

Optimal supply of public transport – subsidising production or consumption or both?

Jan-Eric Nilsson, Joakim Ahlberg and Roger Pyddoke – VTI

CTS Working Paper 2014:27

Abstract

This paper compares two approaches for providing public transport: competitive tendering and vouchers. The functioning of tendering is well known. The voucher alternative means that commercial operators charge passengers a fare and are also paid by the public sector principal for each. The paper establishes that it is also necessary to subsidize quality, here referring to the number of busses. Under full information, the two approaches would result in the same welfare maximizing outcome. Vouchers may be more expensive for taxpayers. The case for vouchers is shown to be stronger under an asymmetric information framework.

Keywords: Tendering, vouchers, Mohring effect, public transport, (a)symmetric information

Optimal supply of public transport – subsidising production or consumption or both?

July 2014

J-E. Nilsson*¹, jan-eric.nilsson@vti.se

J. Ahlberg Joakim.ahlberg@vti.se

R. Pyddoke, Roger.pyddoke@vti.se

Swedish National Road and Transport Research Institute (VTI) www.vti.se

and

Centre for Transport Studies, Stockholm <http://www.kth.se/en/abe/centra/cts/>

Abstract: This paper compares two approaches for providing public transport: competitive tendering and vouchers. The functioning of tendering is well known. The voucher alternative means that commercial operators charge passengers a fare and are also paid by the public sector principal for each. The paper establishes that it is also necessary to subsidize quality, here referring to the number of busses. Under full information, the two approaches would result in the same welfare maximizing outcome. Vouchers may be more expensive for taxpayers. The case for vouchers is shown to be stronger under an asymmetric information framework.

Key words: Tendering, vouchers, Mohring effect, public transport, (a)symmetric information.

¹ Corresponding author.

1. Introduction

In many countries, a range of institutions are in place for ascertaining that the supply of public transport services is higher than it would be under a commercial regime. Competitive tendering, which has been demonstrated to deliver cost savings at implementation, is one such means. The subsequent increase of costs in combination with deteriorating performance generally seems to be less favourable for tendering in a longer perspective (Hensher & Wallis 2005). This paper considers an alternative to tendering for implementing a welfare maximising level of public transport. The idea is to combine a market with free entry with a universal subsidy scheme. Operators meeting minimum fit-willing-and able entry criteria would be free to design both service supply and fares and would in addition receive financial support from the public sector. Although no physical evidence of eligibility in the shape of a piece of paper is issued and transferred, this will be referred to as a voucher approach.

The purpose of the paper is to identify whether tendering or vouchers is the best means for implementing a welfare maximising policy, first in situations with full information and second when information is asymmetric and even incomplete. Our basic finding is that the voucher approach requires subsidies related to two parameters, the number of passengers and the number of departures (buses), in order to ascertain efficiency. It is furthermore demonstrated that the taxpayer costs for the respective systems may, but need not, be higher under the voucher approach than under a full information framework. Under asymmetric information, the operators' information advantage provides motives that tilt the scale to the benefit of a voucher approach.

Some background, including a review of relevant literature, in section 2, introduces the issues in focus for our interest. Sections 3 and 4 compare the policies under a full and asymmetric

information framework, respectively, and the paper is concluded in section 5. Throughout the analysis, only the quality aspect referred to as the Mohring effect is considered as a motive for intervention, i.e. second best motives are ignored.

2. Background

What has come to be known as the Mohring effect is an example of a positive externality: Each additional passenger adds to the socially optimal service supply. Increasing supply by way of more frequent departures would benefit not only the marginal passenger but also all those using the system at the outset; this is what generates the positive externality and the potential market failure, and section 2.1 reviews some aspects of an extensive literature related to the Mohring effect.

Besides public transport, many other social services generate positive externalities. In Sweden, and introduced by governments of both centre-right and centre-left majorities, several of these services are provided by public sector bodies as well as by commercial enterprises on behalf of the tax payers. One example is schools. Most schools are still run by local communities, but parents and children may now choose between schools provided by municipalities and privately provided schools. The commercial provider of schooling is remunerated with the same amount per pupil, as is the provider of municipal services.

The same approach is used for primary health care. Commercial as well as public sector providers charge each patient a nominal amount on arrival and are then compensated according to a system based on the number of people that register with each health centre in combination with the number of visits. In addition, one major hospital has been tendered to an

international provider of hospital services, which then operates on the same basis as public sector hospitals.²

While public transport also generates positive externalities, it is provided today by way of competitive tendering. Section 2.2 describes some aspects of the way in which tendering is organised in order to motivate our interest in a voucher system for bus services as well.

2.1 The literature

In 1975, Michael Spence made an important contribution to the understanding of the detrimental effects of monopolies for social welfare. His attention was directed to both price and quality as decision variables, and to markets where quality is an intrinsic characteristic of the goods or services provided. Spence's core observation is that a market failure problem may materialise not only since the price is set above marginal costs, but also since a single price is not able to convey information about service quality to infra-marginal consumers. He then demonstrates that the monopolist's quality in general may be both higher and lower than the welfare maximising level.

A couple of years before, Herbert Mohring (1973), using an analytically similar approach, made the observation that a profit maximising operator of public transport does not fully account for quality. Quality – in this application the frequency of bus services – may therefore be underprovided. This is driven by the fact that the monopolist's focus when designing frequency is on the possibility of attracting new passengers, while social concerns include everyone using a service, also those that are already on board. For the service to be financially viable, a subsidy is required which is equal in value "... to the stock of consumer-supplied waiting time." (Ibid, p. 593.) The logic is that the average waiting time at stops would be

² Cf. The Economist May 18th 2013, "A hospital case".

reduced if an operator adapted (increased) supply to the demand of all travellers. This provides the basis for the belief that public transport is plagued by market failure.³

van Reeve (2008) questions the Mohring effect as a motive for subsidizing public transport. He argues that since service frequency affects willingness to pay for public transport, a monopolist will take this into account when setting prices. As a result, and in the same way as noted by Spence (1975), frequency provided by the profit maximizing monopolist may under certain circumstances *exceed* the social optimum rather than supply too little quality. van Reeve then concludes that since most public transport is provided by a monopolist, subsidies are not warranted, at least not based on the Mohring effect.

Both Savage and Small (2010) and Basso and Jara-Diaz (2010) argue that the van Reeve result depends on the demand assumptions made. With a slight modification of van Reeve's (2008) model, and introducing heterogeneity in users' reservation price, these authors again show that frequency provision by a monopolist operator is socially suboptimal, justifying subsidies. Karamychev and van Reeve (2010) generalize the demand model of Basso and Jara-Diaz (2010) and show that as users' heterogeneity shrinks, at some point a monopolist will provide socially optimal frequencies and beyond that point will over-supply frequencies compared to the social optimum. Gomez-Lobo (2013) demonstrates that the results of these four papers are special cases of Spence (1975).

Gómez-Lobo (2007) provides a different perspective on the functioning of the public transport market. He shows that due to the passengers' waiting cost, a bus operator in a competitive market faces an inelastic demand around the price charged by other buses; the

³ Parry & Small (2009) estimate the relevance of this and other motives for subsidising public transport in order to ascertain a welfare maximising policy.

extra waiting time at home or at the bus stop is not worth the lower fare charged by the next bus. The incentive to compete in price is therefore reduced.

Using a labour market search model to assess the behaviour of travellers with waiting costs, and assuming random arrival probabilities of vehicles at bus stops – which is the standard assumption in congested cities – he demonstrates that several price equilibria are feasible. Each equilibrium depends on the characteristics of bus frequencies and the chance of several buses arriving at a stop at the same time. In all these cases, the equilibrium fare is, however, above marginal costs (or average costs when there are fixed costs). The conclusion is that it is *not* reasonable to expect tough price competition among operators in competitive urban bus markets. This then explains the empirical observations of prices increasing in these markets after liberalisation. In this model, prices above marginal and average costs will induce socially excessive entry since fixed costs are needlessly duplicated.

Although vouchers have not been used in the market for public transport, Norway has reported the results of elaborate schemes for strengthening the incentives for cost savings and in particular revenue increases for operators; cf. Longva et al (2003). Andersson and Pyddoke (2010) test the idea that a subsidy per passenger could be a possibility for urban transport in a medium sized Swedish town where public transport is supplied by a single bus operator. To do so they simulate demand for a number of price and frequency combinations.

2.2 The provision of public transport in Sweden

Public transport is a system of vehicles such as buses and trains that operate at regular times on fixed routes and may be used by anyone. Focus here is on public transport that would not be provided without taxpayer subsidies. In Sweden, this also includes parts of the railway system, in particular commuter trains in major conurbations. The largest volume of subsidized services is, provided by bus transport, primarily in cities but also in rural areas.

Since the early 1980s, local and regional bus services have been tendered on a competitive basis by regional authorities. Alexanderson et al (1998) show that the reforms initially resulted in cost savings, but the costs of providing tendered services have subsequently increased much faster than consumer prices at large. Although fares have also increased in real terms, this is not sufficient to cover the cost increase. As a result, taxpayers' share of the bill has increased and now covers more than half the costs of these services. In spite of increasing subsidies, the growth of patronage is slow and substantially lower than the growth of car use.

Gross cost contracts between the public sector principal and the commercial service provider have been the standard for a long time. A core feature of these contracts is that the principal specifies fares, and that route and stops, departure frequency, technical vehicle requirements etc. are meticulously specified in the quote for bids. The bidders' main degree of freedom lies in the roster schemes for staff and rolling stock. Some of the contracts are supplemented with incentives linked to increased patronage, timeliness of buses, etc. In the absence of these additions, an operator benefits from having fewer passengers. The combination of a rigid approach to contracting and rapidly increasing costs of providing the services is one reason for considering alternatives to tendering.

The market for public transport, both by bus and rail, was opened for entry in January 2012. Any fit, willing and able operator is now entitled to provide services on a commercial basis. Little happened the first year after the market was opened. An important reason is that a commercial entrant would have to compete with the existing service provider who, on average, charges customers fifty percent of the costs of running buses or trains. There are a few examples where an entrant has been able to run a viable service which complements the existing supply, but no example of the incumbent being challenged.

3. Modeling the choice; full information

The point of departure for the present analysis is that a regional public sector body – the principal – has decided that a bus service is to be operated. Eq. (1) formalizes the objective function for maximization of social welfare (W). Here, the fare or the price is represented by (p) and the demand for trips (x) per unit of time depends on price and service quality; $D(x)=x(p,b)$. Quality is approximated by the number of vehicle kilometers or equivalently the number of buses (b). The inverted demand function $p = p(x, b)$ is assumed to be concave in the number of trips (given the number of buses) and in the same way concave in the number of buses (for a given number of trips). The cost function for producing x trips of quality b , $c(x, b)$, is assumed to be convex in each variable, given the other variable.

Equations (2a) and (2b) give the two first order conditions. The first states the familiar equality between price and marginal cost, given the number of buses. Equation (2b) indicates that the value of the marginal departure or the marginal bus, integrated over all users, should correspond to the cost of providing this marginal service, given the number of trips. The optimal policy is thus characterized by $\{p^w, b^w\}$ with the implied cost c^w for providing the services. The financial outcome of this solution is related to the shape of the cost function and, in a situation with perfectly divisible production costs, service provision break even.

$$\max W(x, b) = \int_0^x p(v, b)dv - c(x, b) \quad (1)$$

$$\frac{\partial W}{\partial x} = p(x, b) - \frac{\partial c}{\partial x}(x, b) = 0 \quad p = \frac{\partial c}{\partial x}(x, b) \quad (2a)$$

$$\frac{\partial W}{\partial b} = \int_0^{x^w} \frac{\partial p}{\partial b}(v, b)dv - \frac{\partial c}{\partial b}(x, b) = 0 \quad (2b)$$

The purpose is now to model the consequences of implementing a welfare maximising policy using tendering and vouchers under a complete information framework. Figure 1 illustrates

the process from having established the social need for public transport to service provision. The outcome of the tendering solution is characterised in section 3.1 and the voucher approach in section 3.2. Section 3.3 summarises the discussion this far and section 3.4 reflects over the way in which the oligopoly outcome is modelled.

<FIGURE 1 ABOUT HERE>

3.1 Tendering

Procurement is assumed to be for a gross cost contract, a basic characteristic being that – based on an assessment of demand and costs – the principal specifies both fares and precisely which traffic is to be operated. To focus thinking, it is assumed that only two inputs or tasks, denoted b_1 and b_2 , are involved in providing the service.⁴ As an example, b_1 may be the number of large buses while b_2 is the use of smaller vehicles. The composition of the vehicle fleet is the outcome of an optimisation which, apart from the vehicle price, inter alia comprises fuel prices, the number of drivers required and the size of the market(s) to be serviced.⁵

For the tendered approach, and having full information about these costs, the principal specifies the input in the quote for bids. Denote these quantities \hat{b}_1 and \hat{b}_2 , the hat indicating that quantities are set by the principal. In the present section it is assumed that implementing activities $\{\hat{b}_1, \hat{b}_2\}$ with certainty yields the desired output. The quote also includes information about the (optimal) fare to be charged.

Each bidder i submits bid ρ^i and the contract is awarded to the supplier offering to operate the service at lowest cost. If the competitive pressure is sufficient, $\rho^i = c^w$ and the optimal policy

⁴ This characterization of the tendering model is based on Mandell et al (2013),

⁵ Lidestam & Abrahamsson (2010) and Lidestam (2014) provide examples of the implications for both costs and emissions from the composition of the vehicle fleet.

will be implemented. The costs of this alternative for society are therefore given by $p^w x^w - c(x^w, b^w)$. Since competitive tendering simply implements the welfare maximizing solution, the financial outcome – i.e. the need for subsidies – is given by the extent of scale economies in traffic production.

3.2 Vouchers

Implementation by way of vouchers is less straightforward since the approach must be designed for a market where the unregulated operator(s) is/are *induced* to behave in a way which implements a welfare maximising policy. As a point of departure, it is therefore necessary to specify the choices of a profit maximizing operator. To this end, eq. (3) defines profits (π), and the marginal conditions if an operator is alone in the market are given by eq. (4a) and (4b).

$$\max \pi(x, b) = p(x, b) x - c(x, b) \quad (3)$$

$$\frac{\partial \pi}{\partial x} = p(x, b) + \frac{\partial p}{\partial x}(x, b) x - \frac{\partial c}{\partial x}(x, b) = 0 \quad (4a)$$

$$\frac{\partial \pi}{\partial b} = \frac{\partial p}{\partial b}(x, b) x - \frac{\partial c}{\partial b}(x, b) = 0 \quad (4b)$$

The first condition is the standard monopolist solution where the service is priced above marginal cost. Equation (4b) indicates that the monopolist will provide frequency in order to relate the income from operating an additional bus to the cost for providing bus services, which establishes the profit maximizing number of buses for a given number of trips.

In the analysis of vouchers, two market situations must be considered. In one, the market collapses and only one operator remains; this seems to be what happened in most English cities outside London after the 1982 market opening (section 3.2.1). The other situation appears if there still is competition in this market, here modelled as a duopoly (3.2.2).

3.2.1 Voucher and monopoly

Assume that there is only one operator in the market. In order to implement a policy resulting in $\{p^w, b^w\}$ the government offers two subsidies, $s(x) = s_m^x$ and $s(b) = s_m^b$ targeting the number of travelers and the number of buses, respectively; sub-index m represents the optimal subsidy under monopoly. The operator's objective function now takes the form indicated by eq. (5) and maximizing profit with respect to x and b ; the two marginal conditions for the profit maximizing operator are given by eq. (6a) and (6b).

$$\pi(x, b) = p(x, b)x + s_m^x x + s_m^b b - c(x, b) \quad (5)$$

$$\frac{\partial \pi}{\partial x} = p + \frac{\partial p}{\partial x} x + s_m^x - \frac{\partial c}{\partial x} = 0 \quad (6a)$$

$$\frac{\partial \pi}{\partial b} = \frac{\partial p}{\partial b} x + s_m^b - \frac{\partial c}{\partial b} = 0 \quad (6b)$$

By combining these conditions with the marginal conditions from the social welfare function (2a and 2b), the monopolist can be induced to behave in a way which corresponds with welfare maximization. Solving for s_m^x and s_m^b results in the following specification of the subsidies; $\varepsilon_x \equiv \frac{\partial x}{\partial p} \frac{p}{x} < 0$ is the price elasticity

$$s_m^x = - \frac{\partial p}{\partial x} x^w = - \frac{p^w}{\varepsilon_x} \quad (7a)$$

$$s_m^b = \int_0^{x^w} \frac{\partial p}{\partial b} dv - \frac{\partial p}{\partial b} x^w \quad (7b)$$

The first condition describes the subsidy necessary in order to make the operator charge welfare maximizing prices, given the number of buses. The optimal passenger subsidy is lower, equal to or higher than the welfare maximizing price depending on the elasticity, i.e. the higher the elasticity of price the higher the subsidy.

$$s_m^x \begin{cases} > p^w & \text{if } \varepsilon_x > -1 \\ = p^w & \text{if } \varepsilon_x = -1 \\ < p^w & \text{if } \varepsilon_x < -1 \end{cases}$$

Figure 2 demonstrates a situation where the monopolist originally charges above marginal costs, resulting in demand below the efficient level. From the monopolist's perspective, a subsidy shifts the demand curve, but the figure only illustrates the new marginal revenue curve. As a result of the subsidy, the operator reduces the fare towards marginal costs, thus increasing patronage. The monopolist will receive $p^w x^w$ from the passengers, and taxpayers have to pay the monopolist $s_m^x x^w = -p^w x^w / \varepsilon_x$.

<FIGURE 2 ABOUT HERE>

The second condition can be rewritten as (7b') which demonstrates that the supply side subsidy corresponds to the difference in value between the average and the marginal traveler.⁶

$$s_m^b / x^w = \frac{\int_0^{x^w} \frac{\partial p}{\partial b} dv}{x^w} - \frac{\partial p}{\partial b} \quad (7b')$$

The important conclusion is that welfare maximization may require not only an intervention in order to correct for the pricing, but also a subsidy in order to induce optimal quality. In the public transport case this is necessary in order to affect the supply of services. The cost of this, $s_m^b b^w$, will be larger the bigger the difference in quality valuation between the average (first term in rhs. in 7b') and marginal (second term) passenger. This is summarized as proposition 1:

Proposition 1: Subsidies as a means for inducing one or more profit maximizing operators to implement a welfare maximizing public transport policy require support to be linked to both patronage and supply.

⁶ Spence (1975) suggests that there is no a priori reason to expect these to be equal and Gomez-Lobo (2011) shows that the recent controversy reviewed in section 2.1 above boils down to what is empirically known about these two values.

3.2.2 Voucher and duopoly

The subsidies for price and quality/number of vehicles in the previous section stem from a situation with only one active operator. Nevertheless, there is no general presumption of natural monopoly and, if the market is opened up, more than one operator may establish services. Suppose therefore that the principal is aware that two equal operators will provide services. In a Cournot equilibrium, each firm maximizes its profits given the quantity chosen by the other. For $i, j \in \{1, 2\}$ profits are now defined by eq. (8) for firm i , along with the marginal conditions in eq. (9a) and (9b), given the optimized quantity of the other operator $\{x_j, b_j\}$; subscript d here represents the duopoly outcome.

$$\max_{x_i, b_i} \pi_i = p(x_i + x_j, b_i + b_j)x_i + s_d^x x_i + s_d^b b_i - c(x_i, b_i) \quad (8)$$

$$\frac{\partial \pi_i}{\partial x_i} = p + \frac{\partial p}{\partial x_i} x_i + s_d^x - \frac{\partial c}{\partial x_i} = 0 \quad (9a)$$

$$\frac{\partial \pi_i}{\partial b_i} = \frac{\partial p}{\partial b_i} x_i + s_d^b - \frac{\partial c}{\partial b_i} = 0 \quad (9b)$$

For each x_j and b_j , let $R_i^x(x_j)$ and $R_i^b(b_j)$ denote firm i 's best response. Two pairs of quantity choices (x_i, x_j) , (b_i, b_j) constitute a Nash equilibrium for $i, j \in \{1, 2\}$ if and only if $x_i =$

$R_i^x(x_j)$ and $b_i = R_i^b(b_j)$ ⁷. Since $\frac{\partial p}{\partial x_i} = \frac{\partial p}{\partial x_j}$ and $\frac{\partial p}{\partial b_i} = \frac{\partial p}{\partial b_j}$, any Nash equilibrium must satisfy

the following conditions:

$$\frac{\partial \pi_i}{\partial x_i} = \frac{\partial p}{\partial x_i} \frac{x_i + x_j}{2} + p + s_d^x - \frac{\partial c}{\partial x_i} = 0$$

⁷ Note that quality is implicit in p and c in this formulation.

$$\frac{\partial \pi_i}{\partial b_i} = \frac{\partial p}{\partial b_i} \frac{x_i + x_j}{2} + s_d^b - \frac{\partial c}{\partial b_i} = 0$$

In the same way as in the monopoly case, the two marginal conditions must be equalized with the corresponding conditions for establishing a welfare maximizing output. Thus, for $x = x_i + x_j$ and $b = b_i + b_j$

$$s_d^x = -\frac{1}{2} \frac{\partial p}{\partial x} x^w = -\frac{p^w}{2 \varepsilon_x} \quad (10a)$$

$$s_d^b = \int_0^{x^w} \frac{\partial p}{\partial b} dv - \frac{1}{2} \frac{\partial p}{\partial b} x^w \quad (10b)$$

The two conditions are identical (in form) to those for the monopolist. Since there are now two operators in the market, the per-trip subsidy in eq. (10a) is half the size relative to a single-operator market.

The first term on the rhs. in eq. (10b) represents the surplus of the optimal number of passengers, x^w , which is exogenously given. The second term is obviously smaller than in the monopoly case (eq. 7b). The principal must therefore increase the support per bus to the duopolists compared to the situation with a single operator in the market. The logic of this cost increase goes back to the downward sloping demand curve. Compared to a monopoly, more people will use the buses. Since the marginal valuation of quality/frequency is lower the higher the consumption of the service, the two operators will each reduce supply, as they will both benefit from the opponent's extra supply. In the absence of a higher per bus subsidy, the supply/quality would be lower than in the single operator solution, where the monopolist will reap all benefits from supply changes.

In a world with complete information about demand and costs, travelers will face the same price and quality $\{p^w, b^w\}$ irrespective of whether a welfare maximizing policy is implemented by way of competitive tendering or by vouchers. Consumer surplus is

therefore $\int_{x^w}^{\infty} p(v, b^w) dv - p^w x^w$ for both outcomes. In the same way, social welfare as defined by eq. (1) above is identical. The only possible difference between tendering and vouchers therefore concerns the operator(s) profits and the principal's costs; this is addressed in the next section. Meanwhile, we have proposition 2.

Proposition 2: Under complete information about demand and costs and if there is no social cost of public funds, both tendering and vouchers will result in an efficient price and quality.

3.3 Complete information and social costs of public funds

There are several motives for being interested in the financial outcome of the arrangements. One reason is related to distributional concerns. In the present context there may be a concern over the fact that taxpayers have to foot the bill for subsidizing one or more commercial enterprises, even if this provides an outcome which is better for society at large than in a situation with no intervention. Another motive is associated with the tax wedge of general taxation, leading to distortions in the use of scarce resources in general. Irrespective of motive, the possibility of a social cost for public funds will affect the comparison of alternatives.

Table 1 summarizes the financial outcome for the principal and for the operator(s) of the different options. It is assumed that the oligopolists have similar cost functions, so that the market is split into two equal parts. Our interest here is less on operator profits but rather on the financial outcome for the principal, i.e. the difference between cells (2), (4) and (6).

To provide an understanding of the financial implications of the voucher approach, eq. (11a) and (11b) spell out the difference in costs for the principal between a tendered solution and the monopoly outcome and between tendering and duopoly, respectively;

Table 1: Financial outcome of tendering and vouchers under full information

	(Each) operator	Principal
Tender	0 (1)	$p^w x^w - c(x^w, b^w)$ (2)
Voucher; monopoly	$p^w x^w + s_m^x x^w + s_m^b b^w - c(x^w, b^w)$ (3)	$s_m^x x^w + s_m^b b^w$ (4)
Voucher; duopoly	$p^w \frac{x^w}{2} + s_d^x \frac{x^w}{2} + s_d^b \frac{b^w}{2} - \frac{c(x^w, b^w)}{2}$ (5)	$2 * [s_d^x \frac{x^w}{2} + s_d^b \frac{b^w}{2}]$ (6)

$$p^w x^w - c^w - (s_m^x x^w + s_m^b b^w) \quad (11a)$$

$$p^w x^w - c^w - (2 * [s_d^x \frac{x^w}{2} + s_d^b \frac{b^w}{2}]) \quad (11b)$$

The financial cost for the principal under tendering is decided by the extent of fixed costs in the provision of bus services: If there are no scale economies in the provision of bus transport, it will be costless for the principal to tender the service, meaning that the first difference cancels out. The term within brackets in eq. (11a) is positive, though, since the operator must be induced to charge the appropriate price and operate the right number of departures. Constant returns to scale and a monopoly or a duopoly outcome of the voucher solution are therefore more costly than tendering for the principal.

Proposition 3a: With constant returns to scale in the provision of bus transport, vouchers will be more expensive for the principal than tendering.

In the presence of scale economies the cost comparison between tendering and a monopoly in eq. (11a) can be made in two steps. The first is to compare the implications for traveler-related support, i.e. $p^w x^w - (s_m^x x^w) = p^w x^w \left(1 + \frac{1}{\varepsilon_x}\right)$. If $\varepsilon_x = -1$ the two approaches are equally costly; at any other value tendering is more expensive.

Second, the difference with respect to the cost of inducing the optimal number of buses is $s_m^b b^w - c(x^w, b^w)$. It is obvious that at optimum the general cost function is $c^w * b^w$ i.e. the welfare maximizing cost per bus (which is exogenously given) times the optimal number of buses. Using (7b) we then get $b^w * \left[\left(\int_0^x \frac{\partial p}{\partial b} dv - \frac{\partial p}{\partial b} x^w \right) - c^w \right]$. If the inner parentheses point to the difference between the average and the marginal passenger's valuation of the service being very large, the value of the whole expression may be positive. If this is so, the difference in the bus-related costs will also be positive. The costs of tendering would then exceed the costs of using vouchers as a means for implementing welfare maximizing services. If the difference in quality valuation between the average and the marginal passenger is small, the term for the bus subsidy will be negative, and if it is large enough it may shift the positive balance of the first term into a negative aggregate outcome.

Turning to the comparison in the duopoly case, the logic is similar. Tendering is now even more costly than vouchers if only the financial costs related to passengers in eq. (11b) are considered. Correspondingly, the chance that the voucher approach is more costly with respect to support to the use of buses is higher. This can be concluded in the following way:

Proposition 3b: In general it is not feasible to predict the financial consequences of the choice between tendering and vouchers under scale economies, but (1) the more price sensitive passenger demand is, the more expensive the tendering alternative will be; (2) the more important quality is to passengers, the more costly the voucher solution will be.

3.4 Choice of oligopoly model

The present analysis is based on a Cournot approach for duopoly analysis in which each firm chooses its output independently, and the market determines the price at which it is sold. It is well known that the outcome of this approach is radically different from a Bertrand model where the firms compete in price. When it comes to selecting a model to fit the market under

analysis, the Kreps and Scheinkman (1983) model provides some comfort. They study a duopoly in which firms play a two-stage game. In the first period, they simultaneously choose capacity levels and then choose price in the second period. The low-price firm sells up to the minimum of market demand and its capacity, while the high price firm may receive residual demand. That model demonstrates that firms will choose the Cournot level of production in the first period and the market clearing price in the second; the outcome is then equal to Cournot competition.

The present approach more or less implicitly has the two operators choose price and supply/quality at the same time. Gertner (1985) addresses this situation, assuming that one operator's choice is made without knowing what the rival's price and quantity will be, but he is unable to establish a pure strategy equilibrium for this model. He also demonstrates that firms earn zero expected profit if marginal cost is constant or decreasing, but that the market price still is above average cost.

If anything, this result demonstrates the significance of modeling the (expected) behavior in a deregulated market with vouchers in a realistic way. Before more is known about how operators actually will act, it is difficult to make any firm predictions. This is all the more obvious in view of the assumptions in the Gómez-Lobo (2007) model reviewed in section 2 above.

4. Implementation under asymmetric information

The results from section 3 assume that relevant information for implementing a welfare maximizing policy is available. Figure 1 pinpoints the information which is required in the respective situations. A quote for bids requires information about demand and costs in order to formulate the welfare maximizing values $\{p^w, b^w\}$. This also includes an understanding of

production costs, in particular with respect to the way in which services are to be provided, here represented as the mix of small and large buses, which is specified in the quote. Using vouchers, information about demand and costs is used to establish $\{s^x, s^b\}$ which in turn is announced well before services are initiated so that operators are aware of the conditions under which services will be operated.

Section 4.1 provides a simple static framework for elaborating on the implications of asymmetric information about these dimensions for the choice between tendering and vouchers. Section 4.2 then takes this a step further by addressing the way in which the administration of the respective approaches may generate differences with respect to social welfare.

4.1 Ex ante specification of preconditions for tendering and vouchers

The input specification for large and small vehicles $\{\hat{b}_1, \hat{b}_2\}$ has so far with certainty been assumed to yield the desired service output X_D , *i.e.* $X(\hat{b}_1, \hat{b}_2) = X_D$ which is required in order to satisfy (expected) demand. Since the mix of vehicle sizes could take many different forms, there may also be other combinations of b_1 and b_2 that could result in X_D . Using the textbook approach to production theory, the bold (convex) isoquant in Figure 3 ties together all combinations of b_1 and b_2 that yield a product with the X_D characteristics. The straight line is an isocost that ties together all combinations of b_1 and b_2 that result in the same total cost. The slope of the isocost is thus given by the relative prices of the respective types of vehicles.

FIGURE 3 ABOUT HERE

The principal's specification of b_1 and b_2 is based on an estimate of the relative prices of the two types of vehicles as well as the rate of substitution between them. Although the principal would like to tender the cost minimizing specification illustrated in Figure 3, one reason for a

different outcome is that prices are poorly known. If the principal knew the unit prices with certainty, there would indeed be little need for procurement. This means that the tasks specified in the contract, i.e. the number of buses of the respective type to use may be unnecessary costly compared to the optimal solution given the actual relative prices.⁸

The outcome may also deviate from an optimal solution due to lack of knowledge with respect to the combinations of tasks which yield the outcome X_D . The quote and the subsequent contract specifying $\{\hat{b}_1, \hat{b}_2\}$ will then be unnecessarily costly if there is another combination that achieves the same outcome at a lower cost.

The risk that the principal's ex ante estimate is not cost minimizing is likely to increase in the complexity of the assignment. The notion put forward in Figure 3 suggests that a contract may be considered less complex the less curved the isoquant is. In the extreme case with Leontief (L-shaped) isoquants, there is no substitutability at all between the different inputs, the cost minimizing combination being the same irrespective of the relative prices, since there is no alternative way to produce the service. To tie complexity to the curvature of the isoquant in this way suggests that a given change in relative prices results in a smaller change in the mix of input factors for less complex projects.

The setting with two inputs may be extended to a row vector \mathbf{B} of arbitrary size without changing the qualitative outcome of the discussion. The principal must make an even more complex estimate of relative prices and the values of the \mathbf{B} vector that are most appropriate to use in the quote. Consequently, in an extended model version as well, the principal may fail to reach a cost-effective outcome since the actual relative prices and/or technology may differ from the principal's estimate.

⁸ Information here refers to the relative prices of each input required to operate services. This includes not only the vehicles but also drivers etc. It is obviously more difficult for a third party, e.g. the principal, to observe or estimate these composite prices rather than the textbook version of input (bus) prices.

A project's multidimensionality suggests an additional layer of complexity. The size of a project may be measured in two dimensions. Implementing $\{\hat{b}_1, \hat{b}_2\}$ where the single values are large – indicating a large number of vehicles – specifies an assignment which is large, for instance with respect to the length of the network. While this does not necessarily add to complexity, the number of input components does. It will be more difficult for the principal to see through the particulars of a project that comprises many activities. It is therefore reasonable to expect that it is more difficult to specify the cost-minimizing $\hat{\mathbf{B}}$ vector the larger the number of activities is. One example could be the distinction between complex networks in larger conurbations compared to a rather straightforward hub-and-spoke system or indeed country-side, low-density services.

Under a voucher approach the implementation of X_D is fully in the hands of the respective operator(s). This means that the agent, who most likely has superior information regarding technology, should be better equipped than the principal to find a cost-minimizing approach resulting in X_D . In addition, there may be an entirely different input mix, $\{b_3, b_4\}$ delivering X_D at lower cost than $\{\hat{b}_1, \hat{b}_2\}$; this would then be an “innovative solution”.

It is also possible that the manager of a transit firm has better information than the principal about the price elasticity associated with different routes of the network at different periods for different types of fares. In particular, this asymmetry can be assumed to increase with the complexity of the network and with congestion. If this is so, the implementation by way of vouchers would result in more demand (more passengers) for the same cost.

To sum up, the outcomes from the two approaches may be identical since the agent may always mimic $\{\hat{b}_1, \hat{b}_2\}$ if given this opportunity under a voucher approach. Since the agent will choose the combination of tasks to conduct and is better informed about relative prices, technology and demand, the total costs will most likely be lower in this case than if the

services are provided under contract. In the wake of formal modeling, this is summarized as a conjecture:

Conjecture 1: Situations with information asymmetries increase the chances that a voucher policy for public transport will result in a higher welfare level than if competitive tendering is used. The more complex the welfare maximizing service configuration, the stronger the benefits of using vouchers.

4.2 Adaptation costs

Contracts for tendered services are signed for a certain period of time; in Sweden typically for seven years. In order to shelter operators from swings in the prerequisites for the activities, a voucher approach would also have to be based on a commitment from the principal to the initial quote of subsidies $\{s^x, s^b\}$ for a period of time. For both implementation mechanisms it is therefore reasonable to assume a similar technique for adaptation of payments to general changes of external conditions.

To illustrate the consequences of this, assume that the principal realizes after one year that demand differs from what was anticipated when the call for bids was specified under tendering or when subsidies were publicized under the voucher approach. To be specific, $p(x, b) \equiv \hat{p} \neq \check{p}$ where the hat represents demand according to the principal's initial expectations, and the inverted hat is realized demand.

One way to handle this under tendering is to ex ante specify a quote for bids in a way which provides some scope for uncertainty. The estimated number of departures or of route kilometers, which is derived from estimates of demand and costs, may for instance be defined for an interval relative to the best guess. The quote could also include a request for "change prices", i.e. the required remuneration for increases or reductions in service frequency. If this

type of uncertainty is not made part of the contract, or if the deviations from \hat{p} are substantial compared to the initial specification, adjustments of the original terms may require a bilateral renegotiation. This provides the operator with an advantage that would make it feasible to secure monopoly rents. Alternatively, the principal could handle the situation by cancelling the contract and opening up a new bidding round. Apart from the risk that the principal may have to pay damages for prematurely breaking the contract, it would also cause administrative costs.

Under a voucher situation, subsidies $\{s^x, s^b\}$ are also based on an estimate of (inter alia) \hat{p} . One or more operator(s) running services on a commercial platform would, however, automatically adjust supply to realized demand, irrespective of whether the difference between \hat{p} and \check{p} is large or not. In parallel with the need to renegotiate cost reimbursement under a tendered alternative, there may be reason to adjust subsidies when deviations from ex ante specifications are seen to be substantial.

These observations are primarily related to problems associated with having appropriate ex ante information about demand and costs. This logic bears over also to exogenous shocks affecting crucial parameters for the activities. One example could be if the market grows by the establishment of a new residential area. A voucher solution would not require any pre-planning to handle this change. This can be seen from formulating the standard demand function, $x = x\{p, b|X\}$, X representing the size of the market measured as the number of inhabitants in a city etc. Doubling the market would not affect the value of $\{s^x, s^b\}$, given that the larger population have similar preferences to the original travelers. On the other hand, if a tendered contract were not based on the possibility of a larger market, the principal would have to call for a new contest.

A radical change in fuel prices may increase costs for operating buses as well as demand for trips to the extent that car users are induced to switch to public transport. Indexation is often used in tendered contracts in order to shield the operator from cost risk.⁹ Substantial increases in the number of travelers may, however, require re-negotiation of contracts based on the same arguments as above. Demand changes could be readily accommodated in a voucher approach and the subsidy could at least in principle be linked to the pricing-relevant marginal cost. We summarize these observations in the following way:

Conjecture 2: In situations with substantial changes in external preconditions, the costs of implementing a welfare maximizing policy for public transport will not become higher under vouchers than under competitive tendering.

5. Conclusions

Gómez-Lobo (2007) provides examples showing that completely unregulated markets for public transport may result in rather different outcomes. This includes elements such as lack of price competition, excessive entry and inefficient frequency decisions, safety hazards, atomized and informal ownership structure and conflict between operators. On the other hand, in many developed-world countries, public transport has been strictly regulated, or bus services have been provided by public authorities on their own account. In order to reduce subsidies and improve performance, England deregulated its market for local bus services outside London in 1986. The outcome of free entry has been that just one operator remains in most major conurbations, and that price is higher than before the market opening.

⁹ It is certainly a double-edged sword in that it eliminates incentives for adaptation to new relative prices.

The situation in Sweden's public transport made a more recent start based on a 1988 Transport Policy Act that *inter alia* established a coherent policy of competitive tendering. While costs to the public sector initially fell, the last 10 – 15 years have seen costs balloon, and Holmgren (2013) shows that average industry efficiency has crumbled during this period.

While there most certainly are several reasons for this development, little is known about the background to the deteriorating performance due to poor data. The rigidity of the standard format of tenders, in particular the use of inflexible gross cost contracts, has prompted the research in this paper. The question is thus whether a more decentralized approach to public-sector support of bus traffic is feasible. Following experiences from other parts of Sweden's public sector, we have sought to explore the pros and cons of vouchers relative to competitive tendering as a means for implementing welfare-enhancing transport.

Mostly out of linguistic convenience, the “voucher” concept is used even though no physical delivery of a voucher, a piece of paper that documents the holder's eligibility for the service, is used. Against the background of the ultimate purpose of a subsidy system for public transport, all potential travelers are eligible without documentation. Nonetheless, the system would have to be rigged in a way which makes it feasible to document ridership in order to reduce the risk of fraud. The current rapid development of electronic ticketing systems provides an obvious solution. The possibility of receiving a subsidy would, thus, have to be conditioned on the operator using a ticketing system that makes it feasible to document ridership in a way that is reasonably robust against deception. Implementing a sophisticated ticketing system could also be used to condition subsidies so that a ticket purchased from one operator is also to be valid for use on a competitor's service. This would reduce the risk for market fragmentation.

The formal modeling has demonstrated that vouchers would have to be linked to both the number of passengers and the number of vehicles in use by operators. While the logic of the passenger subsidy is straightforward, the subsidy for the number of vehicles is a means for reducing the risk that quality may deteriorate in a market where it is not feasible to charge different users for different degrees of quality.

The analysis provides an example of an economist's standard result; the voucher approach may, but need not, be more expensive for tax payers; this may, however, be balanced by the larger flexibility of commercial enterprises to adapt supply to market demand. Rephrasing this in a more constructive way, there is nothing in the analysis pointing to the fact that vouchers, which would be novel to this market, are hazardous for the performance of the market or indeed for taxpayers. The obvious policy recommendation is therefore to proceed by developing the model for a trial solution on the desktop and, if numbers turn out to be favorable, to try out vouchers in a real application.

References

- Alexandersson, G., S. Hultén & S. Fölster (1998). The Effects of Competition in Swedish Local Bus Services. *Journal of Transport Economics and Policy*, Vol. 32, Part 2, pp. 203-219.
- Gertner, R. (1985). Simultaneous Move Price-Quantity Gamed and Non-Market Clearing Equilibrium. Mimeo, Massachusetts Institute of Technology.

Gómez-Lobo, A. (2007). Why Competition Does Not Work in Urban Bus Markets: Some New Wheels for Some Old Ideas. *Journal of Transport Economics and Policy*, Vol. 41, Part 2, pp. 283-308.

Gómez-Lobo, A. (2013). Monopoly, Subsidies and the Mohring Effect: A synthesis and an extension. Universidad de Chile, SDT 336, www.econ.uchile.cl/SDT

Hensher, D. A. and I. P. Wallis (2005). Competitive Tendering as a Contracting Mechanism for Subsidising Transport. The bus experience. *Journal of Transport Economics and Policy*, Vol. 39, Part 3, September, pp. 295-321.

Holmgren, J. (2013). The efficiency of public transport operations. An evaluation using stochastic frontier analysis. *Research in Transportation Economics* 39 (2013), pp. 50-57.

Kreps, D., J. Scheinkman (1983). Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes. *Bell Journal of Economics* 14; 326-337.

Lidestam, H. (2014). Sustainable bus transports through less detailed contracts, *Renewable Energy*, 61 (2014) 141—146.

Lidestam, H. and M. Abrahamsson (2010). Mathematical modeling for evaluation of public procurement for bus transports in terms of emissions. *Management of Environmental Quality*, 21(5) (2010) 645-658.

Longva, F., J-T. Bekken, B. Norheim (2003). Nye avtaleformer for kjøp av kollektivtransport i Telemark. Oslo. TØI-rapport 676/2003.

Mandell, S., J-E. Nilsson, V. Liss (2013). Asymmetric Information and the Choice of Contract Design. VTI Working Paper

- Mohring, H. (1972). Optimization and Scale Economies in Urban Bus Transportation. *American Economic Review*, Vol. 62, No. 4 (Sept), pp. 591-604.
- Nilsson, J-E. & L. Jonsson (2011). Lessons from the tendering of railway services in Sweden. Are some contracts better than others? *International Journal of Transport Economics*, Vol. XXXVIII, No. 1, February 2011
- Parry, Ian W. H. & Kenneth A. Small, 2009. Should Urban Transit Subsidies Be Reduced? *American Economic Review*, vol. 99(3), pages 700-724, June.
- Pyddoke, R., M. Andersson (2010). Increased patronage for urban bus transport with net-cost contracts. VTI-rapport 2010:8.
- Spence, A. M. (1975). Monopoly, Quality, and Regulation. *Bell Journal of Economics*, 417-429.
- van Reeve, P. (2008). Subsidization of Urban Public Transport and the Mohring Effect. *Journal of Transport Economics and Policy*, 42, 349-59.

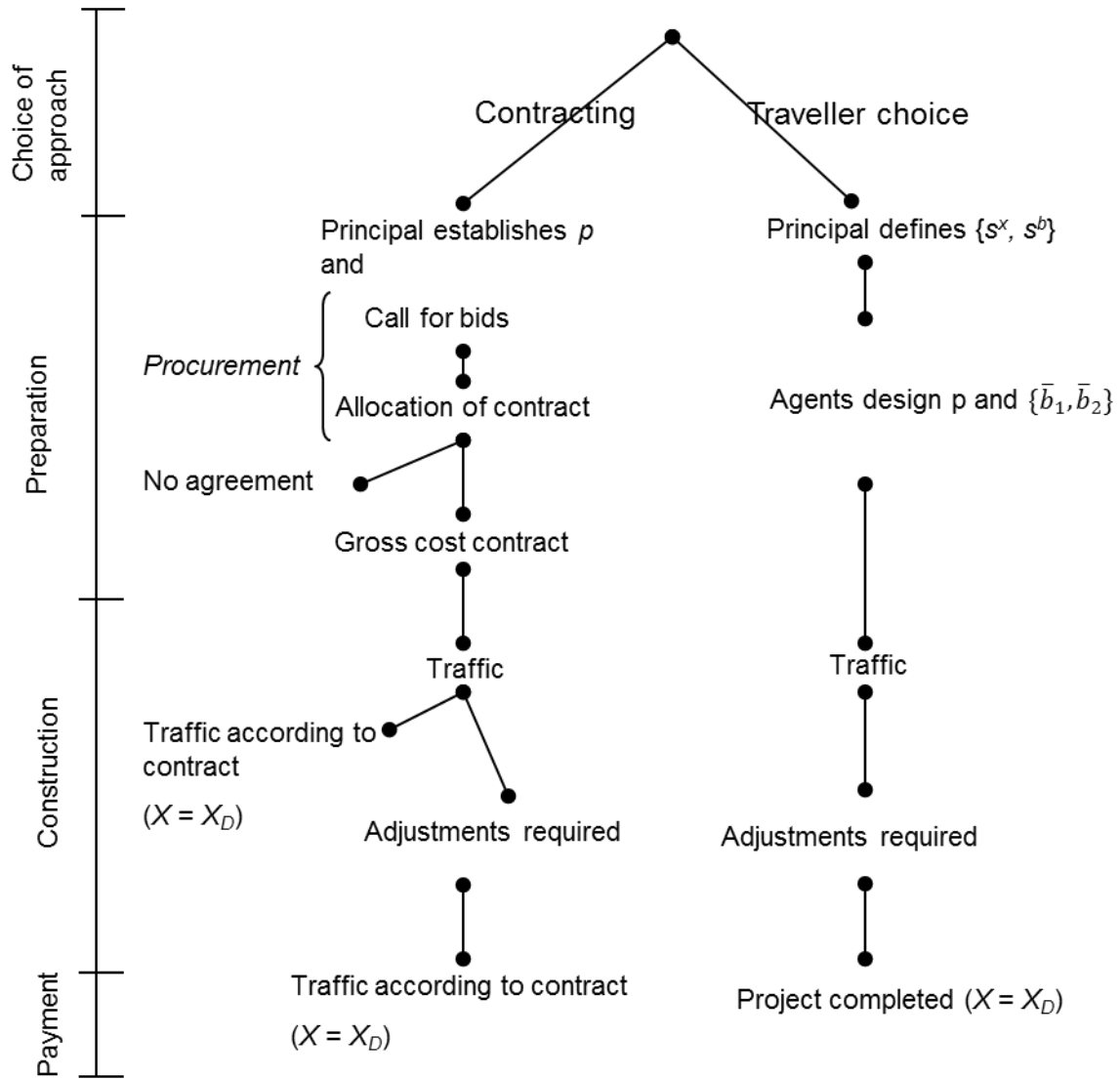


Figure 1. The nature of a tendered and a voucher solution to subsidizing public transport.

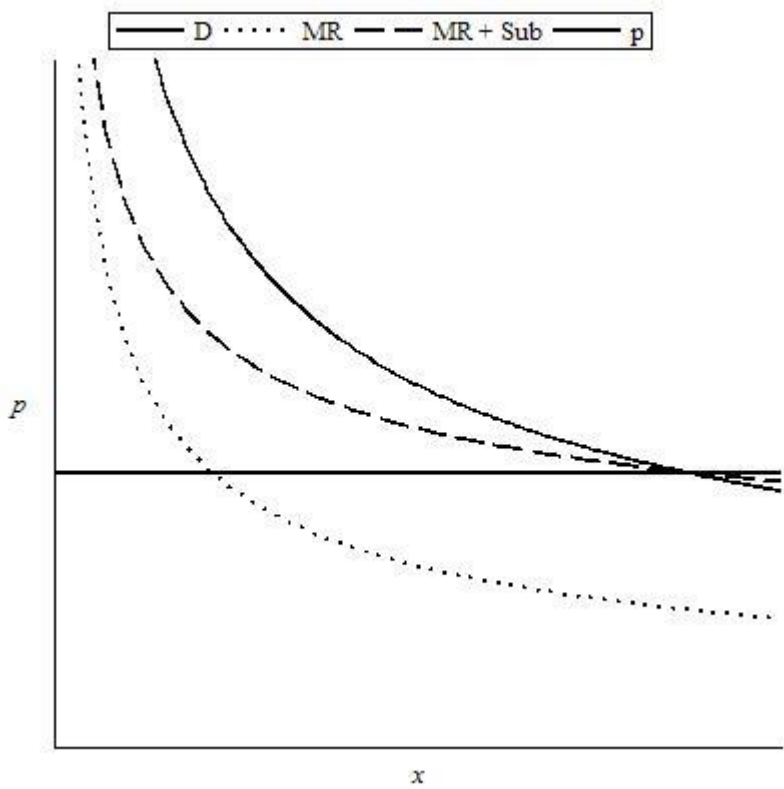


Figure 2: Demand for (thick solid line) and marginal costs of (horizontal line) a bus service. The dotted line represents marginal revenue for a monopolist before, and the dashed line marginal revenue after, the introduction of subsidies.

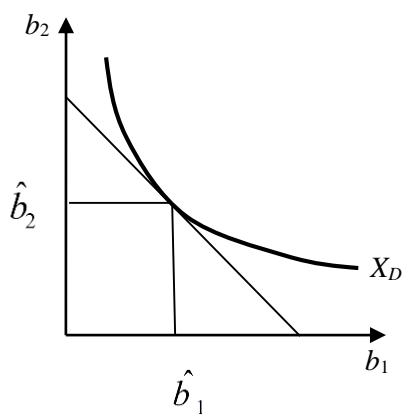


Figure 3

Figure 3. Isocost isoquant diagram for two types of buses (b).