Optimal frequency of public transport in a small city: Examination of a simple method

VTI Working Paper 2021:9

Disa Asplund

Transport Economics, VTI, Swedish National Road and Transport Research Institute

Abstract

This study identifies the welfare optimal frequency of a scheduled public transport service from a methodological perspective, exploring what methods are more suitable for the case of bus services in a small city. The study examines how well various versions of the square-root rule, including established and newly proposed versions, estimate the optimal bus frequencies in the case city of Uppsala, versus estimates generated by a more comprehensive partial equilibrium model. The results indicate that extending the square-root rule by including transfer time, as proposed here, is empirically important. Furthermore, the results indicate that the square-root rule, with this extension, can estimate optimal frequency in Uppsala surprisingly well, and may be suitable for bus lines with two-way demand in the range of 75–200 pax/h.

Keywords

Optimization, Public transport, Service frequency, Square-root rule

JEL Codes

H40; R41

Introduction

Increasing the frequency of a scheduled transport service, such as public transport (PT), is typically expensive in terms of operational costs. However, since people's time is valuable, reducing the frequency of such a service may also be costly in terms of waiting times for passengers. Finding the welfare-optimal PT frequency is an important task that seems to have received too little attention in practice. This study addresses this question from a methodological perspective, exploring what methods are more suitable for the case of bus service in a small city. In the spirit of Occam's razor, the study seeks the simplest model that can adequately describe the given optimization problem. New versions of the existing square-root rule are proposed, and their accuracy is examined and compared with that of previous versions.

Since Mohring's groundbreaking study in 1972, it has been known that the optimal frequency of a transit service under idealized conditions is given by the square-root rule. Furth and Wilson (1981, p. 2) described the situation as follows:

> The best-known theory for setting frequencies on bus routes is the square-root rule, which is based on the minimization of the sum of total passenger wait time costs and total operator cost. In the general case when routes of different lengths exist, the rule states that the service frequency provided on a route should be proportional to the square root of the ridership per unit distance (or time) for that route.

The intuition is that when crowding in PT and external effects are modest, the main benefit of increased frequency derives from decreased waiting times. For a given demand, the total waiting time benefit grows with frequency in a concave fashion. In contrast, the cost of increased frequency is linear in supply, which means that the optimum is well-defined and robust.

According to Furth and Wilson (1981), the square-root rule was not accepted by the PT industry at its introduction, because the formula does not consider capacity constraints and demand responses to changes in service frequency. Furthermore, recent contacts with regional PT authorities in Sweden[1] indicate that the square-root rule is not used in practice. The current official guidance for PT planning in Sweden, Kol-TRAST (2012), does not mention the square-root rule or any other tool to optimize the service frequency; instead, it is claimed (without reference, p. 58) that 6–10 departures per line and hour are optimal for medium-sized towns.

In this study, the performance of various versions of the square-root formula in the setting of a small city will be derived and examined, analyzing how well the square-root rule estimates welfare-optimal service frequencies compared with a more comprehensive model (Asplund and Pyddoke, 2021). This study is similar to that of Jara-Díaz and Gschwender (2003), extending the square-root formula and empirically comparing the results of varied model specifications. While Jara-Díaz and Gschwender (2003) extended the square-root rule in the dimension of capacity constraints, this study tests the effects of multiple extensions in other dimensions, particularly transfer time and including the marginal cost of public funds (MCPF) and wider economic benefits, with the latter being new contributions to the square root formula. The MCPF accounts for the deadweight loss from additional taxation, while the wider economic benefits consist of the positive external effects of reduced generalized travel costs on the labor market. Nash (1978, p. 74) had already identified the MCPF as important when optimizing PT price and supply,[2] and Asplund and Pyddoke (2020) showed that including the MCPF and wider economic benefits is important when determining the optimal PT fare. However, applying the MCPF and wider economic benefits remains somewhat controversial (see, e.g., Jacobs and de Mooij, 2015), and consensus as to whether or not they should be included in cost–

[1] For example, in the Skåne and Uppsala regions.
[2] Wickrey (1955) asserted that the marginal cost of public funds is important when setting public utility prices in general.

benefit analysis has not yet emerged. It is nevertheless interesting to study how important these two factors are for optimal service frequency in practice, as will be done here.

While Jara-Díaz and Gschwender (2003) used calibration data from a large city, this study uses a small city as a case. In fact, much of the later literature on optimal PT provision and pricing has focused on large cities (see, e.g., Asplund and Pyddoke, 2020) and consequently emphasized capacity constraints, which are important in that setting. In the Swedish context, there are a few notable exceptions: Börjesson et al. (2019) and Asplund and Pyddoke (2020)[3] both studied small cities in Sweden from an empirical perspective. The results of these two studies indicate that the capacity constraint is only rarely binding in these settings, so that the square-root rule may be more relevant to small cities. The potential to use a simple formula to optimize service frequencies is attractive, because it means that the calculations are practicable for PT authorities. Vigren and Ljungberg (2018) found that regional PT authorities in Sweden did not use cost–benefit analysis to guide decisions about supply levels. That is, decisions regarding supply increases were not preceded by rigorous weighing of costs and benefits. One of several possible reasons for this is that existing models were too costly to use and needed special expertise not available in house.

Nash (1978) thought that although the simple welfare maximization of frequencies and fares was not very difficult, it could still be too complicated for widespread understanding and support in practice. He claimed that such considerations had led to the adoption by London Transport of a simpler rule: maximizing passenger mileage subject to a budget constraint. Nash further evaluated this decision rule along with the alternative decision rule of maximizing vehicle mileage subject to a budget constraint, concluding that maximizing passenger mileage seemed to be superior. However, such simplified decision rules, not formally based on social welfare, always risk leading to inferior solutions (see, e.g., Frankena, 1983). For example, maximizing passenger mileage may lead to lower fares for long trips than short trips, although the external effects in terms of crowding inflicted on fellow PT riders are greater for long trips. Opting to maximize the numbers of PT trips may, in contrast, lower the price too much for short trips, so that travelers who benefit little from a trip in practice are attracted (e.g., riding only one stop instead of walking), causing crowding and delay for PT riders who really need the service as well as increased costs for operators. From the empirical literature from English-speaking OECD countries, Savage (2004) concluded that transit agencies typically opted to maximize level of service as opposed to social welfare and, as a result, social welfare was lower than it might otherwise be. There seems to be great potential in simple decision rules for optimal frequencies firmly rooted in formal welfare economics, such as the square-root rule with proper extensions.

One reason why simplified decision rules have not been used much may be the perception that supply and prices should be jointly optimized.[4] However, Asplund and Pyddoke (2021) indicated that this may not be necessary for bus services in a small city, since optimal frequency was robust and rather independent of the fare. This result was based on unconstrained welfare maximization. If there is a binding budget constraint, this proposition may no longer hold (see Jara-Díaz and Gschwender, 2009). Keeping this in mind, a first step would be to examine whether it is possible to attain accurate and robust results from a simple formula for unconstrained conditions.

This study examines how well the square-root rule estimates the optimal bus frequency in the case city of Uppsala, compared with the more comprehensive BUPOV model (Asplund and Pyddoke, 2021), a partial equilibrium model with three travel modes and city-specific data. To this end, the BUPOV model is extended by acknowledging the effect of additional boardings on vehicle speed, and hence on operational costs and travel times, since this effect has previously been identified as important (e.g., by

[3] In an elaborated analysis of occupancy rates in Uppsala, Pyddoke (2000) found that for much of the network in Uppsala, crowding only rarely occurs, even in rush hours.
[4] I have not found earlier work dedicated specifically to welfare-optimal frequency from a theoretical perspective. Economic work typically covers both pricing and supply simultaneously, while engineering approaches do not optimize welfare explicitly, but use proxies to acknowledge the most important aspects of welfare optimization.

Jara-Díaz and Gschwender, 2003). Here, BUPOV represents the best practice for estimating the welfare-optimal frequency in Uppsala, to which other practices are compared. One should recall that BUPOV is a model and that estimates from BUPOV do not represent the "true" optimal frequency per se.

The results indicate that extending the square-root rule by including transfer time, as proposed here, is empirically important. The proposed extensions are shown to considerably improve[5] the performance of the square-root rule in a small city. Furthermore, the results indicate that the square-root rule, with proper extensions, explains optimal frequency in Uppsala surprisingly well. A sensitivity analysis with respect to demand indicates that the performance of the proposed formula is not robust with respect to optimal frequency, but still performs well with regard to welfare for an interval of about 75–200 pax/h (two-way demand per bus line). In comparison, the original Mohring formula performs considerably worse for demand levels over 75 pax/h.

Interaction with route density, vehicle size, stop spacing, and pricing is beyond the scope of this study, which should be recalled when interpreting the results, especially when it comes to vehicle size. Adjusting vehicle size is a (first best) substitute policy for the (second best) policy of adjusting frequency compared with the square-root rule in order to mitigate crowding. Only when demand exceeds what can be accommodated by maximum vehicle size available does the square-root rule lose its relevance, because then the capacity constraint becomes the binding condition for the optimal frequency. However, increasing the vehicle size will have a cost, which means that slightly increasing the frequency in addition to vehicle size will be optimal (see, e.g., Jara-Díaz and Gschwender, 2003). Note, however, that in this study, the occupancy after optimizing is well under the absolute capacity constraint in all cases.[6]

The paper is organized as follows. First is a brief literature review focusing on the theory of optimal PT frequency, specifically on the square-root rule. Second comes a theory section, explaining the existing square-root rule as well as proposing two simple extensions. Third is an empirical application, including method, data, and results, followed by the conclusion.

Literature

Jara-Díaz and Gschwender (2003) reviewed the models most relevant to this analysis; most importantly, the optimal frequency of a transit service with idealized conditions was given by the square-root formula (see theory section), which is usually attributed to Mohring (1972). However, Mohring (1972) acknowledged that this formula was communicated to him by William Vickrey.

In the early 1970s, two studies of optimal frequency were presented (Newell, 1971; Salzborn, 1972) using an engineering approach. Newell minimized waiting time given a fixed vehicle fleet size, while Salzborn used a two-step minimization process, with the primary objective to minimize the number of vehicles needed and the secondary objective to minimize waiting time. The tradeoffs between operational costs and passenger waiting times were absent from these analyses.

Mohring (1972) published an economic study of the optimal frequency and fare of a bus service. It included both tradeoffs between operational costs and waiting times as well as endogenous demand (based on fare and service frequency), and discovered a new phenomenon: the Mohring effect. The Mohring effect is a special case of increasing returns to scale, in which higher production reduces the average production costs, not of the producer, but of the consumers, who in Mohring's terms are part

[5] "Improve" here refers to the ability of each square-root formula to replicate the results of BUPOV.
[6] The mean seating capacity in Uppsala is 38 seats/bus (Asplund and Pyddoke, 2020). The optimization resulting in the lowest mean frequency results in a maximum occupancy of 42 pax/bus (in the most central point in peak hours). However, there is considerable spare capacity for standing passengers. For each bus seat in Uppsala there are about 1.4 units of standing capacity (Pyddoke, 2020), i.e., the total mean capacity is about 91 pax/bus.

of the production process by supplying the scarce resource of their own time. Mohring (1972) presented a first version of the square-root rule, which has subsequently been refined and expanded.

Jansson (1979) generalized Mohring's analysis to any scheduled transport service (including freight services), added dimensions, and derived important mathematical properties. One important modification to Mohring's model was that vehicle speed was assumed to be given in the short run, since schedules need to be fixed (and include enough spare time to allow for occasional delays due to demand shocks). Jansson also noted that feeder transport costs (e.g., walking to bus stops) are mainly a function of service network density, while costs relating to waiting times are mainly a function of the frequency. However, Jansson's (1979) main analysis treated supply as a composite good, including both frequency and network density. In a subsequent work, Jansson's (1980) focus was instead on frequency, refining and extending the square-root formula with boarding and alighting time. Jansson's (1980, p. 143) Fig. 5.2 showed that the benefit in terms of reduced waiting times of one extra departure decreased quickly with frequency. Consequently, Furth and Wilson (1981) found the best allocation of buses across routes to be robust with respect to the objectives and parameters assumed. Another important finding was that the results obtained when demand is assumed to be inelastic differ little from those when a more realistic demand model is used, i.e., the static formulation of the square-root rule may not be a major weakness. In contrast, Basso and Silva (2014) showed that under highly congested conditions (e.g., in London and Santiago, Chile) with high bus frequencies (>15 buses/hour), optimal frequencies were not very robust. One interpretation is that, under such conditions, the square-root rule and possibly also the Mohring effect are no longer relevant.

Jara-Díaz and Gschwender (2003) compared various versions of the square-root rule. Fig. 1 in Jara-Díaz and Gschwender (2003, p. 463) shows the optimal frequency as a function of demand per line resulting from various models. The Jansson (1980) version differs little from the Mohring version for passenger demand per line up to 1000 pax/h, but the difference grows with demand up to around a 50% higher frequency for the Jansson model at 9000 pax/h. In this study, demand is 200 pax/h during peak and half of that during off-peak (OP) hours. Jara-Díaz and Gschwender (2003) also extended the Jansson model with crowding inside vehicles, which substantially increased the optimal frequency at high demand levels.

Jara-Díaz et al. (2012) and Jara-Díaz et al. (2018) theoretically examined the implications of transfers for the optimal operation of PT in simple trunk–feeder systems. They derived new square-root formulas for optimal frequencies under such conditions, extending Jansson's (1980) versions with transfers. In this study, this extension is derived under more general conditions and with slightly simpler resulting equations.

For further literature on related issues, see Hörcher and Tirachini (2021). Next follows a theoretical section, mathematically working through the Mohring and Jansson models as well as the proposed new extensions.

Theory: The square-root rule

Mohring (1972) derived the square-root rule, central to this analysis. A slightly refined, generalized, and streamlined version is derived below, largely similar to eq. (5) of Jara-Díaz and Gschwender (2003). Suppose the only costs related to the frequency of a PT service are the operational costs and the delay costs of the passengers. Then the total welfare costs related to frequency can be described as:

$$C_{tot} = C_u \cdot f + P_w \cdot E(t_w) \cdot D = C_u \cdot f + \frac{P_w}{2f} \cdot D, \qquad (1)$$

where:
C_u is the unit cost per round trip of operating the service,
f is the frequency of the transport service,

t_w is the absolute deviation in time between scheduled access and each passenger's preferred timing of usage (e.g., waiting time at bus stop when the user is not using a timetable to plan the trip[7]),
P_w is the shadow price of one unit of t_w,
$E(t_w) = \frac{h}{2}$ is the expected value of t_w, given random, uniformly distributed preferred departure times,
$h = \frac{1}{f}$ is the headway of the transport service, and
D is the (static) demand for the transport service per PT line (summed over both directions), in boarding passengers per time interval (e.g., per hour).

Differentiation of (1) with respect to f gives:

$$f^* = \sqrt{\frac{P_w \cdot D}{2C_u}}. \qquad (2)$$

From eq. (2) it is trivial to see that halving P_w or D will be equivalent to doubling C_u and will result in an approximately 30% reduction in optimal frequency. Since the operator can typically measure both demand and operational costs with rather high precision, the optimal frequency will be rather robust in these dimensions. Fig. 1 shows an example optimum using eqs. (1) and (2).

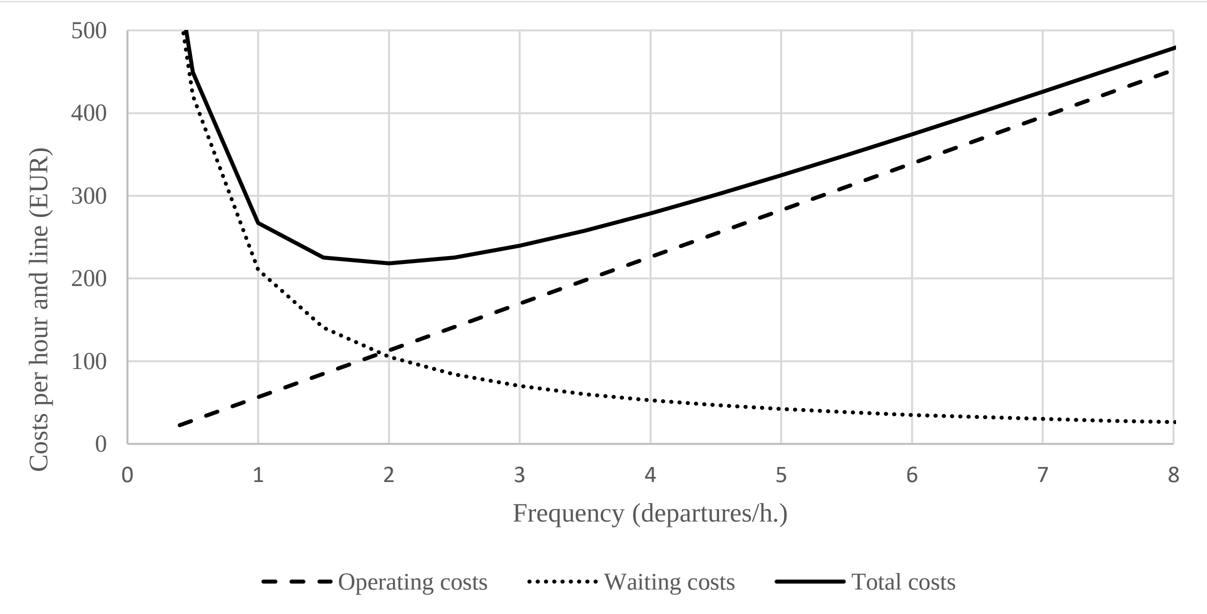

Fig. 1. Costs as a function of frequency. Real data from Uppsala in 2010 (representative line for weekday OP hours, see the section *Square-root rule calculations for Uppsala*).

Boarding time extension
Jansson (1980) extended the square-root rule with boarding and alighting time, i.e., the delay each passenger causes to each vehicle (and fellow passengers) by boarding and alighting. The resulting formula was:

$$f^* = \sqrt{\frac{D}{C_u} \cdot \left(\frac{P_w}{2} + P_v \cdot t_l \cdot D \cdot \frac{t_v}{T}\right)}, \qquad (3)$$

where
P_w is the shadow price of waiting time, corresponding to p_t in eq. (1),

[7] If the user uses the timetable, the waiting time may instead occur at home. In both cases, there may also be waiting time at the destination, if the arrival is planned to be ahead of a scheduled activity.

P_v is the shadow price of in-vehicle time (per transported unit),
t_l is the loading time (i.e., boarding and alighting time) per transported unit (passenger),
t_v is the average time each transported unit (passenger) spends aboard each vehicle, and
T is the total running time of one round-trip cycle per vehicle, excluding loading time.

Jansson (1980) noted that for low to moderately high values of D, the first term $\left(\frac{P_w}{2}\right)$ will dominate the second term within the parentheses. In this study, this assertion will be tested empirically for the small city of Uppsala. Jansson (1980) further noted that when D is comparatively high, the square-root rule will have more serious limitations, since it does not take capacity restrictions into account. Jara-Díaz and Gschwender (2003) extended Jansson's model with capacity restrictions. However, the resulting solution is slightly too complex to apply simply and will therefore be excluded from this analysis. When crowding becomes severe, vehicle size increases in relevance, as do other measures such as increasing the fare. Hence the optimization problem quickly becomes more complex, and a simple formula may no longer be adequate (as indicated by the results of Basso and Silva, 2014). However, the BUPOV model to which the various versions of the square-root rule are compared includes the cost of crowding, resulting in an optimal solution well below the maximum vehicle capacity.

Transfer time extension
If the transport service includes transfers between vehicles (i.e., service lines), the cost of transfer waiting time needs to be considered, but this seems rarely to have been done in earlier welfare optimization studies. Two notable exceptions are those of Jara-Díaz et al. (2012, 2018), who extended Jansson's (1980) model. In the following, Jara-Díaz et al.'s approach is somewhat simplified (excluding boarding times) and, on the other hand, somewhat generalized in two other dimensions. A first generalization is not to restrict the analysis to a specific setup such as trunk–feeder system. A second generalization is to make a distinction between waiting time that may be spent at home (passive waiting time), and time between transfers that cannot be spent at home (active waiting time). Empirical data shows that that the value of these two categories may differ vastly in value, and this is especially relevant for low-frequency services, i.e. for low-density conditions (see Appendix A). However, it should be noted that this generalization is only relevant when two important criteria are met:

1) Frequency is sufficiently low (about 5–12 vehicles/hour or less, see, e.g., Hörcher and Tirachini, 2021) for it to make sense to the passengers to make the effort to time for a specific departure; and

2) There is the possibility for passengers to predict the timing of each vehicle at their stop, i.e., there is a published timetable and reliability is sufficiently high (see Appendix A).

Given that these two criteria are met, let us first define the general problem. The total costs relating to frequency and transferring for a specific transport line are:

$$C_{tot} = C_u \cdot f + C_{passenger}(f), \tag{4}$$

where the total waiting cost of the users is:

$$C_{passenger}(f) = \\ = P_{w1} \cdot E[t_{w1}(f)] \cdot D_1 + (P_{w2} \cdot E[t_{w2}(f)] + P_\tau \cdot t_\tau + \tau) \cdot D_2 \,, \tag{5}$$

and indices 1 and 2 denote boardings to the line in question from new arrivals and transfers, respectively.

P_{w1} is the shadow price of waiting time for new arrivals,
P_{w2} is the shadow price of transfer waiting time,

P_τ is the shadow price of walking during transfers,
t_{w1} is the absolute deviation in time between scheduled access and each passenger's preferred timing of usage for new arrivals,
t_{w2} is the transfer waiting time per person for passengers boarding due to transfers (not including walking time),
t_τ is the necessary walking time for each transfer,
τ is the pure transfer penalty (i.e., the inconvenience of switching vehicles),
D_1 is the number of boardings per time unit for new arrivals, and
D_2 is the number of boardings per time unit for transfers.

If line scheduling is coordinated to minimize transfer time (so that $E[t_2(f)]$ depends not only on f but also on the scheduling of the other lines), it may be impossible to find an analytical solution to the problem, especially if such a coordination process is ad hoc rather than mathematical. However, here a simple approach to approximate the transfer time is proposed: assume that each service line is scheduled independently of other lines, as done by Jara-Diaz et al. (2012). An implicit assumption is that transfer times after optimization are enough to physically perform each transfer, i.e., that $h \geq t_\tau$. If line scheduling is independent, then arrivals at a transfer point can be assumed to be random, as can new arrivals. The headway of the second line used in each transfer determines the expected transfer waiting time for each transferred entity (e.g., passenger):

$$E[t_{w2}(f)] = \frac{1}{2f} - t_\tau \tag{6}$$

where t_τ is the (constant) time needed to physically perform the transfer between vehicles, i.e., alignment, walking, and boarding.

When optimizing the frequency of each line, the number of passengers boarding that line (divided into new arrivals and transfers) is relevant. Thus, the waiting costs related to the frequency of a specific line are:

$$C_{passenger}(f) = \frac{P_{w1} \cdot D_1 + P_{w2} \cdot D_2}{2f} + [(P_\tau - P_{w2}) \cdot t_T + \tau] \cdot D_2. \tag{7}$$

This approach results in the following extended optimal frequency formula:

$$f^* = \sqrt{\frac{P_{w1} \cdot D_1 + P_{w2} \cdot D_2}{2C_u}}. \tag{8}$$

In the special case of a representative line approach, i.e., modelling one line to represent one homogeneous system, D_1 equals the total number of passengers starting a trip in a given time interval (e.g., passengers per hour) divided by the number of lines, while $D_2 = \alpha \cdot D_1$, where α is the average number of transfers per passenger.

Inclusion of linear external effects

PT typically has some external effects beyond the costs and benefits experienced by the operators and passengers. Some of these can be assumed to be approximately linear to the consumer surplus or operational costs when optimizing frequency, while others cannot. Hence the costs and benefits that can be expressed in a linear form may be incorporated into the square-root rule.

PT may or may not have wider benefits for economic development beyond those experienced by users (see, e.g., Eliasson and Fosgerau, 2019). The wider economic benefits may be incorporated as follows. Here the assumption is that there are additional benefits from reduced transport costs, including waiting costs, beyond those accruing to direct users.[8] There may also be additional external effects of reduced car traffic in terms of emissions and noise. If the assumption is that these additional benefits are proportional to the change in the traveler's transport costs, then eq. (7) may simply be extended by an external effect factor to account for this:

$$C_{passenger} = \in \cdot \frac{P_{t1} \cdot D_{T1} + P_{t2} \cdot D_{T2}}{2f}. \tag{10}$$

where ∈ is the external effect factor. Note that external effects in terms of road congestion are typically not linearly related to user benefits arising from changes in PT frequency. This is partly because road congestion is not linearly related to car demand, but, more importantly, is because buses also inflict congestion, and this relationship is not linearly related to user benefits. Appendix B shows an illustration of this for the case of Uppsala.

There are typically also costs related to financing PT extensions. For example, if the additional supply is completely financed by raised taxes, then the operational costs may be multiplied by the MCPF. On the other hand, if increases in supply are financed by increased ticket prices, then demand reductions from this may in practice imply an additional cost (see, e.g., Jara-Díaz and Gschwender, 2009). The financing cost may be incorporated as follows:

$$C_{tot} = \beta \cdot C_u \cdot f + C_{passenger}(f), \tag{11}$$

where β is the financing cost multiplier (typically between 1 and 2). Also, linear external effects from buses (such as emissions and noise) may be incorporated into β.

Then, a modified version of the square-root rule that accounts for external effects and transfer time between lines is:

$$f^* = \sqrt{\frac{\epsilon \cdot (P_{t1} \cdot D_{T1} + P_{t2} \cdot D_{T2})}{2C_u \cdot \beta}}. \tag{12}$$

In the special case of analyzing one representative line for a homogeneous system (e.g. a city), eq. (12) may be simplified to:

$$f^* = \sqrt{\frac{\epsilon \cdot (p_1 + p_2 \cdot \alpha) \cdot D}{2C_u \cdot \beta}}. \tag{13}$$

In the next section, various versions of the square-root rule from this section are compared empirically in the case of the small city of Uppsala.

Empirical analyses

This study examines how well the square-root rule estimates the optimal bus frequency for the case city of Uppsala, compared with the more comprehensive BUPOV model (Asplund and Pyddoke, 2021), which is a partial equilibrium model with three travel modes and city-specific data. The BUPOV model is extended further by acknowledging the effect of additional boardings on vehicle speed, and hence on operational costs and travel times, since this effect has previously been identified as important (e.g., by Jara-Díaz and Gschwender, 2003).

[8] The wider economic benefits reported by Eliasson and Fosgerau (2019) concerned labor market improvements, and in the empirical part of this study, the external benefit factor is restricted to representing labor market improvements, excluding other possible benefits.

The BUPOV model
Uppsala is 70 kilometers north of Stockholm and is home to Sweden's oldest university. In 2010, it had 155,000 inhabitants and its urban area covered 51 square kilometers. Although Uppsala is the fourth largest city in Sweden, in 2016 it had the second most severe congestion problem in Sweden in terms of mean delay, with almost the same delay as the most congested city, Stockholm (Asplund and Pyddoke, 2019). One reason is that the two largest cities at this time had already reduced their congestion problems by implementing congestion charges.

Asplund and Pyddoke (2020) found a substantial and robust oversupply of PT in Uppsala, using the BUPOV model. They modeled welfare-optimal bus pricing and frequency in Uppsala, considering variability in occupancy and using detailed origin and destination data incorporating modal choice and local external effects. In a subsequent study (Asplund and Pyddoke, 2021), the BUPOV model was refined with respect to modeling the car traffic. This later model version is used in this analysis. Formal presentations and complete specifications of the relevant BUPOV models are found in Asplund and Pyddoke (2020) and Asplund and Pyddoke (2021), respectively.

The model is intended to represent the effects of transport policies on mode choice, trip timing, and welfare in a small city with one PT mode (bus only). BUPOV has a nested structure, involving two optimization steps. BUPOV represents traffic demand and is calibrated to variations between peak and OP times in inner and outer parts of the city. As for scope, it attempts to capture the major short-term welfare effects of trips beginning or ending in Uppsala, but only those parts of trips occurring within city boundaries. In the long term, more adaptation may occur due to changes in destination choice, residential and workplace location, private supply of parking spaces, etc.

A social planner is who optimizes welfare, given that she anticipates what the private responses will be. That is, she optimizes welfare via a set of policy variables, given the user equilibrium that will result from such policy changes. Welfare optimality refers to the optimality that a hypothetical social planner who manages all publicly owned assets to maximize citizen welfare would find optimal.

BUPOV is based on a radial spatial representation of a city with two zones: the city center (inner zone) and the outer city (outer zone). The analysis is restricted to workday traffic, divided into two time-period categories: peak and OP. This representation makes it possible to analyze optimal prices and bus frequencies differentiated in time and space. Since the studied policy measures are evaluated at the zone level with trips aggregated, route choices within each zone are assumed to be unaffected, so route choice is not modeled.[9] The mode and timing choices for each trip as well as total demand are flexible.

BUPOV is based on detailed data on current travel behavior in terms of origin–destination (OD) matrices; it implicitly represents the current population density but does not represent changes in population or place of residence. Travelers can choose between three modes of transport: car, bus, and walking/cycling. The choice of travel alternative depends on monetary cost, road congestion, bus crowding, and time gains and losses due to changes in bus frequencies. In addition to the effects of policies on producers and consumers, there are effects on the time cost of freight traffic, effects on health (e.g., of noise and air pollution), and environmental effects primarily in terms of CO_2 emissions. The changes in the PT authority's financial results are evaluated using a MCPF factor. Optimally, this should correspond both to the marginal welfare costs of raising one additional unit of tax revenue and to the marginal value of one additional unit of public funds used for alternative purposes, for example, health care. The BUPOV model also multiplies the consumer benefits by a wider economic benefit (WEB) factor in the welfare estimation, i.e., accounting for better functioning of the labor market arising from increased accessibility, counteracting the effect of the MCPF factor. The WEB factor is calculated by "removing" the MCPF factor from commuting trips, i.e., by also multiplying the consumer surplus by the MCPF factor for the fraction of trips that are commuting trips. This simple approach gives the same WEB factor as calculated by a sophisticated model of the total investment

[9] The network and routing are not included in the model; instead travel distances and times for each mode and OD pair are based on mean values from a separate routing model.

plan for transport projects in Sweden (Anderstig et al., 2018), that is, 12% more benefits from increased generalized accessibility, i.e., consumer surplus.

Three types of OD pairs are modeled: within the inner zone ("inner"), between zones in any direction[10] ("inter"), and within the outer zone ("outer"). Each OD pair constitutes a separate (isolated) demand system, interlinked by sharing space both on the streets and inside the buses. The demand for a travel alternative (mode m and time period t, for an OD pair trip) is modeled as a change from the demand in the reference situation as follows:

$$\Delta D_{m,t,OD} = \Delta D_{m,t,OD}(p, f, o, \delta | \varepsilon), \tag{14}$$

where p is a price vector, f is bus frequency, o is the occupancy level in buses, δ is traffic delay (for buses and cars), and ε is a matrix of demand elasticities. Adjustment to a new user equilibrium caused by a change in a policy variable (e.g., frequency) is done by successively iterating the demand calculations of consumer travel choices, congestion, and in-vehicle crowding in buses. In the baseline case, demand is assumed to be in a steady state, but if a policy reform is introduced, a new steady state is approached through iteration. The levels of congestion and crowding affect the generalized cost of each travel alternative, meaning that some travelers adjust their travel choices when these levels change, so that congestion and crowding will again be updated. This iterative process continues until the model reaches a new steady state.

The value of in-vehicle travel time for PT as function of crowding is based directly on the Swedish national guidelines for the welfare economics of infrastructure investments, ASEK 6 (Swedish Transport Administration, 2016), which are in turn based on international peer-reviewed literature on the crowding multiplier, combined with peer-reviewed Swedish time value studies. The value of in-vehicle travel time for PT (in EUR) is calculated using the equation:

$$VoT^{ivt,PT} = 3.7 + 0.37 \cdot o + 1.33 \cdot o^2, \tag{15}$$

where 3.7 is the value of time in an empty vehicle and o is the occupancy rate (of seat capacity[11]). Asplund and Pyddoke (2020) estimated the optimal frequencies, and the results indicated that the bus frequencies in Uppsala (six buses per hour during peak times and four buses per hour in OP) were on average considarbley too high.

In this study, the BUPOV model presented by Asplund and Pyddoke (2021) is extended by including bus expected delay caused by passenger boardings, in line with Jansson (1980). An additional update is to use the actual seating capacity for buses of 38 pax/bus (used by Asplund and Pyddoke, 2020) instead of the initial assumption of 30 pax/bus. Next follows the empirical basis of the boarding time extension from the international literature.

First one notes that the boarding time per passenger varies considerably depending on conditions. Important determinates are the ticket handling system, bus-stop physical design, bus design, and the number of passengers. The Transportation Research Board (2003) transit planning manual presented rule-of-thumb values for boarding times in the range of 2–4 s/pax. Liu et al. (2017) estimated boarding times in Singapore to be about 1.4 s/pax, when double-decker buses were excluded (Table 7). Simon et al. (2013) estimated boarding times in Beijing. Fig. 5 shows that boarding times increase sharply

[10] Not modeling the direction of trips (i.e., towards and away from the city center) in the morning versus afternoon peak hours is a simplification that may lead to the underestimation of crowding, as we assume that passengers are evenly distributed between the two directions for each line. A sensitivity analysis of this was performed by Asplund and Pyddoke (2020), who tested the extreme alternative assumption that all passengers travel in the same direction—that is, half of the buses run empty and the passengers experience double the crowding versus the reference model. The welfare gains from optimization seem reasonably robust to this alternative model specification.

[11] The total capacity is about 2.4 times the seating capacity, according to Pyddoke (2020).

with the load factor in an exponential fashion, from about 1.7 s/pax at 0% load factor to about 4.0 s/pax at 90% load factor. In BUPOV the load factor is about 27–45% during peak hours and 12–20% after optimization (excluding boarding times). From this, a representative boarding time of 2.0 s/pax is estimated using Fig. 5 in Simon et al. (2013) (see Table 1). A sensitivity analysis is also performed, using 4.0 s/pax (see Table 2).

Table 1
Boarding time estimation. Load factors are estimated by the BUPOV model after optimization (but excluding boarding time). Boarding times are based on Fig. 5 in Simon et al. (2013) as a function of load factor.

Zone	**Time of day**	**Load factor***	**Boarding time (s)**
Inner	Peak	45%	2.3
	OP	27%	2.0
Outer	Peak	20%	1.9
	OP	12%	1.8
		Mean	**2.0**

* Seating occupancy rate divided by 2.4 to account for standing capacity, based on Pyddoke (2020).

Table 2 shows the changes in the welfare optimum, including a sensitivity analysis. When introducing bus delays from boardings of 2.0 s/pax, optimal frequencies are only slightly affected, while the net social benefits of such an optimization are reduced by about 8%.

Table 2
Changes in welfare optimum from updating.

	Frequency optimization resolution		Old model (Asplund and Pyddoke, 2021)	Updated seat capacity per bus, 30-> 38 pax	New model, with delay	
	Zone	Time of day			Main, 2.0 s/p	Sensitivity, 4.0 s/p
Optimal frequency (departures/h)	Inner	Peak	4.0	3.7	3.8	3.9
		OP	3.3	3.3	3.3	3.3
	Outer	Peak	3.1	3.0	3.1	3.2
		OP	2.3	2.3	2.3	2.3
Welfare effects in optimum per weekday (EUR)	Consumer surplus		–24,549	–24,833	–24,511	–24,123
	WEB (≈ CS·0.12)		–3055	–3090	–3050	–3002
	Producer surplus		+39,580	+41,926	+39,951	+38,015
	MCPF (= PS·0.30)		+11,874	+12,578	+11,985	+11,405
	Other external effects*		+667	+709	+690	+670
	Net social benefits		+24,517	+27,290	+25,066	+22,965

* Note that external effects of emissions from car traffic are almost perfectly internalized through taxation in Sweden, which is why the net effect (displayed here) is relatively small.

Square-root rule calculations for Uppsala
Sweden has a well-developed set of national guidelines for the cost–benefit analysis of infrastructure investments, called ASEK (Swedish Transport Administration, 2016), and based on national and international empirical research. Some of the input parameters of the square-root formula are included, notably the MCPF (1.3) and values of waiting time on various occasions. Asplund and Pyddoke (2020) selected the relevant values of waiting time for bus traffic in Uppsala from ASEK 6 (Swedish Transport Administration, 2016) to be EUR 4.16/h for passive waiting time and EUR 11.0/h for transfer waiting time (both as of 2014, see Appendix A).

In Uppsala the most important linear external effects of increased frequency are the wider economic benefits, while other net linear external effects are negligible, net of taxation. Therefore, the main analysis will focus on the wider economic benefits in this respect. The wider economic benefit factor in BUPOV is roughly 1.12,[12] coinciding with a national figure provided by Anderstig et al. (2018).

Other data for the square-root rule for Uppsala can be extracted from the BUPOV model as well (see Asplund and Pyddoke, 2020). The average number of transfers per trip, α, is 0.41,[13] and the average total running time of one round-trip cycle is 74.5 min (excluding loading time). The average in-vehicle time per passenger and trip is 13.5 min. Boarding times are assumed to be 2 s/pax, based on Simon et al. (2013; see Table 1). In this study it is assumed that ticket prices is fixed, so the financing cost equals the marginal cost of public funds. These data imply the following parameter values:

α = 0.41
β = 1.3
$\in$ = 1.12
P_w, P_{w1} = EUR 4.16/h
P_{w2} = EUR 11.0/h
P_v = EUR 4.53/h
t_l = 3.3 s
t_v =13.5 min
T = 74.5 min

The operational costs per round trip of such a line are displayed in Table 3, differentiated in terms of peak (two hours in the morning and three hours in the afternoon during weekdays) and OP (other weekday hours) hours, as well as the demand and occupancy rate. The peak operational costs are higher for two reasons. First, because extra buses are needed in the peak hours, expanding or reducing the service frequency implies expanding or reducing the capital stock. Second, each round trip takes longer in the peak hours due to congestion, implying an increase in both capital costs and driver pay. Table 3 also presents the respective demands as well as the resulting optimal frequencies. Instead of simply examining the results of the two extension formulas, i.e., eqs. (8) and (13), four analyses have been performed. For transfer time, the effects of including transfer time and also using a separate shadow price for transfer time have been separated into two analyses (Transfer times 1 and 2, respectively). For the external economic effects, two versions have again been tested. First a version excluding transfer time and, second, the version in eq. (13) including both external economic effects and transfer times.

[12] Consumer surplus is valued higher (by the marginal cost of public funds) for commuting trips, to reflect the wider economic benefits of reduced transport costs.
[13] From the official national travel model.

Table 3
Optimal service frequency in Uppsala, various models.

<table>
<tr><th rowspan="3">Time of day</th><th colspan="2" rowspan="2">Input data</th><th colspan="7">Optimal service frequency (departures/h)</th></tr>
<tr><th colspan="6">Square-root rule version*</th><th rowspan="2">BUPOV</th></tr>
<tr><th>C_u (EUR)</th><th>D (pax./h)</th><th>Mohring (1972)</th><th>Boarding time**</th><th>Transfer time 1</th><th>Transfer time 2</th><th>External economic effects</th><th>TT2 + EEE ***</th></tr>
<tr><td>Peak</td><td>115</td><td>202</td><td>1.91</td><td>1.94</td><td>2.28</td><td>2.77</td><td>1.78</td><td>2.57</td><td>3.38</td></tr>
<tr><td>OP</td><td>57</td><td>101</td><td>1.93</td><td>1.94</td><td>2.29</td><td>2.79</td><td>1.79</td><td>2.59</td><td>2.61</td></tr>
<tr><td colspan="3">Welfare loss compared with most sophisticated model ****</td><td>–17%</td><td>–17%</td><td>–7%</td><td>–2%</td><td>–23%</td><td>–9%</td><td>0%</td></tr>
</table>

* Using eqs. (2), (3), (8), and (13), respectively. For *Transfer times 1 and 2*, eq. (8) is used with two different values of P_{w2} , respectively. For *Transfer time 1*, $P_{w2} = P_{w1}$ = EUR 4.16/h and for *Transfer time 2*, P_{w2} = EUR 11.0/h. For *External economic effects*, eq. (13) is used with no transfers, i.e., $\alpha = 0$.
** Jansson (1980).
*** *Transfer time 2* and *External effects*.
*** Calculated using the BUPOV model as the difference in welfare compared with the BUPOV optimum, as percent of total operational costs in BUPOV optimum.

Two results emerge from Table 3. First, all analyzed variants of the square-root rule prescribe almost exactly the same optimal frequency in peak as in OP hours, in line with the original observation of Jansson (1980). This result is rather trivial given that both the operational costs and demand are twice as high in peak as in OP hours. However, the BUPOV model, which takes congestion and crowding into account, does not confirm this relationship, but optimal peak frequencies are about 30% higher in the peak hours.

Second, the *Mohring*, *Boarding time*, and *External economic effects* versions seem to considerably underestimate the optimal frequency. However, when transfer time is instead included, the square-root rule seems to perform considerably better, entailing only a 2% loss in welfare compared with the BUPOV model. The somewhat more sophisticated version that also includes external economic effects actually performs worse, but this may not be a general result. One can also see that the two last columns give a rather similar OP frequency of about 2.6 busses/h, but not in the peak case. This indicates that eq. (8) may be accurate enough for cases in which the demand does not exceed 100 pax/h per line (two-way) and the external effects of the car traffic are nearly correctly priced. For a considerably denser traffic situation (i.e., higher PT demand and congestion levels), a more sophisticated model seems to be warranted. This will be studied in more detail in a sensitivity analysis in Fig. 2.

To see how important each extension is, the *Boarding time*, *Transfer time 1*, and *External economic* effects extensions should be compared with *Mohring*, while *Transfer time 2* should be compared with *Transfer time 1* (to isolate the external economic effects). Table 4 shows the percental effect of each extension. One can see that the transfer time extensions are by far the most important empirically, while the other two extensions are small in their effects.

Table 4
Percental effects of each extension.

Time of day	Boarding time*	Transfer time 1*	$P_{w2} \neq P_{w1}$ **	External effects*
Peak	+1%	+19%	+22%	–7%
Off–peak	+1%	+19%	+22%	–7%
Welfare improvement	+1%	+10%	+5%	–6%

* Compared with *Mohring*.
** *Transfer time 2* compared with *Transfer time 1*. A sensitivity analysis was performed on these results, using the travel time multipliers for headway-related waiting time and transfer time (i.e., interchange waiting time) from Wardman et al. (2016) instead of the Swedish guidelines. The headway multiplier was reduced from 0.94 to 0.78 and the transfer multiplier from 2.5 to 1.77. The resulting difference between the *Transfer time 2* and *Transfer time 1* versions was then 17% instead of 22% (for both peak and OP hours).

So far, the results for the *Transfer time 2* formula are encouraging, indicating that the formula provides accurate enough estimates of the welfare-optimal frequencies in Uppsala. However, one may ask whether this result holds in a more general setting. To address this question, a sensitivity analysis was conducted of the accuracy of the *Transfer time 2* and *Mohring* formulas with respect to PT demand (OP),[14] as depicted in Fig. 2. Unfortunately, the accuracy of the *Transfer time 2* formula does not seem very robust with respect to PT demand, but works best at a demand of about 130 pax/h.

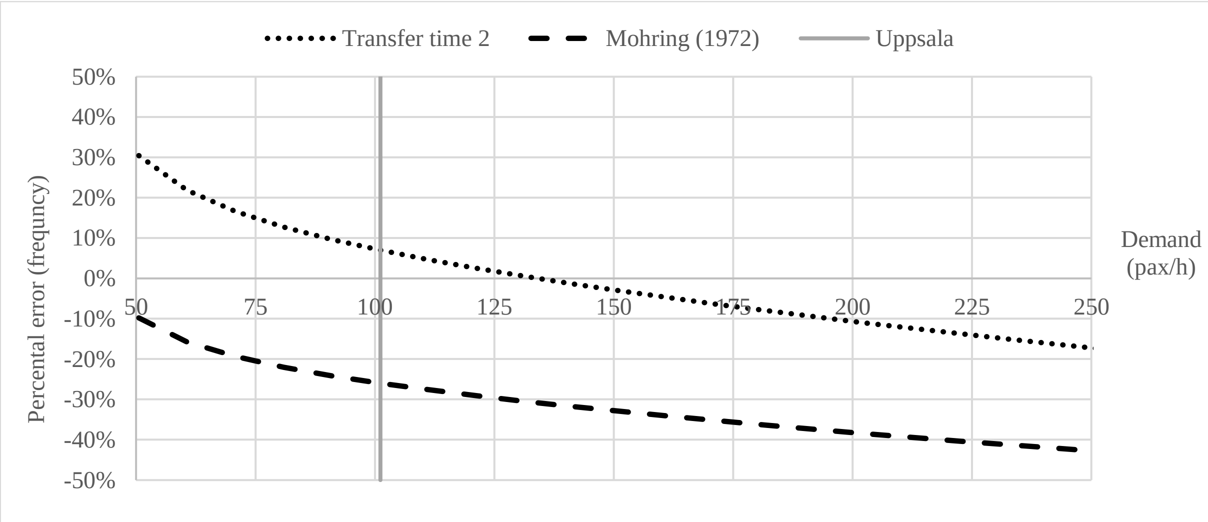

Fig. 2. Sensitivity with respect to PT demand per line (two-way), OP. The ratio of peak trips to OP trips is held constant, i.e., the hourly demand in peak hours is slightly below double the demand in OP hours in each simulation. Percental error refers to the difference in optimal frequency between each square-root rule version (*Transfer time 2* and *Mohring*) and the BUPOV model.

Fig. 2 only assesses the errors in terms of optimal frequencies from using simple formulas, but not the welfare costs, which are most relevant. In Fig. 3, such welfare costs are displayed (but for peak and OP hours jointly, since this is the structure for welfare calculations in BUPOV). Comparing the two

[14] Please note that the bus size is fixed at 38 seats across simulations, although this may not be the optimal bus size for settings with demand differing considerably from that of Uppsala.

figures, there is a striking difference in how well the simple formulas perform. *Transfer time 2* now once again seems to perform well within a certain range (within a demand range of 75–200 pax/h, the welfare loss is less than 4%). In contrast, the *Mohring* formula performs roughly as poorly in both figures. Fig. 1 shows the underlying fundamental relationships explaining these differences, although the numbers do not correspond. Close to the optimal demand, the total cost function is rather flat, i.e., minor deviations from the optimal frequency are not a large problem from a welfare perspective. However, the further away from the optimal frequency we get, the steeper this cost curve becomes, and the relationship is much more pronounced at lower-than-optimal frequencies (as is also the case at frequencies prescribed by the *Mohring* formula).

One additional comment can be made about differences between Figs. 2 and 3. One might think that because, in Fig. 2, the error was 0% at 130 pax/h for *Transfer time 2*, this would correspond to a welfare loss of zero at the same corresponding demand in Fig. 3. However, Fig. 3 includes two demand calculations for each point in the graph, i.e., peak and OP hours (as in Table 3), and it is impossible to simultaneously hit the “zero error” point for these two types of demand.

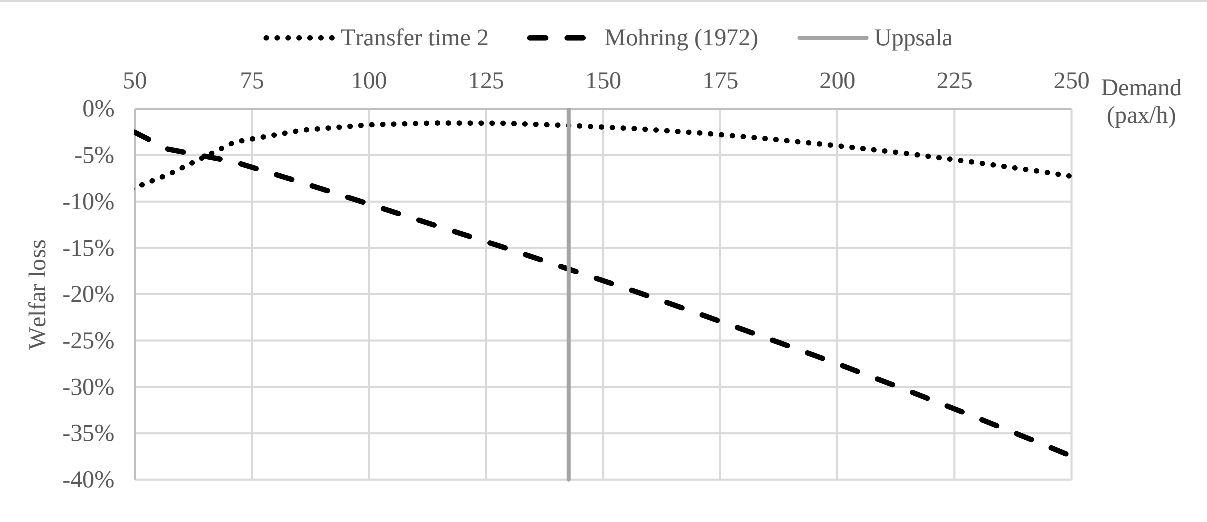

Fig. 3. Welfare loss (as percent of operational costs in BUPOV optimum) from using simple formulas (i.e., *Transfer time 2* and *Mohring*) for optimal frequency, as a function of PT demand per line (two-way, mean of peak, and OP).

Conclusion

This study has been dedicated to finding the welfare-optimal frequency of a scheduled PT service from a methodological perspective, exploring what methods are more suitable for the case of bus services in a small city. To this end, it has been examined how well the square-root rule estimates the optimal bus frequency in the case city of Uppsala, compared to the more comprehensive BUPOV-model (Asplund and Pyddoke, 2019), a partial equilibrium model with three travel modes and city specific data.

The results indicate that extending the square-root rule by including transfer time, as proposed here, is empirically important. Furthermore, the results indicate that the square-root rule, with proper extensions, can estimate optimal frequency in Uppsala surprisingly well. For Uppsala, the present results refute the claim of Furth and Wilson (1981) that wait time minimization accounts for only a minor part of the public benefit of transit service. In fact, it is shown that that a simple square-root rule that takes only operational costs and waiting times into account performs surprisingly well and results in a welfare loss corresponding to only 2% of the total operational costs, compared with a model[15] that takes all relevant factors (such as crowding and congestion) into account.

A sensitivity analysis of demand indicates that the performance of the proposed formula is not robust for optimal frequency, but still performs well for welfare over an interval of about 75–200 pax/h (two-way demand per line). However, one should recall that a sensitivity analysis is a simplified way to assess the generality, since only the PT demand was varied, and in reality PT demand is correlated with other relevant factors, such as car traffic demand and possibly also the number of transfers. Hence, additional similar studies of small cities around the world would ultimately be needed to assess the cases for which the proposed formula is suitable.

Compared with the proposed formula, the original Mohring formula performs considerably worse at demand levels over 75 pax/h. In future studies, it would be interesting to see whether the Mohring formula might be more suitable for low-demand inter-city lines.

Furthermore, the importance of the marginal cost of public funds and the wider economic benefits of optimal frequency have been tested. Applying both these aspects (i.e., marginal cost of public funds and wider economic benefits) remains somewhat controversial, and consensus as to whether or not they should be included in cost–benefit analysis has not emerged. Luckily, this study indicates that they are not critical for determining optimal service frequency, but in fact reduce the accuracy in Uppsala when nonlinear external effects such as crowding and congestion are not included. It has also been demonstrated that including boarding times is unnecessary when optimizing bus frequencies in a small city, as this has only a minor influence on the results.

One contribution of this study was instead to identify the importance of including transfer time in the square-root rule. The valuation of transfer time is important for determining optimal frequencies in terms of both analytical inclusion and empirically estimating the value of time. Relatively few studies examine the value of transfer time, and the ranges implied by these studies are typically vast. Hence, this study reinforces the recent conclusion of other authors (e.g., Hörcher and Tirachini, 2021) that more such studies would be desirable.

Acknowledgements

This study was funded by TRENoP—Transport Research Environment with Novel Perspectives. I am most grateful to: Professor Antonio Gschwender, University of Chile; Associate Professor Erik Jenelius, KTH Royal Institute of Technology; and Roger Pyddoke, Senior Researcher at the Swedish National Road and Transport Research Institute (VTI) for their valuable comments.

[15] One should of course recall that this more comprehensive model, BUPOV, here represents the best practice for estimating the optimal frequency in Uppsala; however, it is also only a model, so BUPOV estimates do not represent the “true” optimal frequency per se.

References

Anderstig, C., Berglund, S., and Börjesson, J. (2018). Regionalekonomiska effekter av planförslagen 018-2029: Beräkningar med Samlok-modellen, Version: 0.1, Trafikverket, Borlänge

Asplund D. and Pyddoke, R. (2021). Optimal pricing of car use in a small city: a case study of Uppsala. Transport Policy, ISSN 0967-070X, https://doi.org/10.1016/j.tranpol.2021.09.008.

Asplund D. and Pyddoke, R. (2020). Optimal fares and frequencies for bus services in a small city. Research in Transportation Economics, Vol 80, ISSN 0739-8859, https://doi.org/10.1016/j.retrec.2019.100796.

Basso, L.J., & Silva, H.E. (2014). Efficiency and substitutability of transit subsidies and other urban transport policies. *American Economic Journal: Economic Policy*, *6*(4): 1–33. http://dx.doi.org/10.1257/pol.6.4.1

Börjesson, M., Fung, C. M., Proost, S., Yan, Z. (2019). Do Small Cities Need More Public Transport Subsidies Than Big Cities? Journal of Transport Economics and Policy, Volume 53, Number 4, October 2019, pp. 275-298(24)

Eliasson, J., & Fosgerau, M. (2019). Cost–benefit analysis of transport improvements in the presence of spillovers, matching and an income tax, Economics of Transportation, 18, 1-9, ISSN 2212-0122, https://doi.org/10.1016/j.ecotra.2019.02.001

Espino, R., Román, C. (2020). Valuation of transfer for bus users: The case of Gran Canaria, Transportation Research Part A: Policy and Practice, Volume 137, Pages 131-144, ISSN 0965-8564, https://doi.org/10.1016/j.tra.2020.05.003.

Frankena, M.W., 1983. The efficiency of public transport objectives and subsidy formulas. Journal of Transport Economics and Policy 17, 67–76

Furth, P. G, Wilson, N. H. M. (1981). SETTING FREQUENCIES ON BUS ROUTES: THEORY AND PRACTICE. Transportation Research Record, Issue Number: 818, pp 1-7, ISSN: 0361-1981 https://trid.trb.org/view/174187

Hörcher, D. and Tirachini, A. (2021). A review of public transport economics. Economics of Transportation, 25

Jacobs, B., and de Mooij R.A. (2015). Pigou meets Mirrlees: On the irrelevance of tax distortions for the second-best Pigouvian tax, Journal of Environmental Economics and Management, 71, pp. 90–108.

Jansson, J. O. (1979). Marginal cost pricing of scheduled transport services: a development and generalisation of Turvey and Mohring's theory of optimal bus fares. Journal of Transport Economics and Policy, Vol. 13, No. 3, pp. 268-294 https://www.jstor.org/stable/20052548?seq=1

Janson, J. O. (1980). "A SIMPLE BUS LINE MODEL FOR THE OPTIMIZATION OF SERVICE FREQUENCY AND BUS SIZE" Transport system optimization and pricing, pp. 127- 162. A Dissertation for the Doctor's Degree in Economics Stockholm School of Economics 1980, ISBN 91-7258-121-2

Jara-Díaz, S., Gschwender, A. (2003). Towards a general microeconomic model for the operation of public transport, Transport Reviews, 23:4, 453-469, https://doi.org/10.1080/0144164032000048922

Jara-Díaz, S., and Gschwender, A. (2009). The effect of financial constraints on the optimal design of public transport services. Transportation, 36, 65–75

Jara-Díaz, S., Gschwender, A. & Bravo, C. Total cost minimizing transit route structures considering trips towards CBD and periphery. Transportation 45, 1701–1720 (2018). https://doi.org/10.1007/s11116-017-9777-z

Jara-Díaz, S., Gschwender, A., Ortega, M. (2012). Is public transport based on transfers optimal? A theoretical investigation, Transportation Research Part B: Methodological, Volume 46, Issue 7, Pages 808-816, ISSN 0191-2615. https://doi.org/10.1016/j.trb.2012.02.002

Kol-TRAST (2012). Kol-TRAST - Planeringshandbok för en attraktiv och effektiv kollektivtrafik, by Trafikverket and Sveriges kommuner och landsting, ISBN: 978-91-7164-842-6

Liu, X., et al. (2017). Impact of Different Bus Stop Designs on Bus Operating Time Components. Journal of Public Transportation, 20 (1): 104-118. DOI: https://doi.org/10.5038/2375-0901.20.1.6

Mohring, H. (1972). Optimization and Scale Economies in Urban Bus Transportation. The American Economic Review, 62(4), 591-604. http://www.jstor.org/stable/1806101

Newell, G.F. (1971). Dispatching Policies for a Transportation Route. Transportation Science, Vol. 5, No. 1, pp. 91-105. https://pubsonline.informs.org/doi/abs/10.1287/trsc.5.1.91?journalCode=trsc

Pyddoke, R. (2020). Analysis of daily variation in bus occupancy rates for city-buses in Uppsala and optimal supply, Working Papers No 2020:8, Swedish National Road & Transport Research Institute (VTI), https://swopec.hhs.se/vtiwps/abs/vtiwps2020_008.htm

Salzborn, F. J. M. (1972). Optimum Bus Scheduling. Transportation Science, Vol. 6, No. 2, pp. 137-148. https://pubsonline.informs.org/doi/abs/10.1287/trsc.6.2.137

Savage, I. (2004). Management objectives and the causes of mass transit deficits, Transportation Research Part A 38 181–199 doi:10.1016/j.tra.2003.10.003

Schakenbos, R., La Paix, L., Nijenstein, S., Geurs, K. T. (2016). Valuation of a transfer in a multimodal public transport trip, Transport Policy, Volume 46, 2016, Pages 72-81, ISSN 0967-070X, https://doi.org/10.1016/j.tranpol.2015.11.008.

Simon, D., Chen, S., Zhou, R., Zhou, Y., Mao, B. (2013). Computation on Bus Delay at Stops in Beijing through Statistical Analysis. Mathematical Problems in Engineering, Hindawi Publishing Corporation. Article ID 745370, https://doi.org/10.1155/2013/745370

Swedish Transport Administration, (2016). Analysmetod och samhällsekonomiska kalkylvärden för transportsektorn: ASEK 6.0.

Transek (1995). 1994 års tidsvärdesstudie, Slutrapport, Del 1: Resultat.

Transportation Research Board, & Kittelson and Associates, Incorporated. (2003). Transit Capacity and Quality of Service Manual, 2nd Edition. Transit Cooperative Research Program (TCRP) Report 100. published by Transportation Research Board.

Wardman, M., Chintakayala, V.P.K., de Jong, G., 2016. Values of travel time in Europe: review and meta-analysis. Transport. Res. Part A Policy Pract 94, 93–111. https://doi.org/10.1016/J.TRA.2016.08.019

Vickrey, W. (1955). Some Implications of Marginal Cost Pricing for Public Utilities. Amer. Econ. Rev. Proc., 45, 605-20. https://www.jstor.org/stable/1823588?seq=1

WSP (2010). Trafikanters värdering av tid – Den nationella tidsvärdesstudien 2007/08, WSP Analys & Strategi rapport 2010:11
https://www.trafikverket.se/contentassets/f250787d665a41f6ad73f76c95b70c9a/rapport_2010_11_trafikanters_vardering_av_tid.pdf

Appendix A: The shadow prices of waiting time

Sweden has a well-developed set of national guidelines for the cost–benefit analysis (CBA) of infrastructure investments, called ASEK (Swedish Transport Administration, 2016), based on national and international empirical research. Asplund and Pyddoke (2020) selected the relevant waiting time values for bus traffic in Uppsala from ASEK (ver. 6.0) to be EUR 4.16/h for journey departure/arrival time displacements (i.e., passive waiting time) and EUR 11.0/h for transfer waiting time (i.e., active waiting time), based on travel time multipliers from underlying empirical studies from Sweden.

The first estimate (i.e., normal waiting time) was based on a travel time multiplier estimated by WSP (2010) based on survey data from 2007–2008. In this study, the relevant travel time multiplier (for in-vehicle time), i.e., departure/arrival time displacements for local/regional trips with service headway intervals of 11–30 min, was estimated to be 0.94. The travel time multiplier for departure time displacements was estimated to be 0.94 with headway intervals of 11–30 min, based on 100 interviews (of 227 contacted) conducted in 2007–2008. This value seems largely in line with the average value of 0.78 (in urban settings) from 225 observations (99 studies) in a meta-study of Europe by Wardman et al. (2016).

The travel time multiplier for transfer time in ASEK 6 was chosen to be 2.5, based on a study by Transek (1995). Transek (1995) estimated the travel time multiplier for transfer time to be 1.4–2.5, and ASEK 6 chose the higher end of the interval to also account for the cost of the transfer itself, which is not relevant to this study. The 2.5 figure is somewhat high compared with the average of 1.8 across nine studies examined by Wardman et al. (2016). However, travel time multipliers can be estimated from Fig. 6 in Schakenbos et al. (2016) to range from 1.7 to 3 (across modes) for excess transfer time. Espino and Román (2020) estimated travel time multipliers for transfer time for bus users in Gran Canaria to be 1.4–5.6. Hence, the ranges are typically large and the ASEK value seems largely in line with these three studies. However, a sensitivity analysis of waiting time multipliers is performed in Table 4, using the Warman et al. (2016) estimates instead of the ASEK values.

Please note that the possibility to enjoy passive waiting time at home for the first bus of each trip (rather than active waiting time at the stop) is dependent the possibilities for passengers to know the bus departure times at their stop in advance, which depends on the regularity of the service and whether there is a publicly available schedule. In Sweden, these conditions are mostly met, but, for example, in Chile, until recently buses have been driven on an ad hoc basis without published schedules, and this is true of many other developing countries, for example, in Latin America (according to personal communication with public transport researcher and professor Antonio Gschwender). Of course, if the possibility to spend the waiting time at home does not exist, the higher value of active waiting time (i.e., EUR 11.0/h in our case) should be used not only for transfer waiting time but also for normal waiting time (i.e., for the first bus in each trip). However, new apps have emerged that can help follow vehicles in real time, making it possible to wait at home even when schedules are not known in advance.

Appendix B: Non-linearity of external effects from congestion

To illustrate the non-linearity of the external effects of congestion due to changes in public transport (PT) frequencies, the external effects on congestion of a 1% change in peak frequency are estimated for two cases. The first case refers to the baseline situation, the actual frequencies in Uppsala in 2016, i.e., six buses per hour in peak and four buses per hour in off-peak (OP) hours. This is compared with a case using the lowest optimal frequencies estimated in Table 3, i.e., 1.78 in peak and 1.79 in OP hours. In the first case, with considerably higher frequencies than optimal (according to the BUPOV model), a 1% increase in peak frequency hardly affects the car traffic, while more buses on the streets imply more congestion. Hence, the net result is an increase in delay due to congestion from 88.96% to 89.04% (and there is also spillover to the OP), implying a welfare loss corresponding to 14% of the change in consumer surplus from decreased waiting times from a 1% increase PT frequency. However, in the second case, with frequencies that are too low compared with what is prescribed by the BUPOV model, the demand effect is much larger, outweighing the negative direct congestion effect of more buses on the streets. In this case, a 1% increase in peak frequency results in a reduction in delay from 90.00% to 89.92%, implying a welfare gain corresponding to 6% of the change in consumer surplus from a 1% increase in PT frequency. From this it is clear that these congestion effects cannot be incorporated into the square-root rule as proportional to user benefits. It may, however, be possible to find a more complex formulation, in which the congestion benefits of reduced car traffic are proportional to the user benefits and congestion costs of more buses and to frequency for a sufficiently small interval. This is, however, outside the scope of this analysis.