not wholly artificial. For example, a passenger who travels on the main haul in the peak hours outside the critical section would cause a price-relevant cost corresponding just to the occupancy charge by lengthening her trip to include the critical section.
charge on those who travel in the critical section. (A different matter is that the price-relevant cost of boarding/alighting can be too small to justify the fare collection costs in some cases.) For backhaul passengers neither occupancy charges nor boarding/alighting charges are warranted. The reason for this seeming oddity is that the time-table is adjusted to the much higher demand on the main haul, which means that some slack should be fit into the backhaul schedule. In order to maintain the same time-table throughout the whole peak period, the headway required on the main haul should be maintained also on the backhaul, which means that a small speed reduction and/or some deliberately prolonged dwell time should be inserted on the backhaul to compensate for the fact that the number of boarding passengers is systematically much less than on the main haul. It is thus predetermined that the occupancy charge is levied only on trips in the critical section and that the boarding/alighting charge should apply to all trips on the main haul, and not be claimed on the backhaul.

The main idea of peak-load pricing is to increase the capacity utilization by levelling out the natural peaks and troughs in demand. This can be supported by some supply adjustments thanks to the possibility to use two types of driver shifts: two consecutive straight shifts for all-day buses and a split shift where the same driver covers both the morning and late afternoon peaks, driving a peak-only bus that stands idle in the off-peak hours. Thereby the peak supply can exceed the off-peak supply. It also means that off-peak supply can be separately increased (decreased) by adding (withdrawing) an all-day bus and withdrawing (adding) a peak-only bus. The day cost of a peak-only bus with driver seems to be negligibly greater than half the cost of an all-day bus with drivers.

The possibility to apply both straight and split shifts makes it feasible to separate the peak period and the off-peak period optimizations. This is done in what follows. It is natural to start by the peak period optimization, and convenient for a start to assume that just peak-only buses are used in the peak period. After that the off-peak supply of all-day buses that maximizes social surplus of off-peak travel is determined. For every additional all-day bus, a peak-only bus is withdrawn which means that the peak supply is unaffected; the marginal conditions for optimality in the peak period obviously refer to adjustments on the margin; hence just peak-only buses are involved.

The peak period optimization concerns both the number and the size of the buses. It is assumed that the peak-only and the all-day buses are of the same size. This seems to be the most common case in reality. Adhering to this practice has the analytical consequence that the
optimality conditions for the off-peak period are of the Kuhn-Tucker variety. When the bus size is a result of the peak period optimization, the optimal off-peak solution can turn out to imply that full capacity utilization will nowhere and at no time be obtained in off-peak. Whether optimal or not, half-empty buses in off-peak can certainly be observed all over the world.

Table 13.3. Day costs (EUR) of different bus-sizes based on Swedish data

<table>
<thead>
<tr>
<th>Bus-type</th>
<th>Number of seats (max capacity)</th>
<th>Day cost of all-day bus (C_{ad})</th>
<th>Day cost of peak-only bus (C_{po})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double articulated</td>
<td>70 (150)</td>
<td>913</td>
<td>494</td>
</tr>
<tr>
<td>Articulated</td>
<td>55 (115)</td>
<td>830</td>
<td>445</td>
</tr>
<tr>
<td>Boogie</td>
<td>45 (110)</td>
<td>775</td>
<td>412</td>
</tr>
<tr>
<td>Normal-sized</td>
<td>35 (75)</td>
<td>719</td>
<td>379</td>
</tr>
<tr>
<td>Service</td>
<td>26 (60)</td>
<td>669</td>
<td>350</td>
</tr>
<tr>
<td>Mini</td>
<td>14 (30)</td>
<td>602</td>
<td>310</td>
</tr>
</tbody>
</table>

As seen in table 13.3, the day costs of peak-only buses are more than half the cost of all-day buses, and a decrease in bus size (from the largest size) by 50 per cent would reduce the costs by 20 per cent for all-day buses and by 25 per cent for peak-only buses. Peak-only buses are used less than a third of the time of operation of all-day buses.

An alternative approach could be to start by calculating a basic level of service of all-day buses, where both size and number of buses are variable. The basic service level would constitute the off-peak supply. It can be anticipated that the optimal size of the all-day buses making up the basic level of service is much smaller than that coming out of the peak period optimization. However, it seems clearly sub-optimal to take this smaller size for given when determining the number of buses for the peak capacity requirement. It is well-advised to allow for a different, bigger size of the peak-only buses. So the question is, what is to be preferred: a common bus size adjusted to peak conditions, where buses would run half-empty also with zero-fares in off-peak, or two widely different bus sizes, which would be more difficult to accommodate in the time-table, but by which a considerably higher rate of capacity utilization would be achieved?

13.4.2 Lagrangian solution for the peak period

A lagrangian expression for the total social surplus is formed of the same kind as (13.16) above from which the derivation of the price-relevant cost in the central city model started.
The main difference is that three different demand categories are distinguished, and the off-peak period is left for consideration in the next section. It can also be noted that the number of trips in the capacity constraint, \( B_1 \) is to be interpreted as \( 2B_{11} \) rather than \( B_{11} + B_{12} \), since the schedule is adjusted to the main haul boarding and alighting time.

\[
\pi = H_1 \left[ \int_0^{Q_1} g(Q_1) dQ_1 + \int_0^{B_{11}} g(B_{11}) dB_{11} + \int_0^{B_{12}} g(B_{12}) dB_{12} \right] - N \cdot C_{po} (S) \\
- H_1 B_1 \cdot AC_{user} (N, B_{11}) - \lambda \left[ Q_1 - \frac{SV(N - tB_1)}{D} \right]
\]

Setting the (partial) derivatives of \( \pi \) with respect to \( N, S, Q_1, B_{11} \) and \( B_{12} \), as well as the lagrangian multiplier, \( \lambda \), equal to zero gives the optimality conditions:

\[
\begin{align*}
\frac{\partial \pi}{\partial N} &= -C_{po} - H_1 B_1 \frac{\partial AC_{wall}}{\partial N} + \lambda \frac{SV}{D} = 0 \\
\frac{\partial \pi}{\partial S} &= -N \frac{\partial C_{po}}{\partial S} - H_1 B_1 \frac{\partial AC_{user}}{\partial S} + \lambda \frac{V(N - tB_1)}{D} = 0 \\
\frac{\partial \pi}{\partial Q_1} &= H_1 g(Q_1) - \lambda = 0 \\
\frac{\partial \pi}{\partial B_{11}} &= H_1 [g(B_{11}) - AC_{user} - B_1 \frac{\partial AC_{user}}{\partial B_{11}}] - \lambda \frac{SV}{D} = 0 \\
\frac{\partial \pi}{\partial B_{12}} &= H_1 [g(B_{12}) - AC_{user}] = 0 \\
\frac{\partial \pi}{\partial \lambda} &= Q_1 \frac{SV(N - tB_1)}{D} = 0
\end{align*}
\]

The generalized costs in the equations above of the three groups of peak period travellers that represent the marginal utilities consist of:

\[
\begin{align*}
g(Q_1) &= P_{occ} \\
g(B_{11}) &= P_{bla} + AC_{user} \\
g(B_{12}) &= 0 + AC_{user}
\end{align*}
\]
13.4.3 The occupancy charge

Bus riders’ claim on the bus transport capacity by occupying space in the critical section justifies occupancy charges that are obtained from (13.29), (13.30a) and (13.30c). The generalized cost is made up of just the monetary price in this case (13.30g). Dividing through in (13.30a) and (13.30c) by $H_1$ and substituting $(N-tB)/Q_1$ for $D/SV$ from (13.29) give:

$$P_{\text{occ}} = \frac{\lambda}{H_1} = \frac{C_{pa}(N-tB_1)}{H_1Q_1} - B_1 \frac{\partial AC_{\text{wait}}}{\partial N} \frac{(N-tB_1)}{Q_1}$$

(13.31)

This expression can be further developed remembering that $dAC_{\text{wait}}/dN$ is equal to $-AC_{\text{wait}}/N$, that $(N-tB_1)/N$ is equal to $R/V$ and that $Q_1/B_1$ is equal to $\phi$.

$$P_{\text{occ}} = \frac{\lambda}{H_1} = AC_{\text{prod}} \frac{R}{V} - AC_{\text{wait}} \frac{R}{\phi V}$$

(13.31a)

A conspicuous but less important difference of this optimal price expression compared with $P_N$ in the central city bus line network case (13.17b) is the factor $R/V$. This is due to the separation of $P_{\text{occ}}$ and $P_{\text{bla}}$ in the present case. The ratio of the overall speed $R$ to the cruising speed $V$ is in the order of 9/10, and as we shall see, $P_{\text{bla}}$ is just about 1/10 of $P_{\text{occ}}$. The most important differences between (31a) and (17b) is that the producer cost is spread over just the trips through the critical section during the peak periods, $Q_1$, and that $H_1$ is a minor part of the total service hours. It is also notable that only the waiting time cost is pricing-relevant among the distribution costs in the present case. This difference is neutralized by the ratio of passenger flow in the critical section to number of trips, $\phi$, that appears in the deductible user cost component of $P_{\text{occ}}$.

The trip length is inconsequential for the optimal occupancy charge in this case; passage of the critical section is the only thing that matters. However, the route distance ($D$) is important for the optimal level of $P_{\text{occ}}$. $AC_{\text{prod}}$ is proportional to $D$ because the number of vehicles of a given size required for capacity to match the passenger flow in the critical section is proportional to $D$.

As seen in Figure 13.11, the occupancy charge (the shaded difference between $MC$ and $AC_{\text{user}}$) is in the region of €2 for a trip across the critical section on a commuter bus line where $D = 16$ km (as compared to a charge of just € 0.20 for a trip of 4 km in the central city bus transport system.) $P_{\text{occ}}$ would cover 60-90 per cent of the producer costs of the peak traffic.
13.4.4 Boarding/alighting charges

The act of boarding/alighting slightly reduces overall speed ($R$) which increases the time costs of the riders on the bus.

The indivisibility problem is admittedly more tractable by considering a slight increase in bus size (rather than another whole bus) to compensate for the capacity reduction due to slightly longer dwell times at bus stops. However, it does not seem worthwhile to penetrate this matter deeper.

From (30d) and (30h) the b/a charge can be written as:

$$P_{bla} = B_1 \frac{\partial AC_{user}}{\partial B_{11}} + \frac{\lambda}{H_1} \frac{tSV}{D} = AC_{user} \cdot E_{ACB} + P_{occ} \frac{tSV}{D}$$  \hspace{1cm} (13.32)

The elasticity of $AC_{user}$ with respect to $B_{11}$ is of the order of magnitude of 0.1 and the size of the fraction $tSV/D$ only about a fifth of that, which means that even if the b/a-charges would cover up to ten per cent of the traffic operation costs, the fact that they are mainly caused by the costs of the fare collection itself makes them hardly worthwhile to cash in, no matter what charging system is applied. They should simply be merged with $P_{occ}$ for peak trips in the critical section, and not be claimed at all for the rest of peak services.
13.4.5 Off-peak fares

The off-peak period is characterized by a rate of travel generation per hour that is much lower than in the peak period. The peak period is about four hours per day, and night travel conditions prevail for, say ten hours. This means that the daily off-peak period also comprises ten hours. Total off-peak travel can thus be of the same size as total peak travel, and could consequently be regarded as equally important as peak travel. However, the willingness to pay per passenger is normally substantially lower in off-peak. In order that off-peak travel could flourish it is necessary that the off-peak fares are kept rather low. That this is in line with the price-relevant marginal cost of off-peak travel is demonstrated in this section.

As was mentioned in the preliminaries above, two alternative approaches to the question of off-peak supply are conceivable: one where the same bus size in off-peak as in peak is postulated, and another where the determination of a basic (minimum) all-day level of service comes first and the bus size is separately optimized without regard to the peak vehicle requirement.

At closer scrutiny, however, it appears that using considerably smaller buses in off-peak would have a serious drawback. The opportunity cost of off-peak buses would rise to the tune of the peak capacity withheld. That seems to be a decisive disadvantage. As mentioned, the off-peak demand per hour is much lower than peak demand, and the possibility to extend the use of peak-only buses that otherwise would stand idle in the off-peak is a great saver, no matter that these buses seem oversized when travel is ebbing.

13.4.6 Kuhn–Tucker analysis of the off-peak pricing problem

The critical assumption is that in the off-peak period the bus size is given by the peak period optimization. Moreover, it is assumed that the systematic variations in demand are negligible. Peaks in demand and backhaul troughs mainly belong to the peak period. The random fluctuations are still considerable. As before, full capacity is defined to be attained when, on average, all seats are occupied, and this leaves quite substantial reserve capacity for standing passengers. The question is if full capacity utilization should be aimed at also in off-peak?

Let us leave this question open for the moment. Irrespective of the answer, it is analytically proper to distinguish off-peak maximum passenger flow per hour, \( Q_2 \) and number of trips per hour \( B_2 \). On the other hand, there is no direction that could be characterised as the “main haul”, so further division of off-peak demand is unnecessary.
A similar lagrangian expression as the previous one for the peak period analysis (13.30) is the starting point. It can be anticipated that on account of the given bus size, full capacity utilization is not necessarily part of the optimal solution. The Kuhn-Tucker conditions (13.33d) and (13.33e) are added to the previous optimality conditions.

\[
\pi = H_2 \left[ \int_0^{Q_2} g(Q_2) dQ_2 + \int_0^{B_2} g(B_2) dB_2 \right] - (C_{ad} - C_{po}) N_{ad} - H_2 \cdot B_2 \cdot AC_{user} - \mu \left[ Q_2 - \frac{\tilde{S}V(N_{ad} - tB_2)}{D} \right]
\]

Taking the derivative of \( \pi \) with respect to \( N_{ad}, Q_2, B_2 \) and \( \mu \), setting them equal to zero gives the optimality conditions.

\[
\frac{\partial \pi}{\partial N_{ad}} = -(C_{ad} - C_{po}) + H_2 B_2 \frac{\partial AC_{user}}{\partial N_{ad}} + \mu \frac{\tilde{S}V}{D} = 0 \quad \text{(13.33a)}
\]

\[
\frac{\partial \pi}{\partial Q_2} = H_2 g(Q_2) - \mu = 0 \quad \text{(13.33b)}
\]

\[
\frac{\partial \pi}{\partial B_2} = H_2 \left[ g(B_2) - AC_{user} - B_2 \frac{\partial AC_{user}}{\partial B_2} \right] - \mu \frac{\tilde{S}V}{D} = 0 \quad \text{(13.33c)}
\]

\[
\frac{\partial \pi}{\partial \mu} = \left[ Q_2 - \frac{\tilde{S}V(N_{ad} - tB_2)}{D} \right] \geq 0 \quad \text{(13.33d)}
\]

\[
\mu \frac{\partial \pi}{\partial \mu} = 0 \quad \text{(13.33e)}
\]

The generalized costs representing the marginal utilities in the equations can be spelled out:

\[
g(Q_2) = P_{occ} \quad \text{(13.33f)}
\]

\[
g(B_2) = P_{bla} + AC_{user} \quad \text{(13.33g)}
\]

The Kuhn-Tucker conditions (13.33d) and (13.33e) leave two possibilities open: the capacity constraint is or is not binding. In the former case \( \mu \geq 0 \) and in the latter case \( \mu = 0 \). What is applicable in a particular case depends, of course, on the relevant parameter values as well as
on the passenger flow \(Q_2\) and the number of trips \(B_2\). Let us start by considering the former case.

**The capacity constraint is binding – the exceptional case**

Combining (33a) and (33b) the occupancy charge is obtained in much the same way as in the peak period case.

\[
H_2P_{occ} = \left(\frac{C_{ad} - C_{po}}{SV}\right)D - H_2B_2 \frac{AC_{wait}D}{N_{ad}SV}
\]

\[
P_{occ} = \left(\frac{C_{ad} - C_{po}}{H_2Q_2}\right)(N_{ad} - tB_2) - \frac{B_2}{Q_2} \frac{AC_{wait}(N_{ad} - tB)}{N_{ad}}
\]

\[
= AC_{prod} \frac{R}{V} - AC_{wait} \frac{R}{\phi V}
\]

\(P_{occ}\) has the same appearance as (13.31) for the peak-period, but \(H_2\) is 2.5 times longer than \(H_1\), and \(C_{ad}\) is less than \(2C_{po}\), which reduces \(AC_{prod}\) correspondingly. Moreover, the value of \(\phi\) is normally smaller in the off-peak than in the peak period, and the value of \(AC_{wait}\) is normally higher in off-peak due to a smaller number of buses in operation. Altogether this makes \(P_{occ}\) relatively low in off-peak conditions, and there is nothing in principle preventing the sum of the two terms of (13.34) to be zero or even negative.

**The capacity constraint is non-binding – the normal case**

Due to the postulated fixed bus size, deliberate excess capacity in off-peak is quite a likely outcome of the social surplus maximizing solution. The question is, under what circumstances will this occur?

This question can be explored by comparing, on the one hand the determinants of the capacity constraint, and on the other hand, the determinants of the user benefits of increased frequency of service. The simple idea is that as long as the additional daily benefits of another bus for the existing bus riders in the form of reduced waiting time at bus stops, exceeds the day costs of an additional bus, it should be added. Total benefits would, of course, be enhanced if also some new riders are taken on, and it would be counterproductive to charge for the bus trips as long as the capacity constraint is non-binding.
The capacity constraint in (13.33) above gives an answer to one part of the question. In a short version it boils down to:

\[ Q_2 = \frac{\hat{S} NR}{D} \]

from which the number of buses required to accommodate a passenger flow, \( Q_2 \) can be found:

\[ N = \frac{Q_2 D}{\hat{S} R} \quad (13.35) \]

Optimality condition (33c) above gives an answer to the other part of the question. Setting \( \mu = 0 \) we have:

\[ \left(C_{ad} - C_{po}\right) = H_2 B_2 \frac{\partial AC_{user}}{\partial N} \quad (13.36) \]

Remembering that \( AC_{user} \) consists of \( AC_{wait} \) and \( AC_{ride} \), and that the latter is independent of \( N \), (13.36) can be developed to give the optimal number of buses, \( N^* \), when the capacity constraint is non-binding.

\[ N^* = \sqrt{\frac{wDH_2 Q_2}{R \left(C_{ad} - C_{po}\right)}} \quad (13.37) \]

If \( N^* \) exceeds \( N \) given by (13.35), it means that the number of buses maximizing the social surplus (\( N^* \)) should be chosen in spite of the fact that some excess capacity would arise. In this situation a zero occupancy charge is indicated. In the opposite case where the capacity constraint requires a larger number of buses than what follows just from considering the net benefit of existing traffic, a non-zero occupancy charge that makes demand match the supply is required.

**13.4.7 Why two cases?**

Which one of the two cases will apply depends on the values of the parameters involved, in particular on the bus size determined in the peak period optimization. Figure 13.12 illustrates the crucial importance of bus size in this connection.

In the diagram where number of buses is plotted against volume of travel the straight rays from the origin represent the capacity constraint assuming different bus sizes. A large number
of buses is obviously required to accommodate a particular trip volume when the bus size is on the small side, and the other way round when the given bus size is large.

The relationship between \( N^* \) and \( Q_2 \) that maximizes net benefits for the existing riders is given by (13.37) above. As seen, \( N^* \) should increase to the tune of the square root of \( Q_2 \), all other things remaining the same. The bus day costs, \( C_{ad} \) and \( C_{po} \) are functions of \( S \), but varying \( S \) in this relationship gives rise to rather moderate shifts, so in spite of the fact that both relationships between \( N \) and \( B_2 \) revolve counter-clockwise with increases in \( S \), it is seen that for small bus sizes, the capacity constraint will be binding already for moderate values of \( Q_2 \), whereas frequency of service considerations will result in excess capacity in a successively wider range of \( Q_2 \) for larger bus sizes.

![Figure 13.12](image.png)

**Figure 13.12.** A zero occupancy charge is indicated when \( N^* \) exceeds the number of buses (\( N \)) required on pure capacity grounds.
13.5. Urban public transport 3: transport system optimization and pricing in conclusion

Urban transport optimization and pricing should be carried out simultaneously. Since the total urban travel market falls into two crucially different parts, the main message has been delivered in somewhat different forms by two separate models: one model which represents mainly quite short-distance all-day travel within the central city of large conurbations, where the focus is on the dominant distribution costs, that is the walking and waiting time costs of the users of the bus transport system. The other model represents a bus or train line mainly for commuting between a suburb and the central city where the peakiness of demand in time and space is the salient feature.

The result of the central city model is that the optimal fare $P_{opt}$ is relatively constant along the expansion path, and because $AC_{prod}$ is falling along the expansion path the ratio of $P_{opt}$ to $AC_{prod}$ is increasing from about 25 per cent where the density of demand ($Y$) is low to 70 per cent where $Y$ is high. The reason for the negative financial result when social surplus maximization is aimed at is basically the pronounced economies of vehicle size. Given the bus size, the financial result of optimal pricing would also be negative due to the Mohring effect, that is, the waiting time savings for existing passengers of the increase in the frequency of service caused by an additional bus should be deducted from the cost of the bus. Calculated in this way, that is, taking the bus size as given, the fare could vary from zero to a value close to $AC_{prod}$. This is inconclusive: only if the given bus size is optimal, the fare calculated in accordance with this rule of thumb would make real sense. The critical importance of the economies of bus size is clear bearing in mind that in the counterfactual case where these economies are non-existing, very small buses should be used everywhere and make the frequency of service high enough to satisfy all wishes, with impunity as regards to total bus costs. An alternative approach is to consider successively larger buses for increasing output. In that case a user cost component also plays a role. As opposed to the Mohring effect, the user cost component with the alternative approach is an addition to $MC_{prod}$, relatively small but sufficient to make $P_{N} = P_{S}$ in optimum.

The second model concerning the commuting of suburbanites is used to analyse the application of peak-load pricing. The peakiness of the demand gives rise to far-reaching fare differentiation in this case. Main haul passengers traversing the critical section should bear practically all traffic operation costs, and there are good reasons to allocate most overheads to
this category of travellers, too. Backhaul passengers in the peak period should travel for free, and that could also go for all off-peak passengers in view of the fact that, save the very busiest lines, the modest price-relevant cost caused by off-peak travel just consists of the delay caused by boarding and alighting. The relatively large buses required for the peak demand, which should also be used for all-day service offered by two straight shifts, would normally not be filled up in off-peak, not even with zero occupancy charges. The price-relevant cost that could justify a boarding/alighting charge up to a quarter of an euro in off-peak would be halved by offering the service free of charge. The fare-collection costs in off-peak do not seem worthwhile in this case.

The peak fares are some ten times higher than the level of fares in the central city model mainly because total bus costs are spread over just the passenger flow through the critical section on the main haul in the peak period. Trip length does not matter, but the route distance plays a role for the level of fares. Given the passenger flow in the critical section, the longer the route, the more and/or larger buses are required to meet a given demand, which raises the ratio $N/Q_{1}$ that is proportional to $AC_{prod}$.

13.5.1 Right for the wrong reason

Depending on different circumstances in each particular town or city, the overall financial result of the social surplus maximizing level and structure of fares would be that fare-box revenue covers 40-60 per cent of the total producer costs. This, by chance, happens to be similar to the actual situation in Sweden and many other countries, too. From a welfare economic point of view, the problem is consequently not the degree of subsidization, but the fare structure. That the structure of urban public transport fares has been permanently out of line with the optimal structure is a clear sign of unawareness on the part of local governments and their advisers of the social economic efficiency conditions. They may be right as far as the degree of subsidization is concerned, but for the wrong reason.

13.5.2 The apoly problem

The very best market for public transport is intra-central city travel as long as the buses are not stuck in car-dominated congestion. In the central city, the potential density of trip demand is the highest conceivable and systematic peaks and troughs are not very pronounced. High demand densities are found also on radial routes between suburbs and the central city in the peak period, but there are great variations in this respect among the radial routes. For small or
sparsely populated suburbs it can be a matter of the very existence of the local bus line. If the demand curve – GC as a function of the volume of travel – does not intersect with the AC curve at any point, apoly would appear if free market forces are ruling, that is, no profit-seeking bus transport operator would care to supply services. In a case where social surplus maximization is the goal, apoly can also arise, although in a still thinner market. Figure 13a and b illustrate under which conditions discontinuation of a particular line could be an option (“corner solution”). The top diagram shows the cost and demand constellation where even peak services could be judged to be non-profitable from a social economic point of view. As long as the lightly shaded welfare triangle is larger than the darker one, the service should be maintained. However, it is easily imagined in this case that an appreciable leftward shift in demand could change this conclusion, although it should be pointed out that estimating the remaining consumer surplus is very difficult. A linear demand curve may underestimate the willingness to pay of people without a car, living far away, beyond bicycling distance.

How would the situation be in off-peak? The bottom diagram giving the corresponding demand and cost picture for the off-peak period suggests that even if a peak supply would be deemed worthwhile, just a split shift could be used for peak services implying that no off-peak services at all are provided. In spite of zero fares, the total benefits of off-peak travel could be insufficient to outweigh the cost of any all-day bus.
Figure 13.13a and b. Threatening a poly in peak and/or off-peak
13.5.3 Future prospects

As long as car ownership is growing and urban sprawl continues, it seems that public transport in towns, as distinct from cities, is on the decline (Holmgren et al. 2008). Is this inevitable? In combination with Baumol’s cost disease, the negative forces are strong.13

Table 13.4. Change in costs per vehicle kilometre in Sweden and Britain 1994 – 2009 relative to RPI

<table>
<thead>
<tr>
<th></th>
<th>Location</th>
<th>Change in costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Great Britain</td>
<td>London</td>
<td>39 %</td>
</tr>
<tr>
<td></td>
<td>GB outside London</td>
<td>50 %</td>
</tr>
<tr>
<td>Sweden</td>
<td>Stockholm</td>
<td>49 %</td>
</tr>
<tr>
<td></td>
<td>Rest of Sweden</td>
<td>52 %</td>
</tr>
</tbody>
</table>

Source: Preston and Almutairi (2013), Holmgren (2013b)

Social surplus maximization would certainly slow down the decline, but could it change the vicious spiral due to Baumol’s cost disease to a virtuous one? This stands and falls with to what extent a wider view of the problem of unbalanced productivity growth is evolving. As long as the productivity growth continues in the “progressive sector” – mainly the material goods sector – we can afford higher quality of public and merit goods by allocating labour from the progressive to the non-progressive sector. These goods should be tax-financed. The benefits of public and merit goods are more difficult to value than marketed private goods, but are increasingly coveted in societies of growing affluence in respect of material goods (Baumol, 2012).

---

13 See also Holmgren (2013a) for a discussion on the effects of increased income and car ownership on public transport.
13.6. Interurban public transport

Perhaps the largest difference between the urban and non-urban preconditions for public transport is the unchallengeable dominance of the private car in most relations outside urban areas. Just on a few, but, of course, very busy and important routes between larger cities, public transport by road, rail and air can challenge the private car for those who have one at their disposal.

As is shown in Figure 13.2 on page 3, for trip lengths in the range from 10 km to 300 km, 85 to 90 per cent of the travel is made by car in Sweden. The remaining 10 to 15 per cent made by public transport are highly concentrated to a relatively limited number of main routes.

13.6.1 Salient economic features of non-urban public transport

To understand the existing pattern of the modal split at the aggregate level, one can start with the elementary fact that transport infrastructure is a necessary condition for public transport, but not a sufficient condition. In Sweden total railways (excluding private sidings) and total road kilometres are:

12 000 km of railways
100 000 km of primary roads
400 000 km of secondary roads

Where there is a rail track there is in most cases a train service in operation, although often a quite infrequent one, and interurban bus services are often to be found on the primary roads, but this is it, and as carless families are becoming more and more rare, the lop-sided modal split will be strengthened over time.

13.6.2 Lop-sided modal split

It is the existence of the fine-meshed secondary road network that makes the large difference for the division of individual and public transport. In Sweden, as elsewhere, practically every house is connected to the road network, and therefore door-to-door transport by car between every pair of houses in the country is possible. The existing road network defines millions of different relations A to B that each could be the basis for public transport services. However, this potential is realized to a very limited degree. Only a small fraction of the possible relations has bus transport, let alone a rail transport connection.
13.6.3 Apoly abounds

The diagram below illustrates the basic reason for this situation, which really is obvious, but surprisingly often misunderstood in transport policy discussions. Given the route distance, the system average cost for public transport is steeply falling in the initial volume range. The inverse demand functions of Figure 13.14 are the products of a representative traveller’s willingness to pay (in GC terms) for another trip as a function of the number of trips per day she makes and the potential number of travellers. This product is in turn more or less proportional to the product of the number of people living at each end of the route in question. If two million-cities are connected by the route, the demand for transport could easily be some million times greater than in a relation of the same distance between two villages, each with a population of just 100 persons. The steepest demand function in Figure 13.14 can represent the latter case and the flattest demand function the former case. On the busy route there would be viable public transport both by road and rail, as well as air if the route distance is long enough. In the other cases there will be none. Apoly is ruling, and apoly is easily the most common market form as regards non-urban public transport, and unfortunately the situation would not change very much by optimal public transport pricing. If the GC-curve in Figure 13.14 would include an optimal price, GC would be lower than in a full-cost pricing case, but the abundance of apolies would by and large remain.

![Diagram](image)

*Figure 13.14. Public transport apoly prevails in most relations.*

The diagram is a warranted reminder that the most urgent problem of non-urban public transport is perhaps not inoptimal fares on the main lines, which is taken up next, but the continuing disappearance of public transport services in the secondary road network.
13.6.4 The level of the price-relevant cost of long-distance public transport

Given that a public transport service is to be maintained in a particular relation, that is, ruling out corner solutions, pricing aimed at social surplus maximization is now briefly outlined. The basic price theory formulated in section 13.2 is intended to apply to both urban and non-urban public transport. When it comes to interurban long-distance travel, an obvious additional aspect is that air transport is increasingly relevant for longer trips. Nonetheless formula (13.6) for the optimal price of public transport remains basically valid also when the trip length is multiplied, and a new mode of transport enters into the picture.

A crucial methodological point is that the second, alternative formulation of the price-relevant cost given in section 13.3.3 above is more practical for non-urban public transport than the first alternative that suits urban public transport better. In long-distance public transport where the headway is many times longer than in urban bus transport – from half an hour to many hours – the frequency delay in the form of disguised (hidden) waiting time is much more difficult to estimate than actual waiting time at stops and stations.

However, when an increase in bus size is considered, the price-relevant cost of scheduled bus transport is much less dependent on the user cost valuation. It is given by (13.18a). This expression can be slightly developed by making the marginal producer cost, $MC_{prod}^S$, explicit in the formula:

$$P_S = MC_{prod}^S + \frac{E_{RS}}{1 + E_{RS}} \left( \frac{1}{2} AC_{distr} + AC_{ride} \right)$$

(13.38)

On long-distance routes, where the elasticity of travel time with respect to vehicle size, $E_{RS}$, is negligibly different from zero\textsuperscript{14}, the preferable formulation of the optimal level of fare boils down to a value close to where only the producer cost component of $P_S$ really matters:

$$Long-distance \quad P_S \approx MC_{prod}^S$$

(13.38a)

The common belief concerning optimal bus fares, conceived from the point of view where an additional bus is the most natural way of increasing capacity, is that long distance coach fares

\textsuperscript{14} Except perhaps for international air transport where the aircraft boarding including check-in time and security controls is generally lengthy and increasing with the aircraft size (number of passengers per aircraft).
should be nearly equal to $AC_{prod}$ because the Mohring effect would be relatively small, that is, the full bus cost (including the cost of the driver and overheads) per passenger should be charged. This is mistaken. To see why, let us first consider the question of how formula (13.38) squares with the efficiency condition that $P_N = P_S$?

Consider a very long and thin coach route where just one departure per day in each direction is the normal case. Additional capacity should be offered by using a successively larger vehicle, and $P_S$ as given in (13.38a) above is a relevant approximation of the price-relevant cost. On a busier route where many vehicles are required, it would be inefficient to propose the same departure time for all the coaches. Departure times should be spread over the day. The Mohring effect cannot be calculated in the same simple way as in an urban context where the density of demand is many times higher, but this does, of course, not mean that the frequency of service is unimportant on long-distance routes: the equality $P_N = P_S$ should be valid in optimum, no matter that the cost of infrequency is difficult to directly calculate or not.

Vehicle size adjustment to increasing volumes of traffic is physically very natural in rail transport in the form of adding carriages to the train, and therefore also mentally obvious when it comes to marginal cost pricing, in particular so far as flexible-formation trains are concerned. Expressing the price-relevant marginal cost of rail passenger transport as the average cost per additional passenger of putting on another carriage to the train is a self-explanatory proxy. It is shown in Jansson et.al. (1992) that this idea yields a virtually constant level of rail fares in a wide range of the traffic volume. The nearly constant level of the optimal fares is well below $AC_{prod}$ in the initial range and close to $AC_{prod}$ in the high volume range. (See also Jansson 1997, 1998, and Ericsson and Jansson, 2001)

13.6.5 The optimal tariff
In the previous models of urban public transport, in particular the first one concerning the intra-central city bus travel, the multi-product aspect was not very prominent. In dealing with interurban public transport this simplification is no longer warranted. The structure of fares is now all-important, so the discussion starts by pointing out the causes of the systematic variations in interurban travel demand.
13.6.6 The structure of long-distance travel demand in time and space

The peakiness of demand in interurban public transport is diurnal as well as weekly. Long-distance leisure travel is at its highest on Friday and Sunday afternoon, and long-distance business trips also tend to go up in the beginning and at the end of the working week.

An example of the structure of interurban rail transport demand is given from the northbound railway line from Stockholm to Sundsvall, the chief town in central Sweden. Table 13.5 shows the average number of travellers leaving Stockholm on this line on consecutive departures different days of the week.

**Table 13.5. Average number of passengers on consecutive trains from Stockholm to Sundsvall different days of the week.**

<table>
<thead>
<tr>
<th>Departure</th>
<th>Monday-Thursday</th>
<th>Friday</th>
<th>Saturday</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td>early morning</td>
<td>62</td>
<td>75</td>
<td>130</td>
<td>_</td>
</tr>
<tr>
<td>Morning</td>
<td>189</td>
<td>160</td>
<td>170</td>
<td>168</td>
</tr>
<tr>
<td>middle of the day</td>
<td>96</td>
<td>210</td>
<td>165</td>
<td>_</td>
</tr>
<tr>
<td>early afternoon</td>
<td>128</td>
<td>250</td>
<td>198</td>
<td>172</td>
</tr>
<tr>
<td>afternoon</td>
<td>210</td>
<td>540</td>
<td>158</td>
<td>505</td>
</tr>
<tr>
<td>late afternoon</td>
<td>200</td>
<td>370</td>
<td>_</td>
<td>380</td>
</tr>
<tr>
<td>Evening</td>
<td>95</td>
<td>325</td>
<td>50</td>
<td>120</td>
</tr>
<tr>
<td>late evening</td>
<td>20</td>
<td>40</td>
<td>_</td>
<td>_</td>
</tr>
<tr>
<td>Total day</td>
<td>1000</td>
<td>1880</td>
<td>877</td>
<td>1345</td>
</tr>
</tbody>
</table>

The demand variations are striking, both within the day and between different days of the week.

On weekdays eight trains leave Stockholm on the line to Sundsvall. On Saturdays and Sundays travel demand is more concentrated and there are just six trains a day. However, two of the four really full trains each week depart on Sunday afternoon.

There are also considerable variations in demand along the route. The number of passengers along three stretches of the railway line between Stockholm and Sundsvall is given in Table 13.6.
As the starting-point of the line concerned is a relatively large city (Stockholm) and the end-point is a much less populous city (Sundsvall) like the stops in-between, the passenger flow is steadily falling along the line.

Table 13.6. Passenger flow per day along three sections of the railway line between Stockholm and Sundsvall.

<table>
<thead>
<tr>
<th>Line section</th>
<th>Fri, Sun</th>
<th>Mon-Thur, Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stockholm – Gävle</td>
<td>3307</td>
<td>1751</td>
</tr>
<tr>
<td>Gävle – Söderhamn</td>
<td>2490</td>
<td>1240</td>
</tr>
<tr>
<td>Söderhamn – Sundsvall</td>
<td>1494</td>
<td>734</td>
</tr>
</tbody>
</table>

On the return from Sundsvall it is the other way round. The passenger flow is steadily increasing towards the final destination (Stockholm).

13.6.5 The principle of joint products pricing
Looking closer at the structure of demand for interurban public transport, it is apparent that the case of urban commuter services discussed in section 13.3 above is a relatively simple special case of the general case where significant systematic variations in demand abound both in time and space, which make interurban public transport a markedly multi-product activity. For urban commuter travel the optimal fare differentiation turned out to be simple: it was concluded that just one flat occupancy charge is required on each particular line. It should be levied during peak hours on the main haul in the critical section. Zero occupancy charges are indicated both outside the critical section in the peak period and generally in off-peak because there is little chance of reaching up to the capacity limit even when travel is free.

This would not be the typical situation in interurban public transport. It is true that the capacity requirement is determined by the demand (passenger flow) across one or more critical sections on each particular route, also in interurban public transport but zero charges are not generally indicated, neither in all other line sections on Friday and Sunday afternoon,
nor all round on normal days, because that could easily reverse the peak and off peak levels of demand.

The main reason for this is simply that the average level of fares in interurban public transport is naturally many times higher than in urban public transport, and the present fare constitutes a much larger share of GC than in urban public transport, which makes the absolute value of the price-elasticity much greater. In addition, the difference in daily demand between the two peak days and the other five days of the week is, relatively speaking, somewhat less than the difference in demand between peak and off-peak hours in urban commuter traffic. This means that the price mechanism can be used to much greater avail in the interurban case and unfortunately also that the calculation of the optimal structure of fares is more demanding. A manageable approach to the pricing problem has the following foundations:

(i) As was repeatedly stressed in the preceding discussion, transport system optimization and pricing should be carried out simultaneously. However, a short-cut is offered in case the optimal price level happens to be constant along the expansion path. The finding that the train lengthening approach, like P₅ in the case of bus transport (and air transport, we dare say), yields a virtually constant price-relevant cost with respect to volume of travel makes it possible in practice to consider the optimal structure of fares without simultaneously determining the optimal number of vehicles and other factor inputs on each line in question.

(ii) On the demand-side the concept of “joint products” is very useful. The theory is traditionally illustrated by sheep-breeding: wool and mutton are joint products in that trade, and given the demand functions of each product, the question is: how many sheep should be raised? The trick of the supply and demand analysis is to add the demand for wool and the demand for mutton vertically. Social surplus maximization is obtained by fixing output (number of sheep raised) where the marginal cost curve intersect with the total demand curve, and set the prices accordingly, that is equal to the marginal willingness to pay for wool and mutton, respectively.

Public transport services are a prime example of joint products. Each vehicle plying a given route and making stops at given places along the route offers joint transport services between every pair of stops in each directions, and the point is that travel along a particular section of the route is non-rival in relation to travel along any other sections. Since a full round is to be completed under all circumstances, each section of the route has to be traversed irrespective
of the demand, and total output can be increased without additional cost until the train is full all round.

13.6.6 Exemplifications of the basic principle

The optimal tariff construction is very involved, so a step-by-step approach is appropriate. Let us start by the simple case of a shuttle-service.

This discussion is mainly about interurban public transport by buses, aircraft and fixed-formation trains, that is, the vehicle size is given once the vehicle in question is in operation. Vehicle size changes can only be made by exchanging the vehicle in operation for a larger or smaller one.

**Fare differentiation by direction: long-distance shuttle–service illustration (trip length = route distance)**

An intercontinental non-stop airline service between two capitals makes a simple illustration of the basic principle. It can be assumed in this example that each aircraft manages to make just one return flight per day and the question is: what seat capacity \( S \) of the aircraft should be chosen, and how is the producer marginal cost of an additional seat to be shared between the two legs of the shuttle service?

The answer to this question is easily demonstrated diagrammatically. The two flights in each direction are joint products and the demands of each should be vertically added as shown in Figure 13.15. The point of intersection between the price-relevant cost of a full circuit, \( P_s \) and total demand gives the optimal aircraft size and the relative fares, \( a/b \) for the main haul and back haul, respectively.
If the expected demand in each direction were the same all days of the week, this would be sufficient fare differentiation, but domestic airlines like interurban rail and coach services are facing pronounced weekly demand variations, as was illustrated for rail transport in Table 13.5 and 13.6 above. In this case the diagrammatic model of Figure 13.15 has to be extended by considering a whole week.

Fare differentiation by day of the week

The weekly cost of a marginal seat is in Figure 16 confronted with the total demand made up of the daily demands in both directions of two peak days and five normal days which are vertically added.
As seen in the diagram, it is suggested that normal day travel may contribute substantially to the sum total of the marginal willingness to pay for an additional seat offered all days by fixed-formation trains, domestic aircraft, or interurban coach services.\textsuperscript{15}

\textit{Fare differentiation by departure}

In case the vehicles on the route concerned manage to make more than one complete round per day, there is a third level of possible fare differentiation.

\footnotesize{\textsuperscript{15} This can be contrasted with urban commuter traffic. Let the demand curves in Figure 16 represent the hourly demands of a weekday, two peak hour demands at the bottom and on top of these, the five off-peak hour demands are given. It easily imagined that } $MC_{prod}$ \textit{in this case would be at a relatively low level that makes it intersect with the total daily demand well below the point where all off-peak demands cease to add to the sum total of the willingness to pay for additional capacity in the form of another seat. In that case zero occupancy charges are indicated during off-peak hours.}
from Stockholm to Sundsvall each weekday. The weekly cost per train should thus be shared between two northbound and two southbound legs of the route in a similar way to that illustrated in Figures 13.15 and 13.16.

*From occupancy charges to ticket prices*

Compared with the shuttle service illustration that started the exemplification of the principle of joint products pricing, the structure of fares of long-distance bus and train transport services should be much more differentiated in view of the fact that both modes make a number of stops en route for alighting and boarding of new passengers. A different fare should apply to travel between each pair of stops along the line that makes for \( n(n-1) \) different ticket prices, where \( n \) is the number of stops.

The analytical problem is that the ticket prices charged on travellers apply to trips comprising at least one section (= stretch of the line between two adjacent stops) and usually several sections. Exactly how this complication should be straightened out would take too long to properly deal with.

Very briefly, the provisional main idea is to begin with the value of the optimal occupancy charge of a trip going the whole way from the start to the end of the line (half the circuit distance). This value can be divided between the constituent line sections, not necessarily equally or in proportion to the length of each section, but in such a way that the total volume of travel measured by total passenger-kilometres is maximized on the side-condition that the sum of the section occupancy charges is equal to the price-relevant cost of a trip going the whole way. The ticket price of a particular journey is obtained by summing up all line section prices involved. The basic rule of thumb is that the line section prices should be determined such that the expected occupancy rate equals the target value (= the practical capacity) everywhere, all the time. The only exception to this rule is when a zero price still leaves the level of demand below this target.

**13.6.7 User-friendly open tariffs**

Tariffs should be open. It would facilitate the planning for travellers and save searching time.

Will the social-surplus maximizing fare structure be very complicated? So complicated that the users (passengers) will have difficulties grasping it? One problem might be that fares would be far from distance-proportional, but depend on critical sections that are changing in a
rather irregular way over time and along each particular route. However, tariffs should be open and user-friendly. This means that the ticket price from A to B by each particular train departure is published in different accessible ways. Another problem for the travellers could be that the price of a trip from A to B could be very different from the price of going back again from B to A; single trip tickets rather than return tickets should be the norm. The optimal single trip fares would vary according to a pattern which passengers at present are not familiar with. First they have to get used to the idea that a return ticket price is normally not the sum of two single tickets of the same price. If passengers learn to think of two separate single tickets, rather than of just one return ticket, it would not be very difficult to remember that being on the train during the morning peak, or late afternoon peak, is more expensive than occupying a seat in the middle of the day or in the evening. By avoiding peak periods and the critical sections, whenever possible, interurban public transport should be generally inexpensive compared with the present situation. This fact should be made clear to the general public, and be the main quality held up in the marketing of bus, rail and airline services. When this fact has become common knowledge, the volume of travel would most likely be greatly boosted.

The present fare differentiation based on the daily demand fluctuations may marginally enhance travel in the short term, but could have a restraining effect in the long term by causing unnecessary uncertainty of whether some present low fares would still hold good in the near as well as the more distant future.

13.6.8 Social surplus maximization versus revenue maximization

The simple idea for a profit-seeking public transport operator that the tariff construction should basically be a matter of revenue maximization, since the total cost (including the railtrack costs) are by and large fixed, has been prevalent for more than a century in the theory and practice of the rail transport industry16.

In modern times this old idea has been developed under the new names of “demand management” or “revenue management” and become highly sophisticated when the Internet can be used. However, it is not the marginal cost unconsciousness that is the main problem from a social economic point of view. The basic problem is wider than that. Solely producer

16 This idea was proposed more than a century ago by the leading railway transport economist, W.M. Acworth
surplus maximization has a number of disadvantages in a decreasing-cost industry producing a large number of joint products.

Compared with present producer surplus maximizing regimes, or full-cost pricing, the greatest difference made by social surplus maximization is the consideration of the joint products aspect. In the normal case of a decreasing-cost, multi-product enterprise the social surplus loss incurred by full-cost instead of marginal cost pricing is measured by the excess burden caused by the average and marginal cost divergence, that is, basically the output-elasticity of the total composite cost function. However, where the multiple products are joint products, a special case appears. For each particular product, with the exception of the capacity-determining one, the marginal cost – in the short and long run alike – is zero up to the capacity limit where it is indefinite. The corresponding optimal price is zero, provided that it does not make demand reach up to the capacity limit. If it does, the final price should be set to avoid excess demand. On an interurban railway line with a good number of stops zero occupancy charges could apply on some sections, since the striving for full capacity utilization follows from the goal of social surplus maximization. The low prices of currently less demanded sections of railway lines would be partly compensated by higher prices in the critical sections. The main feature of optimal pricing is not a generally low level of fares, but a structure of fares, which would be more unequal than at present in order to equalize the occupancy rate and enhance the total volume of travel.

In contrast to this, revenue maximization leaves trains half-empty in many cases, which is a wasteful resource utilization.

13.7. Final methodological conclusion

The methodology for public transport pricing policy is not as well developed as cost-benefit analysis for investments in the transport infrastructure. The present treatise is aiming at eliminating this imbalance and this last section tries to summarize the main methodological points worked out.

The first point is to focus on the role played by the user costs that is the main item of the social cost in urban public transport.

The analytical consequence of taking the user costs into due account in public transport system analysis is that the ordinary price \( P \) in basic microeconomic expositions is replaced by the generalized cost \( GC \), and as distinct from conventional welfare economics where \( P = \)

\(^{17}\) As mentioned before this does of course, not mean that ticket prices would generally be zero.
$MC_{prod}$ is the first-best pricing rule, the fundamental optimality condition is spelt $GC = MC$ where the latter represents the social or system marginal cost.

Figure 13.17 illustrates how the ideal output is determined in a case where demand is a function of the generalized cost. The point of intersection between the inverse demand function, $g(Q)$, and the system marginal cost, $MC$, defines the optimal quantity.

![Diagram showing the generalized cost approach to transport system optimization and pricing](image)

**Figure 13.17. The generalized cost approach to transport system optimization and pricing**

The generalized cost is defined as:

$$GC = P + AC_{user}$$  \hspace{1cm} (13.39)

In the diagram it is seen that in optimum $P$ comes to the following expression when the producer and user marginal costs are distinguished.

$$P = MC_{prod} + MC_{user} - AC_{user}$$  \hspace{1cm} (13.40)
When the number of vehicles is the assumed factor of change in the optimal price derivation, the quality-unadjusted $MC_{prod} = AC_{prod}$. The difference between $AC_{user}$ and $MC_{user}$ is relatively large, and being negative, detracts substantially from $AC_{prod}$. It is shown in section 13.3 that under typical conditions for central city bus transport, the optimal price, $P_N$ comes to the difference between the average producer cost and half the average distribution cost, that is the walking and waiting cost borne by the user of the transport service.

Urban short-distance public transport: $P_N = AC_{prod} - \frac{1}{2} AC_{distr}$  \hspace{1cm} (13.41)

In section 13.6 where long-distance public transport is taken up, an alternative way of optimal price derivation has to be developed because the Mohring effect is much more difficult to estimate when the intervals between departures are exceeding half an hour and can be many hours on thin routes.

Assuming vehicle size as the factor of change, it is shown that the user cost effect to take into account is negligibly different from zero, and the optimal price, $P_S$ can be approximated as shown below, where $MC^S_{prod}$ stands for the quality-unadjusted producer marginal cost, when vehicle size is the assumed factor of change.

Interurban long-distance public transport: $P_S \approx MC^S_{prod}$  \hspace{1cm} (13.42)

However, in section 13.6 it was emphasized that in interurban public transport the optimal level of fares is not the main issue, but the structure of fares. The optimal tariff construction where the right balance is struck between close adherence to the price-relevant costs and intelligibility for the users remains a great challenge.
References


Devanny III, J.W. et al. (1972), Conference rate-making and the West Coast of South America, Commodity Transportation and Economic Development Laboratory, MIT.


SOU 2008:92, Konkurrens på spåret, Swedish governmental report, Fritze, Stockholm (in Swedish)


Tirachini, (2013), Bus dwell time: the effect of different fare collection systems, bus floor level and age of passengers, Transportmetrica, Vol. 9, 28–49.


Webster, V, (1977), Urban passenger transport, some trends and prospects. Transport and Road Research Laboratory, Crawthorne, UK.

Webster, V et.al. (1985), Changing patterns of urban travel. Transport and Road Research laboratory, Crawthorne, UK.