Are Public Goods Public?

A Critical Survey of the Demand Estimates for Local Public Services*

by

Michael Reiter

and

Alfons Weichenrieder

29. July 2003

Abstract

Economies of scale are an important textbook rationale for regulation and government provision of services. The present paper surveys the empirical literature on economies of scale in the consumption of local public services. The majority of studies invoke the median voter model to estimate demand for local public services and reject scale economies. The paper discusses the methodological problems of this literature in detail and comments on the policy relevance of the results.

Kurzfassung


JEL classification: H41, H42

Addresses of the authors:

Michael Reiter
Dept. of Economics
Universitat Pompeu Fabra
Ramon Trias Fargas 25-27
08005 Barcelona
Spain

Alfons J. Weichenrieder
Center for Economic Studies
University of Munich
Schackstr. 4
80539 München
Germany

*The authors thank Hans-Werner Sinn, Marcel Thum and an anonymous referee for useful comments.
1 Introduction

The Samuelsonian view that governments provide public goods which are non-rival in consumption has influenced generations of economics students and continues to be a core element in any principles textbook. If consumers can use a good without inflicting a marginal cost on anybody else they should not be excluded from the good's consumption. Since private markets can be expected to exclude consumers by charging a positive price to cover intramarginal costs, some kind of government intervention is needed to guarantee a first best solution.

In the last 25 years, the validity of this textbook rationale for government activity has increasingly come under attack. The critique is based on an ever-increasing empirical literature on local government expenditures which concludes that, on average, publicly provided goods have roughly the same amount of rivalry in consumption as private goods. Some authors concluded from this that "there is reason to question the entire public goods rationale for government" (Holcombe and Sobel 1995, p. 55). More cautiously, others have asked "why, if there are not increasing returns in the municipal provision of the goods and services we study, is their provision in the public domain?" (Bergstrom and Goodman 1973, p. 293).

A somewhat disturbing feature of the empirical literature on locally provided goods is that most studies take the same approach to the question of scale economies. In every community a median voter is assumed to be in charge of expenditure decisions, and the crowding or congestion properties are then inferred from a cross section analysis of decisive voters' demand functions.1 Unfortunately, this is a very indirect approach to measuring publicness and strong identifying assumptions are needed to separate the crowding aspect from other factors affecting the demand for publicly provided goods.2 In addition, the

---

1This article uses the terms crowding and congestion interchangeably.
2These problems have been acknowledged in the literature. Bahl, Johnson and Wasylenko (1980, p. 67) go so far as to state that "even the best of these studies requires such heroic assumptions that their results are simply not believable".
interpretation of the empirical results raises difficult conceptual problems which have to be solved before any policy conclusions can be drawn. The purpose of the paper is both to survey the empirical results in the literature and to provide a critical discussion of the validity of the approach. While our study highlights a series of simplifications, errors and misinterpretations in the literature, our re-examination of existing work does not suggest a rejection of the privateness result. However, it highlights those methodological problems which deserve further attention before definite conclusions should be drawn.

The plan of the paper is as follows. The next section introduces the median voter approach of estimating crowding effects. It explains how the degree of crowding is derived jointly from the population elasticity, which describes the relation between public expenditures and population, and from the price elasticity of demand. To examine the crowding estimates, Section 3, therefore, discusses the population elasticity results, followed by Section 4 which takes a closer look at the price elasticity estimates. Section 5 is devoted to the question as to which extent the median voter model is an adequate way of modeling public demand. Section 6 reviews the role of the functional form which is imposed on the crowding function in empirical estimates. Section 7 highlights alternative approaches of estimating crowding effects before Section 8 concludes.

2 The Median Voter Approach and the Rejection of Publicness

2.1 The Basic Model

Since the early seventies the median voter model has become the cornerstone for the estimation of demand functions for publicly provided goods. Two seminal papers opened up a literature which now contains numerous studies using roughly the same methodology. While Borcherding and Deacon (1972) consider public provision of goods and services at the US state level, Bergstrom and Goodman (1973) concentrate on the provision of services by New

---

3 An early empirical test of the median voter model which anticipates important aspects of later approaches is included in Barr and Davis (1966).
York state communities. Both papers suggest that local and state governments, contrary to what public economists believe, provide goods which have roughly the same amount of rivalry in consumption as private goods do, and some have even more.

The essential element of both studies (henceforth abbreviated BD-BG approach) is the following model. The median voter who is assumed to decide the amount of a publicly provided good $G$ knows that she will incur a tax price per unit of $G$ which equals her tax share $t$ times the price of the good, $p$. Alternatively, she can use her exogenously given gross income $Y$ to purchase a private good $X$ with unit price one. Her budget restriction is therefore given by

$$X + t \cdot p \cdot G \leq Y.$$ (1)

The amount of public services $q$ which enter the median voter's utility function are a function of $G$. The BD-BG approach assumes the following functional form:

$$q = G \cdot N^{-\alpha},$$ (2)

where $N$ denotes the community's population and $\alpha$ is a measure of the crowding effects (the 'publicness', 'capturability' or 'crowding' parameter). In the case of $\alpha = 0$, the median voter's consumption is independent of the number of users and $G$ is a pure public good in the Samuelson sense ($q = G$). If $\alpha = 1$, the consumer enjoys only an $N$-th part of the amount of $G$ provided, and the good is a private good. Parameter values between zero and one indicate a quasi public good where crowding or congestion effects are present but there are still economies of scale in consumption.

In any community, the median voter exercises her voting power to maximize utility

$$U = U(q, X)$$ (3)

subject to (1) and (2). Since the BD-BG approach makes no attempt to measure the amount of individually consumed services, $q$, it may be unclear what is exactly meant by $q$, and in what units it is defined. A correct way to think of $q$ is to interpret it as a subutility function.
denoting \( N \)-\( G \) combinations which yield the same utility level. An implicit assumption of the model is that these combinations are independent of \( X \).\(^4\)

Assuming that maximization of (3) yields a demand function for \( q \) with constant income (\( \varepsilon \)) and price (\( \delta \)) elasticities, we have the median voter's demand in terms of \( q \) as

\[
q = k \cdot \left[ t \cdot p \cdot N^\alpha \right] Y^\varepsilon, \tag{4}
\]

and in terms of \( G \) as:

\[
G = k \cdot N^\alpha \cdot \left[ t \cdot p \cdot N^\alpha \right] Y^\varepsilon = k \cdot (t \cdot p)\delta \cdot Y^\varepsilon \cdot N^{\alpha(1+\delta)}, \tag{4a}
\]

where \( k \) is a constant. Since the theory assumes that the median voter decides \( G \), the empirical implementation requires the identification of the median voter. Bergstrom and Goodman have shown that, under certain conditions, assuming that the median income household is the median voter is justifiable if the empirical approach at the same time accounts for different voter groups with diverging preferences.\(^5\) Therefore, adding a vector \( Z \) of socioeconomic variables which are expected to influence demand, and multiplying both sides by \( p \) leads to an estimable function of local expenditures, \( E \):

\[
E = k \cdot t \delta \cdot \left( p^{\delta+1} \cdot Y^\varepsilon \cdot N^\alpha \right) \cdot Z^\delta. \tag{5}
\]

2.1 The Estimation Procedure

A main difference between the Borcherding-Deacon and the Bergstrom-Goodman approach lies in the estimation of the price elasticity of demand, d. Bergstrom and Goodman (BG)

\(^4\)The role of (2) is to provide a metrization for \( q \); \( q \) is defined such that it is proportional to \( G \) for given \( N \). We obtain a different metrization, for example, if we choose a natural metric, such as the negative crime rate as a measure of safety. With such a metric, there is no reason to assume (2) to hold. Of course, interpretation and comparison of the BD-BG results with alternative approaches require a crowding measure which is free of arbitrary metrization. Such a measure is given by \( \alpha \) if it is interpreted as the required percentage change in \( G \) necessary to keep \( q \) unchanged when \( N \) rises by one percent. These issues are discussed in much more detail in Reiter and Weichenrieder (1997).

\(^5\)The authors show that this is correct if, among other things, income distributions are proportional across communities for each preference group. Introducing such an income distribution condition to operationalize the median voter model yields what Inman (1979, p. 46) celebrates as "an analytically powerful 'as if' proposition; a proposition which stands as political economy's counterpart to the market economy's supposition that firms are profit maximizers." For a critical discussion of the similarity assumption on income distributions if voters are mobile see Goldstein and Pauly (1981, p. 140).
assume that the price $p$ of the public service is constant across communities and that tax price variations result only from median voters' different tax shares. Since the property tax accounts for more than half the local tax revenues in US communities, BG use the ratio of the tax on the median house value ($H^m$) and on the total property value in the community, $P$, to measure $t = H^m / P$. From (5), they obtain the estimation equation

$$E = k \cdot \left( \frac{H^m}{P} \right)^{\delta} \cdot Y^{\epsilon} \cdot N^{\eta^*} \cdot Z^{\beta}; \theta^* \equiv \alpha(1 + \delta).$$

(5-BG)

Unlike BG, Borcherding and Deacon (BD) assume that households' tax shares are identical ($t = 1/N$) and calculate the demand elasticity by looking at variations of $p$ across US states.\(^6\)

They estimate the equation

$$e = k^* \cdot p^{\delta + 1} \cdot Y^{\epsilon} \cdot N^{\eta^*} \cdot Z^{\beta}; \theta \equiv (\alpha - 1)(1 + \delta),$$

(5-BD)

where $e \equiv E / N$ denotes per capita expenditures.

A convenient property of both equations is that they allow the elasticities $a$, $d$ and $e$ to be measured without having to measure physical units of the publicly provided good, $G$. Instead, the left hand variable is expenditures which are readily observable. However, not measuring $G$ directly also has its drawbacks, an issue to which we will return repeatedly.\(^7\)

\(^6\)More precisely, differences in the voters' tax prices are computed by assuming that public services are produced by using capital, labor and a linear homogenous Cobb-Douglas production function. With the rental rate of capital assumed to be constant across states, differences in public employees' wages determine differences in $p$. Fortune (1983) tries to check whether the Cobb-Douglas assumption made by Borcherding and Deacon is reasonable. Assuming that cost-minimizing local governments use only labor and capital, he estimates an elasticity of substitution $s$ close enough to unity to support Borcherding and Deacon's assumption.

\(^7\)Other variables in (5) are also difficult to measure. Instead of median income, BD and many subsequent studies take average income as a proxy for $Y$. The problems of measuring $t$ and $p$ are discussed in detail in Section 4.
2.2 Results

The standard approach is to apply equations (5-BD) or (5-BG) to either single expenditure categories or general expenditures. Most studies consider local communities, counties or school districts. Typical estimates of the demand elasticities for local public goods are
<table>
<thead>
<tr>
<th>Study</th>
<th>Gov. level</th>
<th>General expenditures</th>
<th>Police and recreation</th>
<th>Local education</th>
<th>Higher education</th>
<th>Health</th>
<th>Fire</th>
<th>Library/culture</th>
<th>Sanitation</th>
<th>Highway/Roads</th>
<th>Legislative expenditures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bergstrom/Goodman (1973)b</td>
<td>C</td>
<td>1.09-</td>
<td>1.07°</td>
<td>1.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borcherding/Deacon (1972)c,d</td>
<td>A</td>
<td>1.019</td>
<td>1.0501</td>
<td>1.0527</td>
<td>0.8161</td>
<td>1.0065</td>
<td>1.0098</td>
<td>0.9320</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clothfelter (1976)</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.797</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edwards (1986)f</td>
<td>C</td>
<td>1.47°</td>
<td>2.01°</td>
<td>0.21°</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edwards (1990)f</td>
<td>C</td>
<td>1.50°</td>
<td>1.47°</td>
<td>3.63°</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gonzales/Means/Mehay (1993)</td>
<td>C</td>
<td>1.106°</td>
<td>1.069°</td>
<td>1.175°</td>
<td>1.061°</td>
<td>2.739°</td>
<td>0.823°</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gramlich/Rubinfeld (1982)</td>
<td>Y</td>
<td>1.010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hayes (1985)</td>
<td>C</td>
<td>1.0229</td>
<td>0.9512</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.9429</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.0087</td>
<td>0.9252</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.7744</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9604</td>
<td>0.6465</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.7862</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hayes (1986)</td>
<td>C</td>
<td>0.9098-</td>
<td>0.9527</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.5015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holcombe/Sobel (1995)</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.34020°</td>
<td>0.77297°i</td>
<td></td>
</tr>
<tr>
<td>Mcgreer/McMillan (1993)</td>
<td>C</td>
<td>1.30-</td>
<td>2.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McMillan (1989)</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.2223°</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McMillan/Wilson/Arthur (1981)</td>
<td>C</td>
<td>0.47-</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.69-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.90</td>
<td>0.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pack/Pack (1978)</td>
<td>C</td>
<td>1.19-</td>
<td>1.63-</td>
<td>1.36-</td>
<td>1.63-</td>
<td>0.93</td>
<td></td>
<td>1.32-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pommerehne (1978)</td>
<td>C</td>
<td>0.83°</td>
<td>1.14°</td>
<td>0.89°</td>
<td></td>
<td></td>
<td></td>
<td>2.30°</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pommerehne/Frey (1976)</td>
<td>C</td>
<td>0.992°</td>
<td>1.07°</td>
<td>0.962°</td>
<td>0.564°</td>
<td>0.887°</td>
<td>2.589°</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santerre (1985)</td>
<td>C</td>
<td>1.35°</td>
<td>1.35°</td>
<td>0.98°</td>
<td>0.965°</td>
<td>1.66°</td>
<td>0.978°</td>
<td>1.25°</td>
<td>1.10°</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vehorn (1979)</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.92°</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zimmerman (1983)</td>
<td>A</td>
<td>0.989°</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Annotations

- Standard error not available.
- Different from 1 at the 95 percent level or higher.
- A: data aggregated at US state level; C: community level data; Y: county level data; S: state level data
- Cf. Borcherding/Deacon, table 3. For some categories the authors divide their sample according to labor share in public service production. The first line reports results for states with a low labor share, the second for states with a high labor share.
- First line: communities < 10,000 inhabitants; second line: > 10,000.
- First line: 1962 data; second line 1977 data.
- The authors' results were rescaled to be comparable to the other studies. The value given in the table equals $\alpha_2 - 1$ in the authors' notation.
- First line: Cities in the Chicago Standard Metropolitan Statistical Area (SMSA); second line: 'downstate' cities within Chicago SMSA; third line: 'downstate' cities outside Chicago SMSA.
- Estimation range depending on year of data and inclusion of additional independent variables.
- Uncorrected estimate.
- First line: school districts with less than 2,000 students; second line: districts with 2,000 – 10,000 students.
- First line: communities with 1,000 – 5,000 inhabitants; second line: 5,000 - 50,000.
- Estimates for Swiss towns with direct democracies computed from Pommerehne's table 4.
- This number is recalculated from information given in Vehorn's table 2. Vehorn incorrectly calculates the crowding parameter as $\alpha = \theta / \delta$ (our notation). In his case, however, $\alpha = (\theta + 1) / (\delta + 1)$ is correct.
- This is a corrected figure for Zimmerman's OLS estimate. Because of an error in his equations (4) and (5), Zimmerman calculates the crowding parameter as $\alpha = \theta + \delta + 1$ (our notation) which explains his finding of $\alpha = 0.625$. In his case, however, $\alpha = \theta / (1 + \delta) + 1$ is correct.
negative and rather inelastic. Estimated income elasticities are usually positive but smaller than unity.

Once d, e, and the population elasticity are estimated, the crowding parameter can be calculated from the definition of q or q*, respectively. Table 1 gives an overview not only of the crowding estimates of BD and BG but also of subsequent related papers. It is worthwhile commenting on some of these papers in so far as the methodology differs from the initial BD-BG approach. Vehorn (1979) looks at public expenditures on fire services. In contrast to BD-BG, he allows the price of private substitutes (insurance, etc.) to have an effect on the demand for public fire protection. While cross price effects between the public service and private substitutes do seem to exist, this extension proves to be largely neutral with respect to the conclusions about the publicness parameter. McMillan, Wilson and Arthur (1981) divide their sample of Ontario communities into those with less than, and those with more than 10,000 inhabitants. They find a significant degree of publicness in smaller communities, while the results for larger communities are more like the BD-BG results. The interpretation of these authors is that replication of certain facilities is only profitable in larger cities. In particular, cities with only one fire station show little crowding, while the publicness disappears for large cities with many fire stations. This comparatively large publicness for small communities, however, is not confirmed by an earlier study (Pack and Pack 1978) which looks at a subsample of communities with fewer than 5,000 inhabitants, as well as by a later paper (Edwards 1986) which uses a cut off level of 10,000. Hayes (1985, 1986), among other things, allows for cross price elasticities between various publicly provided goods. Except for Borcherding and Deacon (BD), Zimmerman (1983) is the only study which considers general expenditures at the US state level. The Holcombe-Sobel (1995) study considers only legislative expenditures at the US state level and is one of the very few studies which report significant economies of scale.8

8There is a very limited number of studies which consider economies of scale at the federal level. Dudley and Montmarquette (1981), who apply the BD-BG approach to national defense by imposing a concrete utility function, conclude that defence is not distinguishable from a pure public good. For a discussion see Borcherding (1981).
Overall, however, the main message derived from table 1 is that most of the follow-up studies agree with the BD-BG result. They provide little evidence that local government provide services which imply economies of scale in consumption.

2.3  How much Publicness Should We Expect?

If optimal city size were solely determined by the optimal sharing of public services, economic theory would lead us to expect a crowding elasticity of unity. A lower crowding parameter would make it worthwhile for a city to attract additional citizens in order to reduce per capita cost, while cities with a crowding parameter larger than unity could reduce per capita cost by shrinking.9

In a more complicated model where communities supply a variety of services, each of which requires a different optimal city size, some of the services supplied will have crowding elasticities higher than unity, and some lower than unity.10 In reality, however, optimal city size is largely determined by complementarities in production and consumption and by negative agglomeration effects such as increased environmental problems. In other words, optimal city size is determined by externalities and public goods in a very broad sense and not only by the goods provided by local governments. It is therefore perfectly possible to construct a spatial equilibrium where pure public goods are supplied.11

Since economic theory fails to make clear predictions, a priori expectations can only be based on plausibility considerations. These suggest that the expenditure categories investigated in the local public good literature are certainly not pure public goods in the Samuelsonian sense. Some are very likely to be private goods, for example local education. Since there are obvious limits to class size, beyond a certain threshold doubling students implies doubling teachers, rooms, etc. This does not mean that there are no externalities

---

9Note that the BD-BG approach is unsuitable for investigating optimal city size. From equation (2), the crowding elasticity is a constant and exogenously given. From the public goods perspective, the optimal city then comprises either the total population (for \( \alpha < 1 \)) or only one person (for \( \alpha > 1 \)).

10This consideration can be generalized to allow for different levels of government. As long as there are fewer levels of government than public goods, the basic conclusion is unchanged.

11See, for example, Fisch (1977). He assumes that living in a city implies travelling to the city center where production takes place. This is increasingly costly as the city grows larger.
related to education.\textsuperscript{12} The literature on endogenous growth strongly argues for positive external effects of education, but these are related to the average level of education in a society or community, and do not provide any economies of scale at the community level. The situation may be less clear for other expenditure categories. Providing police services of constant quality, \( q \), for a city that is twice as large and has twice as many inhabitants obviously requires more resources, but how many more? If the resources required rise by more than 41 percent, which seems likely, from equation (2) the crowding parameter is greater than 0.5, i.e., closer to a private than to a pure public good. We would expect this for all expenditure categories, with the possible exception of legislation and defence. The empirically interesting question is not whether local public services are purely public. It is whether public services imply some economies of scale.

2.4 Understanding the (Non-)Publicness Results

As outlined above, the publicness parameter \( a \) is estimated as a function of the price and the population elasticity. In the BD formulation, for example, we have [cf. equation (5-BD)]

\[
\alpha = \frac{\theta}{1 + \delta} + 1, \tag{6}
\]

where \( q \) is the elasticity of per capita expenditure with respect to population, and \( \delta \) is the price elasticity of demand. The empirical studies usually find that the latter is rather small, \(-0.4 \leq \delta \leq -0.2\), and that population size has only a minor, and more often positive than negative, impact on per capita expenditures (\( \theta \approx 0 \)). As a result, \( a \) is estimated as being close to unity, and often above rather than below unity.

More intuitively, the logic of the argument is as follows. If there were substantial publicness, an increase in the number of inhabitants would lower the tax price of the publicly provided good as the cost is shared among a larger group. Since demand is inelastic, per capita expenditure for the public good would decline with population. The non-publicness results arise because the data do not show this decline.

\textsuperscript{12}For an empirical study of positive externalities of education cf. Wykoff (1984).
An immediate suspicion that arises is that the BD-BG approach estimates a publicness parameter of greater than or equal to one since, for reasons outside the model, per capita expenditures tend either to be unaffected by, or to increase slightly with population. A potentially important issue in this respect is the BD-BG assumption that population has an effect on per capita expenditures only through its effect on voters' tax price (which in turn depends on crowding effects). This excludes systematic relationships between population size and preferences for public goods, spillover effects between communities, and other issues which will be discussed in the next section.

3 The Interpretation of the Population Elasticity

3.1 Agglomeration Externalities and the Relevance of the Crowding Parameter

As mentioned in the introduction, the empirical results of the BD-BG literature tempted several authors to conclude that the public sector is providing goods which could be provided optimally by competitive markets. Such a conclusion is premature. It has been correctly indicated that, even with sufficiently high crowding effects, a pure market solution is unsatisfactory if exclusion is not possible. In this case private providers do not have enough opportunities to learn about consumers' willingness to pay and the free rider problem precludes an efficient solution.

It is important to notice, however, that, even if firms can exclude customers from receiving benefits without incurring costs, a serious problem may remain. It will appear in those cases in which services are needed to compensate for negative agglomeration effects. To make the point, consider the case of police services. The consumption good $q$ which ultimately enters the individual's utility function is safety. The safety level certainly depends on the number of policemen, $G$, the number of citizens, $N_e$, protected by these officers and a

---

community characteristic $V$.\textsuperscript{14} $V$ measures the potential violence of the community and will depend, among other things, on city size, $N$.

Expressing $q$ as a multiplicative function gives

$$q = G \cdot N_u^{-\alpha^*} \cdot V^{-\gamma},$$  \hspace{1cm} (7)

where

$$V = Z^\beta \cdot N^\lambda.$$  \hspace{1cm} (8)

Potential violence, $V$, of a community depends on a vector of socioeconomic factors, $Z$, and population $N$. In the case of public provision of $G$, $N_u$ equals $N$. Hence, from (7) and (8) we have:

$$q = G \cdot N^{-\alpha^* - \gamma \lambda} \cdot Z^{-\gamma \lambda}. $$  \hspace{1cm} (9)

Using the BD-BG approach to measure the crowding parameter $\alpha$ captures the total effect of $N$ and $\alpha = \alpha^* + \gamma \lambda$. Note, however, that in the case of a private firm which considers providing policemen to secure $N_u$ customers, it is $\alpha^*$ rather than $\alpha$ which governs the relationship between a firm’s marginal and average costs. The total number of citizens, $N$, and therefore $V$ will be exogenous to the single firm at least in a competitive market. In case $\gamma \lambda > 0$, economies of scale for any single security firm ($\alpha^* < 1$) apply even though the overall crowding parameter $\alpha$ equals unity. A perfectly competitive market is precluded even though exclusion is possible.\textsuperscript{15}

Of course, the above argument uses a restricted form of exclusion. Customers can be excluded from receiving benefits but citizens cannot be stopped from exerting negative externalities on the production of safety. An internalization of this negative externality is

\textsuperscript{14}Using a terminology developed by Bradford, Malt and Oates (1969), safety is a consumption good (C-good) and $G$ is a (publicly) provided ‘direct’ input in the production of safety (the D-good). The C-good is produced by the D-good and the community characteristic, $V$. This structure is also used in Baum (1986), Schwab and Zampelli (1987) and Dynarsky, Schwab and Zampelli (1989).

\textsuperscript{15}The extent to which efficiency in a market solution will fall short of the first best solution is open to debate. For a very optimistic view see Baumol, Panzar and Willig (1982).
achieved if all citizens belong to the same security provider. Note that this is exactly what we observe in the case of public provision.

Public safety is an intuitive example of the fact that constant economies of scale at the city level do not preclude economies of scale for firms. Similar effects can be expected in a wide range of other cases where public services are provided to reduce the negative effects of agglomeration. Fire protection, sewerage and parks are natural candidates.\textsuperscript{16}

3.2 Economies of Scale, Number of Varieties, and Tiebout Bias

A feature of the BD-BG approach which enormously facilitates its empirical implementation is that it dispenses with a physical measurement of the publicly provided goods. Oates (1988) emphasizes that this is, at the same time, a serious drawback. The problem comes from the fact that the varieties of services provided in a larger town exceeds those in smaller communities.\textsuperscript{17} While the expenditure category 'parks and recreation' will not include a zoo for communities with less than 10,000 inhabitants, this may well be the case for larger cities. In this case, however, the lower expenditure of smaller communities does not come from less crowding but simply from the fact that a zoo needs a certain minimum size to make sense. In other words, there are considerable indivisibilities which preclude provision of certain services in small towns. Interpreting the larger range of services in larger towns as a crowding cost overestimates the crowding parameter.\textsuperscript{18}

\textsuperscript{16}It should be noted that the relationship between population size and the environment variable $V$ is not always a direct one, but may come about through a correlation of $N$ with socio-economic characteristics such as education, race, etc. In this case, the term $\gamma_a$ affects the estimate of the population elasticity only insofar as the relationship is not already captured by the vector of socio-economic variables $\mathbf{Z}$, included in the estimated demand equation. However, large externality effects of population size can be expected to show up even after controlling for socio-economic characteristics. Again considering our example of police protection, it seems plausible that anonymity in larger communities reduces social control mechanisms, reduces the probability of punishment, and thereby increases the tendency to commit crimes. It also induces potential criminals to migrate to larger cities. These effects are not likely to be fully captured by observable socio-economic characteristics.

\textsuperscript{17}Even before the BD-BG literature developed, a similar argument had been put forward by Schmandt and Stephens (1960) who also provide empirical evidence that larger towns provide more varieties. In the German literature, the discussion of why larger communities have larger per capita expenditures dates back to Brecht (1932). The empirical observation of a positive correlation between population and per capita expenditure is known as Brecht's law. For further discussion see Kaebler (1982) and Keller (1969).

\textsuperscript{18}McMillan (1989) claims to provide evidence in favor of Oates's hypothesis, but his results are based on an incorrect application of the bias correction formula in Borcherding and Deacon (1973, p. 899): to make sure that the binomial expansion, on which the bias correction is based, does actually converge (cf. footnote 21 of Borcherding and Deacon), the divisor must be applied to the deviation of $\alpha$ from unity, not to $\alpha$ itself. Without bias correction, McMillan's estimates indicate a relatively high crowding parameter, cf. Table 1.
The fact that larger communities offer a wider range of services may lead to consumers' self-selection. Intuitively, consumers who plan to use public facilities more intensively may be inclined to move to larger cities which provide a more diverse supply of services. Again, if true, this would lead to a reinterpretation of existing crowding estimates. Indeed, the possibility of a Tiebout bias has increasingly found attention among researchers. So far, however, the theoretical concern of the relevant literature\textsuperscript{19} has concentrated on the reliability of the estimated income elasticity and empirical studies which highlight the Tiebout bias problem are exclusively concerned with local education.\textsuperscript{20}

3.3 Spillovers

While a series of authors acknowledge that publicly provided goods are often non-excludable, little of this is reflected in the BD-BG type models. Usually, the implicit assumption is that only local residents benefit from locally provided goods and services.

A simple conceptual way to analyze the impact of spillovers is to distinguish between residents, \( N \), and effective users, \( N_e \). Define \( f \equiv N_e / N \) and assume that there is a systematic relationship between \( f \) and \( N \), \( f(N) = N^n \). The wide literature on the exploitation of central cities by their suburbs suggests \( n > 0 \).\textsuperscript{21} This has an immediate bearing on the effective units available to the single user. Instead of (2), the BD-BG crowding function now has the form

\[
q = \frac{G}{(f \cdot N)^\alpha} = \frac{G}{N^{(r+1)\alpha}}. \tag{2'}
\]

Using the BG assumption that the median voter's tax share \( t \) equals his share of housing value in the community \( (H^m / P) \), this leads to a demand function for public expenditures which is almost identical to equation (5-BG):

\[
E = k \cdot (H^m / P)^\delta \cdot p^{\delta+1} \cdot Y^\varepsilon \cdot N^\theta \cdot Z^\delta.
\]


\textsuperscript{20}Cf. Reid (1987) and Rubinfeld, Shapiro and Roberts (1987). These studies use micro survey data to estimate demand for local school expenditure. Other papers trying to estimate demand for local schooling from micro data are Bergstrom, Rubinfeld and Shapiro (1982), Rubinfeld and Shapiro (1989) and Nold (1992).

\textsuperscript{21}For a survey see Bradford and Oates (1974). While the exploitation hypothesis is disputed, the existence of considerable spillovers is not. For a survey of spillovers between German ‘Mittelzentren’ (medium sized central cities) and their ‘Umlandgemeinden’ (surrounding communities) see Voß (1991, pp. 63-65).
This time, however, the exponent of $N$ is $\hat{\theta} = \alpha(v+1)(1+\delta)$ instead of $\theta^* \equiv \alpha(1+\delta)$ and the crowding parameter is therefore given by:

$$\alpha = \frac{\theta}{(v+1)(1+\delta)}.$$  \hfill (10)

Implicitly, empirical estimates have assumed $v = 0$. According to (10), the impact of a non-zero spillover parameter $v$ on $\alpha$ depends on the price elasticity of demand ($d$). For the empirically relevant case $|\delta| < 1$ we can infer the following. If the ratio of effective users to population ($N_e/N$) is larger the larger the community, then $v$ is positive and the implicit assumption $v = 0$ in previous studies overestimates the crowding parameter $\alpha$.

4 The Estimation of the Price Elasticity

It is widely perceived (Oates 1995, p.8) that one of the robust findings of the BD-BG literature is a very low price elasticity of demand for publicly provided goods, in most cases between -0.2 and -0.4. Folk wisdom says that very low price elasticities are usually related to goods serving basic needs, such as food, and since not all publicly provided goods belong to this category, the empirical results may appear somewhat surprising. It seems justified to examine the reliability of these estimates more carefully.

How does the estimated price elasticity $\delta$ affect the congestion estimates? From the discussion of equation (6) we have seen that the privateness results come from fact that $\theta$ is approximately zero. Therefore, the estimate of the price elasticity $\delta$ has only a small numerical effect unless it is close to -1. The qualitative result that $\alpha$ is smaller or greater than 1 changes if and only if the estimated price elasticity switches from elastic to inelastic, or vice versa.

4.1 Problems of Measuring the Tax Price

Of all the explanatory variables in the BD approach, the tax price is certainly the most difficult to measure. In Section 2.2 we saw that there are two different approaches. In the BD study, differences in tax price between communities arise through variations in the tax share
of the median house owner, \( t \). In the BG approach, price differences are due to different wage rates.

Interestingly, the two approaches lead to very different results. The above mentioned stylized fact of low price elasticities is found by studies in the BG tradition, while studies in the BD tradition come up with wildly fluctuating estimates of price elasticities. Tables 1 to 3 of BG provide estimates of tax share elasticities for 30 different states and several services. All except three estimates are smaller than 0.6 in absolute value, the maximum being 0.81. In contrast, the price elasticities in Tables 2 and 3 of BD fluctuate wildly between -4.7 and +2.4. The same pattern is found in subsequent studies. Hayes (1985, 1986) and Perkins (1977), using cost data to measure the tax price, come to varying, and often high, estimates of price elasticities. Most other studies listed in Table 1 follow BG and measure the variation in the median voter tax share. Almost all of their estimated price elasticities are below 0.6 in absolute value, with the interesting exception of Pommerehne (1978) which we will discuss below. Studies which consider both wage and tax share differentials [Gramlich and Rubinfeld (1982), Santerre (1985), Vehorn (1979) and Zimmerman (1983)] derive results which are similar to the BG-type studies. It seems that the variation in tax shares drives the results, and that the specification error from neglecting differences in wage rates is of minor importance.

A possible explanation for these conflicting results are measurement error problems. Bergstrom and Goodman (1973) have already investigated this problem. They provide estimates based on different assumptions about the size of the measurement error, but conclude that the absolute price elasticity is smaller than unity even under rather extreme assumptions. As shown above, such changes would only have small numerical effects on the estimate of the crowding elasticity. Bergstrom and Goodman's results were obtained, however, under the assumption of "classical", unsystematic measurement errors. Effects of systematic errors may well be much stronger.

\[22\] The application of this approach is problematic in countries where the property tax accounts for only a small fraction of total community revenues. In a German study, Bothe (1989) therefore uses the general income tax to calculate the median voter's tax share. Unfortunately, German communities have no control over income tax rates. An earlier study of German municipalities (Pross 1982) assumes that all voters bear identical tax shares.
Problems of this kind may be responsible for the erratic results of studies in the BD tradition. For example, their price measurements are based on very stringent assumptions about the production technology (cf. fn. 3) and the quality of labor differentials across communities (BD, fn. 16). Moreover, the costs of publicly provided goods include the rental costs of land and property used in the production process. To the extent that these factors are publicly owned, the costs are opportunity costs and do not appear in the expenditure data, i.e., on the left hand side of the estimated equation. Neither do differences in the rental costs of land appear in the cost measures of BD or BG. If land prices are correlated with city size, this may tend to confound price and population effects. The overall effect of the land price omission will depend on whether governments tend to own or rent land.

The higher consistency of the BG results makes their price elasticity estimates appear more reliable than those of the BD approach. However, the measurement of the median voter's tax share poses a host of methodological problems, many of which have been addressed in the literature. We organize our discussion under three headings: measuring the median voter's tax share, finding the right tax base, and measuring tax distortions.

The median voter's tax share

It is useful to write the median tax share as \( t = \frac{\tau H^m}{\tau(SH + C)} \). It equals the tax on the median house value, as a proportion of all property tax, \( \tau \times P \), which is the sum of the property tax on private homes, \( H \), and on commercial and industrial buildings, \( C \). Variations in \( t \) come from three sources: variations in the number of residents, variations in the median home relative to the average home, and variations in the fraction of commercial and industrial property in total property. The first effect is already captured by including \( N \) in the regression, which leaves the second and the third effect. The second effect is rather subtle: higher order characteristics of a distribution, such as skewness, on which the difference between median and average depend, are difficult to measure. Measurement errors are likely

---

23For simplicity, we assume here a common tax rate on house value and commercial property. Allowing for different tax rates would not change the argument.
to arise and they tend to bias the estimated price elasticity towards zero (Schneeweß and Mittag 1986, p.140).

The third effect is the most important empirically. According to Bergstrom and Goodman (1973, fn. 8), variations in $t$ stem mainly from differences in the amount of commercial and industrial property. In communities where industrial property bears a high share of the property tax, the median voter faces a low tax price since she can “exploit” the owner of the industrial property, who may be a resident or a non-resident. The consistently and significantly negative estimates of the tax share elasticity demonstrate that there is a robust effect of this kind.

The problem with this approach is that there are obvious limits to the exploitation of capital owners which are not properly taken into account. There are three ways in which owners of nonresidential property may avoid paying for public goods they don't want. First, and contrary to the median voter model, they may try to influence the political process through their connections with the local bureaucracy. The study of Aragon et al. (1988) has found empirical support for this hypothesis. For French cities in the Toulouse area, it separates data on the number and income distribution of people who pay a “professional tax” in addition to property taxes on housing and those who do not. They reject the “democratic” hypothesis that both types of residents have the same influence on the choice of the tax rate: those who pay the professional tax have a greater influence, and this effect is stronger in smaller cities.

Secondly, if owners of commercial buildings have some monopoly power they may pass on part of their property taxes to local customers. This would mean that this part of the tax is actually borne by the median voter. Ladd (1975) accounts for this effect by assuming that the median voter bears a constant fraction of the taxes on commercial and industrial property. Estimating shift parameters to identify this fraction, she obtains price elasticities that are about twice as large as those without the corrections for tax shifting. This is also found by Vehorn (1979), who adopts her procedure.

The third, and probably most important obstacle to the exploitation of owners of industrial property, is tax competition. High rates of property taxes, which are used to pay for
public goods that do not increase the productivity of these capital assets, may lead to
migration of capital. Ladd (1975) claims to have captured this effect, called the price-
migration effect, through the tax shifting parameters. Ladd's model, however, is inadequate
for this purpose, since it assumes that the median voter bears a constant fraction of taxes on
nonresidential buildings. This misses the most significant effect of tax competition, which is
to make the budget share of the median voter a function of budget size. A larger budget
implies higher tax rates, which leads to migration of capital, so that a higher share of the
budget has to be borne by the median voter. As a result, the tax price of public goods is
nonlinear for the median voter.

The effect of tax competition on estimates of price and crowding elasticities is not
easy to work out. In the appendix, we show that tax competition does not affect the estimates
if the tax share depends iso-elastically on the budget size, but the iso-elasticity assumption is
critical. The empirical relationship between budget size and the tax share of the median voter
is also difficult to establish. Contrary to our theoretical expectations, Zimmerman (1983)
finds a negative relationship and explains it using political economy arguments.24 The
situation becomes even more complex if we take into account the fact that tax competition
affects smaller communities more than larger ones, since their smaller tax base allows them to
attract relatively more capital through low tax rates. This effect would lead to a bias in the
estimation of the population elasticity. Summing up, the effect of tax competition on the
crowding estimates depends on the exact nature of the relationship between tax rates and
capital movements, and has to be regarded as unknown at the moment.

Finding the right tax base

Property taxes are not the only source of revenue for local communities, and a comprehensive
measure of the median voter's tax costs have to include other types of taxes, in particular the
income tax. Vehorn (1979) finds that taking income taxes into account leads to higher price
elasticities, but his estimates remain well within the inelastic range. Tax exportation effects,

24The demand function estimates in Zimmerman (1983) do not take into account the way the relationship
between budget size and tax share affects the budget constraint of the median voter.
for example those arising through the deductibility of local taxes from the federal tax base, have been investigated, e.g., by Zimmerman (1983). They lead to only minor changes in the results.

*Tax distortions*

Another problem of measuring the adequate tax price was addressed by Wildasin (1989). Since the property tax may distort the demand for housing, the observed tax price underestimates the cost of public services. The true costs also include the excess burden of taxation. Wildasin shows that the omission of this excess burden leads to an overestimation of the price elasticity for public services if the measured elasticity is inelastic as it is in most studies. In this case, higher prices imply a more than proportional increase in the (omitted) excess burden. Although Wildasin does not discuss it explicitly, such an overestimation of $\delta$ may lead to an overestimation of the crowding factor. However, since the upper range of overestimation for $\delta$ according to Wildasin is some 16% to 25%, the impact on the crowding parameter is also very small, as our discussion above has shown.

Crane (1991) builds on Wildasin's analysis and considers the additional effect that changes in government spending might induce migration and affect property prices. This is the idea we have discussed above, only applied to home owners rather than owners of nonresidential buildings. Crane finds (p.102) that estimates of the price elasticity are biased, and that the bias can be positive or negative.

Summarizing the discussion, we can say that the exact measurement of the price elasticity of demand for public services poses difficult problems which have not yet been fully solved in the literature. At the same time, there is no clear indication that further advances in any of these directions will change the basic non-publicness results that we are mainly interested in.

*4.2 Fiscal Illusion*

The BD-BG approach assumes that voters understand the mechanics of public finances in their community perfectly. There is a substantial literature, partially overlapping with the BD-
BG literature, which investigates this assumption. Among the reasons why voters may systematically misperceive the costs of public services, the three most important ones are probably renter illusion, the flypaper effect and revenue complexity.

The renter illusion hypothesis argues that renters, who bear property taxes only indirectly, underestimate the tax price of public services and, therefore, opt for higher public expenditures. BG and some other articles in the same tradition therefore included the proportion of renters as an additional explanatory variable in the regression equations, mostly confirming the hypothesis.

The flypaper effect states that lump-sum grants tend to reduce the perceived tax price, since these grants are used to pay for public services and bureaucrats tend to conceal their lump-sum nature from the public. Finally, the revenue-complexity hypothesis states that a more complex revenue system makes voters underestimate tax prices and leads to higher demand for public goods.

The empirical relevance of all three effects is debated and their importance for the measurement of publicness cannot be assessed without knowledge of the exact nature of the induced distortions. For example, if the perceived tax price is proportional to the true tax price, the estimated price elasticity and therefore the crowding parameter would not be affected. A detailed discussion of fiscal illusion is beyond the scope of this paper. The interested reader is referred to the survey by Dollery and Worthington (1996).

4.3 Direct versus Indirect Democracies

While the low price elasticity estimates may be due to problems of measuring the tax price variable, the study of Pommerehne (1978) brings up the more fundamental question of whether these estimates are biased because the political process is not correctly described by the median voter model. We will discuss the validity of the median voter framework more broadly in the next section, and consider here only the implications of different political systems for the price elasticity measurement.

Investigating seven expenditure categories for four types of Swiss communities and using tax share data, Pommerehne finds price elasticities of about 1 or greater in absolute
value in about one third of all cases. This study is different from the US studies in two respects. First, tax shares refer to personal income tax shares “because this tax represents the main fiscal source of Swiss municipalities” (Pommerehne 1978, p. 278). Arguably, these data are more accurate than tax share estimates based on median house value estimates. Second, he distinguishes four different categories of cities according to their institutional structure (cf. Section 5 for details). Cities with direct democracy and obligatory referendums tend to have higher (absolute) price elasticities.

5 The Validity of the Median Voter Model

The discussion of the Swiss example has shown that institutional differences matter for the estimation of price elasticities. In this section we consider the extent to which the simplifying median voter assumption is justified, and how it affects the estimates of crowding elasticities.

5.1 Statistical Tests

5.1.1 Consistent Demand Systems?

Since by assumption the median voter decides on quantities which maximize her utility, one way to check for consistency is to analyze whether estimated demand systems actually comply with utility maximization.\(^{25}\) Important requirements for demand systems derived from utility maximization are the absence of money illusion (homogeneity in all prices and income) and the symmetry of substitution effects. To test for these properties, demand systems must be estimated which allow for cross price effects between publicly provided goods. A few papers have actually done this. The results are mixed and suffer from the fact that only police and fire protection services imply significant cross price effects. Grosskopf and Hayes (1986) show that estimates for a local public good demand system are not fully consistent with utility maximization. Most requirements for utility maximization are violated (e.g. absence of money illusion). In another study, Hayes and Grosskopf (1984) show that a model which does not

\(^{25}\)Unlike in markets for private goods, this precludes aggregation problems which may prevent total demand from being consistent with utility maximization.
presume utility maximization does a better job in explaining empirical budget shares. Deacon (1978) cannot reject the homogeneity assumption but finds that the symmetry postulate is not satisfied in his data. McMillan and Tuffour (1988) reject both symmetry and homogeneity. DeBoer (1986) tests for consistency of revealed preferences and finds demand broadly consistent with constrained utility maximization.

5.1.2 Median versus Average Income
Evidence in favor of the median voter model was found by Pommerehne and Frey (1976). These authors compare demand estimates for the BD-BG approach with estimates derived from using average income and average tax price instead of median values. Indeed, the BD-BG type estimates statistically outperform the alternative if both approaches are applied to the same data set of Swiss communities with direct democracies.26

5.2 The Political Process and the Measurement of Publicness
The validity of the median voter model can be questioned for communities where the quantity of public goods is determined in a bureaucratic process rather than by direct referendums. Walzer (1972) provides anecdotal evidence that bureaucratic decisions frequently rely on simple per capita rules which tend to keep per capita expenditures constant. If this were true on a grand scale, it would explain why $q$ in equation (6) is estimated close to zero and a close to unity.

Theoretical models of bureaucratic behavior, however, do not predict constant per capita expenditures. Borcherding, Bush and Spann (1977) argue that bureaucrats can influence the amount of goods provided. Because of voters’ rational lack of interest, this effect is stronger, the larger the city.27 Therefore, the larger public expenditures in larger cities may result not only from crowding effects, but also from the greater influence of the bureaucrats and their efforts to maximize the budget.28

---

26For papers which test whether the median income household is the decisive voter see Inman (1978) and Aronson and Wikström (1996). However, recall that the BD-BG approach does not claim that the median voter necessarily earns the median income (cf. fn 4).

27Such an assumption is in line with the political theories of Downs (1957).

28Additionally, the authors (p. 216) note that bureaucrats might wish to switch toward the provision of private goods since their monopoly power is larger than with public goods.
Unfortunately, the situation is not so simple. While the literature on bureaucracy is unanimous in its prediction that reduced democratic control over bureaucrats in larger communities tends to lead to higher expenditures, it is not clear what this means for the estimation of the crowding parameter. Gonzales and Mehay (1985) and Wyckoff (1988) explore this issue in a model where governments maximize slackness, defined as the tax revenues minus the minimum cost of providing the public services, rather than trying to maximize their budget. The local government is seen as a monopolistic provider of services which is only restricted by the citizens' outside option to migrate, and which can extract the citizens' total consumer surplus.\(^{29}\)

In the model of Gonzales and Mehay, the political process causes public expenditures on private goods to grow proportionally with population, while expenditures on a public or quasi public good (i.e. \( < 1 \)) will grow more than proportionally. The intuition is the following. Assume that A is a community with twice as many inhabitants as B. If A raised the same per capita tax as B and provided exactly twice as many goods as B, the imperfect crowding would leave town A's citizens with a rent. It is optimal for A's bureaucrats to react by raising the tax rate to appropriate this rent, thereby increasing per capita expenditures. The empirical estimates of Gonzales and Mehay do not support the existence of quasi public goods, since per capita expenditures do not significantly rise with population. The Gonzales/Mehay argument, like the BD-BG approach, sees constant per capita expenditures across jurisdictions as a confirmation of privateness, although for totally different theoretical reasons.

Wyckoff's approach follows Borcherding, Bush and Spann in assuming that bureaucrats in larger cities have greater power over the budget. This idea is incorporated in Wyckoff’s model by assuming that citizens in larger communities have a lower reservation utility. This is equivalent to an outward shift of the demand function for public services and leads to higher expenditures in larger communities. Wyckoff then continues by deriving the surprising result that, if his model is the true one, estimation of the BD-BG model will lead to

\(^{29}\)Total rent extraction comes from the bureaucrats' ability to make an all or nothing offer. Governments are not assumed to price discriminate.
an underestimation, rather than an overestimation of the crowding parameter \( \alpha \). However, this effect is closely related to the fact that, in his model, the elasticity of expenditures with respect to population, with a constant tax share of the median voter\(^30\), is negative. This elasticity conforms to the expression \( \alpha(1+\delta) \) in equation (4a). Since this term is typically estimated as positive in the BD-BG studies, the correct interpretation is not that the crowding parameter is underestimated in these studies, but that the results reject Wyckoff’s model.

The existing literature, therefore, does not give a definite answer on how bureaucracies influence the BD-BG results. It remains unclear whether bureaucrats tend to maximize slackness or the budget size. Presumably, the stylized models of bureaucratic behavior are just as extreme in their assumptions as the median voter model. In representative democracies, where large bureaucracies exist but are monitored by politicians and the public and expenditure decisions are subject to complicated budgeting procedures, a hybrid model is probably more realistic than either of the extreme assumptions.

Since the theoretical situation is unclear, an important avenue for evaluating the impact of the political process on crowding estimates is to look at empirical estimates under different institutional environments. If bureaucracy is responsible for an upward bias in the crowding measurement, then communities with direct democracies should exhibit a higher degree of publicness. Swiss municipalities, displaying a wide variety of institutional structures, provide an ideal object of study. Pommerehne (1978) groups the 110 largest Swiss cities into the following four categories: (a) direct democracies with obligatory referendums, (b) direct democracies with optional referendums, (c) indirect democracies with obligatory or optional referendums, or (d) indirect democracies with no referendums. While Pommerehne does not list crowding parameters, these can be calculated from his price and population elasticities. Unfortunately, no clear pattern emerges. For general administrative expenditures, the crowding elasticities for the four categories are a) 0.83, b) 1.07, c) 1.04 and d) 0.84. Crowding estimates are small in direct democracies with obligatory referendums, but also in indirect democracies with no referendums. For other expenditure categories, the picture is

\(^30\) Since the tax share of the median voter decreases inversely with population size, expenditures, of course, still rise with population.
even more blurred. This does not mean, however, that institutional differences did not matter. The study shows that income and price elasticities are significantly higher in direct democracies than they are in indirect ones. Also, parameter estimates tend to have a higher significance level in direct democracies. Direct democracies with obligatory referendums are the only category of cities with a consistently significant price elasticity.

Together, these results raise doubts about the extent to which the median voter model, if applied to indirect democracies, really measures voters' demand characteristics. The doubts are reinforced by a series of studies which use estimated price, income and population elasticities to explain budget increases over time. Borcherding (1977, p. 56) finds that the cross-section elasticities can account for only some 50% of US public sector growth between 1900 and 1970. Gramlich (1982) also finds an explanatory gap between cross-section estimates and time series developments, but he concludes that this is not necessarily a reason to mistrust the BD-BG results. Preferences for publicly provided goods may have changed over time, and Borcherding's use of local demand parameters to explain total public sector growth is rather crude.

These objections are overcome in a study by Pommerehne and Schneider (1982). Using 1965 data, they derive BD-BG type demand estimates for Swiss cities with direct democracy. In a second step these estimates are used to predict local expenditures ten years later. Unlike Borcherding, they obtain strikingly accurate estimates. Repeating the same approach for Swiss cities with indirect democracies, however, gives very imprecise predictions for 1975 expenditure. Again, price elasticities prove to be much higher in direct democracies. Moreover, if demand elasticities derived from direct democracies are used to predict expenditure growth of indirect democracies, expenditure growth is systematically underestimated. The results suggest that bureaucrats use their political power to increase expenditures.

\[^{31}\text{Unfortunately, it is not possible to calculate the usual crowding parameter from the information provided by Pommerehne and Schneider.}\]
6 Congestion: The Role of the Functional Form

The studies in the BD-BG tradition use the congestion function \( q = XN^{-\alpha} \) which implies a constant elasticity of congestion, \( (\partial q / \partial N) / (q / N) = -\alpha \). This specification is obviously very restrictive. One of its implications that has been questioned in the literature is that, with rising population, marginal congestion \( \partial q / \partial N \), measured as the (negative) change of the units consumed by an intramarginal inhabitant, is decreasing in absolute terms: \( \partial^2 q / \partial N^2 > 0 \).\(^{32}\)

This is unacceptable for Craig (1987, p. 338-9), who considers increasing congestion as an essential aspect of public goods. He bases his argument on Buchanan's theory of club goods, and on empirical studies in transportation economics (Inman 1978). We have argued elsewhere (Reiter and Weichenrieder 1997) that increasing marginal congestion is an unnecessarily stringent assumption that is not strongly related to economically interesting aspects of crowding functions. The discussion of the constant elasticity crowding functions should therefore not be tied to the question of increasing or decreasing marginal congestion. Constant elasticity is unsatisfactory for more general reasons, because it a priori excludes the possibility that economies of scale, and therefore publicness, depend on the population size. It thereby implies that the optimal city size is either infinite (\( a < 1 \)) or infinitesimally small (\( a < 1 \)). It seems much more plausible that certain goods may be public for small cities and private for large ones, while between the two regimes there is an optimal city size which delivers public goods at minimal average costs. It is therefore necessary to study less restrictive functional forms.

Edwards (1990) systematically investigates the issue of functional form within the median voter framework. He estimates five different congestion functions for three expenditure groups (general administrative expenditures, parks and recreation, and police protection), using data on 78 New York State municipalities with populations between 10,000 and 150,000. For each of the three expenditure groups, the BD-BG congestion function and a

\(^{32}\)Note that the marginal (physical) externality, \( d\left( N \cdot \frac{dq}{dN} \right) / dN \), to all existing inhabitants also decreases in \( N \) if \( 0 < \alpha < 1 \).
flexible exponential function seem to outperform the alternatives. While neither specification can be rejected by specification tests, the exponential function seems to explain the data slightly better. Both functional forms lead to very similar estimates of price and income elasticities. However, the exponential function leads to different estimates of the population elasticity, and thereby to different estimates of the crowding parameter. While the results with the BD-BG function indicate that all three goods are private goods in all the communities considered, the exponential function indicates the exact opposite, that is, none of the three expenditure groups is a private good in any of the communities. Another important finding is that the estimates with the exponential function - which in principle allows decreasing as well as increasing marginal congestion - confirm the BD-BG assumption of decreasing marginal congestion and contradict Craig's argument.

This sensitivity of crowding parameter estimates to functional form has been confirmed by other studies. McKinney (1987, p. 319) finds that the estimated publicness of fire protection services is considerably lower with the BD-BG form than with an alternative specification. Hayes and Slottje (1987) find statistical evidence favoring the BD-BG formulation over a simple exponential function, but they also confirm the sensitivity to functional form. McGreer and McMillan (1993) estimate the BD-BG function, Edward's exponential function and Craig's function (cf. Section 7) for three expenditure categories (total expenditures, recreation and culture, and road maintenance) with data on Australian municipalities. While their results are sensitive to functional form, they cannot confirm Edward's finding that the exponential function yields systematically lower crowding parameters.

Edward's results are most seriously challenged by a study by Means and Mehay (1995), who draw attention to the fact that the exponential form, general as it is, does not encompass the case of a purely private good. They show that Edward's findings may well be an artefact of imposing this functional form on a good that, in reality, is purely private. In their own empirical application, using data on four expenditure categories for Californian municipalities with 25,000 to 250,000 inhabitants, they slightly generalize the exponential form so as to include the private good case. They find that the more flexible functional forms
do not have significantly more explanatory power than the traditional specification. They cannot reject the pure public good hypothesis for any of the expenditure categories except one (sanitation).

These results do make it appear unlikely that the non-publicness findings are mainly an artefact of imposing a too restrictive functional form. The flexible functional forms studied so far did not prove superior to the basic formulation. However, since the assumption of constant elasticity of congestion seems unappealing on a priori grounds, as we have argued above, we hope that the search for the most appropriate functional form of the crowding function will continue. In particular, it would be important to know whether there is an optimal city size for the provision of a specific public good, and what this optimal size is.

7 Alternative Approaches

7.1 Measuring Service Quality

We have seen that the BD-BG literature relies on strong behavioral assumptions to identify economies of scale. If the quality of public services could actually be measured, many of these assumptions could be avoided, since the task is then reduced to estimating a production function relationship. Presumably, more direct and reliable estimates of congestion effects could thereby be obtained. This approach raises difficult measurement issues, but promising attempts along this line have been made for fire protection and police services.

Data on house insurance premiums and the evaluations of municipal fire protection services allow Brueckner (1981) to compute the reduction in expected fire losses achieved by the fire protection services. This is a direct measure of the individual consumption of the publicly provided good fire protection, expressed in monetary units. Brueckner estimates significant positive returns to scale in consumption. A one percent increase in population only requires a 0.602 percent increase in expenditures\(^{33}\) to keep fire ratings constant. This figure is the analogue to the parameter \(a\) in the BD-BG studies, but the estimate indicates a much

\(^{33}\)Computed, in Brueckner's notation, as \(\gamma / (\alpha + \beta)\); cf. p.54.
higher level of publicness than is usually found in the BD-BG studies. This result is all the more remarkable in that Brueckner only considers large communities with populations exceeding 30,000, which include Chicago and Detroit.

The congestion effects in the provision of public safety were estimated by Craig (1987) and Craig and Heikkila (1989). These papers deviate from the standard approach in several respects. First, they simultaneously estimate a production function for police services and a police allocation function. This avoids the simultaneity bias that might result if larger police forces are allocated to neighborhoods with higher crime rates. Second, they use a functional form for the crowding function which assumes increasing marginal congestion. Third, and most interestingly, they use data on police input, clearance rates, which is the intermediate output produced by the police force, and safety, which is the good consumed by users, rather than expenditure on police services. According to their estimates, there is no significant crowding in the production of public safety if we hold the clearing rate constant. Crowding only occurs in the production of the intermediate output clearing rate. Unfortunately, their estimated crowding elasticities cannot be compared to the usual crowding parameter, due to differences in metrization (cf. Reiter and Weichenrieder, 1997, for details).

It is also unclear whether the difference in the results comes from the use of output data or from the use of a different functional form of the crowding function. McGreer and McMillan (1993) find that Craig's formulation tends to lead to low estimates of the crowding function.

The above studies approach the problem of measuring the public services provided in the most direct way. In addition, there is an older tradition in the literature which tries to estimate cost or production functions for public services. The output variable is usually not a public good enjoyed by consumers, but rather a measure of intermediate goods (D-goods in the terminology of Bradford, Malt and Oates (1969)). We only cite a few studies which seem most relevant for our discussion here.

Walzer (1972) estimates a production function for police services and finds small but significant economies of scale. His output measure is a combination of the number of offences cleared, number of accidents investigated, and miles travelled by police vehicles.
This study is only one example (a relatively recent one) in a larger literature, and provides further references.\textsuperscript{34} Hambor, Phillips and Votey (1973) estimate a technical relationship between teacher per student, number of students and other environmental characteristics as inputs and students’ test achievements as output. Unsurprisingly, they do not find returns to scale using state level data.

Loehman and Emerson (1985), building on Emerson and Loehman (1976), estimate the costs of the provision of public services (a combination of schooling and fire protection services). Quality is measured by students’ test rates and by fire casualty rates, combined with a hedonic price technique. Holding the quality of the service constant, they find that average costs as a function of population are U-shaped, with a minimum at about 500,000 inhabitants. For smaller city sizes, there are (minor) economies of scale.

At the moment, the small number of studies that use direct data on output does not yet allow definite conclusions to be drawn. It seems nevertheless noteworthy that many of those studies tend to find significant publicness in the data. Since this approach does not require stringent assumptions about the political choice mechanism, we think that it is the most promising method for measuring the publicness of publicly provided goods. More research along these lines is clearly needed.

7.2 \textit{The BD-BG Approach and the Use of Micro Data}

A different way to check the BD-BG estimates has been introduced by Gramlich and Rubinfeld (1982). They estimate an equation which is similar to equation (5-BD)\textsuperscript{35} by using both the usual community level data as well as micro survey data. Where micro data are used, the dependent variable is desired expenditure rather than actual expenditure. Like the macro estimates, the micro estimates also show a crowding parameter very close to unity. At the same time, however, the estimated price elasticity for the micro estimate hovers between – .004 and – .011 and, while statistically significant, is both unbelievably small and contradicts

\textsuperscript{34}A good survey is also given by Hirsch (1970, pp. 178-184).
\textsuperscript{35}Among other things, Gramlich and Rubinfeld allow benefits derived from public services to be dependent on income. This reflects the possibility that the distribution may be pro-rich or pro-poor. Other studies which discuss income dependent benefits are Denzau and Mackay (1976), Blecha (1982), Gonzales and Means (1991), and Gonzales, Means and Mehay (1993).
estimates derived from direct democracies. Since this micro estimate and the follow up studies are all based on surveys which ask citizens whether they prefer more, the same, or less public expenditure than the current expenditure in their community, a simple explanation for the low elasticity might be that respondents find it most convenient to respond 'the same'. Although recent papers provided additional micro estimates, most of these do not consider crowding estimates or concentrate on local school expenditure where a priori little publicness can be expected.

8. Conclusion

This paper has reviewed the literature which investigates the publicness of services provided by local governments empirically. The large majority of these studies do not find significant economies of scale, but conclude that the crowding of publicly provided services is so high that these goods are essentially like private goods. The aim of this paper was to discuss critically the conceptual and methodological problems of these approaches and to see whether the results are reliable or if they seem to be an artefact of the methodology used. While we have argued that the standard approach rests on very strong assumptions and gives rise to serious measurement problems, we concluded that most of these problems are unlikely to be responsible for the privateness results.

Nevertheless, we think there are reasons to be sceptical about the estimated crowding elasticities for public services. First, the BD-BG approach does not measure the demand for public goods but only for public expenditures. If expenditures in larger cities reflect a larger variety of public services, as is argued by Oates, then the high expenditures in larger cities may reflect economies rather than diseconomies of scale. Some studies which try to measure the quantity and quality of public services directly find substantial publicness. Second, spillover effects may enable small cities to exploit larger ones and these are then induced to

---

36See also Schokkaert (1987, p. 182). To overcome this problem, later studies which deal with local public school finance introduce a threshold value. Only citizens whose preferred budget deviates from the actual budget by more than this threshold are assumed to be expressing their dissent.

37Cf. the papers mentioned in fn. 20.
spend more if demand is inelastic. Finally, up to now there are no studies which try to pin down the potential effects of spillovers and fiscal externalities between communities. If smaller jurisdictions are subject to more intensive tax competition, which is a standard result in the theoretical literature (cf., e.g., Kanbur and Keen 1991), then larger communities may have larger budgets because of their greater power to tax, not because of crowding effects.

Essentially, there are two policy problems for which empirical crowding estimates of local public services are of interest. The first is the privatization of public services. The empirical crowding results for local public services have tempted several economists to conclude that most services can be efficiently privatized. This conclusion is premature. A frequent counter argument is that excludability is rather costly for many services. This paper has provided another argument which concerns the interpretation of estimated crowding elasticities. If the crowding is not induced by rivalry in the consumption of public services, but rather by negative agglomeration externalities, crowding elasticities close to unity do not imply that these services can be efficiently supplied by private competing clubs.

The second policy problem is whether competition between jurisdictions is beneficial. It has been argued that if production of public services implies economies of scale, competition between jurisdictions might lead to a suboptimal provision of public services. The literature on community expenditures surveyed suggests that this concern is not justified. However, the concern about underprovision was mainly formulated in connection with competition between central governments, to which the crowding results for local public services do not necessarily apply.
Appendix: Tax Competition and Non-linear Budget Constraints

Assume each of a community’s identical households owns a house whose value is normalized to 1. Let total property value be given by \( P = NQ \), where \( Q \) is the ratio of total property value to house values. \( Q \) depends positively on the exogenous community characteristic \( \phi \), and negatively on the property tax rate \( \tau \). The latter reflects the tax competition effect. The higher the tax rate, the more firms leave the community. \( \phi \) does not depend on the amount of public goods \( G \), since by assumption these are useful for residents, not for capital. Taking an isoelasticity assumption yields

\[
Q = \phi \tau^{-\omega}.
\]  

(A1)

If we normalize the price of the public good to 1, the government budget constraint is \( G = \tau NQ \) and from (A1), the tax rate \( \tau \) necessary to finance \( G \) is

\[
\tau = \left( \frac{G}{N\phi} \right)^{1/(1-\omega)}.
\]  

(A2)

The tax price for the individual household of one unit of the public good is \( t = 1/(NQ) \). Using (A1) and (A2), we can express the tax price \( t \) as an increasing function of \( G \):

\[
t = \left( \frac{G^\omega}{N\phi} \right)^{1/(1-\omega)}.
\]  

(A3)

Consider now the representative household’s decision problem. To keep things as simple as possible, consider a utility function that generates a constant price elasticity of demand:

\[
U = X + q^\varepsilon = X + (GN^{-\alpha})^\varepsilon.
\]  

(A4)

Using (A3), the budget constraint (1) becomes

\[
XY = - - G^{1/(1-\omega)} (N\phi)^{1/(1-\omega)}.
\]  

(A5)

Then, the maximization of (A3) with respect to \( G \) yields

\[
G^\omega = \left[ \varepsilon(1-\omega) \right]^{1-\omega} N\phi N^{-(1-\omega)\alpha \varepsilon} (\varepsilon-1)^{1-\omega}.
\]  

(A6)

From (A3), \( N\phi = G^\omega t^{\alpha-1} \). Inserting this into (A6) and solving for \( G \) we obtain

\[
G = kt^{1/(1-\epsilon)} N^{\omega/(\varepsilon-1)} = kt^{\delta} N^{\omega(1+\delta)},
\]  

(A7)
where $k$ is a irrelevant constant and $\delta = 1/(\epsilon - 1)$. For the variables of interest, equation (A7) is identical to (4a). Therefore, applying the BD-BG approach to the new situation does not lead to a bias in the estimation of $\alpha$.

**Bibliography**


DeBoer, L., 1986, State and local government utility maximization according to GARP, Public Finance Quarterly 14, pp. 87-99.
Inman, R.P., 1979, Testing political economy's 'as if' propositions: is the median income voter really decisive?, Public Choice 33, pp. 45-65.


