



TI 2009-077/1

Tinbergen Institute Discussion Paper

A Monopolist in Public Transport: Undersupply or Oversupply?

Vladimir A. Karamychev

Peran van Reeven

Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute.

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam

Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam

Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at
<http://www.tinbergen.nl>.

A Monopolist in Public Transport: Undersupply or Oversupply?

Vladimir A. Karamychev and Peran van Reeven

Erasmus University Rotterdam

Abstract. A monopolist in public transport may oversupply frequency relative to the social optimum, as van Reeven (2008) demonstrates with homogeneous consumers. This result generalizes for heterogeneous consumers who know the timetable. Whether a monopolist oversupplies or undersupplies frequency depends on the degree of consumers' heterogeneity as reflected in the distribution of consumers' reservation prices. Oversupply is likely to occur when this distribution is peaked, and undersupply is likely to occur when this distribution is rather flat. In particular, monopoly production results in the oversupply of frequency when consumers' reservation prices are concentrated around the entry costs of the private car, being the main alternative to public transport.

This version: July 2009

Correspondence addresses. Peran van Reeven, Department of Applied Economics, Erasmus University Rotterdam, Burg. Oudlaan 50, NL-3062 PA Rotterdam, The Netherlands, e-mail: vanreeven@ese.eur.nl.

1. Introduction

Current levels of subsidization in public transport make it difficult to accept that no generally valid economic justification exists for these subsidies. Traditional economic arguments for subsidizing public transport, such as economies in production, second-best, and distributional equity, have proven to be weak (see, *e.g.*, De Borger *et al.*, 2002, Hencher, 1998, Paulley *et al.*, 2006, Small, 1992). The final remaining argument is due to Mohring (1972), who argues that public transport exhibits user economies that result in underproduction. However, van Reeve (2008) shows that the modeling assumptions in Mohring (1972) do not result in underproduction and, therefore, subsidization of public transport is not an imperative.

In *this* journal, Savage and Small (2009) and Basso and Jara-Díaz (2009) argue that the consumer homogeneity assumption is critical in van Reeve (2008), and that a monopolist will undersupply frequency when consumers are heterogeneous and have downward-sloping demand. Basso and Jara-Díaz (2009) demonstrate this for the uniform distribution of consumers' reservation prices, which results in a linearly decreasing demand function. Savage and Small (2009) show that the shape of the distribution is irrelevant in this respect. As long as consumers do not know the timetable, any distribution that leads to a downward-sloping demand schedule results in the under-provision of frequency. Savage and Small argue that this result becomes ambiguous if consumers know the timetable, as both undersupply and oversupply can occur.

Van Reeve (2008) shows that consumers have an incentive to get to know the timetable. Learning the times of departure reduces their waiting costs. This unambiguously increases the demand for transportation so that operators have an incentive to publish the timetable. Hence, in a general equilibrium framework, consumers must be informed. In this paper, we generalize van Reeve (2008) to heterogeneous consumers who know the timetable. The results of this generalization are as follows.

Whether monopoly oversupplies or undersupplies frequency depends on a trade-off between two effects. The first effect is the traditional result that a monopolist prices above marginal cost, which results in the production of a lower quantity than is socially desirable. In an equilibrium context, this creates an incentive to undersupply frequency relative to the social optimum.

The second effect is pointed out by Spence (1975). A social planner looks at the effect of an increase in quality (read: frequency) on all consumers in average, whereas the monopolist considers this effect on the marginal consumer. In public transport, for optimally chosen departure times (see van Reeve, 2008, for the uniform distribution, and Janssen *et al.*, 2005, for non-uniform distributions), the marginal consumer, *i.e.*, the consumer who is indifferent between two consecutive departures times, has a waiting time that is twice as large as the waiting time of an average consumer. This creates an incentive to oversupply frequency relative to the social optimum.

The balance of these two forces is determined by the distribution of consumers' reservation prices. In van Reeve (2008), consumers are homogeneous, and the distribution is degenerate and concentrated at the utility level of an outside option, *e.g.*, private car, in the spirit of Salop (1979). This eliminates the incentive to undersupply frequency, and the monopolist oversupplies with informed consumers, and is socially optimal with uninformed consumers. In Basso and Jara-Díaz (2009), to the contrary, consumers' reservation prices are distributed evenly, which is the opposite case. This eliminates the incentive to oversupply frequency, so that the monopolist strictly undersupplies.

Using the family of the power distribution functions, which has the homogeneous consumer case of van Reeve (2008) on one hand, and the uniform distribution of Basso and Jara-Díaz (2009) on the other, as two opposite limiting cases, we show that there is a critical degree of heterogeneity with the following property. If the heterogeneity in consumer reservation price is stronger than the critical level, the undersupply result occurs. If, to the

contrary the heterogeneity is weaker, the monopolist oversupplies. This illustrates the argument of Savage and Small (2009) that undersupply and, therefore, subsidization of public transport is not an imperative.

Without an empirical investigation of the shape of the distribution function of consumers' reservation prices, theoretical modeling will hardly shed a new light on the likelihood of undersupply/oversupply outcomes. Considering private car as the main alternative to public transport, it is reasonable to assume that consumers' reservation prices are concentrated around the entry costs of the car. This concentration makes oversupply just as likely as undersupply.

The rest of the paper is organized as follows. Section 2 sets up the model which is then analyzed in Section 3. Section 4 concludes.

2. The model

We build the model of a monopolist that provides a public transport service on van Reeven (2008). The monopolist schedules f equidistant departures in a unit time circle $y \in S^1 \equiv [0, 1)$ so that the time gap between two consecutive departures is $1/f$. A unit measure of consumers decides whether to use a public transport service on a fixed route. Each consumer i is characterized by his most preferred departure time $y_i \in S^1$, and by his reservation price $v_i \in [0, \bar{v}]$, where $\bar{v} > 0$. When this consumer uses the public transport service and departs at time y , he gets the following utility (surplus):

$$u_i = v_i - p - \tau_i,$$

where p is the price charged by the monopolist, $\tau_i = tw_i$ is the disutility of waiting, $w_i = |y - y_i|$ is the consumer's waiting time, and $t > 0$ is the consumers' waiting cost per unit

of time.¹ When consumers decide not to use the service, their surplus is zero. Thus, we interpret the reservation price v_i as a gross utility level from using the monopoly public transport service net of the opportunity cost.

Consumers' most preferred departure time y_i is evenly distributed in time, so that the waiting disutility τ_i is uniformly distributed over the interval $[0, x]$, where $x = t/(2f)$ is the maximum waiting disutility that consumers can experience. Consumers' reservation price v_i follows a distribution $F(v) = (v/\bar{v})^k$, where $k \in [1, \infty)$. This family of the power distribution functions has the following property. For $k = 1$, v_i is uniformly distributed over the interval $[0, \bar{v}]$, as in Basso and Jara-Díaz (2009). In the limit when k unboundedly increases, all v_i become equal to the highest level \bar{v} with probability one. This is the uniform reservation price as in van Reeve (2008). Without loss of generality we assume $\bar{v} = 1$.²

In order to abstract from the effect of returns to scale in the transportation industry, we assume that monopoly operates at zero marginal cost and pays constant fixed cost $c > 0$ per unit of frequency. For the sake of convenience, we assume that the choice variable is not the frequency but the maximum waiting disutility x , from which the frequency can be obtained by $f = t/(2x)$. When $x > \bar{v} = 1$, no consumers travel, so that the feasible range of x is $x \in [0, 1]$. Thus, the monopoly profit is given by

$$\pi(p, x) = pD(p, x) - cf = pD(p, x) - \frac{ct}{2x},$$

where $D(p, x) = \Pr(v_i - tw_i - p \geq 0)$ is the demand for public transport, *i.e.*, the measure of consumers who get non-negative surplus from using the service at price p . When $p > 1$, no

¹ When $y < y_i$, τ_i can be interpreted as a schedule delay cost.

² This only has the effect on measurement units of all the costs, prices, and utility levels.

consumer uses the service from the monopolist, so that the feasible range of p is $p \in [0, 1]$.

Profit maximizing values of x , p , and f are denoted by x^π , p^π , and f^π .

In order to compare f^π with the first best frequency, we compute the social welfare function, which is:

$$SW(p, x) = S(p, x) + \pi(p, x),$$

where $S(p, x) = D(p, x) \cdot E(v_i - tw_i - p | v_i - tw_i - p \geq 0)$ is the aggregate consumer surplus.

Social welfare maximizing, *i.e.*, the first best, values of x , p , and f are denoted by x^{FB} , p^{FB} , and f^{FB} . The purpose of the analysis is to compare f^{FB} and f^π for different values of the parameter k .

3. The analysis

We begin with computing aggregate demand and consumer surplus. Let us fix values of $x \in [0, 1]$, price $p \in [0, 1]$, and a single departure time y , and only consider consumers with most preferred departure times from the interval $y_i \in [y - x/t, y + x/t]$. All such consumers, if they opt for using the service, will depart at time y because it is their closest departure time.

Suppose that x is chosen optimally so that monopoly profit or social welfare is strictly positive.³ Then, when $p \in (1 - x, 1]$ all consumers with

$$y_i \in \left[y - \frac{x}{t}, y - \frac{1-p}{t} \right) \cup \left(y + \frac{1-p}{t}, y + \frac{x}{t} \right]$$

³ The case when exogenous parameter c , t , and k lead to zero maximum values of the profit function and social welfare function is non-generic.

do not travel irrespective of v_i as $u_i = v_i - p - \tau_i < v_i - p - (1 - p) \leq 0$. In this case, there are gaps in demand in the time space, and profit or social welfare values can be increased by decreasing x , *i.e.*, by having more departures at smaller intervals from each other. Thus, the case when $p \in (1 - x, 1]$ cannot happen with optimal value of x , and we restrict the domains of the profit function $\pi(p, x)$ and social welfare function $SW(p, x)$ so that they still contain the corresponding optimal values (p^π, x^π) and (p^{FB}, x^{FB}) by $x \in [0, 1]$ and $p \in [0, 1 - x]$.

For fixed values of $x \in [0, 1]$ and $p \in [0, 1 - x]$, we divide all consumers with $y_i \in [y - x/t, y + x/t]$ into the following three categories.

- a) Consumers with reservation price $v_i \in [0, p)$ get negative surplus from the service:

$$u_i = v_i - p - tw_i < 0.$$

These consumers do not use the public transport service and get zero consumer surplus:

$$s(v_i) = 0.$$

- b) Consumers with reservation price $v_i \in (p + x, 1]$ get positive surplus from the service irrespective of their most preferred departure time y_i :

$$u_i = v_i - p - tw_i \geq v_i - p - x > 0.$$

These consumers use the public transport service and get consumer surplus:

$$s(v_i) = \int_{y-x/t}^{y+x/t} u_i dy_i = 2 \int_0^{x/t} (v_i - p - ty_i) dy_i = \frac{x}{t} (2(v_i - p) - x).$$

- c) Consumers with reservation price $v_i \in [p, p + x]$ get positive surplus from the service only when $u_i = v_i - p - tw_i \geq 0$, that is, when their most preferred departure time y_i belongs to the following interval:

$$y_i \in \left[y - \frac{v_i - p}{t}, y + \frac{v_i - p}{t} \right].$$

Such consumers use the public transport service and get consumer surplus:

$$s(v_i) = \int_{y-(v_i-p)/t}^{y+(v_i-p)/t} u_i dy_i = 2 \int_0^{(v_i-p)/t} (v_i - p - ty_i) dy_i = \frac{(v_i - p)^2}{t}.$$

Aggregate consumer surplus $S(p, x)$ can be found by integrating $s(v)$ over $v \in [0, 1]$ with respect to the distribution function $F(v)$, and then by multiplying by the number of departures $f = t/(2x)$:

$$S(p, x) = f \int_0^1 s(v) dF(v) = \frac{t}{2x} \left(\int_p^{p+x} \frac{(v-p)^2}{t} kv^{k-1} dv + \int_{p+x}^1 \frac{x}{t} (2(v-p) - x) kv^{k-1} dv \right)$$

In order to find aggregate consumer demand, we differentiate (the negative of) the aggregate consumer surplus with respect to price:

$$D(p, x) \equiv -\frac{\partial S}{\partial p} = \frac{1}{x} \int_p^{p+x} (v-p) kv^{k-1} dv + \int_{p+x}^1 kv^{k-1} dv = 1 - \frac{(p+x)^{k+1} - p^{k+1}}{x(k+1)}$$

It is worth to note at this stage that when consumers are uninformed, they all experience waiting disutility $x/2$, and the aggregate consumer surplus $S^{UN}(p, x)$ will be written as

$$S^{UN}(p, x) = \frac{t}{2x} \left(\int_{p+x/2}^1 \frac{x}{t} (2(v-p) - x) kv^{k-1} dv \right),$$

which will lead to the following aggregate demand from uninformed consumers

$$D^{UN}(p, x) \equiv -\frac{\partial S}{\partial p} = 1 - \left(p + \frac{x}{2} \right)^k$$

For $k=1$, $D^{UN}(p, x)$ coincides with $D(p, x)$ as argued by Basso and Jara-Díaz (2009).

However, this is just a coincidence and demands $D^{UN}(p, x)$ and $D(p, x)$ are different for $k \neq 1$.

The monopoly profit function $\pi(p, x)$ is:

$$\pi(p, x) = pD(p, x) - \frac{ct}{2x} = p \left(1 - \frac{(p+x)^{k+1} - p^{k+1}}{x(k+1)} \right) - \frac{ct}{2x}$$

First, we look at the social welfare function $SW(p, x)$. Using $D = -\partial S / \partial p$, it is easy to check that $SW(p, x)$ strictly decreases in p for all $p > 0$:

$$\frac{\partial SW}{\partial p} = \frac{\partial S}{\partial p} + \frac{\partial \pi}{\partial p} = \frac{\partial S}{\partial p} + D(p, x) + p \frac{\partial D}{\partial p} = p \frac{\partial D}{\partial p} = -p \frac{(p+x)^k - p^k}{x} < 0.$$

Hence, the first-best price is zero, $p^{FB} = 0$, so that the social planner chooses $x \in [0, 1]$ in order to maximize the following (reduced-form) social welfare function:

$$\begin{aligned} SW^*(x) &\equiv SW(p^{FB}, x) = S(0, x) + \pi(0, x) = \frac{t}{2x} \left(\int_0^x \frac{v^2}{t} k v^{k-1} dv + \int_x^1 \frac{x}{t} (2v-x) k v^{k-1} dv \right) - \frac{ct}{2x} \\ &= \frac{1}{(k+2)(k+1)} x^{k+1} - \frac{1}{2} x + \frac{k}{k+1} - \frac{ct}{2x} \end{aligned}$$

The resulting socially optimal values of the maximum waiting cost x^{FB} and the corresponding frequency f^{FB} are:

$$x^{FB} \equiv \arg \max_{x \in [0, 1]} SW^*(x) \text{ and } f^{FB} \equiv \frac{t}{2x^{FB}}.^4$$

The monopolist chooses $x \in [0, 1]$ and $p \in [0, 1-x]$ in order to maximize $\pi(p, x)$. Profit maximizing values of price p^π , the maximum waiting cost x^π , and the corresponding frequency f^π are:

$$(p^\pi, x^\pi) \equiv \arg \max_{\substack{x \in [0, 1] \\ p \in [0, 1-x]}} \pi(p, x) \text{ and } f^\pi \equiv \frac{t}{2x^\pi}.$$

As the first best and profit maximizing values depend on the distribution of reservation prices through the parameter k , we often write them as functions of k explicitly.

⁴ Generically, for the chosen class of distribution functions, the problems of maximization of social welfare and monopoly profit always have a unique solution, which explains the chosen definitions.

Whether monopoly profit and social welfare are positive or negative depends on the exogenous value of ct . Apart from k , this product of fixed cost c and unit waiting cost t is the only exogenous entry in the objective functions $\pi(p, x)$ and $SW^*(x)$.⁵ When the monopolist gets non-negative profits, social welfare is non-negative as well. In what follows, we assume that the value of the product ct is sufficiently low so that a monopoly is operational for any $k \geq 1$. In particular, we assume that $ct < C = 4/27$. Next, we will relax this assumption and discuss the case $ct > C$, for which the monopoly profit might become negative for some values of k .

In the following proposition, we generalize the result of van Reeve (2008), which is obtained for the limiting case when k unboundedly increases, and which we formally (with slight abuse of the notation) denote by $k = \infty$. We show that the monopolist always oversupplies frequency when the distribution density function of consumers' reservation prices sufficiently increases, which is measured by parameter k .

Proposition 1. For all large values of k , the monopolist over-produces frequency, *i.e.*, there exists $k^* \geq 1$ such that for all $k > k^*$: $x^\pi(k) < x^{FB}(\infty) < x^{FB}(k)$ and $f^\pi(k) > f^{FB}(\infty) > f^{FB}(k)$.

Proof. Van Reeve (2008) shows that in the limit when $k \rightarrow \infty$:

$$x^\pi(\infty) = \sqrt{\frac{ct}{2}} < x^{FB}(\infty) = \sqrt{ct}, \text{ so that } f^\pi(\infty) = \sqrt{\frac{t}{2c}} > f^{FB}(\infty) = \sqrt{\frac{t}{4c}}.$$

We first show that there exists a $\delta > 0$ such that $x^\pi(k) > \delta$ for all $k \geq 1$, and we can restrict the domains of the functions $\pi(p, x)$ and $SW^*(x)$ by $x \in [\delta, 1]$. Next, we show that there exists $k^* > 1$ such that $x^\pi(k) < x^{FB}(\infty)$ for all $k > k^*$. Finally, we show that $x^{FB}(k) > x^{FB}(\infty)$ for all $k \geq 1$, and the statement of the proposition follows.

⁵ For a fixed level of ct , the value of t has its own and independent of ct effect on the frequency choice through $f = t/(2x)$.

Both functions $\pi(p, x)$ and $SW^*(x)$ have the following property:

$$\lim_{x \rightarrow 0} \frac{\partial \pi}{\partial x} = \lim_{x \rightarrow 0} \frac{ct}{2x^2} = +\infty \text{ and } \lim_{x \rightarrow 0} \frac{dSW^*}{dx} = \lim_{x \rightarrow 0} \left(\frac{ct}{2x^2} - \frac{1}{2} \right) = +\infty,$$

where the convergence is uniform with respect to k . Thus, there exists a $\delta > 0$ such that for any k, p , and $x \in (0, \delta)$:

$$\frac{\partial \pi}{\partial x} > 0 \text{ and } \frac{dSW^*}{dx} > 0.$$

Consequently, $x^{FB} > \delta$ and $x^\pi > \delta$, and we restrict domains of the functions $\pi(p, x)$ and $SW^*(x)$ by $x \in [\delta, 1]$.

Maximizing $\pi(p, x)$ w.r.t. (p, x) yields the following first-order conditions:

$$\begin{cases} 0 = \frac{\partial \pi}{\partial p}(p^\pi, x^\pi) = 1 - \frac{p^\pi (p^\pi + x^\pi)^k}{x^\pi} - \frac{(p^\pi + x^\pi)^{k+1} - (p^\pi)^{k+1}}{x^\pi (k+1)} + \frac{(p^\pi)^{k+1}}{x^\pi} \\ 0 = \frac{\partial \pi}{\partial x}(p^\pi, x^\pi) = \frac{ct}{2(x^\pi)^2} - \frac{p^\pi (p^\pi + x^\pi)^k}{x^\pi} + p^\pi \frac{(p^\pi + x^\pi)^{k+1} - (p^\pi)^{k+1}}{(x^\pi)^2 (k+1)} \end{cases}$$

By subtracting the first equation from the second one and rearranging, we get:

$$x^\pi(k) = \sqrt{\frac{ct}{2 - G_1(k)}},$$

where

$$G_1(k) = 2 \left(\frac{(p^\pi + x^\pi) \left((p^\pi + x^\pi)^{k+1} - (p^\pi)^{k+1} \right)}{(x^\pi)^2 (k+1)} - \frac{(p^\pi)^{k+1}}{x^\pi} \right).$$

It is easy to see that $\lim_{k \rightarrow \infty} G_1(k) = 0$ uniformly for $x^\pi \in [\delta, 1]$ and $p^\pi \leq 1 - x^\pi$. Hence, there

exists $k^* > 1$ such that $G_1(k) < 1$ for all $k > k^*$. Therefore, for all such $k > k^*$:

$$x^\pi(k) = \sqrt{\frac{ct}{2 - G_1(k)}} < \sqrt{\frac{ct}{2 - 1}} = \sqrt{ct} = x^{FB}(\infty).$$

Maximizing $SW^*(x)$ w.r.t. $x \in [\delta, 1]$ yields the following first-order condition:

$$0 = \frac{dSW^*}{dx}(x^{FB}) = \frac{ct}{2(x^{FB})^2} - \frac{1}{2} + \frac{1}{(k+2)}(x^{FB})^k,$$

which leads to the following expression:

$$x^{FB}(k) = \sqrt{\frac{ct}{1-G_2(k)}} > \sqrt{ct} = x^{FB}(\infty),$$

where

$$G_2(k) = \frac{2}{(k+2)}(x^\pi)^k > 0.$$

Thus, $x^\pi(k) < x^{FB}(\infty) < x^{FB}(k)$ for all $k > k^*$. This ends the proof. ■

Proposition 1 demonstrates that the monopolist oversupplies frequency not only when consumers are homogeneous and have unit demand but also when consumers are heterogeneous and demand is downward-sloping. Two effects determine whether monopoly oversupplies or undersupplies frequency. The first effect follows from the traditional result that a monopoly under-produces due to pricing above marginal cost. This creates a strict incentive to decrease the frequency. The second effect is that the monopolist has a strict incentive to oversupply frequency relative to the social optimum because the waiting time of a marginal consumer is twice as large as the waiting time of an average consumer.

When the reservation price distribution density function sufficiently increases so that its variance

$$\text{var}(v_i) = E(v_i^2) - (E(v_i))^2 = \frac{k}{(k+1)^2(k+2)},$$

which reflects consumer heterogeneity, is relatively small, the first effect is weak and the desire of the monopolist to increase consumer demand by choosing higher level of frequency drives the oversupply result. When, to the contrary, the reservation price distribution is more even, the first effect is strong, monopoly output is low, and, consequently, frequency will be undersupplied.

The latter is the case for the uniform distribution, when $k = 1$, as Basso and Jara-Díaz (2009) show. Also for $k = 2$ the monopolist undersupplies frequency, as the following proposition demonstrates.

Proposition 2. For $k \in \{1, 2\}$, the monopolist undersupplies frequency, *i.e.*, $x^\pi > x^{FB}$ and $f^\pi < f^{FB}$.

Proof. For $k = 1$, the proof is given in Basso and Jara-Díaz (2009). When $ct < C$, the monopoly is profitable. Indeed, by choosing, *e.g.*, $p = 1/3$ and $x = 2/3$, monopoly gets:

$$\pi\left(\frac{1}{3}, \frac{2}{3}\right) = \frac{3}{4}\left(\frac{4}{27} - ct\right) > 0.$$

Thus, monopoly is operational for $ct < C$.

For $k = 2$, functions $SW^*(x)$ and $\pi(p, x)$ are given by:

$$SW^*(x) = \frac{1}{12}x^3 - \frac{1}{2}x + \frac{2}{3} - \frac{ct}{2x}, \text{ and } \pi(p, x) = p\left(1 - p^2 - px - \frac{1}{3}x^2\right) - \frac{ct}{2x}.$$

Maximizing $SW^*(x)$ yields the bi-quadratic equation $x^4 - 2x^2 + 2ct = 0$. A unique suitable solution is $x^{FB} = \sqrt{1 - \sqrt{1 - 2ct}}$. Maximizing $\pi(p, x)$ w.r.t. p yields $p^\pi = (\sqrt{3} - x^\pi)/3$, and the bi-quadratic equation $2x^4 - 6x^2 + 9ct = 0$ (this equation is only valid for $x^\pi \leq (3 - \sqrt{3})/2 \approx 0.63$ so that $p^\pi + x^\pi \leq 1$; if this is not the case, then the corner solution $p^\pi = 1 - x^\pi$ must be taken).

The unique suitable solution is $x^\pi = \sqrt{3(1 - \sqrt{1 - 2ct})}/2 = \sqrt{3/2}x^{FB}$, so that $f^\pi = \sqrt{2/3}f^{FB} < f^{FB}$. This ends the proof. ■

Propositions 1 and 2 provide a partial characterization of the frequency chosen by the monopolist relative to the first best. In accordance with Proposition 2, at least $k^* > 2$. In

order to find out the lowest value of k^* , which ensures oversupply $f^\pi(k) > f^{FB}(k)$ for all $k > k^*$, we have performed numerical simulations.⁶ It turns out that the maximization (of the profit and social welfare functions) problems have nice analytical properties (monotonicity and quasi-concavity), so that $x^\pi(k)$ and $x^{FB}(k)$ are monotone and the numerical algorithms used in the simulations rapidly converge.

Figure 1 shows functions $f^\pi(k)$ and $f^{FB}(k)$ for $ct = 0.1 < C$ and $t = 1$. The value of k^* is defined by $f^\pi(k^*) = f^{FB}(k^*)$. It turns out that $f^\pi(k^*) > f^{FB}(k^*)$ for all $k < k^*$ and $x^\pi(k^*) < x^{FB}(k^*)$ for all $k > k^*$.

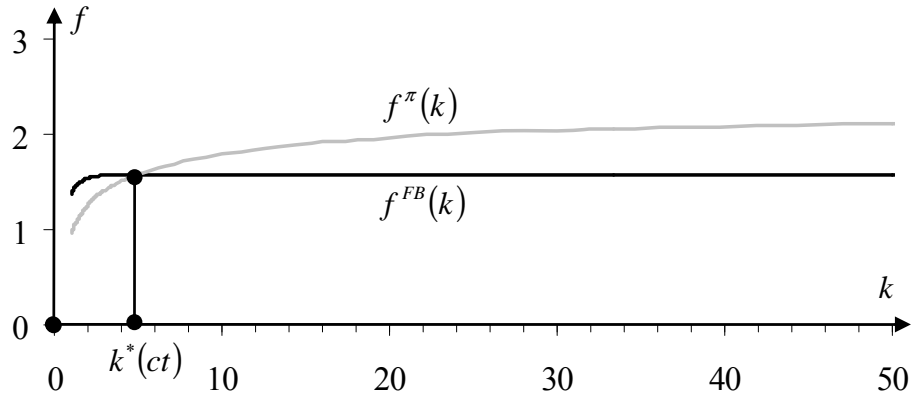


Figure 1. Profit maximizing frequency $f^\pi(k)$ and socially optimal frequency $f^{FB}(k)$ as functions of k for $ct = 0.1$ and $t = 1$.

Figure 2 shows how the critical value k^* continuously and monotonically depends on ct for $ct < C$. When ct converges to zero, the value of k^* unboundedly increases. It can be seen that the frequency oversupply result holds only for $k > \underline{k} \approx 4.7$.

⁶ The simulations are performed within MS Excel on VBA, and this tool is available from the corresponding author upon request.

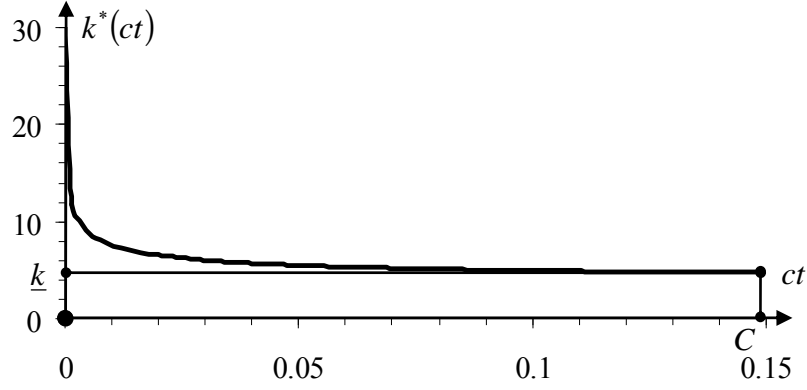


Figure 2. Critical value function $k^*(ct)$.

It is important to note that $k^*(ct)$ decreases with cost parameter ct , so that an increase in ct makes oversupply more likely. The intuition is that for larger cost values, the monopoly frequency f^π decreases at a much lower pace than socially optimal frequency f^{FB} . This is so because the condition $x^\pi + p^\pi \leq 1$ becomes binding, and the monopoly price is a corner solution $p^\pi = 1 - x^\pi$. As a result, the monopolist has a much lower incentive to reduce frequency as the required decrease in price has a strong adverse effect on monopoly profit.

In the analysis, we have assumed $ct < C$ so that monopoly is operational for any k . For a given value of k , there is a range of ct values where monopoly does not produce even though it would have been socially optimal. Savage and Small (2009) point out that subsidizing the monopolist in this case is a solution to this market failure. We now turn to this market failure and have a closer look at the impact of subsidization.

Figure 3 presents an example of the reduced-form social welfare function $SW^*(x)$ and the reduced-form profit function $\pi^*(x) \equiv \max_p \pi(p, x)$ for $k = 4$ and $ct = 0.5 > C$. The maximizing values of x for $SW^*(x)$ and $\pi^*(x)$ are $x^{FB} \approx 0.75$ and $x^\pi \approx 0.67$ respectively. For the chosen parameter values, the monopoly is inactive because its maximum profit level is negative: $\pi^*(x^\pi) \approx -0.141 < 0$.

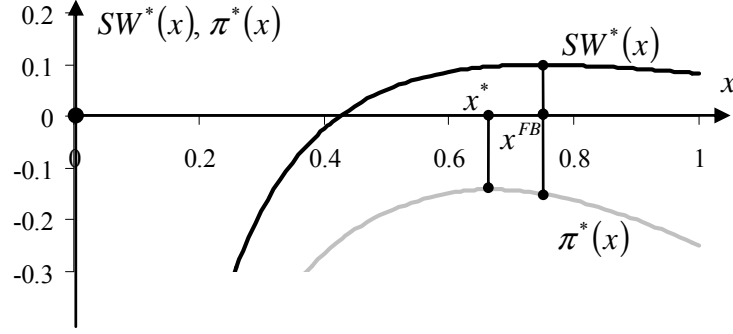


Figure 3. Functions $SW^*(x)$ and $\pi^*(x)$ for $k=1$ and $ct=0.3$.

Subsidizing the monopolist with the amount of $\Delta = 0.141$ guarantees him just a non-negative profit, and the production starts (we assume the moral hazard problem away). The main point of this example is that such subsidy results in oversupply as $x^\pi < x^{FB}$. In order to sustain the socially optimal frequency $f^{FB} = t/(2x^{FB})$, the subsidy has to be at least $\Delta = -\pi^*(x^{FB}) \approx 0.149$, *i.e.*, the monopoly must receive additional subsidy in order to able to produce this lower frequency profitably. However, such additional subsidy does not give any incentive to do so unless the first best level of frequency is explicitly imposed on the monopolist.

We are now able to relax our assumption $ct < C$ and extend our computations by assuming that the monopolist gets a lump-sum subsidy if its profit is negative. Figure 4 represents the parameter space with $ct \in [0,1]$ on the horizontal axis and $k \geq 1$ in the log scale on the vertical axis. We have also drawn a line that represents $k = \infty$. The bold line represents the generalization of the critical value function $k^*(ct)$ for $ct \in [0,1]$. The shaded area to the left of $k^*(ct)$ reflects those parameter values for which the monopolist undersupplies. For all other parameter constellations, the monopolist oversupplies. In case $ct > \bar{C} \approx 0.381$, monopoly operation always yields oversupply regardless the value of k . In order to see whether this oversupply is generated under profitable conditions or under

subsidization, the other two curves $k^\pi(ct)$ (we do not have a closed form expression for it) and $k^{FB}(ct) = 2ct/(1-ct)$ represent those parameter values for which the maximum values of monopoly profit $\pi^*(x^\pi)$ and social welfare $SW^*(x^{FB})$ equal zero.

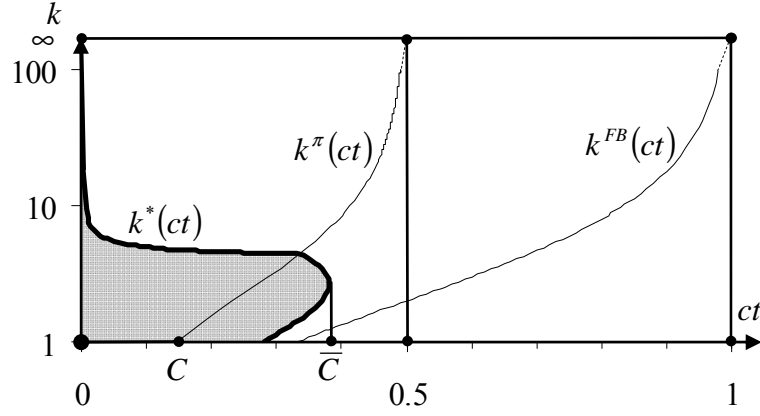


Figure 4. Parameter's space (ct, k) , regions of undersupply (shaded) and oversupply.

In the region below $k^{FB}(ct)$, it is not socially optimal to produce a transportation service. In the region between $k^{FB}(ct)$ and $k^\pi(ct)$, production is socially optimal but requires a subsidy. Finally, in the region above $k^\pi(ct)$, service provision is optimal and the monopoly gets positive profit. Hence, figure 4 shows that under subsidization oversupply may occur for all $k \geq 1$, and therefore is even more likely to happen.

Finally, we admit that the original model of van Reeve (2008), as well as this generalization, is very stylized. Savage and Small (2009) argue that a richer model is required for the proper analysis of the monopoly frequency choice. Nevertheless, even these simple models make it clear that theoretically oversupply is equally likely to happen as undersupply. Extending the model will certainly change the balance of the effects presented here. However, it is not obvious that such extensions will favor undersupply. For example, increasing marginal production cost (*e.g.*, due to increasing boarding and alighting costs when vehicles become fuller) strengthens the incentive to increase frequency. Similarly, increasing

marginal waiting cost increases the difference in utility between the average and the marginal consumers and, therefore, further increases the incentive to oversupply. Hence, oversupply is even more likely to happen under these assumptions.

4. Conclusion

Van Reeve (2008) shows that monopoly does not undersupply frequency if consumers are homogeneous with respect to their reservation price. This result generalizes to the case of heterogeneous consumers when the distribution density function of consumers' reservation price peaks. In contrast, when consumers are very heterogeneous, so that reservation prices are distributed evenly and the distribution function rather flat, the monopolist undersupplies frequency. Hence, the question whether a monopolist oversupplies or undersupplies frequency critically depends on how heterogeneous the costs of alternative means of travel are.

Since the private car is the main alternative to public transport, consumers' reservation prices are likely to be concentrated around the entry costs of the private car. The distribution density function may include some other peaks for other alternatives to public transport as well. These concentrations in the distribution function make oversupply just as likely as undersupply. It very much depends on the particular case at hand and requires empirical research on the exact shape of the distribution of consumers' reservation prices, as well as other effects that are omitted in our analysis for simplicity.

The main implication from the analysis is that the oversupply of frequency by a monopolist in public transport cannot be excluded, so that subsidization in public transport is not an imperative. For many years, the transport economic literature has been preoccupied with aligning economic theory with increasing subsidization of public transport in practice. It generalizes and improves Mohring's original model with a better representation of user costs.

Every improvement has come with even higher levels of optimal production, as shown, *e.g.*, by Jara-Díaz and Gschwender (2003), justifying ever-increasing subsidies to public transport operations. In this respect, van Reeve (2008) may have put the cat among the pigeons, whereas this paper provides a more balanced view on the topic. However, the main conclusion remains the same: just the fact that consumers' waiting costs are external to the public transport monopolist does not constitute a generally valid economic justification for subsidization of public transport.

References

- Basso, L. and S. Jara-Díaz (2009): 'The case for subsidization of urban public transport and the Mohring effect', *Journal of Transport Economics and Policy*, this issue.
- De Borger, B., K. Kerstens, and A. Costa (2002): 'Public transit performance: What does one learn from frontier studies?', *Transport Reviews*, 22, 1-38.
- Hencher, D. (1998): 'Establishing a fare elasticity regime for urban passenger transport', *Journal of Transport Economics and Policy*, 32, 221-246.
- Janssen, M., V. Karamychev and P. van Reeve (2005): 'Multi-store competition: Market segmentation or interlacing?', *Regional Science and Urban Economics*, 35, 700-714.
- Jara-Díaz, S., and A. Gschwender (2003): 'Towards a general microeconomic model for the operation of public transport', *Transport Reviews*, 23, 453-469.
- Mohring, H. (1972): 'Optimization and scale economies in urban bus transportation', *American Economic Review*, 62, 591-604.
- Paulley, N., R. Balcombe, R. Mackett, H. Titheridge, J. Preston, M. Wardman, J. Shires, and P. White (2006): 'The demand for public transport: The effects of fares, quality of service, income and car ownership', *Transport Policy*, 13, 295-306.
- Salop, S. (1979): 'Monopolistic competition with outside goods', *Bell Journal of Economics*,

10, 141-156.

Savage, I. and K. Small (2009): ‘A comment on “Subsidization of urban public transport and the Mohring effect”’, *Journal of Transport Economics and Policy*, this issue

Small, K. (1992): *Urban Transportation Economics*, Harwood, Chur.

Van Reeve, P. (2006): ‘Subsidization of urban public transport and the Mohring effect’, *Journal of Transport Economics and Policy*, 42, 349-359.