Intergovernmental Grants, Urban
Congestion and the Provision of Local Public Goods

Thesis submitted for the Degree of
Doctor of Philosophy
at the University of Leicester

by

Paulo Trigo Cortez Pereira

Public Sector Economics Research Centre
Department of Economics
University of Leicester

October 1993
To Guida, Mariana, Catarina
and the memory of my brother José
Contents

List of Figures ................................................................. v
List of Tables ................................................................. vii
Acknowledgements ......................................................... viii
Preface ................................................................. x

Introduction ......................................................................................... 1

1 Positive and Normative Approaches in Public Economics 5
  1.1 A critical rationalist approach to economics ............. 6
    1.1.1 The Positivists ...................................................... 6
    1.1.2 Methodological pluralism ................................ 8
    1.1.3 Critical Rationalism ........................................... 11
  1.2 On the Use of the two Paretian Criteria ................. 13
  1.3 Normative and Positive Approaches in Public Economics 17
  1.4 The Problem ............................................................ 23

2 Equilibrium and Efficiency in an Economy with Local Public Goods 25
  2.1 Seminal papers: Tiebout, Buchanan and Hamilton .... 27
    2.1.1 Tiebout and the Importance of Migration ........ 27
    2.1.2 Communities as Clubs ..................................... 30
    2.1.3 Hamilton and The Role of Fiscal or Exclusionary Zoning . . 32
  2.2 Equilibrium and optimality in an economy with local public goods ........................................ 36
  2.3 Equilibrium and optimality in a system of clubs ... 41
2.3.1 Communities providing only one service .................. 42
2.3.2 Communities providing several services .................. 45
2.4 Communities as clubs? ................................................... 51
   2.4.1 Communities, clubs and entrepreneurs ..................... 52
   2.4.2 Is the Property Tax a "Benefit" Tax? ..................... 53
   2.4.3 How excludable are new residents? ....................... 55
2.5 Conclusion .................................................................. 59

3 Economies of community size, congestion and the role of
   intergovernmental grants .................................................. 63
   3.1 The median voter model .............................................. 65
      3.1.1 The Classical Median Voter Approaches ............... 66
      3.1.2 Technology of Production of Local Public Goods . 75
      3.1.3 Economies of scale in the production of local services 78
      3.1.4 Extensions on congestion ................................. 81
      3.1.5 Extensions on the tax-price .............................. 87
      3.1.6 Median Voter model: summary ........................... 88
   3.2 Bureaucratic Approaches ........................................... 90
      3.2.1 Theoretical approaches ...................................... 90
      3.2.2 Empirical results ............................................ 95
      3.2.3 Summary ..................................................... 98
   3.3 Positive Approaches to Intergovernmental Grants ............ 99
      3.3.1 A Typology of Grants ...................................... 99
      3.3.2 Intergovernmental grants, the flypaper effect and the
              demand framework ........................................ 101
      3.3.3 Institutional framework and Intergovernmental Grants 108
      3.3.4 Positive approaches to grants' design .................. 110
   3.4 Conclusions .......................................................... 114
### 4 Economies of community size and the quality of services within centralized fiscal structures

- **4.1 Decentralized versus Centralized fiscal structures**
- **4.2 Properties of the Crowding and Congestion Functions**
  - 4.2.1 Crowding and congestion with a unique facility per jurisdiction
  - 4.2.2 Crowding and congestion when each community provides several facilities
  - 4.2.3 Crowding function, increasing marginal congestion and "publicness"
  - 4.2.4 Economies of scale
- **4.3 Community Size and Fiscal Position of Communities**
  - 4.3.1 Local public good produced under constant returns to scale
  - 4.3.3 Extending the model to two local public goods
  - 4.3.3 Towards a positive approach of lump-sum grants
  - 4.3.4 Equalization versus Political Hypotheses on Intergovernmental Lump-Sum Grants
- **4.4 Conclusion**

### 5 Anatomy of a Fiscally Centralized Country: Portugal

- **5.1 Fiscal Centralization in European (OECD) Countries**
- **5.2 Fiscal centralization in Portugal: the revenue side**
  - 5.2.1 Introduction
  - 5.2.2 Local Tax Revenues within Centralized Governments
- **5.3 Intergovernmental Grants, Fiscal Equalisation and Community Size**
  - 5.3.1 The fiscal equalisation hypotheses
  - 5.3.2 A politico-economic hypothesis on intergovernmental grants
- **5.4 Local Government Expenditures**
- **5.5 Conclusions**
List of Figures

2.1 Tiebout’s optimal community size ......................................... 28
2.2 Optimal sharing groups for different goods .......................... 48
3.1 Congestion and Crowding Functions ................................. 83
4.1 Equalization of fiscal capacity and fiscal positions ............. 121
4.2 Break down of scale and density effects. ............................ 127
4.3 The crowding function and the “privateness” result .......... 134
4.4 Intergovernmental Grants, Crowding and the Quality of Services .................................................. 140
4.5 Intergovernmental Grants and Communities’ Rank-Size .... 154
5.1 Ratio of Intergovernment Grants to Local Governments’ Receipts in OECD-European countries ............................. 166
5.2 Per Capita Intergovernmental Grants and Local Governments’ Expenditures when Communities are Ranked by Increasing Population Size (1989) .............................................. 180
6.1 Average Annual Growth Rates of Population (1980/91) for Portuguese Municipalities (ranked by increasing population size (1991) ................................................................. 194
6.2 Average Annual Growth Rates (1980/91) of Primary Schools’ Enrollments for Portuguese Municipalities (ranked by increasing population size (1991) .................................................. 195
6.3 School Investment Costs per Classroom ............................. 197
6.4 The Utilization of Capacity of Primary Schools and Community Size ........................................... 201

6.5 Proportion of Primary Schools' Classes with Normal Timetables .................................................. 204
List of Tables

5.1 Final Consumption Expenditures by levels of Government in OECD-European countries (1985) .......... 163
5.2 Gross Fixed Capital Formation by levels of Government (1985) 165

5.3 The Structure of Revenues of Portuguese Local Governments 169
This thesis would not have been possible without the help of several people and institutions.

First of all, I would like to thank Peter Jackson for his guidance, comments, expertise and friendliness all of which contributed to the successful completion of this thesis. I am also grateful to him for the excellent working facilities in the Public Sector Economics Research Centre (PSERC) where I was an Honorary Research Associate from October to March 1991.

This research started in my Lisbon University (Instituto Superior de Economia e Gestão / Universidade Técnica de Lisboa) where I considerably benefited from exchanges with Manuel Brandão Alves. To him and Francisco Pereira de Moura I would like to thank for helping me in pursuing my research in the United Kingdom.

I also would like to thank Robert Bennett who kindly enabled me to be an Academic Visitor at the London School of Economics from April until September 1991. At the LSE I also had constructive conversations with Michael Hebbert and Christine Whitehead.

In the Economics Department of Leicester University I received helpful comments from Ali-al-Nowaihi, Seilesh Tanna and Woycek Charemza.

The task of getting data is frequently a difficult issue and this case was no exception. In this context I am grateful to Daniel Santos and António Leitão from the Instituto Nacional de Estatística for providing me with the budget accounts for all (mainland) Portuguese local governments. At the Gabinete de Estudos e Planeamento I would like to thank Maria José Rau and Mafalda Leónidas for providing access to unpublished data. In the present Departamento de Programação e Gestão Financeira I also benefited from valuable conversations with Maria Clara Sousa Lopes and António Catarino.
I express my appreciation for the support of Madalena Abrantes and Barbara Allen who helped to introduce data and references. I also would like to acknowledge the friendly support of Ian Cross in improving my graphics and Maria Salerno for improvements in my written English.

I am very grateful to my University (ISEG/UTL) for having offer me leave for three years, and also to a grant from the Junta Nacional de Investigação Científica e Tecnológica which also covers the binding costs of this thesis.

Other persons helped by creating a favourable environment for research during this period.

In the PSERC I would like to thank Diego Piacentino for lively and stimulating discussions and Rasik Fyodorovitch for his exuberant temperament which helped me to get through the first cold winter. I would like to mention Dean Garratt who as well as being a very good company during the two years we shared the same office, also helped in improving the English of a substantial part of this thesis. The staff at the PSERC was also very friendly and I would like to thank in particular Pat Greatorex for her kindness and helpfulness.

To all friends made in Leicester, I express gratitude for the light-hearted moments that made life easier and more enjoyable.

In Portugal I am deeply grateful for the support extended by my parents and parents-in-law who in addition to dealing with personal matters were very helpful in enthusiastically caring for my children during their holidays.

A dear note of thank goes to my wife, Guida, for her loving support and contribution to an invaluable homely environment.

Finally, I owe a great deal to the lively, entertaining company of our children Mariana and Catarina who often relieved their parents of their academic worries. When we first arrived, they were too young to express their feelings about coming to England. I do believe, however, that they too benefited from the experience of living abroad.
Preface

The basic motivation for the development of this research was to understand why we observe a degradation of local services in urban and suburban areas in Portugal.

There is, I believe, an excess of demand for local services in relation to those supplied by local governments. The sprawl of illegal housing construction and the fiscal pressures associated with rapid and unplanned urbanization\(^1\) suggested that a "solution" to the problem should be investigated on the demand side of the problem.

Therefore, when I went to the L.S.E to meet Robert Bennett and Christine Whitehead my topic of research was defined as: "Zoning Regulations, Urban Growth and Congestion of Local Public Goods". Zoning could be, I thought at that time, the main instrument to control urban growth. As a consequence, I turned to the extremely interesting Law and Economics literature which has been developing as a new branch of economics following the works of Richard Posner (Chicago) and Guido Calabresi (Yale).

It was only when I asked Michael Hebbert (L.S.E.) what happens in the U. K. when someone decides to build a house without planning permission that I suddenly realized I was on the wrong track. His surprise at even posing the question was the best answer I could get to understand how different is the legal and institutional framework in this country and how different is the behaviour of the economic agents regarding the Law. This might be due to the greater effectiveness of law enforcement in the U.K.\.

At the same time I was becoming aware of the ethical problems associated with zoning. Basically, zoning can be an instrument used by the residents of a community so that their preferences rule over the preferences of those who want to enter the community. It is not easy to decide over which preferences should rule. Famous court cases in the U.S.A. have dealt precisely with this issue.

\(^1\)The situation is slightly better now since most municipalities have Planes Directores Municipais. However, the enforcement of these planning rules is still a problematic issue.
Zoning would be expected to be ineffective in countries like Portugal because people do not obey the Law in the same way as in Britain or the United States. Considering also that zoning raises difficult ethical issues, I realised I should start looking at the supply side of the problem.

Assuming that urban growth is what it is, why do local governments not increase supply accordingly?

The analysis of the particular case of Portugal clarified that part of the answer to this question relies on the centralized nature of government, where the ability of local governments to realize discretionary changes in their revenues (and therefore expenditures) is severely constrained.

Therefore, this thesis can be understood as an inquiry into the implications on the quality of local services of a centralized system of government.

However, most of the economic literature on local governments' decision making assumes a decentralized government and therefore the issue that naturally arises is whether this research has only a parochial scope (the Portuguese case) or a more broad range of interest.

There are two main reasons to justify a broader scope for this thesis. Firstly, we might use Tullock's argument that people usually write about democracy although the majority of political regimes in the world are still autocracies. The same applies to decentralized and centralized countries if the conjecture that autocracies have usually politically centralized systems of government is accepted. Economists usually use models assuming decentralized governments (e.g. the median voter model) when most of the countries in the world are autocracies and therefore most likely centralized.

Moreover, even within democratic countries there are centralized systems of government and on the other hand those which are decentralized are always subject to centralization trends.2

Secondly, some issues addressed in this thesis are not confined to centralized governments. This is the case of the analysis on the economies

---

2 See the preface of Tullock's book Autocracy.
3 This is what happened in the U. K. after the introduction of the "poll tax" which had an effect of increasing centralization of public sector revenues.
of community size and the "privateness" versus "publicness" of local public goods which is valid independently of the system of government. This analysis is extremely important when considering intergovernmental grants with equalization purposes.

The institutional and fiscal rigidities associated with centralized governments undermine the idea that local governments provide services according to the preferences of a representative voter within each jurisdiction (the median voter). On the other hand, it was an unexpected conclusion that there are affinities between the centralized governments' approach developed in this thesis and the approach developed by Tiebout (1956). Citizens unable to influence local decision-making through the vote will, ceteris paribus, migrate to jurisdictions where the fiscal "package" (local services and taxes) is more in line with their preferences. Alternatively, citizens who have more ability to pay may simply "exit" from public towards private provision.

The analysis that follows is essentially the diagnosis of the problem. The remedies would certainly involve fiscal and institutional reform which lie beyond the scope of this thesis.4

---

4 Nevertheless, we will address these issues in the conclusion.
Introduction

The essential problem addressed in this thesis is whether there are economies of community size, i.e. whether local governments in more populated communities are able to provide local public goods at a lower per capita cost keeping constant the quality of services provided to their residents. This issue can only be fully understood when it becomes clear whether there are economies of sharing in the consumption of local public goods and/or economies of scale in production of these goods.

The former aspect has been considered in the literature of local public economics as the problem of crowding, congestion and the "publicness" of local public goods and will be an important topic developed here. The latter issue has not received much attention in the literature, due to severe methodological problems and will not be considered much in this thesis.

In most countries intergovernmental grants play a decisive role in financing local authorities, and there is frequently a fiscal equalization aim associated with these grants. Whatever the concept of fiscal equalization that central governments may have (or practice), it can not be specified without addressing the issue of economies of community size. Therefore, there are important normative implications to be drawn from such an analysis.

On the other hand, a positive approach to intergovernmental grants will clarify whether or not they are expected to offset the existing economies of city size.

It is necessary to define the concepts of centralized governments and

\(^5\) This term is sometimes referred to as economies of agglomeration.
economies of community size before introducing the main argument of this
thesis.

Centralized governments or centralized fiscal structures are those multi­
tier systems of government where the main sources of local governments' revenues are exogenous to them. This arises because central government can determine either directly (through intergovernmental grants) or indi­rectly (through the set of local tax rates) the great majority of the resources available to the lower level governments.

On the other hand, economies of scale is defined as a relational concept. There are economies (diseconomies) of community size in the provision of a standard bundle of local public goods whenever, ceteris paribus the combined effect of economies of scale in production and “congestion” in consumption leads to a decrease (increase) in the per capita cost of provision as the population of communities increase.

The main argument of this thesis applies to centralized governments and can be developed as follows: intergovernmental grants’ design leads, ceteris paribus, to urban communities being in a relatively worse fiscal position than medium size communities. This leads to a decrease in the quality of local services provided in urban areas.

This “hypothesis” will be established through three main steps. Firstly, an analysis of the economies of community size. Secondly, a positive app­proach to intergovernmental grants. Thirdly, an analysis of “congestion” in local public goods provision.

We will conclude that albeit there are some economies of agglomeration, intergovernmental grants are predicted to be more regressive regarding the population size of communities. Since within centralized governments communities are not able to increase their tax effort to compensate for a deterioration in their fiscal position, congestion in local service provision will increase.

Improvements in the situation are not only a matter of reallocation of funds between central and local governments and within local governments,
but also a question of institutional design. Rules and institutions play a decisive role in the analysis being developed at different dimensions.

Firstly, the analysis and the solution of the problems stated above change according to the specific institutional framework. Namely, fiscal problems within decentralized democratic governments, have a different nature from those within centralized governments.

Secondly, the relevance of models of local government fiscal behaviour is dependent on the particular institutional and legal framework in place in a particular country. The criticism made of the dominant median voter model is precisely based on this approach.

Thirdly, it is necessary to maintain the methodological distinction between the positive and normative analysis of fiscal behaviour within a given set of rules and institutions, and on the other hand, the normative analysis of changing the rules and institutions.

Therefore the thesis is structured as follows.

Chapter 1 introduces the forthcoming chapters at three distinct levels. Firstly, it clarifies the methodological judgements that are employed throughout the thesis and how we stand from an epistemological point of view. Secondly, it clarifies our approach to public economics, the distinction between normative and positive approaches and the concept of local public goods. Finally, it elucidates the nature of the problem under research and the strategy developed to tackle it.

Chapters 2 and 3 are mainly surveys of the literature, concerning decentralized systems of government. Chapter 2 addresses the issue of the existence of an equilibrium in an economy with local public goods and mobile individuals. This is simultaneously an interjurisdictional spatial equilibrium and an intrajurisdictional political equilibrium. It also considers whether communities might be seen as clubs that provide services to their residents.

Chapter 3 analyses intrajurisdictional decision-making and the role of intergovernment grants. Firstly, it considers the median voter model which
has been the main theoretical framework employed for empirical analysis. Particular attention is given to congestion in local public goods. Secondly, it considers other views of local decision-making namely the bureaucratic approaches. Finally, the role of intergovernmental grants and the impact on local government expenditures is considered.

Chapter 4 develops a positive approach to economies of community size. The meaning of a centralized government is clarified and a new approach to crowding and congestion is introduced. Moreover, a positive approach to intergovernmental grants is also developed. Several hypotheses result from the analysis which will be the subject of empirical scrutiny in the following two chapters.

Chapter 5 considers the Portuguese case. It is shown why Portugal has a centralized government both in terms of the meaning of the concept and in comparison with the other European OECD countries. The reasons why local revenues are mainly exogenous to local governments and therefore determine the level of local expenditure are clarified.

It is a corollary of our argument that we should be able to observe a decreasing quality of services in more populated areas. This is true even if they are not capacity constrained. Chapter 6 addresses this issue looking at the specific case of primary school facilities, not in terms of expenditures but directly in terms of output.

Finally, chapter 7 concludes with three types of considerations. Firstly, there are more theoretical conclusions regarding the nature of local public goods, of crowding and congestion and the role of intergovernmental grants within centralized governments. Secondly, suggestions are made regarding fiscal reform in Portugal towards a more decentralized government where local expenditures are more likely to meet the requirement of satisfying the demands of the citizens. Finally, the provision of educational facilities in the context of institutional design is addressed.
Chapter 1

On the Positive/Normative Approaches in Public Economics

The aim of this chapter is twofold. Firstly, to make explicit the methodological judgments that are employed in this thesis and which shape its structure. Blaug (1980) emphasizes the existence of judgments involved in methodological statements to clarify that there are always personal values attached to scientific research. These are often implicit rather than explicit, but we prefer to make them clear. In particular we will deal with the nature of economic reasoning, the distinction between normative and positive economics, and the role of observation (or empirical "evidence") in assessing competing theories.

Secondly, it clarifies the development of positive public economics, and the critical role played by rules and institutions. This clarifies the approach developed in this thesis which is mainly a positive approach to local governments' behaviour within a centralized government.

Therefore, section 1.1 clarifies the meaning of a critical rationalist approach to public economics as opposed to a positivist approach. It argues in favour of a pluralist approach to methodology and for the advantage of keeping the heuristic distinction between normative and positive approaches in
economics.

Section 1.2 clarifies the normative/positive distinction, and the problem of choice among and within institutions. It also clarifies how Paretian criteria and distributional issues can be dealt with in either normative or positive frameworks.

The section 1.3 uses the former discussion to clarify the upsurge of the "new" political economy and the main recent contributions.

Finally section 1.4 clarifies the main topic of the thesis taking into account all the discussion in former sections.

1.1 A critical rationalist approach to economics

The objective of this section is to clarify the meaning of a critical rationalist approach to economic science in opposition to a positivist approach. The different epistemological status's of the theory are then clarified and the role of observation in each approach is considered. It is also aimed at understanding differences in the meaning of the distinction between normative and positive approaches in Economics.

1.1.1 The Positivists

Among the authors who consider that the distinction between positive and normative economics is relevant, there are two fundamental perspectives arising from different conceptions about the scope and method of Economics.

Briefly, the first approach that can be labeled positivist defends the methodological monism between all the sciences. Thus, Positive Economics should develop the same methodological procedures as the physical sciences.

The second approach defends precisely the opposite and is labeled methodological pluralism. It is argued here that Economics could not and should not develop in the image and resemblance of the natural sciences.
The positivist approach has its roots in the classical positivists of the nineteenth century (Comte and Mach), the logical positivists of the early twentieth century (the Vienna Circle) and the logical empiricists of the middle of this century. They adopted the following methodological judgments: (i) thesis of the methodological unity of all sciences, (ii) Hypothetical-deductive method as the correct method to structure a theory, that gives rise to theorems or predictions (without cognitive significance or empirical relevance) (iii) rules of correspondence that make possible that those predictions are indirectly tested and (iv) confirmationism (rather than falsificationism) as the essential criterion to validate a theory. Explanation in science must proceed by logical deduction from an explanans, a group of sentences including at least one law, to an explanandum, the phenomenon to be explained. The referred law can be derived using a nomological-deductive process or an inductive-probablistic one.

Even after the decline, during the sixties and seventies, of the positivist approach within the philosophy of science, positivism is still important in Economics due to the influential paper of M. Friedman (1953). He accepts the fundamental principles of the positivist approach.

Even acknowledging the Popperian influence that the successful rejection of falsificationism (and not confirmationism) is the relevant criterion for assessing a theory, Friedman is essentially a verificationist. In fact, he adopts the instrumentalist point of view that it does not matter if the assumptions of a theory are unrealistic in so far as its predictions are accurate, because "the ultimate goal of a positive science is the development of a 'theory' or 'hypothesis' that yields valid and meaningful (i.e. not truistic) predictions about phenomena not yet observed." ¹ Thus, Friedman makes a clear distinction between positive and normative economics based on scope and method. Regarding scope, he adopts J.N. Keynes’ distinction between "positive science", a body of systematic knowledge about what is, "normative" science about what "ought to be" and "art", a set of rules

to obtain certain ends. On method, Friedman considers that positive economics should have the same method as the physical sciences and that in principle it is independent of the ethical positions of the economists and so of any normative judgment. Following from this, Friedman conjectured that economists diverge in their prescriptions on economic policy, not because they have different value judgments but because they have different conceptions on the consequences of certain measures of political economy. As a consequence he believed that the development of Positive Economics will generate a greater consensus among economists. 

As Friedman’s paper is on the methodology of positive economics, he does not address the problem of method in normative economics. However, it is implicit that according to the object of normative economics, it must use some normative criterion, which reflects the ethical position of the economist and that prescriptions are valid under the assumption that the normative criteria is accepted. As a consequence, normative economics is not a “science” and therefore there is no place for “scientific method”. Criticism of Friedman’s approach uses the same kind of arguments as those used against positivists and empiricists. Empirical “evidence”, by itself, can not validate or even refute any theory. Firstly, because there is no evidence; as Popper states, all observation is directed by a theory. Secondly, any theory can be protected against potential falsifiable observations, with \textit{ad hoc} assumptions. This is the approach developed by Lakatos (1978) about the “protective belt” that protects the core of each research program.

1.1.2 Methodological pluralism

The criticism of the methodological role of observation and empirical “evidence” is simultaneous with the rejection of the thesis of the methodological unit of science. This rejection was essentially dual in the nineteenth cen-

\footnote{From 1953 onwards, this consensus did not occur (and gave no signs of increasing), so that two “hypotheses” can be suggested: (i) “positive” economics has not developed since then, (ii) economists diverge fundamentally on ethical issues.}
tury, since some authors, like J.S. Mill, recognized the differences between the method in Economics and the method in the natural sciences. Nowadays the rejection of the methodological unit of science is plural because the majority of methodological economists and philosophers of science advocate the methodological pluralism even within Economics.

J.S. Mill (1836) inquired about the method of Political Economy and acknowledged that there is a significant difference from the method in physical sciences. The reason why Mill adopted the “a priori” (deductive) method was precisely the problem of the role of observation stated above. He argues that in Economics as in other “moral sciences” there is no experimentum crucis, i.e. the crucial experiment that isolates a hypothetical causal factor from all the other factors intervening in a specific phenomenon, in order to understand if this hypothetical factor is the cause of a certain effect.

This is the argument Mill uses to deny the possibility of the a posteriori method, and also to justify that Economics is an abstract science in a double sense. Firstly, because it uses assumptions about human nature that do not need to be in accord with reality. Secondly, the conclusions which are logically deduced from the assumptions are only abstract truths. The conformity of these abstract truths with reality (observation, empirical evidence) depends on the effects of all other factors that were excluded from the analysis. This means that observation can be contradictory to the abstract truth without denying it.

Mill does not seem very worried about the problem of assessment of the abstract truths in Economics. However, he addresses the issue saying that a way out of the problem is to come back to the initial assumptions and verify, e.g. by introspection, if they are true. In this case, propositions

---

3 Which are a small minority within the profession and so do not necessarily reflect the methodological preferences of the "median voter" economist.

4 We follow here J.S. Mill’s “On the Definition and Method of Political Economy”. However, as J.N. Keynes clarifies, the method that Mill advocates in that paper is one thing (the deductive method), another is what he employs in his Principles of Political Economy (viz. both deductive and inductive methods.)
logically deduced from true assumptions are also true.

This idea seems to be in contradiction with the one referred to earlier that the assumptions do not need to be realistic. Within the positivist interpretation, there is in fact a contradiction because the assessment issue is a crucial problem. However, in a non-positivist framework both ideas are compatible using the following interpretation: assumptions do not need to be realistic to derive abstract truths; however, those who consider the assessment issue a problem, should use realistic assumptions so that the abstract truths are also true. This is, we believe, Mill's position in his paper (1836). Nevertheless, in his Logic Mill has a different view, more palatable to Positivists in general and Friedman in particular, which is that "the ground of confidence in any concrete deductive science is not the a priori reasoning itself, but the accordance between its results and those of observation a posteriori." ⁵

The pluralist approach to methodology in Economics was clearly supported by J.N. Keynes (1917). The important contributions of Keynes were: (i) to clarify the distinction (mentioned above) between positive science, normative or regulative science and art, (ii) to accept that there are different methods valid in different branches or sections of Economics, and finally that (iii) what frequently appears to be one method is frequently a combination of two or more, e.g. the hypothetical-deductive model needs induction in the beginning (to derive the initial assumptions) and in the end (to verify the propositions derived from the initial assumptions).

In this sense J.N. Keynes anticipates what is currently the mainstream methodological pluralist approach of disparate authors such as: K. Popper, P.K. Feyerabend, L. Boland, B. Caldwell and D.M. Hausman. ⁶ The most influential author during the twentieth century was and still is Karl Raimund

⁵See Mill's Logic vi. 9 quoted in J.N. Keynes (1917).
⁶Blaug is not included because his approach focuses on: The Methodology of Economics (italic ours), as the title (and content) of his book (1980) illustrates. This does not diminish the merit of his book, which is perhaps the best book on methodology of economics written by an economist.
1.1.3 Critical Rationalism

One of the central issues in Popper's research has been the problem of demarcation, i.e. to get a criterion to distinguish propositions belonging to the empirical sciences (Theories, Hypotheses) from other kinds of propositions such as pseudo-scientific, metaphysical but also mathematical and logical. He developed the criterion of falsifiability but he recognizes that this is only a necessary (and not sufficient) criterion.

In fact, Popper assumes himself to be a rationalist critic, in the sense that he advocates that the appraisal of competitive theories should be based on critical reasons informed by the results of experience. It is in this relative context that his falsificationist theory should be understood.

Popper distinguishes the terms falsifiable and falsifiability. The first term has only to do with the logical structure of statements. A statement or a theory is falsifiable if there exists at least one potential falsifiant i.e. a basic statement in logical conflict with it. One useful example that Popper gives of an unfalsifiable statement is the common postulate used by economists that human actions are motivated by self-interest. He argues, that every example of an altruistic action can be interpreted as having an egotistical motive behind it.

In this sense the falsifiability of a theory must be understood using this logical-technical concept of falsifiable. However, there is another meaning of falsifiability (or falsification) which relates to the possibility of a theory being definitely and demonstratively refuted on the basis of experimental

---

7 This problem is different from the problem of the truth. See Popper's Introduction (1983).
8 See his Realism and The aim of Science (1983) vol.I of the Postscript to the Logic of Scientific Discovery.
9 The same seems to apply to the reverse proposition according to which it is impossible to refute the proposition that all individual actions have altruistic motivations. It can be argued that individuals have the intuition of the "Invisible" Hand so that their apparently egotistic actions are driven by the general interest.
data. Popper defends that no theory can be falsified on this basis. 10

As will became apparent throughout the thesis economists have extem- 
reme difficulties in testing their hypothesis and very seldom (if ever) is 
one hypothesis supported against competing ones on the basis of empirical 
evidence.

That empirical evidence is very often consistent with alternative hypoth­ 
thesis was emphasized, for instance, by Musgrave11 in discussing Brennan and 
Buchanan’s (1980) Leviathan Hypothesis, and supported by Oates (1985) 
and Pereira (1991). A more comprehensive discussion of several hypothe­ 
theses within the neoclassical research program and how the tentative tests 
of them face up to extreme difficulties and the observation that they are 
more exercises in verificationism rather than falsificatonism, can be found 
in Blaug (1980 chs. 6 to 14).

This strongly supports the rationalistic critical approach of Popper’s 
later works i.e. that the assessment of competing theories or hypotheses 
should be carried out according to critical reasons and that the results 
of observation might be additional (not crucial) reasons to support a the­ 
ory. Particularly in cases where data are not produced consistently and 
coherently with the purposes of research 12 the theory should be supported 
especially by theoretical reasons rather than empirical evidence.

Friedrich Hayek (1989) in his Nobel memorial address presents an epistemological approach similar to Popper’s. Since we fully subscribe to his 
approach,12 it is worth quoting him. He labels the approach that economics 

10See Popper (1983) for a clarification of his concepts and reply to the misunderstandings 
of interpretations.
11Musgrave made a comment referred to in Oates’ paper (1985). We will return to this 
issue later in section 3.2.
12This is typically the case in Public Economics when data is produced by a government 
agency. In this case, administrative and technical criteria are the dominant ones to direct 
data collection. Very often time-series data are not consistent since there are changes 
in governmental departments, new legislation that introduce new taxonomies (e.g. on 
public expenditures) without the appropriate rules of correspondence between old and 
new variables and so forth.
13We do agree with F. W. Hayek (1989) on epistemological grounds. This does not mean
should imitate the physical sciences as "scientistic" and comments that it,\textsuperscript{14} is decidedly unscientific in the true sense of the word, since it involves a mechanical and uncritical application of habits of thought to fields different from those in which they have been formed.

On the role of mathematics in economic reasoning he adds:

I regard it in fact as the great advantage of the mathematical technique that it allows us to describe, by means of algebraic equations, the general character of a pattern even if we are ignorant of the numerical values which will determine its manifestation.\textellipsis It has led to the illusion, however, that we can use this technique for the determination and prediction of the numerical values of those magnitudes.

This is in fact the way we approach the role of mathematics for economists and it is the way it will be used in this thesis.

1.2 On the Use of the two Paretian Criteria

Assuming that empirical evidence plays a much more minor role in assessing competing theories than positivists would like us to believe, at least two consequences can follow. One is to drop the distinction altogether. Another is to adopt it with a different meaning. This latter approach will be considered in this section.

It will be argued here that the essential distinction between positive and normative approaches in Economics lies in the different kinds of \textit{problems} that are addressed and the ways value judgments are introduced in the analysis, rather than differences in method.

\textsuperscript{14}Hayek quotes himself from "Scientism and the Study of Society" Economica, 9(35) 1942.
As J.N. Keynes explained, positive economics tries to predict and explain what is and Normative Economics what should be according to a certain normative criterion. The difference is in the aim of research and not necessarily in method, as Friedman considered it. In fact, if we consider Normative and Positive Economics within the neoclassical research program, there are flagrant similarities in method. Both start with the same postulates or axioms regarding the behaviour of the different economic agents (those unfalsifiable propositions as Popper would say). Both proceed with an essentially hypothetical-deductive reasoning. The difference seems to be in the end. Normative propositions or Theorems (e.g. the Two Fundamental Theorems of Welfare Economics) are not tested, while propositions (hypotheses) in Positive Economics are generally submitted to the "test" of empirical evidence. However, for a critical rationalist there is not a significant difference, because this last step\(^15\) may be understood as an exercise in verificationism, since falsifiability on empirical grounds is extremely difficult\(^16\).

The assertions that the difference between Positive and Normative Economics is essentially about object and not method, and also that they may employ the same method, could lead to the conclusion that it is an irrelevant distinction. We argue that the distinction should be maintained as far as it can go, because it distinguishes the different nature of the problems under analysis.

The use of the Paretian criterion in both Positive and Normative Economics can now be discussed.

We also want to argue the polemical proposition that the Pareto criterion can be consistently used in Positive Economics. The use of the Pareto criterion in Positive Economics, has nothing to do with what some authors label "Positive Paretian welfare economics" which we consider to be a con-\(^15\)In reality they are two as we referred earlier: the derivation of rules of correspondence and the empirical test.
\(^16\)We would like to suggest the conjecture that a great part of econometrics is in fact an exercise in verificationism.
tradiction in terms. There are good reasons to accept the former and reject
the latter as will be considered below.

There are few economists who claim that the Pareto criterion might be
used in Positive Economics. Perhaps Buchanan (1968) is the first of the
few economists to understand and sustain this point of view. Since it is an
uncommon proposition it must be better explained.

Let us consider the Edgeworth Box in which there are drawn the indif­
ference curves of two rational utility-maximizer individuals, regarding two
commodities. Up to this point what was introduced, were the unfalsifiable
postulates without which the deductive reasoning could not follow.

As is known there are two kinds of points (allocations) within the Edge­
worth Box that can be distinguished using two Paretian concepts. The
points on the contract curve (points where the indifference curves of both
individuals are tangent) known actually as Pareto efficient points, and on
the other hand points off the contract curve, from where a Pareto Improve­
ment is possible. So, the first task of both Paretian concepts is descriptive:
it provides a taxonomy of all the possible allocations. Let us consider that
the initial allocation is a point off the contract curve. In this case, with
the initial postulates and the further assumption of absence of transaction
costs it can be acknowledged that individuals have incentives to engage in
trade until all the gains from trade are exhausted, i.e. until having reached
an allocation on the contract curve.

This is a proposition in Positive Economics, due to the nature of the
problem under research, and is an unfalsifiable proposition.

However, as Samuelson (1967) pointed out when discussing monopolistic
competition a non-Pareto optimal outcome can not be ruled out since there
is a problem of deciding a priori how the gains from trade are to be shared.
Coase (1988 pp. 159-163) discussing this point argues that although it is
not correct to state that in the absence of transaction costs individuals
must end on the contract curve, the probability that they will is great.

On the other hand let us assume that there are transaction costs. The
statement that transaction costs should be diminished to improve efficiency or that Pareto-optimal allocations should be enforced are statements typical of a normative approach.

The distinction between both approaches is clear although sometimes it is difficult to maintain. In the former, Pareto optimal allocations are seen as stable equilibrium points to which the result of voluntary exchange between individuals has a tendency to converge. In the latter, Pareto optimal allocations are norms.

However, it is not isolated propositions that must be analyzed to understand the nature of the approach. In fact, several propositions, theorems, or hypotheses can be considered either in a normative or positive approach.

To give an example, the Tiebout model\(^\text{17}\) was built for normative purposes. In this sense the unrealism of the assumptions is not necessarily a problem and the fundamental conclusions derived from the Tiebout model\(^\text{18}\) are as valid today as they were in 1956. A different thing is to consider the Tiebout model as an hypothesis to be confronted against empirical data. As will be discussed this is a different problem which is typical of a positive approach.

The whole systematic body of abstract knowledge in relation to the problem under research must be considered to identify whether the approach is positive or normative.

In this context, Paretian welfare economics is normative because of the nature of the problems addressed and because it is mainly within the Utilitarian framework. It is Utilitarianism, and not the use of the Pareto criterion, that makes welfare economics a research program within normative Economics. Utilitarianism, following Sen and Williams (1982) "is a theory of the correct way to assess or assign value to states of affairs" (welfarism) and a theory that "claims that actions are to be chosen on the basis of

\(^{17}\text{See Tiebout (1956). The model will be discussed in chapter 2.}\)

\(^{18}\text{That economic policies that facilitate residential mobility and increase the information of citizens on the packages of revenues/services of local governments, will improve the efficiency in the allocation of resources at a local level.}\)
states of affairs which are their consequences" (consequentialism). 19

One final point should be made. We stated that the distinction between positive and normative approaches within Economics is useful as far as it can go. The values of the economist are translated into the normative approaches through the methodological judgements and the criterion chosen as a norm for prescribing or judging public policy. The values of the economist get inside positive approaches through the methodological judgements and the choice of the problem under research. 20 It is in this last sense that a positive approach is always “normative”.

1.3 Normative and Positive Approaches in Public Economics

Public Economics is perhaps more than other branches of economics subject to this mix between normative and positive approaches. It is the aim of this section to contribute to a clarification of the different approaches and in particular to clarify our understanding of positive public economics.

The source of the confusion between the two approaches can be situated in the beginning of this century with the development of the voluntary exchange theory of public finance. There is no doubt nowadays, that the works of the Italian school in particular De Viti de Marco (1934) and Pantaleoni, and also the works of Lindahl (1919) and Wicksell (1896) were a big step forward in the public finance literature with two fundamental contributions. Firstly, that taxation should be analysed simultaneously with expenditures, i.e. it is necessary to consider the opportunity cost of public expenditures. 21 Secondly, that public finance is elaborated in

---

19 Sen and Williams (1982) pp. 3-4
20 This is assuming an interpretation of the economist, not as a technocrat that deals with problems settled by the decision-makers, but having an important role on the definition of the relevant problems.
21 Taxation was the main bulk of analysis in Public Finance and therefore its name. The increased role to expenditure analysis and with it the changing nature to Public Economics
a political process under specific institutions and procedures of decision-makin

The mix of positive and normative issues can be seen in the writings of several authors. Lindahl who gave the title of "Just taxation - a positive solution" to his important paper developed essentially a normative analysis to derive the preferences for public goods and at the same time an optimal form to finance them.

De Viti de Marco pointed out that the problems faced by public finance within a monopolistic state grounded on slavery and absolutism are not the same as those in free and parliamentary states. In the former the taxpayer should not be approached in terms of a receiver of public goods' benefits since a great part of the public expenditures are for supporting an absolutist state. Therefore, fiscal principles such as the benefit principle of taxation are meaningless. On the other hand, in democratic states it has much more relevance. Furthermore, the decision-making process of an executive which must answer before a parliament is completely different from the procedures in absolutist monarchies.

Wicksell introduced more positive insights, since he tried to analyse institutions in democratic states. In particular he stated that the executive should be considered as nonbenevolent, but that the parliament was tendentially benevolent. In this sense his theory was mainly addressed to the members of parliament, suggesting that legislative majorities should not replace the old oligarchies, and that fair taxation solutions should be unanimous or based upon an ultraqualified majority. This is the only guarantee that it is consensual, does not impose involuntary coercion, i.e. in modern terminology it assures a Pareto improvement.

22 The Lindahl equilibrium is still an important concept in public economics (see Lindahl (1919) and Johansen (1963)). An auctioneer proposes a set of different tax-"prices" and the individuals respond with the quantities of the public good demanded. When a given quantity is unanimously preferred an equilibrium is reached, that has the property of being Pareto-efficient.

23 This approach was extended in Buchanan and Tullock's Calculus of Consent. The
sights Wicksell's analysis was essentially normative. His problem was "just taxation" and his advice was that rules must be chosen so that coercion is minimized (assuming that originally there is a reasonably fair distribution of wealth in society).

On the other hand, Wicksell used what seems today a normative approach to democracy. The new approach to democracy was only developed later on by Schumpeter (1943) and formalized by Downs (1957). Instead of looking at what democracy "should be" - the process through which the general public interest is transformed into political decision-making - they emphasized how democracy should be modelled for positive analysis - a process of political competition for political leadership and power. It might happen that the "public interest" (whatever that may mean) is partially satisfied since politicians must compete for votes. However, this is a subproduct of political competition, as the delivery of goods and services in private markets is a result of competition between firms and the prosecution of profit.

In spite of these developments in democracy theory, the emergence of welfare economics, firstly in the pigouvian ("old")^24 tradition and afterwards in the ("new") Samuelson (1954) tradition reinforced the normative aspects of the theory. There is an implicit assumption of a benevolent decision-maker and the role assigned to the public economist is to give advice on three fundamental problems: (i) the revelation of preferences for public goods (or how to overcome the free-rider problem), (ii) the problem of the aggregation of preferences (the social welfare function problem ^25) and finally (iii) the optimal taxation problem (or how to minimize distortions in the economy). Mainstream public economics continues this topic.

^24See Pigou (1920).

^25In fact there are two problems here. The first, introduced by Arrow (1961) is the problem of the existence or not of such a function that satisfies a given set of axioms. The other is the choice of the specific form of the social welfare function, the seminal paper being Bergson (1988).
giving more emphasis to norms that might drive public policies, particularly using and developing efficiency and equity criteria.

This thesis develops a positive approach to public economics and more attention will be given now to this topic.

A fundamental contribution to positive public economics was given by the public choice research programme. The basic features of it are generally conceived as being the application of methodological individualism and of the homo oeconomicus postulates to the agents having public roles, i.e. to consider that the same behaviour postulates assigned to economic agents behaving in private markets should be extended to agents behaving in "political markets".

In spite of the actual existence of an important subdiscipline of normative public choice, the public choice research programme developed initially as a criticism of normative approaches in public economics. The contribution of Downs was already been referred to and this can also be seen along with another of the "founding fathers"- Mancur Olson. He analysed the conditions under which a group that has a strong common interest regarding the provision of some public good will not organize its provision (or under provide) because relying on voluntary contributions in large communities or groups leads to a free-riding behaviour and a prisoners' dilemma outcome. His analysis is mainly positive, because it is driven by the prediction of the conditions under which collective action will take place and the conditions where it is not expected to arise.

One of the main contributions of public choice theory is to have devoted considerable theoretical (more than empirical) attention to the role of rules and institutions. Particularly important is the distinction made between the "constitutional" and "post-constitutional" stages of analysis. The former applies to the choice of the rules/institutions (electoral, fiscal

\[2\text{0}\]
and others) that structure the behaviour of the economic agents. The latter refer to the in-period choice within a given set of rules.

Any economic agent behaves within a specific institutional framework. To predict his behaviour it is always necessary to assume some motivational postulate and also to know the institutional framework that frames his conduct, since it can be predicted that the same motivational postulates will generate different patterns of behaviour under different institutional frameworks.

More emphasis has also been given to supply-side issues. An individualistic approach to governments and public agencies was introduced, instead of an organic approach and more attention started to be given to the motivational postulates of politicians and bureaucrats and their role in the processes of decision-making.

Positive analysis is always done within a specific set of rules and institutions, while normative approaches are concerned with the procedures of choice of these rules and institutions. Therefore while Paretian welfare economics suggests remedies to "market failures" with the aim of improving Pareto efficiency in-period, i.e. within the existing set of rules, Paretian public choice scholars try to derive new rules and institutions at the constitutional stage to face market and political failures.

Notwithstanding the contributions of the public choice research programme, there are two major shortcomings. Firstly, the Leviathan model of government developed by public choice scholars from Virginia lead by Buchanan is only the opposite polar to the idea of the benevolent philosopher-king to which economic advice seemed to be addressed in earlier times. The will to fiscally exploit citizens as a governmental motivational postulate suffers from the same weaknesses as the will to serve the public interest. Secondly, the analysis is developed within a hypothetic-deductive framework but in general public choice scholars fail to submit their hypotheses to empirical tests. There are essentially theoretical critical reasons that

27 See Brennan and Buchanan (1977), (1980).
support the theory, but there are minor empirical studies to give additional support. In particular, it is not clear the method to be used in comparative institutional analysis, that should constitute the core of the research programme.

Buchanan set the agenda for the development of a positive theory of public economics:²⁸

There are two separate parts of this institutional stage of a positive theory of public finance. The first consists in the development of a logical theory of individual choice among alternative institutions... The second part of a theory of institutional choice... is essentially empirical. Perceptive observations of real world fiscal structures are needed, and the analyst must try to isolate the central elements in such structures that serve best to explain and predict.

In spite of the agenda being set twenty five years ago, there have been few analyses that consider the particular cases of fiscal structures outside North America. This is a consequence of the dominant literature in this field coming from U.S. scholars.

An interesting side effect, associated with the development of the public choice research programme, was the increased interest in positive public economics. This has occurred both in the mainstream public economics literature with a greater emphasis on the effects of different constraints on the behaviour of governments ²⁹ and in the emergent "new" positive political economy [Inman (1987)]. The main difference with traditional normative analysis is that instead of focusing attention on what should be according to equity or efficiency criteria, emphasis is on what is or is predicted to be under a particular institutional and fiscal context.

Positive analysis also gives more attention to equilibrium analysis than

²⁸See Buchanan (1968) pp. 200/201.
²⁹For instance the literature on the impact of different sorts of intergovernmental grants on the behaviour of the recipient local governments.
a normative approach generally does. In fact a precondition for studying
the normative properties of a given system is to know whether it is in
equilibrium or not. If there is no stable equilibrium there is not much
scope for normative analysis. However, even in disequilibrium it is still
relevant to develop a positive approach.

Extending the views of De Viti de Marco we could say that the prob­
lems of public economics within parliamentary democracies are different
depending upon the centralized or decentralized nature of the state. Here
we could also define two ideal-type situations. One where central (or fed­
eral) government has all the power of raising revenues and spending either
directly or through lower tier governments that are their passive agents,
and a second polar situation where the local (or regional) governments are
the essential agents regulating the collection and uses of public funds and
central government has only residual tasks to perform.

Real world fiscal structures define a spectrum between these two polar
situations. This thesis analyses local governments' fiscal dilemmas within
a centralized democratic State.

1.4 The Problem

There is essentially one problem addressed: the relationship between com­
munity population size and the provision of local services within a centralized
democratic State.

The "hypothesis" to be developed states that intergovernmental grants
are likely to be more regressive regarding the population size of communities
than economies of agglomeration would suggest. A consequence of which
is a lower quality of congestable local public goods in urban areas.

This problem characterizes a positive approach to the provision of local
collective goods. It will be analysed in five steps.

Chapter 2 addresses the problem of existence, stability, uniqueness and
optimality of an inter and intrajurisdictional equilibrium with mobile in­
dividuals in an economy with private and local public goods, and many
jurisdictions. It will be argued, following a critical survey of the literature, that an equilibrium is unlikely to exist. This gives support to the emphasis that is made on locational disequilibrium. On the other hand it gives less importance to Pareto optimality issues within a given fiscal and institutional context.

Chapter 3 reviews the issue of intrajurisdictional political equilibria and the positive role of intergovernmental relations. A critical review of theoretical and empirical results using the median voter model will be considered here.

Chapter 4 develops a model of local governments' fiscal "behaviour" under a centralized democratic State. The model generates several predictions regarding the effects of community size on the provision of local services, under different assumptions concerning intergovernmental grants. These predictions are critical reasons in the Popperian sense clarified above, that support our theory. Several predictions can be supported by empirical evidence that fails to refute them as the next chapter clarifies.

Chapter 5 analyses empirically a subset of critical reasons introduced in chapter 4 for which an empirical analysis may (not) support the hypothesis. Firstly, it is shown why Portugal can be considered as having a centralized fiscal structure and secondly, it analyses empirically some implications of the model.

Chapter 6 considers a particular congestable local public good: primary school education provided by Portuguese local governments. It will be shown that congestion is higher and quality is lower in urban areas and that this is a result of a continuing policy of central government negative discrimination against urban areas.
Chapter 2

Equilibrium and Efficiency in an Economy with Local Public Goods

Introduction

This thesis analyses the relationship between population size of communities and the provision of local public goods within a centralized fiscal structure. It will be shown that even in reasonably decentralized fiscal structures, equilibrium (both interjurisdictional (spatial) and intrajurisdictional (political)) is unlikely to occur. It will also be argued that, even assuming that such an equilibrium exists and prevails, provision of local public goods is inefficient.

However, the main literature on local public economics has stressed either equilibrium or optimality or both. These problems were initially addressed in the fifties by Losch (1954) who studied locational equilibrium in a spatial economy and by Samuelson (1954) who solved analytically the problem of the necessary conditions for optimality in an economy with private and pure public goods.

The problems which remained to be solved were the existence, uniqueness and stability of an equilibrium (or equilibria) in an economy with local public goods, and on the other hand the practical mechanisms to get infor-
mation on the preferences for public goods since they were not revealed in
the market. Tiebout's paper (1956) joined the two problems and suggested
an approximate solution which became known as the Tiebout Hypothe­
sis. The extensive literature following this seminal paper, has addressed
equilibrium and optimality issues.

A decade after Samuelson and Tiebout several authors started to dedi­
cate particular attention to non-market decision-making processes develop­
ing a distinct research program known as theory of public choice. Here the
"benevolent" government of neoclassical welfare economics which implicitly
or explicitly wants to maximize a social welfare function was challenged on
a number of different grounds. Within the public choice research program
a more positive approach to economic policy was given and the existence
and stability of a "political" equilibrium was discussed.

Even if these research programs ran side by side some interface between
them arose. Scholars started to integrate Tiebout with politics, and the
research issue became a study of the properties of an equilibrium which is
simultaneously an interjurisdictional (locational) equilibrium and an intra­
jurisdictional (political) equilibrium.

This chapter primarily addresses the problem of why such an equilib­
rium is not likely to arise under reasonable assumptions, and why even if
it exists it is almost certainly inefficient. The non-existence of an equilib­
rium carries with it the irrelevance of the Pareto optimality problem. This
conclusion will give support to a more positive approach to the effects of
community size on the provision of local public goods.

In section 2.1 the seminal contributions of Tiebout, Buchanan and
Hamilton will be considered. Buchanan’s paper (1965) was the first analyt­
ical approach to the theory of club goods (with private and public charac­
teristics) which gives a better understanding of the nature of local services.
Hamilton’s paper (1975) emphasizes the disequilibrium process likely to
occur in a Tiebout "world" and the necessary legal constraints (zoning) to
establish equilibrium. Section 2.2 surveys literature which stresses inter-jurisdictional equilibrium and optimality in an economy with local public goods and mobile individuals. Section 2.3 addresses the same problem when local public goods are considered as being club goods. Section 2.4 discusses the analogy between communities and clubs, and the validity of applying the club goods model to communities and section 2.5 offers some conclusions.

2.1 Seminal papers: Tiebout, Buchanan and Hamilton

Three papers (Tiebout (1956), Buchanan (1965), Hamilton (1975)) had a major influence on the development of the literature that addresses equilibrium and optimality in an economy with local public goods. In this section these contributions will be analysed.

2.1.1 Tiebout and the Importance of Migration

Tiebout's paper (1956) was the first important contribution to address the problem of efficiency in consumption and production of local public goods, in an economy with mobile individuals.

He developed a model which assumes that (i) consumers have no costs of mobility (ii) have full information about the revenue-expenditure "packages" of each of a large number of communities, (iii) employment opportunities are not considered in migration decisions and (iv) public services generate no spillovers between communities.  

Tiebout further assumed that (v) the average cost function of the "composite" local public good is U-shaped due to the existence of a fixed factor,  

1This means that the benefits of the local services provided in one community do not extend to residents of neighbouring communities.

2In fact Tiebout considered that each community provides a bundle of local services, but since this bundle is fixed it can be considered a composite good.
so that there is an *optimal community size* for each pattern of community services, and that (vi) communities below the optimal community size try to attract new residents, and those temporarily over the optimum size will return to the “right” size due to an auto-exclusion of some residents.

From these strong assumptions Tiebout concluded that consumers will “shop around” the communities and go to the community in which the predetermined pattern of revenues and services best satisfies their preferences. As a consequence they reveal their preferences for local public goods, and the problem of preference revelation stated by Samuelson (1954) has a conceptual solution for (at least) local public goods. Efficient use of public resources will be achieved because each community with optimum size will produce public services at minimum average cost. Efficiency in consumption will *not* be reached as a result of institutional rigidities but “like a general equilibrium solution for a private spatial economy, is the best that can be obtained given preferences and resource endowments”.

Tiebout’s optimal community size can be illustrated with figure 2.1 where $AC$ is average costs $N$ is population size and $N^*$ is the optimal size.

The Tiebout article gave rise to a stream of literature which will not

---

^See figure 2.1 where population size is on the horizontal axis and average costs on the vertical axis.

^Tiebout (1956) p. 424. This quotation was introduced to clarify that there is no presumption in the Tiebout model regarding overall efficiency (in production and consumption) as it is so often stated.
be reviewed here. However, some misunderstandings must be avoided so that a full comprehension of the present relevance of the model is achieved. Confusion arose as a result of a mixture of normative and positive interpretations of the model.

The model was built with normative purposes in mind and the assumptions were consciously unrealistic. From these strong assumptions, conclusions were deduced about production and consumption efficiency in the local public sector. If efficiency is assumed as the objective of public policy, measures should be taken so that reality becomes closer to the assumptions. In fact, one of the major conclusions for policy was that residential mobility and increased consumers' information about the packages of services/revenues of several communities will improve the allocation of local governments' expenditures. Tiebout wanted to emphasize the importance of migration on the allocation of resources in the public sector.

Therefore, criticisms of the Tiebout model based on the unreality of the assumptions show a misunderstanding of the nature and purposes of the model. He did not suggest any hypothesis subject to empirical refutation.

Having said this, we do not want to dismiss the positive interpretation of the Tiebout model which has also been made frequently. This approach tries to understand whether the local sector in the real world is sufficiently "Tiebout-like" to enable the use of the model not only for normative purposes (for which it was built), but also to make predictions. This is a different problem and will be considered later in section 2.1.3

Tiebout's paper (1956) was the first to deal with the problem of urban growth controls, congestion costs and local public goods from the perspective of optimal community size.

The optimal dimension of the community was defined in terms of production efficiency in the public sector and zoning was referred to as an instrument that could be used by communities which already have the op-

---

4See Zodrow (ed.) (1983). Developments relevant to our topic will be introduced in section 2.2 below.
timal size. This gives for the first time an economic rationale for zoning, which was only formally explored twenty years later by Hamilton.

It is worth emphasizing that Tiebout considered the situations where communities have a population below and over the optimal size as asymmetrical. This asymmetry is justified on political grounds. In the case of a less than optimal city size, communities may attract residents through Chambers of Commerce or other means. In this case local governments have to play a role if efficiency is to be achieved. On the other hand, when communities are bigger than the optimum he considers that it is not feasible politically to get rid of residents. In this case economic forces will intervene so that residents will choose to leave the community because other communities will become more attractive.

2.1.2 Communities as Clubs

The club goods model was introduced in an analytical form by Buchanan (1965) and the initial formulation addresses the problems of the optimal level of provision and membership size of facilities in which there is shared consumption and the possibility of excluding non-members from consumption. The model considered that homogeneous individuals (i.e. same tastes and endowments) derive utility not only from goods but also from the number of individuals that share the consumption of the goods. Buchanan considered that utility may increase initially with the number of individuals sharing the consumption of a given facility (due to camaraderie) but that after some level of utilization is reached additional members will impose congestion costs on others.

As a consequence there is an optimal membership size of the club good for each level of provision of the facility. Optimal membership size is reached when the marginal benefit of having a new member (i.e. the decrease in per capita cost due to the marginal member) just offsets the marginal (congestion) cost that he imposes.

On the other hand, for a given sharing group there is an optimal level of
provision, so that the individual marginal rate of substitution between the club good and a private numeraire good must be equated to the individual's marginal provision cost (payment).

Optimality for the members of the club (also known as within-club optimality) requires that the membership and provision conditions are simultaneously satisfied.

The idea that local governments could be analysed as clubs and that the Tiebout model could be understood as a "nonspatial world of voluntary clubs" was formulated by Buchanan and Goetz (1972).

Their paper emphasizes the limits of the Tiebout model; namely, that the use of the assumption of individuals living with dividends income is a way of neutralizing the problem of efficient allocation of resources in the private sector.

Regarding jurisdictions as fiscal clubs means that communities can be considered as clubs that provide local collective goods, and residents may be analysed as members of the club, who pay for local services through local taxes.

The model of one club developed in Buchanan's paper (1965), was generalized to two, and migration from one jurisdiction to another as moving from one club to another. The migrant imposes net costs or net benefits in both the jurisdiction where he enters and from which he leaves. If local collective goods were purely public, in the sense of perfect nonrivalness of consumption, the receiving jurisdiction would have net benefits with new residents, because the same collective goods would be provided at a lower average cost. However if congestion exists, a new migrant will impose congestion costs on the existing residents and this effect may offset the decrease in the average cost.

Buchanan and Goetz concluded that to achieve efficiency, fiscal discrimination between old and new residents would be necessary. Nevertheless, since this goes against the principle of equality of individuals before the

---

4The authors attribute this idea to M.Polinsky in a Working Paper of the Urban Institute. For a discussion of this approach see section 2.4 below.
Law and against the idea of free migration, they considered that local governments should not make such a discrimination. In reality they do not make fiscal discrimination and, even if they wanted to and no legal or constitutional constraint existed, they would be unable to do so because of informational problems.

The important point of Buchanan and Goetz's paper is that it represents a departure from the appeal of the efficiency criterion as a norm for public policy, even if fundamentally in practical terms.  

In this sense this paper is more advanced, from a positive political economic point of view, than Hamilton's paper (1975) produced three years later, which again emphasizes efficiency issues. However, Hamilton's paper had the merit of clarifying the importance of zoning regulations.

3.1.3 Hamilton and The Role of Fiscal or Exclusionary Zoning

Tiebout did not explicitly introduce the system of revenues of local governments and his remark about zoning was a brief one. It was Hamilton who filled this gap.

Hamilton (1975) extended the market-like analysis of the local public sector with the introduction of "prices" of local services. His approach was a positive interpretation of the Tiebout model since he considered it as a good model for describing consumers' behaviour in suburban sections of urban areas.

The model he developed used the same assumptions of the Tiebout model and he adds that (i) the only revenue source of local governments is a proportional property tax (variable between communities) and that (ii) each community has a "zoning ordinance" establishing the minimum

---

7 It is useful to distinguish between practical and substantive reasons for rejecting a normative criterion. The former refers to the fact that it might not be feasible to apply it in practical terms (e.g. the inability to use fiscal discrimination to achieve Pareto efficiency) and the latter to a criticism of substance.
amount of housing to be consumed by each household wishing to live in the community.

With these assumptions several results were derived: Pareto optimal levels of the local public good are provided; the property tax is an efficient price; communities are homogeneous with respect to house values; income redistribution by local governments is impossible and tax and expenditures of local governments are fully\(^6\) capitalized into house values.

The crucial role of zoning to obtain these results can be better understood if assumption (ii) is dropped. In this case an endless disequilibrium process will occur as can easily be shown. High income households will try to locate together since a high residential tax base may provide either a high level of local public goods provision \textit{given} the property tax rate or a low property tax rate \textit{given} the levels of public services provision (or a combination of the two). Low income households will migrate to high tax base jurisdictions for the same reasons and also because they will pay much less in taxes than they receive in local services (assuming the property tax is proportional).

The zoning ordinance brings about the possibility of achieving an equilibrium because it establishes the minimum of housing consumption (which means also the minimum of property tax given a property tax rate) within the community. Therefore, it is not rational for a household to live in a house with a value higher than the minimum value settled by the zoning regulation of his community.\(^9\)

This gave rise to another “economic” rationale for land use controls, known in the literature as fiscal or exclusionary zoning. It means that

\(^6\)There was some misunderstanding between no capitalisation and full capitalization as we will clarify in the next section. Hamilton presented his results (1975, 1976b) as achieving no capitalisation even if he now prefers the concept of full capitalization (1983), and this is the reason why we use the latter concept instead of the former one.

\(^9\)Irrationality here means non utility maximisation. This household can increase its utility by moving and buying a house of the same price in another jurisdiction where the minimum housing value exactly matches the value of the house. Of course, this assumes the existence of multiple communities.
communities may use zoning regulations as a way of obliging new residents to pay a reasonable share of the local budget. In the extreme case (the Hamilton model is in equilibrium), communities are homogeneous regarding house values and each resident pays exactly the average cost (equal to marginal cost) of provision of the local public services.

Hamilton's approach also considers the jurisdiction as a club, zoning being the necessary exclusion device, so that only households willing to pay the "price" of local public services will consume them (and will enter the jurisdiction).

This analysis was extended in a later paper (Hamilton 1976a) he considers a metropolitan area with many communities, each one providing a given level of public service financed by a proportional property tax and having balanced budgets. He further considers heterogeneous communities with two kinds of houses, high income (HIH) and low income houses (LIH), and also (for didactic purposes) two homogeneous communities, one with HIH only and the other having LIH.

The capitalization equations play a crucial role in his argument. They can be expressed as $V_i = V' + (X - V_{iti})D$ 10 They state that the house value in a mixed community equals the value of the house in the homogeneous community plus the present value of the "fiscal surplus". The fiscal surplus (positive or negative) is the difference between public service benefits per household ($X$) and tax liability ($V_{iti}$). Low Income Houses are more expensive in mixed communities than in the LIH homogeneous community because the fiscal surplus there is positive and capitalized into the house value. The reverse happens in HIH in mixed communities.

On an interjurisdictional basis it means that (ceteris paribus) a household is indifferent between living in a LIH (HIH) belonging to a mixed or homogeneous community even if it has a positive (negative) fiscal surplus in the former and a nil fiscal surplus in the latter. In fact, if we consider

10There are two identical equations referring to HIH and LIH. $V'$ is the value of the house in the homogeneous community.
a model where households derive utility from a composite private good, housing and a local public good, and if it is further assumed that there are identical demands for public services in all communities a \( LIH \) (\( HIH \)) will pay less (more) in the mixed communities for the same level of public services and (due to full capitalization) will pay more (less) for housing. As a consequence, it will consume the same amount of the private good and obtain the same level of utility that it would have had if it were located in the homogeneous community.

Hamilton uses the capitalization equation to argue that even in heterogeneous communities property taxes may constitute average cost pricing for local public services. This is because even if households in \( LIH \) pay less property taxes (than those in \( HIH \)), they pay also a positive capitalization premium. Hamilton can not establish this result for each household individually but only for the aggregate households (in \( LIH \) and \( HIH \)) of a mixed jurisdiction. In aggregate, since the sum of fiscal differentials is nil, property taxes are just able to finance the local public service.

Average cost pricing of local services is only efficient if it also equates to marginal cost. If it is assumed that the average cost function is U-shaped regarding population size, there is only one point where that equality holds, which is the minimum of the average cost function as we have seen above in figure 2.1.

Since Hamilton wants to argue that average cost pricing is efficient irrespective of community size he must assume constant average costs. In fact Hamilton recognizes that although the average cost function is U-shaped in the short run, it might be considered "roughly" horizontal in the long run, after a minimum size of population is reached where economies of scale are exhausted.  

---

12This statement (see Hamilton (1976a) footnote 2 on page 744) is crucial in Hamilton's argument. See also the discussion in Hamilton (1983) pp. 92-94. The shapes of cost functions and the role of congestion are critical issues that will be explored in chapter 4. As will become apparent then, our approach has similarities with Hamilton's approach.
Hamilton's paper (1976a) is interesting because it explores the possibility of Pareto efficiency being compatible with heterogeneous communities. The importance of considering heterogeneous communities lies in the fact that homogeneous communities may be too small to fully exploit the economies of scale in the production of local services.

However, Hamilton's model depends critically on the zoning assumptions.

The literature developing the seminal papers of Tiebout, Buchanan and Hamilton followed two main strands. On the one hand there is the research on equilibrium and optimality in a decentralized economy with local public goods and migrant individuals. The emphasis here is on interjurisdictional efficiency. This approach will be considered in section 2.2. The same problem has been addressed using club goods theory, and the relevant literature will be introduced in 2.3. Finally, the issues of whether communities can be considered as clubs and the property tax can be seen as a benefit tax will be reviewed in section 2.4.

On the other hand another strand of literature has focused on theoretical approaches to intrajurisdictional efficiency and empirical analysis of the demand for local public goods related to different approaches to local political processes. This will be considered in chapter 3.

2.2 Equilibrium and optimality in an economy with local public goods

The problem of the existence, uniqueness and stability of an equilibrium in an economy with local public goods is crucial and logically prior to the problem of optimality. In fact if existence or stability of an equilibrium

---

Footnotes:

13 On the distinction between interjurisdictional and intrajurisdictional approaches see Rubinfeld (1987).
can not be established, the optimality issue becomes irrelevant. In spite of this (and perhaps due to the complexity of general equilibrium models) there is a comparatively small literature that addresses equilibrium issues in comparison to the literature that addresses optimality issues.

There have been two main strategies to deal with these problems. The first approach considers "local public goods" as public provided private goods or club goods so that there is no reason (on efficiency grounds) why they should not be provided privately. As a consequence the problem is shifted into the issue of general equilibrium in an economy with only private market goods (club or "normal" private goods). This approach will be discussed in the next section.

The second approach, to be introduced now, considers "local public goods" as being pure public goods (in Samuelson-Musgrave terms) within the boundaries of the community, but generating no benefits to individuals living outside the jurisdiction. Authors are either sceptical about the existence and stability of an equilibrium or they prove the existence of an equilibrium under very restrictive assumptions. The problematic nature of the existence and stability of an equilibrium when both interjurisdictional locational choices and intrajurisdictional political choices are considered is that either locational equilibrium or voting equilibrium are problematic when considered separately.

One of the more important disequilibrium processes considered in the literature is related to the problem of exclusion illustrated by the issue of the "poor chasing the rich" on fiscal grounds. This is the case, referred to above, where low income individuals are attracted to jurisdictions where high income individuals live in order to benefit from a higher tax base and consequently lower tax bills for the same level of public services.

14On club goods see section 2.1.2 above.

15Stiglitz (1983) defends this concept saying that it is the "pure case" of local public goods and is similar to the pure (national) public goods concept introduced by Samuelson (1954).
The solution that Hamilton suggested to this issue relies on strict and binding zoning regulations which are far from easy to enforce in many communities. Zoning establishes locational equilibrium and, as a result of the segregated homogeneous solution, there is no need to consider political equilibrium since in homogeneous communities each individual is representative of the community. Hamilton kills two birds with one stone by using a rather unrealistic model.

Another approach to solve the disequilibrium problem is to consider whether local governments may discriminate fiscally between local residents, and in particular may use Lindahl taxes. The Lindahl equilibrium\(^{16}\) is very attractive theoretically but also very unrealistic. A Lindahl equilibrium is obtained as a result of an iterative process where an auctioneer calls out tax shares for every individual. Each individual then responds by declaring the utility-maximizing quantity of the public good at that tax-share. The process ends where every individual chooses the same quantity of the public good. Eventually, an equilibrium is reached and it is Pareto efficient.

As a consequence, any distribution of preferences and income between members of a community can be compatible with an internal (political) equilibrium if an appropriate set of Lindahl-taxes and lump-sum redistribution is in place. Greenberg (1983) presents a model where equilibrium is compatible with mobility. However, he allows for Lindhal taxes, which not only have never been used but also are not feasible in practice due to informational problems.

Stiglitz (1977, 1983) analyses the problem assuming that the local services are pure public goods and considers whether the two fundamental theorems of welfare economics are valid in an economy with local public goods. The problem is to know if, in such an economy, every competitive equilibrium is Pareto efficient and under which conditions Pareto efficient

\(^{16}\)See Lindahl (1958) and Johansen (1963)
allocations can be sustained by a competitive equilibrium (with "correct" lump-sum redistributions). As referred to earlier Samuelson showed that with pure public goods, such an equilibrium would not be Pareto efficient and Tiebout suggested that the problem could be partially solved in the context of an economy with local public goods and costless migration of individuals.

Stiglitz showed that only under very special and unreasonable conditions would Pareto efficiency emerge. In fact, three conditions should be met simultaneously: (i) the right number of communities must exist, (ii) individuals must be distributed correctly among jurisdictions and (iii) each jurisdiction must provide the right supply of local public goods. None of these conditions hold under reasonable assumptions.

Stiglitz (1977) analyses the case of two communities ("islands") with total population 2N, where the optimal size of the jurisdiction \( N^* \) is such as: \( N < N^* < 2N \). He shows that there are Pareto inferior equilibria that are stable and that a Pareto efficient equilibrium may be unstable.

Rose-Ackerman (1979) develops a model showing that an equilibrium in a metropolitan area with competing fiscal jurisdictions may not exist and even if it exists may not be stable. She points out that any model of local government fiscal behaviour should contain three distinct features: (i) multiple competing local governments, (ii) a model of the political process and (iii) a land market model since local fiscal choices have an impact on land values. Her model applies to a metropolitan area with a fixed number of communities, with fixed land areas where the price of land is determined competitively. Furthermore, individuals can vote on the tax-expenditure of their community and also migrate costlessly between jurisdictions. She assumes that local political choices are driven by the median voter.

Rose-Ackerman's paper clarified the difficulty in proving the existence and stability of an equilibrium and shows that if it exists it is likely to be inefficient. She considers a model with two communities each producing a local public good with identical U-shaped average cost functions. Assuming
that total population equals the sum of the optimal population of each community (at the point where average costs are minimum) an efficient equilibrium would be achieved when population is split in two between communities.

However, if the starting point is an unequal distribution of population and the average cost function is assymetric this will imply higher average costs in the smaller jurisdiction and lower average costs in the larger. Migration will continue towards the larger jurisdiction in so far as average costs remain lower and eventually equilibrium will be achieved with all the population in the larger jurisdiction.

Rose-Ackerman also argues that sometimes residential patterns in metropolitan areas are stable, but this does not imply the existence of an equilibrium of a Tiebout type. On the contrary, it might mean that there are frictions such as: migration and commuting costs, elections are not frequent and there is some loyalty to neighbourhoods from residents-voters.

Epple, Filimon and Romer (1984) agree with Rose-Ackerman that a model analysing a local public economy should include at least a housing (or land) market and a political market. In such a model equilibrium would mean intercommunity equilibrium, i.e. no individual can increase his utility by moving to another jurisdiction, each community has positive population and is in internal equilibrium. This means that in each community the housing market is cleared, the community budget is balanced, there is political equilibrium and residents consume utility maximising consumption bundles. To prove the existence of an equilibrium they had to introduce strong (and unusual) assumptions regarding consumer preferences and production functions.

On the other hand Bewley’s model (1981) considers the case where local public goods are publicly provided private goods financed with income taxation, and shows that no equilibrium is Pareto optimal.

In several papers market failure is associated with the integer problem.\(^{17}\)

\(^{17}\)See subsection 2.3.1 below for a clarification of the integer problem.
The general conclusion of the literature that addresses the problem of existence, stability, uniqueness and optimality of an equilibrium (or equilibria) in a decentralized economy with competing local jurisdictions, local public goods and migrant individuals is that equilibrium may not exist under reasonable assumptions. It exists and is stable if strong assumptions are made (e.g. perfect zoning or the ability to use Lindahl taxes). Even if it exists it is very likely to be inefficient, i.e. only under extremely unrealistic assumptions will the Pareto efficiency conditions (on consumption and production) be satisfied simultaneously.

Nevertheless, this conclusion is not necessarily pessimistic. With Tiebout the majority of the economists agree that lower costs of mobility and a better understanding by residents of the package of tax-benefits offered by competing local jurisdictions improves efficiency in the local public sector and that local government provision is preferable to central government provision, since it allows a better satisfaction of preferences for local collective goods.

Another approach that has reached similar conclusions using a different path is based on club goods theory. This approach will be considered next.

2.3 Equilibrium and optimality in a system of clubs

After the seminal paper of Buchanan the literature on club goods has had a great development. Buchanan's model considered homogeneous individuals, with a fixed utilization rate of the club where the optimality conditions were derived assuming that only the utility of individuals who share the club is relevant.

Subsequent literature developed different approaches corresponding some-
times to different problems under analysis and on other occasions to different strategies that authors use to overcome complexities in models.

The first important dichotomy is related to the "within-club" or "total economy" points of view. The former considers only the utility of the members of the club, while the latter consider also the non-members’ utilities.

The existence of non-members is related to the partition issue. If total population is partitioned into clubs, then every individual belongs to a club and there are no non-members. If additionally clubs are equal in membership and capacity sizes and each individual belongs only to one club, then the "within-club" approach becomes tantamount to the "total economy" point of view.

Another important distinction relates to homogeneous versus heterogeneous clubs. Homogeneous clubs are formed with identical individuals in tastes, endowments and personal characteristics. Heterogeneous clubs have a mixed composition of attributes described above. An interesting case is when an heterogeneous population is "discretely homogeneous", i.e. can be partitioned into homogeneous subgroups.

### 2.3.1 Communities providing only one service

Initial models considered population homogeneous in tastes and income and only one local collective good. Pauly (1967,1970) analysed the problem of the optimal dimension of the club \( (N^*) \) in a similar way as Buchanan, defining it as that which maximizes the average net benefits of the members.

He demonstrated that a Pareto optimal solution exists and is a stable equilibrium, if the total population \( N \) is a multiple integer of \( N^* \) and if there is equal cost-sharing between the members of the club.

Heterogeneous population (either in tastes or income) was introduced through what we have labelled as discrete homogeneity. In this case, and considering one local public good 19, it was demonstrated [Pauly (1970),

\[19\] This means that there is only one kind of service under analysis. This is compatible with the existence of several clubs providing the same service and having the same size or
McGuire (1974), Henderson (1979), Berglas and Pines (1981)] that efficiency will require that individuals are segregated into homogeneous jurisdictions.

The simplest case is when there are two types of individuals $a$ and $b$ with total population $N$ such that $N = N_a + N_b$ and the optimal size of the clubs, $N^*_a$ and $N^*_b$, is different. In this case, efficiency will require that there is no integer problem within each subset of homogeneous population, i.e $N_a/N^*_a = k_a$ and $N_b/N^*_b = k_b$ with $k_a$ and $k_b$ integers. Two alternative interpretations of this condition might follow.

If each community provides one club, $k_a$ and $k_b$ are the numbers of homogeneous communities having individuals of type $a$ and $b$ respectively. This approach implies that the number of communities is variable, which is only consistent with a long-run approach, and can be replicated, which assumes that existing communities do not possess special locational attributes.

On the other hand, if each community can supply more than one identical club, the number of communities is uncertain. Let us define $H_a$ and $H_b$ as the two sets of communities with homogeneous individuals $a$ and $b$ respectively. Let $a^i$ belong to $H_a$ and $b^j$ to $H_b$ and having population $N^i_a$ and $N^j_b$ respectively.

Necessary conditions for efficiency will require that:

$$\forall a^i \in H_a, \frac{N^i_a}{N^*_a} = k^i_a$$

and also

$$\forall b^j \in H_b, \frac{N^j_b}{N^*_b} = k^j_b$$

with $k^i_a$ and $k^j_b$ (positive integers) being the number of identical clubs existing in jurisdictions $a^i$ and $b^j$.\(^{21}\)

---

\(^{20}\)This assumption has been made very often and gave rise to a (misleading) identification between "communities" and "clubs". The distinction will be explored in the next section. We will use quotation marks whenever we want to emphasize the difference between those concepts.

\(^{21}\)A total partition of population into homogeneous communities would imply also that
It becomes apparent why the relaxation of the homogeneous population assumption leads to a growing difficulty in achieving Pareto optimality.

The discussion so far has assumed that optimality in "discretely homogeneous" populations requires partition into homogeneous clubs. This has been supported by Pauly (1970), McGuire (1974), Berglas (1976), Stiglitz (1977), Henderson (1979) and Berglas and Pines (1981).

Some of these authors recognized that heterogeneous "clubs" might be preferable to homogeneous ones (with respect to optimality) if economies of scale can not be exhausted with homogeneous clubs. Moreover, heterogeneous "communities" are also preferable if they provide more than one local public good or if the needs of production of private goods (that need "skilled" and "unskilled" workers) are considered.

However, Sandler and Tschirhart (1980,1984) and Cornes and Sandler (1988) have been arguing that heterogeneous clubs may be optimal independently of the reasons just stated. They argue that heterogeneous clubs might be provided efficiently if (i) provision (the facility capacity) affects utility only through congestion, (ii) congestion functions are similar for everyone (iii) utilization rates of the club good can be monitored at a low cost, and (iv) individual cost shares may be discriminated according to utilization rates.

They also point out that the (correct and apparently contradictory) result that Pareto optimality requires homogeneous clubs is derived from a "second best" constraint that members share costs equally. Under this constraint it becomes obvious that heterogeneous clubs where members consume different amounts of the club good and pay the same price are Pareto inferior to equal size homogeneous clubs. In so far as transaction costs (monitoring and "price" discrimination costs) are low they argue in favour of heterogeneous clubs.

Nevertheless, we will argue that when considering "communities" and

\[ \sum_{i} k_{i} = k_{0}, \sum_{j} k'_{i} = k_{0}, \sum_{i} N_{i} = N_{0} \text{ and } \sum_{i} N'_{i} = N_{0}. \]

Berglas and Pines's model (1981) to be presented below assumes provision affecting utility directly.
local public goods these transaction costs are likely to be high.

The models considered so far, deal with only one local public good in the sense of one service. However, to become more realistic regarding its application to local communities, the provision of several services must be considered. Only a few authors have attempted to extend the club goods model to several goods. We will consider next the approach of Berglas and Pines (1981) and Berglas (1984) in more detail since it enables a better understanding of the general applicability of the club goods model to “communities” and local public goods.

2.3.2 Communities providing several services

Berglas and Pines (1981) made one of the few attempts to integrate the club goods approach with local public goods in a quasi-Tiebout framework.

Following Buchanan (1965) they distinguish private, club and local public goods according to the optimal size of the sharing group. If it is the “smallest possible” a good is private, if it is finite but small relative to community size there is a club good, and if the optimal sharing group is the community itself there is a local public good.

They consider a model where individuals’ utility in an homogeneous population is given by

\[ u = u(x, v, y, n v) \]  

Utility of an individual increases with the consumption of a private good \( x \), the number of visits to the club \( v \), the size of the facility \( y \), and decreases with the groups’ overall consumption of the club good \( n v \) where \( n \) is the membership of the club.

This formulation was criticized by Sandler and Tschirhart (S-T) on the basis that utility should not be dependent directly on facility size but in-

\[^{23}\text{They do not acknowledge Buchanan's pioneer approach, in defining endogenously the “nature” of goods in terms of optimal dimension of the sharing group, but it is evident their debt to Buchanan.}\]
directly through the congestion function. The usual form of the utility
function assuming homogeneous individuals that has been used in the lit-
erature (also by S-T) is

\[ U = U(x, v, c(k)) \]

where \( c \) is the congestion or crowding function, and is assumed to depend
on the average utilization rate \( k \), where \( k = nv/y \).

Berglas and Pines (1984) reply saying that their formulation is more
general and enables \( y \) to be interpreted has denoting a vector of characteristics of the club good.

They also consider a resource constraint given by:

\[ nx + c(y, nv) = nI \]  \hspace{1cm} (2.2)

where \( I \) is individual income and \( c(.) \) is a cost function. The reason \( nv \)
appears as an argument in the cost function is because maintenance costs
are assumed to increase with overall consumption of the club good. The
private good is assumed as a numeraire good so that its price is equal to
one.

Maximization of the utility function (2.1) subject to the resource con-
straint, and assuming \( n \) fixed, gives the following necessary conditions for
an optimum: \(^{24}\)

\[ u_2/u_1 = c_1 - nu_4/u_1 \]  \hspace{1cm} (2.3)

\[ nu_2/u_1 = c_1 \]  \hspace{1cm} (2.4)

Equation (2.3) equates each member's marginal rate of substitution be-
tween club visits and the private good \(^{25}\) to the social cost imposed by
an additional visit: the marginal maintenance cost \( (c_1) \) plus the marginal
congestion cost imposed on all members.

Equation (2.4) is the Samuelson condition that the sum of marginal
rates of substitution should equal the marginal cost of provision.

\(^{24}\)Subscripts denote partial derivatives regarding the \( i \)th argument of the utility or cost
functions.

\(^{25}\)Since the private good is a numeraire good \( u_2/u_1 \) gives the marginal monetary valuation of each visit.
A second step of the maximization problem can be done allowing \( n \) to change. Since \( u^*(n) \) is the optimal level of utility regarding \( n \), when equations (2.1) through (2.4) are satisfied, Berglas and Pines derive (using the envelope theorem) the additional condition:

\[
\frac{du^*(n)}{dn} = \frac{u_1}{n^2}[c - nv(c_2 - n\frac{u_1}{u_1})] \tag{2.5}
\]

This condition can be rewritten, assuming that both utility and cost functions are homogeneous of degree \( r_u \) and \( r_v \) respectively, as:

\[
\frac{du^*(n)}{dn} = \frac{u_1}{n^2}r(1 - r_v) + \frac{u_1}{n}r_u \tag{2.6}
\]

A particular case is when the utility function is homogeneous of degree zero regarding capacity and overall consumption \( r_v = 0 \) and production exhibits constant returns to scale \( r_u = 1 \). In this case

\[
\frac{du^*(n)}{dn} \equiv 0
\]

This means that maximum utility is invariant to the size of the community. This is an important result and we will return to it in section 4.3.

Equation (2.5) gave rise to some controversy in the literature as can be seen in: Berglas and Pines (1981), Sandler and Tschirhart (1984) Berglas and Pines (1984) and Cornes and Sandler (1986). The expression \( c_2 - nu_4/u_1 \) should be the price per visit of the club, since with price equal to marginal social cost we will have (see equation (2.3)), \( u_2/u_1 = P_v/P_s \) to induce utility maximizing consumers to consume the "right" mix of the private and club goods. The expression in square brackets in (2.5) is costs minus revenues of the club good. Equation (2.5) provides a taxonomy of goods (private, club and local public goods) according to the optimal sharing group, as is illustrated by figure 2.2.

With two services (1 and 2) and homogeneous individuals, optimal dimension of the community \( (n^*) \) is obtained, through the maximization of the utility function, under the budget constraint and also the constraint \( \sigma_1 = \sigma_2 = N \). This constraint (referred below as HPC) means that the residents that consume service 1 \( (\sigma_1) \) and service two \( (\sigma_2) \) should be identical.
Figure 2.2: Optimal sharing groups for different goods

to the population size of the community ($N$). This is a consequence of the homogeneous population (HP) assumption.

Assuming that the optimal sharing group of the facilities ($n_1^*, n_2^*$) exists, is finite and is different ($n_1^* \neq n_2^*$), the HPC constrains the maximum utility achievable. In fact, if say $n_1^* > n_1^*$, this implies $n_1^* > n^* > n_2^*$. On the other hand the HPC implies $n_1 = n_2 = n^*$, so that membership in one is too small ($n_1 < n_1^*$) and membership in two is too big ($n_2 > n_2^*$). Under this condition B-P show that facility one would be produced under a deficit and facility two under a surplus, so that the surplus just finances the deficit (which suggests both goods should be provided publicly).

Considering now an heterogeneous population, i.e. if the homogeneous population assumption is dropped (and consequently the HP constraint), it is possible to derive necessary conditions, where optimum community size is compatible with optimum sharing groups of both facilities.

As a conclusion, in a model where two or more facilities are provided by local communities, the optimal solution under an heterogeneous population will Pareto dominate the optimal solution with an homogeneous population.

In another paper Berglas (1984) develops what is the more realistic approach to local collective goods in local communities using a Buchanan/Tiebout framework. He considers that “local services” possess mixed private and public characteristics and should be distinguished from “pure local
public goods" in the sense that the Samuelson condition of zero marginal cost of sharing does not hold.

Therefore, he specifies the utility function of an individual as:

$$ u = u(x, y, Q) $$

Where $x$ is a private (possibly composite) good, $y$ is the quantity of a "local service" (the private characteristic) and $Q$ is the quality of such service (the public characteristic). As a consequence an important distinction arises between differences in tastes with respect to quantities and to qualities of the local service.

Berglas assumes that (i) each service is produced and consumed within the community (ii) residents can vary the quantity consumed (iii) cost of provision depends on its utilization (iv) the average cost per unit is dependent on total utilization and can be increasing, decreasing or constant.

If additionally, (v) residents have different preferences regarding the quality of the local service and if (vi) there is only one local service provided, he concludes that the Tiebout hypothesis holds, i.e. an efficient equilibrium with segregated (homogeneous) communities would arise.

On the other hand if (v') preferences regarding quantities differ (and are identical regarding qualities) and (vi') several public services are considered, efficiency will generally require heterogeneous communities. This result is reinforced if economies of scale in production of the local service are considered.

---

Generally the literature has considered only one of these aspects to identify local public goods. Emphasizing the first characteristic lead Stiglitz (1974) to consider education as a "public provided private good" and stressing the Samuelson condition to consider local services as "pure local public goods" (Stiglitz (1983)).

He introduces Tiebout's hypothesis as meaning that with costless mobility among communities of utility maximizing individuals, communities would be of the optimal size, supplying the optimal mix of public services and each community would have individuals with similar tastes. As we discussed earlier (see supra 2.1.1) Tiebout did not formulate an hypothesis. However, Berglas' approach is consistent with the later interpretation of Tiebout's paper (1956).
This means that different preferences regarding qualities and quantities of the local services have opposite effects regarding the composition of communities in the equilibrium solution (if it exists). If differences regarding qualities are "sufficiently strong" efficiency may still require segregated communities. However, if they are not, efficiency will require heterogeneous communities.

The proof of existence of equilibrium when several local services are considered is done under very restrictive conditions. Constraints have to be imposed on the form of the cost functions and individuals are assumed to have identical tastes regarding the quality of local services (or this quality is assumed to be fixed within each community). Under these constraints, an equilibrium must satisfy three conditions: (a) all communities are identical and mixed, (b) each community is efficient with respect to size and provision of goods and services (c) communities do not redistribute income.  

It is sufficient to appraise the necessary and sufficient conditions for the existence of an equilibrium to understand how unlikely it is that such an equilibrium will arise. Moreover Berglas (1984) does not consider a "political sub-model" to explain the level of provision of local services in heterogeneous communities.

Summary

The club goods approach to equilibrium and optimality in a system of local communities providing local services, developed mainly in a Buchanan-Tiebout framework, has given additional insights to the literature that has a local public goods' approach to local services, and on the other hand has reached conclusions that are consistent with this literature.

The further insights given by the club goods approach are related firstly, to a much more realistic approach to the nature of the local services, in which each "local public good" might be seen as a club; secondly, the em-

---

This is Berglas' proposition 4 (page. 313-314).
phasis on the simultaneous importance of utility and cost functions, and finally, the different institutional contexts where these goods might be provided since it is clear that they do not necessarily require public production and (or) provision.

The consistent results with the non-club goods literature, are related to the difficulty in establishing the existence and stability of an equilibrium.®®

Nevertheless, the club goods approach, that considers each community as a club has severe shortcomings as will be considered below.

2.4 Communities as clubs?

This section addresses the problem of the applicability of the club goods model to each community.

This issue must be distinguished from the club goods approach to local services. The difference is whether the club goods model is applied to each community or to each local service. This distinction is sometimes obliterated in the literature, since several models consider that each community provides one local public good. In this particular case both approaches are identical. However, when several local public goods are considered, it is necessary to discuss if the club goods approach still applies to each community, as suggested by Buchanan and Goetz (1972).

This problem is relevant since to establish the analogy between each community and a club is a strong argument towards an efficient provision of local public goods, since it is acknowledged that clubs might provide services to their members efficiently if appropriate user charges are introduced.

Henderson (1979) pointed out correctly that there are some necessary adaptations to the club goods model so that it might be applied to juris-

®®And not only the cost functions as in Tiebout (1956).

®®The implications of this result will be explored in the last sections of this chapter. However, it can be anticipated that a fundamental corollary of this conclusion is that there is an over emphasis on Pareto optimality issues.
dictions. Firstly, membership is now reached through the act of buying a house within the jurisdiction. Secondly, the provision of public services is not done under direct benefit charges but is generally done through "indirect" property taxes. Finally, excludability is no longer a straightforward issue.

To consider each community as a club it is necessary to discuss (i) Communities/clubs and entrepreneurs, (ii) the nature of payments made by the members of each community/club and (iii) the possibility (or not) of excludability.

2.4.1 Communities, clubs and entrepreneurs

The fundamental appeal of club good theory is the ability to overcome the preference revelation problem. This is possible because club goods are not "pure local public goods", but have mixed private and public characteristics. In particular it is possible to exclude non-members from consumption. Optimal provision and membership can be obtained in a competitive market\(^\text{31}\), because it is assumed that entrepreneurs are always ready to create new clubs, if the existing ones do not have optimal dimension and/or do not set efficient charges. On the other hand it must be assumed that clubs can be replicated, i.e. it must be possible that an optimal dimension club is duplicated, triplicated or replicated \(n\) times in identical clubs. This is generally assumed possible in the long run, where all factors of production are assumed variable.

However, when considering communities several complications arise, even not taking into account the important problem of excludability.

Firstly, city managers can not exactly be identified with entrepreneurs. Secondly, the number of communities is generally fixed in the short/medium run and even in the long term when new communities might be created they can not be identical to previous ones. As a consequence, reaching a

\(^{31}\text{As referred to earlier, in so far as the optimal dimension of the club is small in relation to total population.}\)
competitive market equilibrium of the sort proposed by Berglas (1984) is problematic.

Thirdly, if communities finance local public goods through the budget and provide them without direct charge, efficiency can not be achieved in heterogeneous communities. In fact heterogeneous residents can neither adjust their levels of consumption of local public goods, nor pay different tax-prices according to their marginal valuations of the public goods.

Furthermore, there are two types of members. Owners of property and tenants. Owners do pay property taxes but whether this tax has anything to do with the benefits received from local services or not is a controversial issue in the literature. Tenants pay rents and the fact that rents include a partial shift of the property tax is also a debated issue on the incidence of the property tax. We will turn now to the issue of whether the property tax can be considered a "benefit" tax.

2.4.2 Is the Property Tax a "Benefit" Tax?

A central and ongoing theme in the analysis of efficient allocation of resources in the local sector has been the discussion regarding the benefit nature of the property tax.

As was referred earlier, Hamilton (1975),(1976a), assuming that the property tax is the only source of revenues of local governments and that they use appropriate zoning ordinances, argued that the property tax could be considered a benefit tax in homogeneous communities and could also be considered as such in heterogeneous communities if perfect capitalization of fiscal differentials into house values is considered.

In this sense, he argued, consumers face the correct "prices" for local public goods and there are no distortions in the housing market.

The further conclusion that could be derived if property taxes were benefit taxes is that the exclusion from the benefits of local services could be made efficiently, and the distinction between club and public provision

---

32 Assuming also that Lindahl taxes are not possible to implement.
of local collective goods would be irrelevant regarding efficiency. In other words, taxes would became user charges.

There are two important aspects that have been considered in the literature. One is the empirical problem of knowing what is the level of capitalization of fiscal differentials into house values. The other is theoretical i.e. whether perfect capitalization gives support to a "benefit" approach to the property tax or is consistent with other approaches.

Empirical research on capitalization surveyed by Bloom, Ladd and Yinger (1983) leads to two major conclusions: the existence of serious methodological problems, and a significant level of inter and intrajurisdictional property tax capitalization. 33

The "benefit" view of the property tax has been challenged in recent years by the "new view" developed mainly by Zodrow and Mieszkowski. 34

Using a general equilibrium approach, they concluded that the burden of the property tax is placed on the owners of capital (residential or not) and thus, is a progressive tax on capital instead of a "benefit" tax.

Mieszkowski and Zodrow assume that the total stock of capital of the nation is inelastic, that each jurisdiction sets its public expenditure levels to maximize the welfare of its residents and finances them either with a lump-sum tax or a property tax. Each local government perceives a completely elastic supply of capital. They derive the conclusions that optimality requires lump-sum taxes, the availability of which is dependent on perfectly binding zoning regulations. 35

33There seems to exist a certain contradiction in these two conclusions since if the methodological problems are relevant, it is difficult to reach any definite statement on the issue.


35In this sense Hamilton's theory can be considered a particular case of their theory when zoning is perfectly binding and the head taxes may be interpreted as nondistortionary property taxes. However this approach depends crucially on the unrealistic assumptions about zoning.
As they relax this assumption, local governments need to rely on property taxes and capital bears the “average” burden of the property tax in the nation due to the inelastic supply of capital. Jurisdictions with residents who have relatively high demands for public services will have higher property taxes and so a higher cost of capital which result in higher housing prices, lower wages and land rents. The reverse happens in low tax jurisdictions.

This has important consequences regarding the provision of local public goods. Local governments financed by property taxes will under provide public services because they are reluctant to impose taxation on mobile capital that might leave the jurisdiction. However, as grants from central government are introduced as an additional source of local governments’ revenues, the case for under provision cannot be established. Moreover Mieszkowski and Zodrow argue that capitalization is not discriminatory between Hamilton’s “benefit” approach and their “new view” of the property tax because both are consistent with capitalization.

As a conclusion the “new view” on the property tax, does not give theoretical support to the community/club analogy since property taxes are no longer seen as the payment for the services received by local governments. At the same time it does not support the hypothesis that local governments provide local public goods efficiently.

2.4.3 How excludable are new residents?

The other main difference between local governments and clubs is the issue of excludability. In the case of clubs the optimal dimension of the club can be achieved easily because there are direct excludability devices to forbid consumption to nonmembers.

When communities are considered, there is usually no mechanism to exclude non-resident individuals from the consumption of certain local public goods. 

---

55 This same approach that “tax competition” might lead to under provision of local services was developed earlier by Oates (1972) and more recently by Wilson (1987).
goods such as public parks, the local road network. This gives rise to spillovers of benefits between jurisdictions and is a known source of inefficient provision of local public goods.

On the other hand there is no direct mechanism to exclude entry of new residents. Anyone willing to buy or rent a house within a jurisdiction will benefit from the local public goods provided.\footnote{The existence of an optimal community size depends on the shapes of the cost functions of local public goods and the congestion functions, as discussed above.}

In so far as there is an optimal community size and an optimal composition of the community the problem of excludability becomes relevant.

Three types of excludability have been developed in the literature. It is worth considering if they apply only to non-residents or to non-residents and residents alike, and in the former case whether it discriminates between individuals or not.

Firstly, there is excludability due to congestion (Tiebout \(1956\)). As utility is generally assumed to be a decreasing function of congestion, increased population leads to self-exclusion of some residents and excludability of some non-residents. Secondly, there is exclusion based on zoning regulations. It is directed to non-residents and generally it discriminates regarding housing/land values. Thirdly, there is fiscal exclusion: overconsumption of public services is an indirect way to exclude individuals who have lower preferences for local public goods and/or lower incomes, since they will also have to share the costs of provision (Stiglitz \(1977\), Henderson \(1979\)).

In spite of different forms of exclusion, zoning has been in certain countries\footnote{Mainly in the U.S.A., where the main body of literature on zoning has been produced.} the main instrument used by local governments to constrain urban growth and consequently for excluding non-residents. Zoning regulations have assumed two main forms, namely those which regulate the allocation of land for different uses (residential, industrial, environmental) and those which directly affect the housing market through density and lot-size regulations.
There are four main reasons that have been suggested in the literature that explain why selfish local residents would like to adopt zoning regulations.

Firstly, lot-size zoning enables local residents to exclude low-income citizens in order to avoid subsidizing the newcomers through the local budget. In fact, if local services are provided equally to all residents and financed through a proportional property tax, new low-income residents will impose a redistribution in kind through the local budget. Zoning obliges new residents to “pay their way” and this is the reason why it is known as fiscal zoning.

Secondly, if local public services are produced using public purchased inputs and also community residents' characteristics, zoning may act as a filter selecting the characteristics that have greater "productivity" in the production of local services. This approach was introduced in Mills and Oates (1975) and developed in Oates (1977), (1981) Schwab and Oates (1991). It seems to apply particularly to education, police, and other services where residents appear both as consumers and producers of local services.

Thirdly, zoning might have an effect on land and house values. According to this approach zoning and other land-use controls can be seen as a local residents' monopoly power on the local housing market. Constraining the supply of new houses is a way to increase the value of the old ones, so that owners of property will be better off, but potential residents are worse off.

Finally, zoning regulations are much cheaper methods of excluding non-residents and achieving optimal community size than alternative ones such as

---

35 See Mills and Oates (1975), White (1978) and the applications made by Hamilton (1975),(1976a) and discussed supra in this chapter.

40 Zoning as a monopoly power as been advocated by Hamilton (1978) and Fischel (1985).
as compensation due to expropriation under eminent domain \(^4\), or a direct purchase of land.

All these reasons seem to support the conclusions that (i) local residents are better off with the control of urban growth through zoning or other land-use controls, therefore (ii) local governments (mainly in suburbs) do control urban growth. However, exclusion is "excessive" when overall efficiency is considered, i.e. the welfare of both residents and non-residents is considered. These conclusions are subscribed to by the dominant legal and economic literature\(^5\).

In spite of an overall consensus on (i), the logical step from (i) to (ii) is controversial since it implies a model of decision-making of the median voter type. In fact, zoning being a "public good" for local residents does not imply that it will be provided by local governments, unless a median voter model is implicitly or explicitly assumed.

In this context, authors who consider a "monopolistic" approach to local governments emphasize the rent-seeking behaviour of developers in order to obtain favourable regulations. Moreover, in many countries and regions, illegal construction (without planning permission) is a common fact particularly in suburbs of metropolitan areas. To name but a few in Europe: Portugal, Spain, southern Italy, Greece and Turkey all face the problem of a more or less chaotic suburban growth.

Therefore, excludability is not a straightforward issue in local communities and is certainly a much more complex issue than exclusion in clubs, where an exclusion device and a fee are enough to exclude non-members.

\(^4\)This is also known as "taking", and is the situation where local governments establish that there is a "public interest" reason that justify a transfer of property rights from the private sector to the public sector.

\(^5\)See respectively Ellickson (1977) and Fischel (1985).
Summary

This section analysed critically arguments put forward in the literature suggesting that each community might be seen as a club supplying local services to its members (residents). This approach has some normative appeal, since residents live voluntarily in the same jurisdiction, and make payments to contribute to the provision of shared consumption facilities. However, on positive grounds the analogy seems to be taken too far. "Entrepreneurs" do not establish communities easily, "members" generally pay property taxes that are not directly related to the benefits received and "excludability" might be achieved but it is complex and not easily enforced.

If communities were clubs, the achievement of optimal community size would be easy, and the reality of suburban growth (a temporarily disequilibrium) would only mean that new residents were imposing lower marginal congestion costs (on actual residents) than marginal benefits (through the enlarged tax base of the jurisdiction). Not being clubs, communities can not achieve intrajurisdictional optimal size, so that a continuous disequilibrium process might be in place initially assuming the form of suburban growth and afterwards the form of changes in the composition of population with a reasonably stable size of population.

2.5 Conclusion

This chapter addressed two main problems in the literature of an economy with local public goods, migrant individuals and local political processes of decision making. The first one asks whether an equilibrium (or equilibria) exists and is stable when residents reveal their preferences for local public tax-expenditures by "voting with their feet". The second questions whether such an equilibrium (assuming it exists) is Pareto efficient or not.

The general conclusions asserted in the literature are that an equilibrium fails to exist under reasonable conditions, and that when an equi-
librium exists it is likely to be inefficient. These conclusions might seem rather pessimistic, but some insights were reached since Tiebout advanced his "hypothesis". It is generally acknowledged, with Tiebout, that lowering the costs of migration and widening the information available to residents on the fiscal packages of several local governments, improves allocative efficiency.

The main implications of these conclusions are twofold. Acknowledgment that disequilibrium is likely to occur, particularly when there are no practical processes to control community's "membership" implies that disequilibrium processes are worthwhile considering. Therefore communities are not able to achieve optimal community size even in decentralized fiscal regimes. In fact, with respect to residential mobility there are two fundamental aspects that reflect the existence of disequilibrium: urban growth and changes in the composition of an heterogeneous population.

The changing nature of the size distribution of population across communities, the fact that these communities can not achieve optimal population size even if they wish to do so, has three major implications. Firstly, too much attention has been given to optimality, which in the absence of an equilibrium is a normative property without relevance. Secondly, the implementation of policies aiming at horizontal fiscal equity across communities becomes problematic since the locational disequilibrium means also that the housing market is not stable and therefore it is difficult (if not impossible) to study the long run incidence of a set of intergovernmental grants to local authorities.

Thirdly, it suggests the need for a positive inquiry into the effects of population size of communities on the provision of local services. Before proceeding to this analysis within a centralized fiscal structure, we will consider the literature that has addressed this problem within a decentralized fiscal structure.

As clarified in the beginning of this thesis within centralized democratic states, central (or federal) governments have a major role in setting the ex-
penditures of lower level governments because: (i) intergovernmental grants are a substantial share of local revenues, (ii) the more important local tax rates are set by central government and (iii) there is a control on local borrowing. This introduces special and peculiar problems that constitute the core of this thesis and will be dealt with in chapters 4, 5 and 6.

It is useful to compare at this stage three different institutional frameworks: the centralized approach (CC), the decentralized Tiebout approach (DC) and the decentralized median voter approach (DD).

In the Tiebout framework the fiscal structure is decentralized (there are many local governments with independent powers) but within the jurisdiction it is centralized (the mayor sets the revenue-expenditure “package”). As a consequence citizens do not reveal their preferences within the jurisdiction. They “vote with their feet” which means that they have to migrate to the jurisdiction that best suits their preferences.

On the other hand, in the classical median voter framework citizens reveal their preferences only within the jurisdiction. It is decentralized both at the overall government fiscal structure and at the local government level.

The approach developed in this thesis assumes an institutional structure that shares something with the Tiebout framework and shares nothing with the institutional framework assumed in the median voter framework. It assumes centralization at both intergovernmental and local government levels. The point of contact with Tiebout is that it is also a centralized approach from an intrajurisdictional point of view. Therefore, some of the conclusions of the Tiebout approach can be extended to centralized

---

43 This approach will be developed in the next chapter. As we will see initial approaches consider the demand for local public goods as a function of intrajurisdictional “tax-price” and income variables of the median voter in each community. Therefore, it is exclusively intrajurisdictional variables that matter. Only later on developments of the model started considering the effects of migration and how they biased the estimates of the parameters obtained in the classical approaches.
governments.\textsuperscript{44}

The relevance of the median voter literature, in the context of our thesis, is that most of the empirical literature uses this framework. Therefore, it provides some empirical content to the problem of economies of community size regarding population. Moreover, it has an intrinsic normative interest as far as comparative institutional economics is concerned.

\textsuperscript{44}See section 4.4.
Chapter 3

Economies of community size, congestion and the role of intergovernmental grants

This chapter surveys empirical literature on economies of community size and positive approaches to intergovernmental grants.

We will search for an answer in the existing literature to the following two problems:

1- How does the population size of communities affects the local resources needed to provide a standard quality of local services?

2- Are intergovernmental grants expected to offset differentials in the fiscal position of communities resulting from economies of community size?

Answering the first problem presupposes dealing with the "privateness" or "publicness" of local public goods and has been a major topic of economic debate. The second issue, set in a positive framework, has had a much smaller importance since it implies some sort of endogenization of intergovernmental grants.

These issues have been discussed mainly in terms of the median voter approaches to the demand of local public goods. The conclusions obtained

\footnote{This concept, which is central in our thesis, will be clarified in section 3.1.6.}
are, among other things, dependent on the model of local government as­sumed. Therefore, a brief review of these models will be introduced in this chapter.\footnote{However, some of these models will not be much developed since our interest in local decision making models is marginal (in the context of this research). As clarified in chapter 4.3 within centralized fiscal structures there is no need of such a model to determine local government expenditures.}

Models of local government fiscal behaviour can be grouped in three main categories, according to the types of agents relevant in the process of decision-making, their preferences, the information available to the agents and the resource constraints they face.

Firstly, there is the “full-information” median voter model. There is only one type of agent relevant in the analysis - the median voter in each jurisdiction. Only their preferences over private and local public goods are relevant. Local governments are not autonomous agents, they are mere instruments of the preferences of this decisive voter in each jurisdiction. This voter does not need to have full information. On the contrary, he actually might be rationally ignorant. The important issue regarding information is that he does not systematically misrepresent the true nature of either costs or revenues in the local public sector. Herein after this model will be called the classical median voter model.

Secondly, there are bureaucratic and fiscal illusion models within the median voter framework. These models have at least two kinds of agents with different sets of preferences. Their common characteristic is that there is an asymmetry of information between one kind of agents at the local government level (usually the bureaucrate) and the median-voter. In general bureaucratic models assume that politicians have the same set of preferences as the median voter (due to the wish to be re-elected) so that the asymmetry of information is between politicians and bureaucrats. In both models, although for different reasons, the preferences of the median voter are not satisfied.

Thirdly, there is a more heterogeneous group of models where the me-
dian voter does not play any role. These models emphasize more supply-side aspects of the local decision-making process, give more importance to the institutional constraints faced by local governments and also to the status quo. They assume, in general, that local governments have a considerable degree of monopoly power, that there are strong barriers to entry in local political markets and that there is a very strong asymmetry of information between agents at the local government level and the voters. All this contributes to neutralize the voters' power in general and the median voter in particular. There is typically one agent (the politician or the bureaucrat) with a maximizing behaviour or not. We will label these models as pure supply side models to emphasize that the median voter (and voters in general) do not play any role in these models.

Therefore, section 3.1 gives particular emphasis to the results regarding congestion and economies of city size using the classical empirical approach to demand for local public goods within the median voter model. This model has been used very comprehensively in the literature, and most of the available results were produced in this framework. Section 3.2 considers theoretical and empirical bureaucratic approaches within the median voter framework, and also the pure supply side models. Section 3.3 surveys positive approaches to intergovernmental grants, and the flypaper effect. Finally, section 3.4 summarizes and concludes.

3.1 The median voter model

It is understandable that when the economic approach started to be translated to the political “market” that the same methodology was employed and as aggregate demand for a private good is obtained by aggregation of individual demands for that good, that (with some qualifications) demand for local public goods could be considered as reflecting the preferences of the citizens of the respective polity.

The fundamental qualifications are (i) that local public goods are provided uniformly to all the members of a jurisdiction, so that citizens (withe-
out migrating) can not adjust the quantity consumed unilaterally, and (ii) that prices do not play any role when these goods are provided publicly. In spite of these qualifications, within the demand-side approach it is considered that local governments are institutions that aggregate, in some way, the preferences of the respective constituents.

The median voter model has been the more widely used and gives a crucial role to voting as the main form of political participation of the citizens in a local political process which is assumed to be competitive.

The median voter theorem helped this task, since it establishes that, in a context of uni-dimensional issues where the preferences of the voters are single-peaked \(^3\) and a majority rule is employed to achieve the collective decision, the alternative preferred by the median voter will never loose against any other alternative.\(^4\)

In a context where political parties (entrepreneurs) set the agenda before the electorate, it assumes perfect information of utility-maximizing voters and of vote-maximizing parties.

It is not surprising that the median voter model was introduced to derive the demand for local services, even if the institutional context was not the most appropriate.\(^5\) A more detailed approach to the empirical studies using the model will be given in the next section.

### 3.1.1 The Classical Median Voter Approaches

In order to simplify the exposition of different models, a uniform notation will be used. Some concepts will become clear during the critical survey of

\(^3\)Single-peaked preferences when considering preferences over the level of provision of a public good, means that there is an utility maximizing level (say \(g^*\)) and that utility decreases for levels increasingly distant from \(g^*\) (be it higher or lower levels).

\(^4\)The initial formulation is due to Duncan Black (1948).

\(^5\)In fact, as Romer and Rosenthal (1979) clarified, the only institutional context that seems relevant to an empirical analysis is to local school expenditures in the United States. This paper is perhaps the best critical appraisal of the median voter. Our criticisms of the model can be found in Pereira (1991).
the literature. The notation is as follows: 

$y$ = private good

$X$ = aggregate provision (in ‘capacity’ units) of local public goods (or clubs) in a community

$X^i$ = dimension of a unique local public good (club)

$x$ = local public service or local public good (LPG) in ‘units of consumption’

$E$ = total expenditure on LPG

$e$ = per capita expenditure on LPG

$N$ = population of the community

$n$ = population size of a club

$\tau$ = tax-share or the fraction of the total cost of LPG’s that his paid by one individual of the jurisdiction

' (superscript) refers to individuals, whenever it is assumed that the population is not homogeneous

' (superscript) refers to communities

$b$ = individual tax base (e.g. assessed value of property)

$B$ = Tax base of the jurisdiction (including all sorts of property)

$B_1$ = Residential local tax base

$B_2$ = Non-residential local tax base (commercial, industrial, etc.)

$p_X$ = “Tax-price” of local public goods (production units)

$p_x$ = “Tax-price” of local public goods (consumption units)

$I$ = gross income

$\alpha$ = population elasticity of demand

$\delta$ = price elasticity of demand

$\varepsilon$ = income elasticity of demand

$C(.)$ = Cost of local public good

$c = C(.) / X$ = average or unit cost of LPG

$\beta$ = elasticity of labour in the Cobb-Douglas production function

$\rho$ = elasticity of capital in the Cobb-Douglas production function

$\eta$ = Technological parameter which is the inverse of the degree of homogeneity of the production function ($\eta = 1/(\beta + \rho)$)
\[ \gamma = \text{crowding parameter} \]
\[ \hat{\gamma} = \text{estimator of the crowding parameter} \]
\[ \cdot \text{(hat) refers to estimates of parameters} \]

In two seminal and widely quoted papers Bergstrom and Goodman (BG) (1973) and Borcherding and Deacon (BD) (1972) established what we call the classical empirical approach to demand for LPG. They assume that local public facilities are produced with constant returns to scale within each jurisdiction, i.e. that the production function is homogeneous of degree one. BG do not specify the particular form of the function but use the hypothesis that the commodities are supplied at a constant unitary cost within and across jurisdictions. BD explicitly assume a Cobb-Douglas function with constant returns to scale within each jurisdiction.

Therefore both authors assume

\[ C(X) = c \cdot X \quad (3.1) \]

They also consider the same “technology of consumption”. Each individual only “captures” a fraction of the provided LPG given by:

\[ z = \frac{X}{N^\gamma} \quad (3.2) \]

If \( \gamma = 0 \) the facility is a pure public good, while if \( \gamma = 1 \) it is “as if” the individual only receives and benefits from a fraction \( 1/N \) of the local public good. In the latter case the LPG has “privateness” characteristics.

This specification is a constant elasticity (\( \epsilon \)) of “crowding” since:

\[ \epsilon z N = \frac{\theta (X/N^\gamma)}{X/N^\gamma} = -\gamma \]

Therefore, a one percent increase in population is associated with a \( \gamma \) percent decrease in “captured” or subjective consumption independently of the level of population.

This model is “as if” the production of local services is done in two stages. The first one is the production of facilities (public sector output). The second one is the “production” of services. Services \( x \) are facilities \( X \) weighted by the factor \( 1/N^\gamma \). What is relevant for the consumer (what
appears in his utility function) is $x$ services and not $X$ facilities. As a consequence of the public good not being pure ($\gamma > 0$) population size acts as a "deflator" of public facilities.

**Tax-share and "Tax-price".**

The individual tax-share $r^i$ is the fraction of total cost of municipal expenditures that is paid by the individual. It is exogenously defined in the model of BG which considers it to be independent of the size of municipal expenditures and of preferences regarding municipal services. In spite of this they refer to the tax share as being related to population size and the size of commercial and industrial tax bases.

How the tax-share is related to population size is not explained in their paper but can be easily shown with the following alternative assumptions.

If all taxable property is residential and assuming homogeneous consumers (in incomes and preferences), the tax-share is inversely proportional to population:

$$r^i = \frac{b^i}{\sum_{i=1}^{N} b^i} = \frac{b^i}{N b^i} = 1/N$$

The higher the population of the jurisdiction the lower the tax-share.

However, if households have different preferences regarding housing then

$$r^i = \frac{b^i}{\sum_{i=1}^{N} b^i} = \frac{b^i}{N b^i} = \frac{1}{\bar{b}}$$

In this case individual tax-share is a function of jurisdiction size and the relative value of his house compared with the average house value of the jurisdiction.

Finally, assuming homogeneous individuals but now the existence of residential ($B_1$), commercial and industrial tax-bases ($B_2$) which finance the local budget, the individual tax-share is now

$$r^i = b^i/(B_1 + B_2) = b^i/(N b^i + B_2) = 1/(N + B_2/b^i)$$

Tax-share is a function of the population size of the jurisdiction and also the relative size of the non-residential tax base.

---

*We assume that effective tax rates on all kinds of property are the same.*
These examples show that whenever residential property tax base is a considerable part of the tax base of the jurisdiction, the population size of the community has a significant effect on the tax-share.\(^7\)

The "tax-price" \( p \) of the local public goods is the "price" each household pays of the unit cost of the public goods: \( p_X = r c \) where \( c \) is the unit cost measured in units of production \((X)\).

Therefore, they define the tax-price in community \( i \) as:

\[
p_X^i = r^i c
\]  

(3.3)

Assuming constant returns to scale in the production of \( X \) within and across jurisdictions, \( c = C = C/X \).

It is also possible to consider the tax-price in terms of "units of consumption" or the subjective tax-price. In this case \( p_x = r.C(X)/x \) and assuming \( x = X/N^7 \),

\[
p_x = r \frac{C(X)}{X} N^7
\]

Introducing the constant returns to scale assumption yields:

\[
p_x = r c N^7
\]

(3.4)

A particular case arises when property is residential and individuals are homogeneous (i.e. \( r = 1/N \)) so that

\[
p_x = c.N^{(r-1)}
\]

This result is similar to the approach of BD.\(^8\) They label it "subjective marginal tax-price" to indicate that it is not the marginal tax-price per unit of public good \( provided \) but per unit of good consumed or "captured".

The formulation of the subjective tax-price has the disadvantage of being a function of the "crowding" parameter \( \gamma \). If \( \gamma = 0 \) (pure public good), each individual pays \( 1/N \) of the average (and marginal) cost of the public

\(^7\) This point will be further discussed below.

\(^8\) For BD \( c \) is marginal cost instead of average cost which are identical under the assumption of constant returns to scale.
good \((p_t = c/N)\). If \(\gamma = 1\) each individual pays a price equal to marginal cost as in private goods \((p_t = c)\).

The empirical results of BG/BD were obtained with the assumption of constant returns to scale in production and suggest that the crowding parameter is not significantly different from one. This has important implications regarding economies of community size. It leads to the conclusion that there are no economies of community size. The fact that communities are bigger does not allow a reduction on per capita costs since provision must increase in proportion.

The relationship between tax-prices in production and consumption units is given by

\[
p_t = N^\gamma p_x
\]  \hspace{1cm} (3.5)

Some remarks should be made at this stage relating to the “tax-price” concepts.

The tax-price in “units of production” has a clear meaning of being the individual share of the unitary cost of production of LPG. Whatever the definition of tax-share, it is always an inverse function of population. Moreover, if the existence of constant returns to scale is assumed, the tax-price is not dependent on the level of provision of the LPG.

The subjective tax-price is a more problematic concept since it depends on the specification of the crowding function. Such a function potentially introduces two additional problems. Firstly, the subjective tax-price might be a function of the level of provision even assuming constant returns to scale in production. Secondly, the effects of increasing population can be either positive, negative or nil. To see this, consider the example where the crowding function is given by \(g = g(X/N) = (\frac{X}{N})^\gamma\) with \(0 < \gamma \leq 1\).

Therefore, we would have

\[
p_t = \tau c X^{1-\gamma} N^\gamma.
\]

\(^{9}\)This is an example of an homogeneous crowding function. The argument for using an homogeneous crowding function (not this one in particular) will be developed in chapter 4. Other specifications of the congestion function will be discussed in 3.1.4 below.
Since \( r \) is a function of \( N \) the effect of population on the subjective tax-price would be given by:

\[
\frac{\partial p_x}{\partial N} = cX^{1-\gamma}N^\gamma \left[ \frac{\partial r}{\partial N} + \gamma r N^{-1} \right]
\]

The term within brackets can have any sign. A particular case where it equals to zero is when \( r = 1/N \) and \( \gamma = 1 \), which is the same case seen above where each citizen pays a subjective price equal to marginal cost so that the size of the population in his community is irrelevant.

**Demand, Tax-price and the Congestion Parameter**

Let the utility function of an individual be:

\[
U = U(y, z)
\]  

(3.6)

and his budget constraint \( y + p_x X = I \).

Using the BD/BG crowding function we can have \( p_x X = p_x N^\gamma z = p_x N^\gamma (X/N^\gamma) = p_x X \), so that the budget constraint can also be expressed as

\[
y + p_x z = I
\]  

(3.7)

Maximizing \( U \) subject to the latter budget constraint, it is possible to obtain the demand function for \( x \).

However, BD and BG postulate a log-linear demand function for \( x \) (services) with constant "price" and income elasticities :

\[
x = Ap_x^I I^e
\]  

(3.8)

For Bergstrom and Goodman the quantity \( X \) demanded by the median voter can be obtained substituting equations (3.4) and (3.2) into (3.8) so that:

\[
X = N^\gamma Ap_x^I I^e = A(rcN^\gamma)^I I^e N^\gamma = A(rc)^I N^\gamma \gamma^{(I+1)} I^e = A^c(r)^I N^\gamma \gamma^{(I+1)} I^e
\]

The demand equation estimated by BG is therefore :

\[
\ln X = \ln A' + \delta \ln r + \gamma(\delta + 1) \ln N + c \ln I + u
\]  

(3.9)
where $A' = Ac'$.

They assume that total expenditures are a proxy for output.\(^{10}\) Therefore, they obtain the following expression:

$$\ln E = \ln A + \delta \ln r + \delta \ln N + \delta \ln I$$

The estimates of the elasticities of demand regarding population ($\delta$) and the tax-price ($\delta$) enable the indirect computation of the crowding parameter:

$$\hat{\gamma} = \frac{\delta}{\delta + 1}$$

The approach of Borcherding and Deacon is similar, but they introduce explicitly a Cobb-Douglas production function with constant returns to scale and assume that the rate of return of capital is constant across jurisdictions but wages differ between jurisdictions so that

$$C_X = \frac{\partial C}{\partial X} = d'w^\beta$$

where $d'$ is a constant\(^{11}\)

$w$ is the wage rate and $\beta$ the elasticity of output with respect to labour in the Cobb-Douglas production function. Therefore, tax-prices of the local public goods are not constant across jurisdictions and are given by:

$$p_x = C_X N^{(1-\gamma)} = d'w^\beta N^{1-\gamma}$$

Expenditure in BD is not a proxy of output but given by

$$e = p_x x = p_x X$$

\(^{10}\)To be more precise BG should do what BD do, and consider $E = cX$, so that we would have $E = Ac^{x+1}(c)^{\gamma} N^{(x+1)}I^\gamma$ and taking logs, $\ln E = \ln A + \delta \ln r + \gamma(\delta + 1) \ln N + \delta \ln I + u$ with $A'' = Ac^{x+1}$. This correction changes the intercept but does not change the elasticities in so far as $c$ is assumed constant within and across jurisdictions.

\(^{11}\)See BD equations 3' and 3'' bearing in mind differences in notation. $d'$ is a constant since the rental rate on capital is assumed constant.
It is given in per capita terms since the tax-price is a per capita tax-price due to the tax-share being \( r = 1/N \). Using the same demand function and the definition of tax-price it is possible to obtain an expression for expenditures per capita:

\[ e = p_x x = A p_{x+1} I^* = A \left( \frac{1}{N} d' w^\delta N^{\gamma+1} I^* \right) = A d^{\delta+1} w^\delta \delta^\gamma (\gamma+1) N^{(\gamma+1)(\delta+1)} I^* \]

For estimation purposes they realize a transformation which consists of dividing both terms by \( w^\delta \) which yields 12:

\[ \frac{e}{w^\delta} = A' (w^\delta)^\delta N^{(\gamma-1)(\delta+1)} I^* \]

where \( A' = A d^{\delta+1} \).

Taking logarithms and adding an error term yields:

\[ \ln(\frac{e}{w^\delta}) = \ln A' + 6 \ln(\delta^\delta) + (\gamma - 1)(\delta + 1) \ln N + \varepsilon \ln I + u \quad (3.10) \]

This was estimated using ordinary least squares:

\[ \frac{\hat{e}}{\hat{w}^\delta} = \hat{A}' (\hat{w}^\delta)^\hat{\delta} N^{(\gamma-1)(\hat{\delta}+1)} I^* \]

The estimator of the crowding parameter is now given by:

\[ \hat{\gamma} = \frac{\alpha}{\hat{\delta} + 1} + 1 \]

which enables one to calculate \( \gamma \). 13

Note that \( w^\delta \) is proportional to marginal cost so that the series \( w^\delta \) can be used as a proxy of the series of prices \( d' w^\delta \) for local public goods. However, the price variable that appears in the estimated equation is \( N \) times the price of local public goods \( (X) \). 14

The essential distinction between BD and BG regarding the formulation of the tax-price variable should be clarified because failure to do so seems to

12 The objective of the transformation is to get direct estimates of \( \delta \) instead of \( \delta + 1 \) (see BD (1972) footnote 20 in page 898). Note that the expression for \( \gamma \) given below would be different. Without this transformation \( \gamma = \frac{\delta}{\delta + 1} \).

13 Since the ratio of two unbiased estimators is not unbiased BD used an approximation of the relationship given above (see BD (1972) pages 898-899).

14 \( N p_X = N N^{-1} C_X = C_X \).
have led to some controversies in the literature regarding the introduction (or not) of population into the tax-price variable.\textsuperscript{16}

Bergstrom and Goodman consider, both in their model and in the empirical application, a tax-share variable which is used for the definition of the tax-price variable. In spite of being \textit{exogenously} defined they say that it is related to population. In other words if it was endogenized it would be a function of population among other variables.

It follows that the population elasticity of expenditures measures the effect of population given incomes and tax-prices. This has two main implications. Firstly, this elasticity does not fully measure the effect of population on expenditures since when population changes so does the tax-price. On the other hand since the tax-price is the "correct" tax-price\textsuperscript{16} the estimate of the "tax-price" elasticity is also the "correct" one.

Borcherding and Deacon implicitly assume that the tax-share is a function of population size. Since they endogenize the tax-share, when they simplify the reduced form demand equation, the tax-share does not appear into the tax-price variable. Therefore, the population elasticity incorporates the full effect of population on per capita expenditures. On the other hand the tax-price elasticity is not the "correct" one because it does not include the tax-share of the median voter.

\textbf{3.1.2 Technology of Production of Local Public Goods}

Technology within a reduced form demand approach

Due to the difficulty in measuring public sector output the empirical analysis of the demand for local public goods has generally assumed a technology with constant returns to scale either across jurisdictions or only within jur-

\textsuperscript{16}See Gonzalez and Mehay (1985), Wyckoff (1988) Gonzalez, Folsom and Mehay (1989), Wyckoff (1989). One of the points of debate is precisely whether the tax-share is a function or not of the population size of communities.

\textsuperscript{16}In this context we mean that it incorporates the tax-share variable. As seen below there has been a continuing debate on the correct specification of the "tax-price" variable.
Two pertinent problems regarding technology are to know (i) what are the consequences for the other parameters of the demand model if we drop that assumption and (ii) whether it is a reasonable assumption or not.

Inman (1979) addresses the first problem in the context of a Cobb-Douglas technology without the requirement of constant returns to scale. This specific technology implies that

$$C(X) = cX^n$$  \hspace{1cm} (3.11)

He assumes the same log-linear demand function and the same technology of consumption considered by BD and BG. However, Inman’s definition of the tax-share ($\tau = (b/B)\pi$) is not the share of local taxes paid by the median voter (as in BG) but by all residents in the jurisdiction. In fact if residential property is the only tax base of the jurisdiction ($B = B_1$), if there is no deductability of local taxes from federal taxes or local tax credits ($\tau = 1$) and individuals are homogeneous, the tax-share is equal to one ($\tau = 1$). If either or both assumptions do not hold $0 < \tau < 1$. Therefore population size does not enter into the tax-share variable.

Inman introduces population ($N$) into the definition of the subjective tax-price variable:

$$p_x = \frac{C(x)}{N}$$  \hspace{1cm} (3.12)

since he considers the tax-price as the share of the per capita average cost. This definition is similar to BD’s approach, although with a special difference. Inman considers implicitly an average cost pricing rule to finance the provision of local public goods since marginal cost pricing with increasing (decreasing) returns to scale would generate deficits (profits). Therefore, the tax-price is a function of the level of provision whenever there are not

---

17In fact any technology homogeneous of degree $k$ has a cost function given by that equation where $\eta = 1/k$. The Cobb-Douglas specification is therefore an unnecessary restriction made by Inman.
constant returns to scale.

With the production technology equation, the pricing rule equation, the consumption technology equation and the log-linear demand function it is possible to derive a reduced form equation of expenditures (assumed to equal costs) of local public goods:

\[ C(X) = \hat{A}t^\gamma p^\delta N^\alpha \]

The expenditure parameter estimates: \( \hat{\kappa} \), \( \hat{\delta} \) and \( \hat{\alpha} \) are a function of the demand parameters \( \kappa \) and \( \delta \) and of the technological parameters \( \gamma \) and \( \alpha \).

The following relations hold.\(^\dagger\)

\[ \hat{\kappa} = \frac{\kappa \eta}{\eta \eta} \]
\[ \hat{\delta} = \frac{\delta \eta}{\eta \eta} \]
\[ \eta \gamma = \hat{\alpha} + \hat{\delta} \]

There is one degree of indeterminacy since there are four structural parameters and only three reduced form parameter estimates (and equations). In general the only thing that can be known is \( \hat{\kappa} / \hat{\delta} = \kappa / \delta \) and \( \eta \gamma \) i.e. the ratio between income and price elasticities of demand and the product of the two technological parameters.

It is important to emphasize the meaning of the third equation. Assuming constant returns to scale \( \eta = 1 \) it is possible to obtain estimates for the "crowding" parameter \( \gamma \) and also for \( \hat{\kappa} \) and \( \hat{\delta} \) since \( \eta = 1 \) implies \( \gamma = 1 \)\(^\dagger\). However, with increasing (or decreasing) returns to scale it is not possible to obtain an estimate for the crowding parameter.

As referred above, several authors\(^\ddagger\) have concluded that \( \gamma \) is not significantly different from 1, when considering jurisdictions with populations

\(^\dagger\) See appendix A.3 for calculations. Our results are identical to Inman's except the third equation below which appears in Inman's paper (1979) as \( \eta \gamma = \frac{\delta}{\kappa} + 1 \). Since Inman does not introduce the calculations, for the time being, we accept our result as the correct one. Nevertheless, the discussion below is valid in any case.

\(^\ddagger\) Bergstrom and Goodman (1973), Borcherding and Deacon (1972), MacMillan et al. (1981) among others.
between 10000 and 150000 citizens. This seems to suggest that local public goods use a "private" good technology of consumption \( x = X/N \). Therefore each citizen would be indifferent to the size of the community where he lives insofar as the level of *per capita* provision of LPG's is equal across jurisdictions.

MacMillan et al. (1981) find that in communities with less than 10000 persons the "crowding" parameter is significantly lower than one. Since they assume constant returns to scale in production, they interpret this result as indicating "economies of sharing" in small communities. However, this interpretation is misleading since either it can truly mean that there are "economies of sharing" \( (0 < \gamma < 1) \) to be explored, that there are economies of scale \( (\eta < 1) \) or that both situations happen.

Returning to the problem addressed in the beginning of this section, when demand is considered in the reduced form equation, that the main implication of allowing increasing or decreasing returns to scale in is that it does not allow for estimation of the demand parameters and the "congestion" parameter. In particular the "crowding" estimates found in the literature should be read, according to the existing literature, as indicating the combined effect of sharing consumption and economies of scale.\(^{21}\)

A different approach to economies of scale is to consider directly the production function of local public goods.

### 3.1.3 Economies of scale in the production of local services

In discussing economies of scale in the local public sector two main problems arise. Firstly, is the notion of output that should be considered and the proxy that should be used in empirical analysis. Secondly, is whether the production function is a relevant conceptual tool and if so, the problem of

\[^{21}\text{In chapters 4 and 5 we will argue that the population elasticity should be taken as the combined effect of quality changes in output and economies of scale.}\]
the specification of the production function.

As we have seen above the median voter approach distinguishes the local public sector output from the services "captured" by the median voter. The mediation between the two concepts is done through a crowding function. This is a function of the size of LPG and the number of citizens sharing the consumption.

LPG output might be measured by the capacity of sewage and water systems, provision of schools, quantity of arrests or fines enacted by local police, number of books issued by local libraries and other indicators. Therefore, a production function might be said to exist that relates public purchased inputs to outputs.

Nevertheless, the output that is relevant in a demand framework (or more broadly for the citizen) is the level of services received through the existence of these goods and services.

It is the subjective service rather than output that appears in the citizens' utility functions. Theoretically, these "services" were not considered in the classical median voter approach as a different sort of output. Yet, algebraically they were almost formalized as such. Insofar as a crowding function exists and is continuously differentiable it is possible to shift from services to output and vice-versa.

This created, perhaps, a path to distinguish two concepts of output which have received some attention in the literature recently.

The seminal paper was Bradford, Malt and Oates (1969) that considered public sector output as distinct from what is really of interest to citizens. Their line of argument is that citizens are not interested in the arrests made by police but in safety standards. They appreciate school and staff qualities but they are more interested in the education that their children are able to receive in public schools. Both services are jointly produced through public sector inputs and the characteristics (or "environment") of the community.

Therefore it is possible to distinguish, at least in theory, a "first" stage of production of a public sector output (using inputs purchased by the
public sector) from a "second" stage of production where this output is combined with the characteristics of the community in the production of local services. Bradford, Malt and Oates are cautious in considering this a two-stage production function since they admit that the "environment" variables might not be continuous and differentiable.22

Several authors have written papers within this framework. Hamilton (1983a) used it to partially explain the flypaper effect. Craig (1987) to analyse the impact of "congestion" on local public goods production. However, they do not address the issue of economies of scale in production.

The approach of a two-stage production function might be appealing, but has its problems as pointed out by Schwab and Zampelli (1987), since it mixes consumption with production of local public services.

To clarify this, consider the case of education. It is an empirical result that educational services are a function of teachers, schools' quality and the socio-economic background of the families the students belong to. The "production" of educational services is entangled with the consumption of the same services and it is difficult to separate the two concepts.

Some authors argue that the concept of production functions is not appropriate for the reality of several local services.23

Due to these problems we prefer to limit the concept of economies of scale to the output of public sector production. This will be the concept applied in chapter 6.

Although the analysis of economies of scale is essential for the understanding of economies of community size, "publicness" of local public goods and fiscal equalization at a local level, there has been a lack of empirical research on this topic.

DeBoer (1992) analyses economies of scale in public libraries and concludes that there are economies of scale for small libraries (with circula-

---

22 They do not use these calculus' expressions but it is this idea.
23 For the case of education see the papers from Hanushek (1979, 1986).
tion below 55000 books per year) but they are exhausted after this level is reached. It is important to note that DeBoer considers only current expenditures.

Gyimah-Brempong (1989) analysed economies of scale in municipal police departments and concludes that there exists diseconomies of scale which he attributes to larger police departments.

Hirsch (1984) reviews prior research on economies of scale and finds that the main conclusions that can be drawn relate to severe methodological problems and to the existence of economies of scale in capital intensive services (e.g. infrastructures and public utilities) whereas this is not the case in labour intensive services (e.g. police, education).

We will adopt in chapter 4 the conclusions taken from Hirsch, and we will also introduce critical reasons to support them.

Now we will turn to developments in the congestion function.

### 3.1.4 Extensions on congestion

The congestion function $g$ has at least two arguments, population size and capacity size

$$x = g(X, N)$$

The argument that will be put forward in this section and developed in chapter 4 is that there are three distinct congestion related concepts that most of the literature has failed to acknowledge: congestion itself (as defined in club goods theory), crowding ("congestion" in communities) and "economies of scale in consumption".

The approach of congestion within the demand framework \(^{24}\) is that the individual demands services, not facilities, and that to obtain the service level from the facility level it is necessary to introduce, in some way, the number of persons sharing the consumption of the given facility.

Borcherdin and Deacon, Bergstrom and Goodman specified this relationship as $x = X/N^\gamma$, and that the elasticity of $x$ (services) regarding $N$

\(^{24}\)Which is different from the club goods approach as clarified in chapter 4.
(population) is constant and equals $-\gamma$.

Let us define with Brueckner (1981) returns to scale in consumption the relationship between services $x$ and the scale of provision (capacity), keeping per capita levels of provision constant \(^{28}\). It is important to note that assuming the congestion function is homogeneous, constant returns in consumption implies the degree of homogeneity is zero (and not one as in constant returns in production) i.e.

$$g(\lambda X, \lambda N) = \lambda^\gamma g(X, N) = g(X, N)$$

where $\lambda > 1$. Increasing returns means that the degree of homogeneity is positive.

This concept clarifies another feature of the BG and BD formulation. Since:

$$\frac{\lambda X}{(\lambda N)^\gamma} = \lambda^{(1-\gamma)} \frac{X}{N^\gamma} = \lambda^{(1-\gamma)} x$$

services are homogeneous of degree $1-\gamma$ regarding population and capacity sizes. If $\gamma = 1$ there are constant returns to scale in consumption, if $0 < \gamma < 1$ there are increasing returns to scale in consumption since $\lambda^{(1-\gamma)} > 1$ and if $\gamma > 1$ there are decreasing returns to scale.

BD and BGs approach to crowding or congestion has been criticized on the grounds that although $g_s$ and $g_n$ have the correct signs, $\frac{\partial^2 g}{\partial N^2}$ has the wrong sign. The second derivative in BD/BG is positive which means that marginal congestion becomes less negative with the level of population. This means that local public goods are asymptotically pure public goods when $N$ increases.

However, Craig (1987), Edwards (1990) and others have criticized this approach saying that

$$\frac{\partial^2 g}{\partial N^2}$$

"should" be negative because, according to club goods theory, marginal congestion increases with population.

\(^{28}\)Of course it could be seen as the relationship between services $x$ and the population sharing them, keeping per capita provision constant.
Their argument can be clarified with the help of Figure 3.1. The continuous lines represent the BD/BG *crowding* function for different values of the parameter $\gamma$. They show how the level of services change according to different values of the crowding parameter. The dashed line shows the *congestion* effect common to club goods theory when the capacity of the *local public good is fixed*.

For instance Craig (1987) made the above criticism of BD and BG and argues that cities should be of sufficient size to exploit regions of low congestion and not too big so that congestion is very high and consumption almost nil. This would be true if the capacity of LPG was *fixed*. However, allowing changes in the level of provision (which is the case in cross-section studies) there is no case to sustain that point of view. Club goods theory has already demonstrated that there is *no* optimal community size under the (not unreasonable) assumptions of homogeneous cost function (degree one) and utility functions (degree zero).

Similarly, Edwards (1990) discusses several models of congestion within the club goods framework. His increasing marginal congestion specification

\[ \text{Figure 3.1: Congestion and Crowding Functions} \]

---

“is designed to satisfy the more usual assumption that congestion accelerates with the intensity of use”\textsuperscript{37}.

He considers several congestion function specifications and analyses the sensitivity of the non-publicness results obtained by BD/BG to these different specifications. Apart from the BD/BG specification he introduces a “more realistic” decreasing marginal congestion specification:

\[ z = X[2 - e^{1(N-1)}] \]

a camaraderie function

\[ z = X N^\gamma \sigma^{1-\gamma} \]

where it is possible that \( \frac{\partial z}{\partial N} > 0 \) depending on \( N < \frac{\gamma}{\sigma} \) (where \( a \) is a constant). Finally, he also considers a generalized congestion function taken from Inman.

It is intuitive and not really surprising that Edwards' concludes that the non-publicness result is sensitive to the specification of the congestion function.

His functions are as \textit{ad hoc} as the ones used by BD/BG. They have the advantage over the latter, of allowing increasing marginal congestion. However, they have the strong disadvantage of ruling out the possibility of “non-publicness” as a particular case. This is a very important shortcoming as will be discussed later on chapter 4.

Both criticisms are misleading since they fail to understand the difference between crowding (“congestion” in communities) congestion and economies of scale in consumption.

They also seem to forget the point that clubs might be replicated at any size with any possible level of individual congestion. Therefore, the general statement that the “crowding” function should be concave and not convex to the origin is correct, but the point that this limits optimal community size is incorrect. This aspects will be fully discussed in section 4.2.

\textsuperscript{37} Edwards (1990) page 84.
A different and interesting approach to congestion was developed by Brueckner (1982) who estimated directly the congestion function and therefore the crowding elasticity and the degree of returns to scale in consumption.

Applying to the fire protection case he assumes that the congestion function is given by:

$$z = g(X, N, s)$$  \hspace{1cm} (3.13)

where $s$ is a vector of environmental variables which are supposed to affect the consumption level of fire services. Since he does not have direct measures of each community's fire suppression system capacity ($X$) he assumes it is a function of department pumping capacity ($F$) and the water supply system capacity ($W$):

$$X = f(F, W)$$  \hspace{1cm} (3.14)

Assuming that expenditures on pumping and water systems are functions of the respective wages ($y$) and capacity then we have respectively:

$$E_F = \pi(F, y)$$  \hspace{1cm} (3.15)

and also

$$E_W = \mu(W, y)$$  \hspace{1cm} (3.16)

These functions can be inverted and hence, capacity becomes a function of expenditures and wages. Department pumping capacity is therefore given by:

$$F = \hat{\pi}(E_F, y)$$  \hspace{1cm} (3.17)

and water supply system capacity by

$$W = \hat{\mu}(E_W, y)$$  \hspace{1cm} (3.18)

Introducing these two equations in the capacity function (3.14) yields:

$$X = h[\hat{\pi}(\cdot), \hat{\mu}(\cdot)]$$  \hspace{1cm} (3.19)
Substituting equation 3.19 into the congestion equation (3.13) we get:

\[ x = g[\bar{\theta}(\cdot), \bar{\mu}(\cdot), N, s] \]  

(3.20)

It is interesting to note here that congestion will depend on the production functions of department pumping capacity and water supply system capacity. Under the crucial assumptions of constant returns to scale in production in both services it is possible to write:

\[ x = g[E_f, E_w, N, y, s] \]  

(3.21)

Therefore he estimates the following function:

\[ z = aE_f^{\alpha_1}E_w^{\alpha_2}N^{\alpha_3}y^{\alpha_4}s^{\alpha_5} \]

with \( E_f \) and \( E_w \) indicating expenditures on pumping capacity and water supply system respectively.

That equation can be interpreted as:

\[ z = g(X, N, s) = AX^{\beta_1}N^{\beta_2}s^{\beta_3} \]

Brueckner concludes that \( \alpha_1 + \alpha_2 + \alpha_3 \) is significantly greater than zero so that there are increasing returns to scale in consumption. He also finds that \( \alpha_3 \) (elasticity of services regarding population) is negative and significantly different from zero and minus one. This suggests the "publicness" of fire services.

Two remarks should be made concerning Brueckner's approach. Firstly, his conclusions are dependent on some critical assumptions, namely the existence of constant returns to scale in production. Secondly, fire services seem a very peculiar local service, since capacity is idle most of the time in small and medium communities and demand depends heavily on hazards. This is not the case in the majority of local services where there is a continuous flux of services and capacity is always being used.\(^{28}\)

\(^{28}\)Brueckner's paper could constitute empirical evidence supporting the rejection of the approach developed below. However, for the reasons stated above we do not read Brueckner's paper in this way.
Oates (1988b), followed by MacMillan (1989), also questions the "privateness" result obtained by BD/BG arguing that there might be measurement problems in the analysis. They argue that bigger jurisdictions supply either a more diversified output than smaller ones do, or the same type of output with improved quality. Failure to acknowledge that output is heterogeneous and treating expenditures in big and small communities with the same label, leads to an upward bias in the crowding parameter.

3.1.5 Extensions on the tax-price

Let us recall here that the "tax-price" variable is a construction of economic theory to put the provision of local public goods in the demand framework common to private goods. It was always a problematic variable and it has become still more controversial with the development of two different approaches.

Firstly, because the median voter demand framework is essentially an intrajurisdictional approach assuming that the population of the community is given and therefore that citizens can only reveal their preferences by voting and not by leaving the community. However, as clarified in chapter 2 citizens migrate between jurisdictions, at least partially, as a response to the differential in the fiscal "packages" offered by different communities. They exercise their "voice" within the jurisdiction through the vote, but they can also use their "exit" option. As a consequence several authors have pointed out that migration biases the tax-price estimator. This is known as Tiebout bias and was initially considered by Goldstein and Pauly (1981). Wildasin (1989) and Crane (1989) also identify different sources of bias in the "tax-price" variable.

The second problem associated with the "tax-price" variable is the fact that citizens do not have the proper information on costs and benefits of local services. This is the problem known as fiscal illusion.

The issue of the "privateness" of local public goods will also be considered within the bureaucratic approaches (section 3.2)
Fiscal illusion means that citizens do not have perfect information about the true nature and incidence of either taxes or expenditures and that this misinformation leads to systematic bias in preferences. This means that to be consistent with the proposition that fiscal illusion enables a permanent and stable underprovision (or overprovision) of local public goods, it is necessary to have a theory that explains why information is systematically biased either towards underprovision or overprovision, since fiscal illusion is compatible with both.

As we will discuss in section 3.3.2 it is not easy to test the fiscal illusion hypothesis.

3.1.6 Median Voter model: summary

Empirical analyses using the median voter model to study demand for local public goods were clearly a step forward in the understanding of the nature of these goods and the importance of crowding in consumption.

In the introduction we made the following definition of economies of community size:

There are economies (diseconomies) of community size in the provision of a standard bundle of local public goods whenever, ceteris paribus the combined effect of economies of scale in production and "congestion" in consumption leads to a decrease (increase) in the per capita cost of provision as the population of communities increase.

The empirical literature considered in the first part of the chapter only partially addressed this issue. In fact all empirical studies have “neutralized” the problem of economies of scale through the assumption of constant returns to scale. This assumption seems valid for the limited scope of goods which are provided through current expenditures, but not realistic for

---

20 On this see Oates (1988a).
21 Most of the empirical analysis in the median voter framework has considered only current expenditures.
capital expenditures.

As a consequence of the constant returns (in production) assumption, the issues of economies of agglomeration can be reduced to the issue of the "privateness" or "publicness" of local public goods. In other words, whether or not local public goods have congestion characteristics so that provision has to increase proportionally to population to keep levels of service constant.

The literature reviewed is divided on this issue. The classical median voter approach used a decreasing marginal congestion specification and reached the "privateness" result. More recent approaches have considered different specifications of the congestion function and showed that the "privateness" result is dependent on that specification.

Apart from the new insights using the median voter model it seems clear that it is not a good model of local decision-making and does not enable an understanding of citizens' preferences regarding local public goods. The main reason is because the institutional framework within which most local governments behave is far from the assumptions of the model. This, at first sight, would not seem a problem if we adopted Friedman's position on methodology. In short, the unreality of the assumptions is not a problem if the predictions of the model are accurate. However, some major predictions are not supported by the empirical evidence as will be clarified below when introducing the flypaper effect.

Moreover, there are several methodological problems in using the model with the objective of estimating "tax-price" and income elasticities of demand for local public goods. The main problems are the identification of the median voter, the use of expenditures as a proxy for demand and the specification of the "tax-price" variable.

All these methodological problems led to a decreasing use of the "full information" median voter model in empirical analysis during the eight-
ies, after being intensively used in the seventies following the papers of Bergstrom and Goodman (1973) and Borcherding and Deacon (1972).

Alternative approaches to preference revelation for local public goods were developed either using qualitative survey data\(^4\) or direct market data on private goods the demand for which is supposedly a function of a public good characteristic, among other variables.\(^5\)

More interestingly from the point of view of the present research is another strand of literature that emphasizes a supply side approach to local government behaviour.

### 3.2 Bureaucratic Approaches

There has been an extensive literature on bureaucratic, or more generally, government supply-side approaches which will not be reviewed here.\(^6\) We are particularly interested here in analysing how the bureaucratic approaches criticize and reappraise the median voter model conclusions and, on the other hand, how they introduce further insights to understand local governments' behaviour.

#### 3.2.1 Theoretical approaches

A landmark on theoretical analysis of bureaucracy was made by Niskanen (1971). His analysis, using an economic methodology, is based on the similarities and dissimilarities between bureaucrats in public agencies and managers in private corporations.

The basic asymmetric information that exists between shareholders and


\(^5\)This is the case of the hedonic approaches which usually try to infer from property values an implicit price for one public characteristic (e.g., pollution). See for instance Brookshire et al. (1982).

\(^6\)For a very good review see Jackson (1982), Mueller (1989) chap. 14 and also Blais and Dion (eds.) (1991) for developments of the Niskanen approach.
managers exists between citizens and politicians and also between politicians and bureaucrats (or the funding agency and the bureau).

However, there are some peculiarities of public agencies that distinguish them from most private organizations. They supply (in general) levels of activities, rather than measurable output, so that monitoring their activities becomes very difficult. On the other hand, salaries of the bureaucrats are not performance related, so that incentives to increase efficiency are very weak or non-existent.

Niskanen assumes that the objective of the non-elected bureaucrat is to increase the size of the budget since this is related to salary, public reputation and personal power. His model considers that a bureau receives a budget from an agency, which is a function of the perceived output of the agency. He assumes also the existence of a cost function known by the bureau but not by the funding agency.\(^\text{35}\) This essential asymmetry in information leads the bureau to receive funds in excess of those needed to reach a Pareto-efficient provision of the good, i.e. that the marginal benefit of an extra unit of output equates the marginal cost to the bureau. According to Niskanen the bureau maximizes the size of the budget under the only constraint that the budget must not be inferior to costs. This leads the budget to be expanded so that marginal benefits are less than marginal costs. As a consequence the general effect of bureaucrats' behaviour would be an increase in the overall size of the public budget.

Niskanen's approach to bureaucracy led several authors to approach local governments as monopolists and to study the impact that competition has on the provision of local services. Emphasis shifted from allocative efficiency to X-efficiency, i.e. to analyse whether bureaus provide a given level of service at a minimum cost.

During the sixties it became acknowledged (Williamson (1964), Lieben-\(^\text{35}\)This is the general problem that is considered in the principal/agent literature, which had a great development since Niskanen's work. The seminal paper was Ross (1973) and an application to intergovernmental relations can be found in Levaggi (1993).
stein (1966), Commanor and Liebenstein (1969)) that managers of the modern corporations have a critical role in acquiring and processing information in an uncertain environment, and therefore pursue objectives alien to the basic interest of their stockholders (profit maximization). Apart from a sufficient level of profits (enough to keep their jobs) a set of non pecuniary objectives were pursued that fall under two basic categories: in-kind “payments” and excessive growth in size of the immediate staff and corporation as a whole.

Liebenstein considers that if a firm has a permissive budgetary environment, the looser is the effort responsibility consequence for each individual (i.e. the weaker is the relation between individual effort and the consequences of his behaviour) the greater the degree of X-inefficiency, i.e. the excess of present over minimum cost to produce a given output. This approach led to several empirical approaches as we will see in the next subsection.

Other bureaucratic approaches were developed aiming at explaining why local decision-making is not driven by the median voter’s preferences.

The essential criticism of the “fully informed” median voter model is that the model’s assumptions are far from reality. In fact, it assumes unidimensional issues, i.e. that local agencies deal and submit to the electorate single issues (only one service: e.g. expenditures on education or tax-expenditures on the sewage system, etc.) and also that this good enters only in the utility function of the citizens and does not affect their wealth. However, neither of these conditions hold. Firstly, in general, local governments decide over a wide range of issues simultaneously and secondly, several services have an impact both on the utility function and on wealth, so that preferences are no longer single-peaked.

Even in the few cases where the institutional context fits the assump-

---

38 For a full criticism see Romer and Rosenthal (1979).
39 This is the case of all expenditures that might have an impact on house values as the literature on capitalisation emphasises.
tions of the model it has been argued that a manipulation of the referendum by the agenda setter might originate in an over supply of expenditures. Provision of local education in many States of the U.S.A is done by educational authorities which face referendum processes. This sort of "direct democracy" seems to be the institutional context closest to the assumptions of the median voter model.

Romer and Rosenthal developed a model where even in this context the preferences of the median voter are not satisfied since the agenda setter gives an option between a higher level of expenditures and a too low reversion level that would be adopted in case the expanded budget is not approved. By fear of the reversion level being approved (which could have as a consequence the closure of some schools) voters "choose" the higher level of expenditure.

The bureaucratic and monopolistic approaches to government depart from the traditional analysis in that the behaviour postulate regarding public agencies or individuals belonging to these agencies is not driven either by allocative efficiency and X-efficiency or some equity criterion.

However, the budget maximizing and agenda setter models, are still within the median voter framework albeit departing from the classical approach where the "fully informed" median voter drives the local political process. They introduce additional agents on the process of decision-making, but local government is still constrained by the voting behaviour of the median voter.

Other approaches in the literature do not give any role to the "voice" (vote) of the median voter. They consider that local governments have a monopoly power which is constrained only by the exit option of the citizens, who can move to another jurisdiction. In some models there is a maximising behaviour and in others there is not.
A theoretical approach that emphasizes supply side factors and is rather sceptical regarding political competition as a way to reveal voters’ preferences is the Leviathan approach to government, developed by Brennan and Buchanan (1977), (1980). They consider that governments behave like monopolists wanting to maximize revenues and political competition being insufficient to constraint their proclivities, constitutional limitations on the power to tax should be introduced.

Their analysis focuses essentially on central government, yet they extend it to local governments. The issue is whether federalization of the political structure can be a substitute for direct fiscal constraints on the fiscal exploitation from governments.

Brennan and Buchanan argue that in an idealized Tiebout world with costless migration between jurisdictions, no locational rents and local public goods, local governments are unable to exploit the fiscal surplus of their residents. On the other hand, if there are locational rents and (or) costs of migration are positive the prospects for fiscal exploitation are higher. In any case, they conclude that competition between local governments always exert a positive effect on the subnational units of government which does not exist when considering central government. Thus, they derive a positive conclusion that local governments will try to secure arrangements to moderate competition, which might assume the form of a uniform tax system across all jurisdictions and also a greater share of revenue raising powers to central government.

Epple and Zellenitz (1981) also consider governments as maximizing expenditures and they conclude that if increases in the number of jurisdictions limit the ability of discretionary government power, it does not eliminate it completely since fixed jurisdictional boundaries enable local governments to exercise monopoly power as a result of land being immobile.

Competition between local governments has a contradictory effect on

\footnote{See Brennan and Buchanan (1980) chapter 9.}
allocative efficiency. Improvements are expected in local expenditures to enhance local services and thereafter attract new residents and firms. Moreover, the ability of "fiscal exploitation" by local governments is constrained. On the other hand, negative effects can be associated precisely with excessive tax competition that leads to an under provision of local public goods.

As a whole the approaches considered in this section are much more sceptical about the ability of the median voter revealing his preferences through the political process. This would have necessary implications for the impact of intergovernmental grants, as will be discussed below.

3.2.2 Empirical results

Empirical analysis based on bureaucratic approaches to local governments' decision-making has not been very frequent due to several methodological problems, some of them highlighted in this section.

The first approaches merely reinterpret the median voter results trying to give an answer to the following question posed by Bergstrom and Goodman (1973): "why, if there are not increasing returns to scale in the municipal provision of the goods and services, is their provision in the public domain?". This question, somehow introduced the debate regarding the "publicness" or "privateness" of local public goods and the public versus private provision of LPG.\(^{41}\)

Borcherding, Bush and Spann (1977) try to give answers, suggesting two possible explanations for the "privateness" result. Firstly, assuming local governments have a considerable degree of monopoly power, they will prefer to provide goods with a high degree of congestion because this will have a much higher impact on increased expenditures (their maximand) than would be the case if they provide goods with a greater degree of publicness.

\(^{41}\)These two questions have been frequently treated as only one, as if "publicness" would imply public provision and "privateness" private provision. Of course this is not so, because there are publicly provided private goods.
Secondly, assuming also that bureaucratic power is positively correlated with population size, economies of collective supply are appropriated by budget maximising local governments.

This idea was previously introduced by Niskanen (1975) who summarizes the problem: either the goods are private and "divisible" at the margin in which case there is no argument for government supply, or they are public but the existing economies are appropriated by monopolistic governments.

More recently Gonzalez and Mehay (1985) using a (non median voter) monopolistic model of local government, reach the "privateness" result and blame bureaucracy as the explanatory cause.

Wyckoff (1988a) develops a slack-maximizing model of bureaucracy. Bureaus want to maximize "slack" i.e. X-inefficiency (production above minimum costs) instead of maximising output (allocative inefficiency) the typical outcome within Niskanen's model. Wyckoff's approach is within the median voter framework and shows that bureaucrats, if they can choose, will prefer goods with a higher degree of "publicness" so that bureaucracy can not be responsible for the privateness result. Thus, Wyckoff contradicts both Gonzalez and Mehay (1985) and also Borcherding, Bush and Spann (1977).

Curiously, Wyckoff (1984) tries to discriminate, for purposes of empirical analysis, the slack-maximizing model from the budget maximizing model and the median voter model. However, in a later paper (Wyckoff 1988b) he only confronts the Niskanen model with the median voter model, since the predictions\textsuperscript{43} of the slack and budget maximizing models are too similar.

In relation to the Leviathan approach to government, there have been several empirical papers\textsuperscript{43} but they were directed to overall government (public sector size) rather than local government. Moreover, the empirical tests done were of the "fiscal decentralization" hypothesis which is only a

\textsuperscript{42}In terms of acceptable ranges for the parameters in a median voter demand function.

corollary of the Leviathan hypothesis.

Brennan and Buchanan (1980, pg.185) stated the decentralization hypothesis as: “Total government intrusion in the economy should be smaller, ceteris paribus, the greater the extent to which taxes and expenditures are decentralized”.

As Musgrave pointed out to Oates there are approaches to government consistent with the fiscal decentralization hypothesis other than the Leviathan. He gives the example of the redistribution function (e.g. social welfare programs) which, if done at a local level, will be much smaller than if centralized. This is because of the Tiebout effect of individuals being sorted out in homogeneous communities therefore making intrajurisdictional redistribution difficult. Moreover, local governments would not be active in such programs with the fear of attracting low income people and therefore lowering the per capita tax base of the jurisdiction.

Finally, regarding X-inefficiency, some empirical studies analysed the impact of competition within the public sector, when services provided by local governments are contracted out from neighbour cities. R.Deacon (1979), Mehay (1984) and Mehay and Gonzalez (1985) all present evidence suggesting that when having to compete for contracts local governments provide their services at a lower cost. Similar evidence was suggested from many studies that aim at comparing the costs of supplying local services between public and private firms, and that in general conclude that private provision is done at a lower cost. Nonetheless, some authors pointed out that it is competition rather than ownership that is relevant, since a private monopoly can be as bad as a public monopoly.46

---

44 See footnote 2 in Oates (1985).
45 For a survey of these studies see Borchering, Pommerehne and Schneider (1992) and also Parker (1990). This issue is particularly relevant in the U.K. where the 1988 Local Government Act extended compulsory competitive tendering to the provision of a wide range of local services.

46 See Jackson and Price (forthcoming) for an excellent review.
3.2.3 Summary

The short review made here of the bureaucratic approaches was intended to clarify how the issue of economies of agglomeration could be understood within this framework and on the other hand to gain further insights into local governments’ fiscal behaviour.

Curiously, the bureaucratic approaches introduced an argument that can be used in favour of the “publicness” of local public goods and (assuming constant returns to scale) supporting the existence of economies of agglomeration. In fact, the “privateness” result found by Bergstrom and Goodman, Borcherding and Deacon can be only a veil to cover a different reality. Local public goods might have “publicness” characteristics, but the economies of agglomeration might be appropriated by monopolistic governments.

On the other hand, authors developing the classical median voter framework argue that the “privateness” result might not be accurate due to an incorrect specification of the congestion function.

Therefore, what can we conclude?

Our position is that this controversy, which is likely to remain in the forthcoming years, can not be settled mainly on empirical grounds. We would need to have a clear understanding of local public goods’ production functions, which we do not have. We would also need to have precise measures of output to be able to distinguish X-inefficiency from allocative inefficiency.

However, following Hayek’s methodological stance, we can devise simple models that show us the crucial elements of the problems we are dealing with. This will be the approach developed in chapter 4.

Beforehand, we will briefly review the literature on positive approaches to intergovernmental grants which is a necessary preliminary for the development of a positive approach to grants which will also be considered in chapter 4.

See section 1.1.
3.3 Positive Approaches to Intergovernmental Grants

There are mainly two issues that have been addressed in the context of positive approaches to intergovernmental grants. One, is to study and predict the impact of different sorts of grants design on the behaviour of local governments. This is the issue that has received more attention in the literature. The impact of central governments' grants on local governments' fiscal behaviour depends mainly on three distinct sort of factors: (i) the specific type of grant, (ii) the assumptions regarding local government behaviour and the (iii) institutional context where that behaviour takes place.

A different problem is to search for a rationale (and therefore predict) central governments' behaviour regarding intergovernmental transfers. A modified version of this issue is to consider whether intergovernmental grants are actually designed in accordance with the normative criteria which are supposed to give a rationale to grants' design.

Therefore, section 3.3.1 clarifies the different types of grants, section 3.3.2 considers approaches within the demand framework and in particular the flypaper effect, section 3.3.3 comment on the role of grants in a broader institutional context, section 3.3.4 surveys papers on positive approaches to grants' design and finally section 3.4 summarizes the conclusions of the chapter.

3.3.1 A Typology of Grants

Early in the development of the literature (Bradford and Oates (1972), Gramlich and Galper (1973)) it was acknowledged that the specific legal requirements attached to the grants were important to understand their differential impact on local governments' behaviour. The basic distinction was between lump-sum grants and matching grants. The former are fixed amounts of money to be spent by lower level jurisdictions at their own dis-
cretion. The latter, are variable grants where central government matches local governments' expenditure in a fixed proportion. If there is no limit for the amount of these grants they are usually known as open matching grants, if this upper limit exists they are closed matching grants.

Another type of grant considered is the categorical or special grant. These are grants provided to fund special local governments' services. They can either be of a matching or lump-sum form.

In those initial papers, as well as in some empirical median voter approaches, it was assumed that lump-sum grants have only income effects (on local government or the median voter) and matching grants, an income and price effects.

These initial approaches were soon developed in quite different perspectives. Some authors argued that lump-sum grants have also price effects due to fiscal illusion (Oates (1979), Courant, Gramlich and Rubinfeld (1979) )

Others have argued that the resources that jurisdictions receive through matching and/or special grants are, at least in some considerable part, fungible and used for general purposes. This has been labelled the "fungibility hypothesis". This approach has been developed by McGuire (1975), (1978),(1979) and extended by Zampelli (1986). McGuire argues that there is a distinction between the nominal conditions imposed by central governments and the effective conditions perceived by local governments. Therefore, the legal provisions of the grant programs are sometimes misleading regarding their effects on the resource constraints of the receivers.

From the point of view of modeling local governments' fiscal behaviour the primary criterion to distinguish grants is whether the distribution is dependent or not on their fiscal behaviour. Therefore we distinguish primarily

---

48 This will be developed in a subsection below.
49 As Zampelli (1986 p.31) puts it: "when conditional aid is intended to apply to increases in local public output, beyond current levels, the local government can reduce its own normal funding of the subsidized output, use a project which was going to be undertaken anyway as the means for securing the matching grant, re-define budget categories, or re-allocate overhead costs to transform some or all of the aid into fungible resources."
lump-sum grants (not effort related) from behaviour induced grants.

Lump-sum grants can be either general or specific in scope. Given the fungibility hypothesis, this distinction is not so dramatic as it may seem. The concept of lump-sum grants used here is somewhat more restrictive than the one generally used in the literature. Namely, revenue-sharing grants will not be considered lump-sum grants whenever there is a significant tax effort element in the allocation formula. This is the case of the USA where the general revenue sharing formula includes a tax effort index, together with other indicators.

Behaviour induced grants are mainly of two types: effort-related grants and matching grants. The former ones are related to some indicator of fiscal effort applied to the grantee and all other jurisdictions, since fiscal effort is more a relative than an absolute concept. The latter were described earlier. In both cases the amounts of the grants depend on the behaviour of the grantees.\footnote{There is a taxonomy of grants in King (1984) which differs from ours in one essential aspect. King considers that the primary difference is between specific and general grants. Specific grants are divided into lump-sum specific grants and matching specific grants (open and closed) while general are divided in lump-sum general grants and effort-related grants. The fungible character of specific grants, even partial, suggests that this distinction might be elusive.

However, there might be a case to group together lump-sum general grants and effort-related grants. Insofar as the tax effort element in the allocation formula has a very low weight, the two types of grants became roughly identical. In this case there is no incentive for local governments increasing their tax effort to obtain a greater share of grants.}

3.3.2 Intergovernmental grants, the flypaper effect and the demand framework

It is not our aim to discuss the rationale for grants nor the optimal design or the optimal properties that a system of grants should have. These are questions addressed by normative approaches.

Grants will be treated as exogenous to local governments. However,
in general, grants are informed by normative criteria set by the central government and the type of criteria used is relevant for the way grants are introduced in the models.

Early empirical literature on intergovernmental grants developed in parallel with the empirical median voter demand literature. At the beginning of the seventies, the papers of Bergstrom and Goodman, Borcherding and Deacon analysed demand in a median voter framework without considering intergovernmental grants. On the other hand Gramlich and Galper developed an empirical approach of grants (discriminating matching and lump-sum grants) in a non-median voter framework.

At the same time the link between the approaches was done (Oates (1972)) and soon the intergovernmental grants literature was "plunged" into the median voter framework.

In short, lump-sum grants were considered to increase the "full" fiscal income of the median voter by the share he captures from these grants, while matching grants were assumed to diminish the "tax-price" he pays.

Authors sympathetic with government provision and growth generally argue that citizens underestimate the benefits of public expenditure programs, and as a consequence, there is an undersupply of public goods. On the other hand, critics of increasingly expanded budgets, consider that citizens fail to recognize the full extent of the fiscal burden associated with increased expenditures.

Fiscal illusion is usually explained by one of the following factors: complexity of the revenue system, renter illusion, income elasticity of the revenue system and the flypaper effect. The more complex the revenue structure (the greater the number of sources of revenue and the smaller their visibility) the less voters are aware of the fiscal burden. Renters might not be aware that in spite of the property taxes being the liability of the landlords, there is a partial shift of the property tax into higher rents. The elasticity of the tax revenue is related to the fundamental asymmetry between discretionary and automatic increases in taxes. The more income
elastic is the tax revenue structure the less voters will be aware that their
tax burden is increasing with time. Finally, there is a misperception of
either the tax-price or the "full" fiscal income due to intergovernmental
grants, which became known as the flypaper effect.

Particular theoretical attention has been given to the flypaper effect,
both theoretically and empirically.\textsuperscript{51}

The flypaper effect is essentially a falsifiant\textsuperscript{52} to one major prediction
of the full-information median voter model. This prediction was stated
in Bradford and Oates (1971), who argued that under certain conditions
of local decision-making, a lump-sum intergovernmental grant to a local
government should have the same effect on local public goods' expenditure
as a set of lump-sum intergovernmental grants given to the individuals
of the jurisdiction. Each individual would capture his share of the grant
according to his tax-share and he would spend it in private and local public
goods according to the respective income elasticities of demand.

Contrary to this prediction, empirical evidence has shown that intergov­
ernmental grants do increase local governments' expenditure much more
than would be the case if they were directly given to the individuals taking
into account their tax-share and their income elasticity of demand.

As referred to above a lump-sum grant to a local government can be
modelled as a transfer of income to the median voter of the jurisdiction.

\textsuperscript{51}Some theoretical papers are Bradford and Oates (1971) Courant Gramlich and Ru­
Hamilton (1983b), some discussion in Inman (1979), and Turnbull (1992). Empirical ap­
proaches were usually made within the median voter framework even when different mod­
els (including bureaucratic approaches) were considered. Gramlich (1977) makes a useful
survey of the empirical literature. There are also papers with an empirical analysis from
Gramlich and Galper (1973) (in a non-median voter framework), Wychoff (1988b) (who
confronts bureaucratic and median voter models (within the median voter framework)),
Barnett et. al. (1991a) (who compare two models where local governments perceive ( or
not ) correctly their budget constraint) and Heyndels and Smolders (1993) (who analyse
different sorts of fiscal illusion including the flypaper effect).

\textsuperscript{52}See the discussion on Popper in chapter 1.
More precisely, let us assume that a local public good $X$ is produced under constant returns to scale, and financed through taxes $T$ obtained through an effective tax rate $t$ which applies uniformly to a diversified tax base $B$ including residential, commercial and industrial property. Local government receives a lump-sum grant $G$ from central government and as a balanced budget. Therefore we have the following relations:

$$C = cX$$  \hspace{1cm} (3.22)

$$T + G = cX$$  \hspace{1cm} (3.23)

$$t = T/B$$  \hspace{1cm} (3.24)

Individuals have pre-tax incomes $I$, net income $y$, and individual tax base $b$. Individual tax-shares can be given by:

$$
\tau = \frac{b}{\pi} \hspace{1cm} (3.25)
$$

Where $\pi$ allows for the fact that there might be tax deductions and tax credits. The individual resource constraint is:

$$y = I - tb\pi$$  \hspace{1cm} (3.26)

Assume also that there is "private" technology of consumption so that:

$$x = \frac{X}{\pi}$$  \hspace{1cm} (3.27)

The taxes paid by the individual ($tb\pi$) are by definition identical to his share of the costs (net of intergovernmental grants) of provision of local public goods. Therefore:

$$tb\pi = \frac{\tau}{\pi}(C - G)$$  \hspace{1cm} (3.28)

---

23 This exposition follows Inman (1979) corrected by Fischer (1982). Note that the tax-share is not the individual tax-share as it "should" be considered (see discussion in the beginning of the chapter). In fact if $\pi = 1$ and $b = B/\pi$ then $\tau = 1$. Therefore, Inman's tax share is $N$ times the usual individual tax share. If we label the usual tax-share concept as $\tau'$ then $\tau' = \tau/N$. However, this is not a problem in the following developments since it is correctly introduced in the tax-price variable.
Introducing equations 3.28, 3.22, and 3.27 in 3.26 we can obtain:

$$y = I + r \frac{G}{N} - rex$$  \hspace{1cm} (3.29)

In general we will have net income as a function of "full" fiscal income and "tax-price" :

$$y = y(I + r \frac{G}{N}, rc)$$  \hspace{1cm} (3.30)

Expenditures $E$, considered a proxy of demand, will be a function of the same variables so that:

$$E = E(rc, rg + I)$$  \hspace{1cm} (3.31)

where $g$ are per capita grants.

It can be demonstrated that the two following relations should hold:\textsuperscript{54}

$$\varepsilon_{E,q} = \frac{r G}{I N} \varepsilon_{E,I}$$  \hspace{1cm} (3.32)

which can also be expressed as:

$$\frac{\partial E}{\partial q} = r \frac{\partial E}{\partial I}$$  \hspace{1cm} (3.33)

Therefore the absence of the flypaper effect can either be treated in marginal or elasticity terms.

A marginal increase in per capita grants should have the same effect on expenditures as a marginal increase on private income weighted by the median voter tax share $r$. The tax-share is precisely the marginal amount of grants that goes (as extra income) to the median voter.

In elasticities terms, the elasticity of expenditures regarding per capita grants should be identical to the elasticity of expenditures regarding income weighted by the factor $\frac{r}{I}$.

Empirical research has found (see Gramlich (1977), Inman (1979) and Fischer (1982)) that the left hand side of those equations is consistently higher than the right hand side, i.e. that the impact of intergovernmental grants is much higher than an equivalent increase on incomes weighted by the individual tax-share.

\textsuperscript{54}See Fischer (1982).
This result leads to an explanation of the flypaper effect through the fiscal illusion approach. The difficulties in approaching the flypaper effect through the "misinformed" median voter can be illustrated by the papers of Courant, Gramlich and Rubinfeld (CGR1979) and Oates (1979).

The argument developed by CGR is that voters misperceive their true resource constraints as well as the tax-price of local public goods. They illustrate the effect, assuming that there is a proportional income tax collected by central government and that the tax yields return to the local government on a lump-sum grants form. Although community resources remain constant, they assume that there are two sorts of misperception involved when the tax rate rises.

Firstly, voters will consider that their income was reduced by the amount of taxes paid, since they do not realize that their full fiscal income remains constant. Secondly, they will consider that their tax-price diminishes as well. The average tax-price is tax-share times the unitary cost of the public good (assumed constant). The individual tax-share is the ratio of individual income over community income. Since the latter is constant while the former declines, tax-share (and therefore tax-price) decreases.

These two forms of illusion have contradictory effects on demand for public goods. The income effect (assuming public goods are normal) implies a reduction in demand, while the tax-price effect implies an increase. However, the authors argue that the income effect illusion is only a short-run misperception while the tax-price illusion is more persistent. As a consequence the long term effect of fiscal illusion is a level of expenditures and employment in the public sector above optimal levels obtained with a consideration of marginal tax-prices.

The argument in Oates is similar, but he only considers the price-effect since he ignores the problem of financing the grants which is responsible for the income effect in CGR. It is also the average tax-price that decreases but this time due to a decrease on the unitary costs to the community of
the provision of local public goods due to the lump-sum grants.\textsuperscript{56} The median voter tax-price is considered to be a fixed fraction of community's tax-price.\textsuperscript{56}

The problem with the fiscal illusion hypothesis is that modelling the nature of the asymmetry in information is, very often, \textit{ad hoc}.

Why is the income effect transitory and the price effect permanent in CGR? If it is not so, it is possible to devise a model where these effects offset each other so that even with "fiscal illusion" the outcome is the same as in the "fully-informed" median voter model.

Other authors (e.g. Hamilton (1983a)) argue that the flypaper effect is somewhat elusive. He considers that there is a misspecification of the income variable in the model. Income, Hamilton argues, is a proxy of the socio-economic characteristics of the median-voter, and enters the production function of local public goods. \textit{Ceteris paribus} higher incomes mean lower unitary costs on the provision of local public goods.

Therefore when the "full" fiscal income of the median voter is higher due to intergovernment grants, demand and expenditures will increase according to the income elasticity of demand. However, when it is due to increases on own income, demand will increase as well in line with the income elasticity but public expenditures will increase less than that\textsuperscript{57}. Therefore, intergovernmental grants have a greater impact on expenditures (but not on output) than equivalent increases on the median voter's own income (weighted by his tax share). This would partially explain the flypaper effect.

Rather than elusive, the empirical evidence supporting the flypaper ef-

\textsuperscript{56} Unitary costs to the community are simply taxes divided by public sector output.\textsuperscript{56} Gates considers that local officials are budget maximisers, which could lead us to believe his model was essentially a bureaucratic model. However, if they just satisfy median voter preferences according to the information they possess the same result applies.\textsuperscript{57} Public expenditures are no more a proxy of output. Higher expenditures are necessary to achieve the same level of output when the income of the median voter is lower.
fect is clear and one of the (few) consensual findings of empirical analysis in the local public goods literature.

The flypaper effect is one additional and strong argument against the "full information" median voter model.

### 3.3.3 Institutional framework and Intergovernmental Grants

The flypaper effect, or the empirical constatation that "the money sticks where it hits" is in fact a problem for the full-information median voter model but it is perfectly normal under alternative assumptions of local government fiscal behaviour.

If we consider a budget maximizer model of local government not constrained by the median voter, it is clear that the effect of intergovernmental grants will be an increase in local expenditure largely superior to what is expected taking into account the median voter tax-share and income elasticity.

Moreover it seems that it is very important to take into account the institutional structure which frames local governments' behaviour. For instance, under the **centralized** system of government considered in the following chapters, if grants increase, expenditure will increase as well in approximately the *same* amount, since local governments can not reduce their tax rates.

However, there is a different point we would like to stress here. The logic and rationale of intergovernmental grants is sometimes in conflict with the median voter demand framework. This is particularly clear with lump-sum general grants with equalization purposes.

Lump-sum intergovernmental grants have generally a revenue sharing and an equalization purposes. Equalizing grants have usually a redistributive purpose across jurisdictions.

---

*See for instance King (1984).*
In terms of their fiscal position jurisdictions differ regarding the structure of their tax bases, their tax rates, the "needs" of provision of local services and the unitary costs of provision of these services. The typical objective of equalization is that jurisdictions are able to provide a similar level of services with an identical "tax-effort" although having different tax-bases, "needs" or costs.

Revenue sharing grants with an equalization component tend to equalize the "full" fiscal income of the median voter across jurisdictions. Higher tax base jurisdictions are net payers and the median voters in these jurisdictions see their net incomes depressed. The opposite happens in low-income jurisdictions.

Therefore, the redistributive nature of revenue sharing grants operates from a quite distinct framework than the median voter demand framework. In the former, the aim is to neutralize differences in per capita incomes so that jurisdictions are able to provide similar levels of services although they have different fiscal capacities. In the latter (particularly when grants are not considered as a component of the "full" fiscal income of the median voter) it is considered that where the income of the median voter is higher (ceteris paribus) "demand" will also be higher due to the assumption that local services are normal goods.

Intergovernmental grants, when they have a strong fiscal equalization component, have their rationale in terms of supply according to centrally defined (minimum) standards of services. This is particularly important for meritorious "local public goods". This is at odds with the traditional median voter approach which considers expenditures as reflecting demand which is a function of income and the "tax-price" of the median voter.

\[^{10}\text{The typical case is basic educational services which, as widely acknowledged, should not depend on the households' incomes.}\]
3.3.4 Positive approaches to grants’ design

The majority of positive approaches to grants’ design come from political science scholars, as can be seen in Rich (1989) and references herein. Economists usually think in terms of the efficiency and equity rationales for intergovernmental grants and only seldom they test in which measure are these normative criteria operative in grants’ design. Exceptions to this “rule” can be found in Faith (1979), Alperovich (1984) and Inman (1988). All three papers argue that political factors are important in explaining grants’ design. We will look now in greater detail at these papers.

Faith (1979) approaches intergovernmental grants as a way of redistributing income between richer and poorer communities. He considers that we are living in a benevolent society where charity can be seen as a normal good. Benevolence is both interjurisdictional and intrajurisdictional from rich to poor.

Assuming that grants are a function of the perceived level of “needs” of each community, jurisdictions competing for fiscal aid have an incentive to reduce their levels of provision of local services in order to increase the frequency and or magnitude of actual or potential fiscal crises so that they might obtain further grants in the future.

Faith develops a simple model where communities are divided in two groups (rich and poor) and citizens in each community are further divided into a group of rich or poor citizens. Each rich community is endowed with a homogeneous good which is allocated to their residents (rich and poor) and also to other poor communities. Therefore, it is assumed that each benevolent rich community derive benefits from the consumption of services by their residents and also from the consumption levels of residents in poor communities.

60 The terms “rich” and “poor” need not be associated with wealth but “to those individuals whose levels of utility or those services the levels of output of which, when reduced, evoke charitable responses from others.” (Faith (1979) pg 321)
On the other hand, poor communities have their own endowments of the homogeneous good as well as incoming transfer from richer communities. Since richer communities cannot (e.g. asymmetric information) or do not want to make transfers contingent on the behaviour of the recipient jurisdictions, poor communities are able to develop a strategy of producing smaller benefits for their poor without being punished by a decrease in grants. Conversely, transfers will actually increase since the perceived needs of poor jurisdictions increase.

The conditions under which such a strategy might be successful are the existence of a monopolistic local government where there is a strong political “machinery” and/or a strong bureaucracy and the mayor has strong executive powers. All these conditions enable local governments to decrease the quantity/quality of local services without being removed from office by voters’ preferences.

Faith’s hypothesis seems somewhat machiavellian and unable to explain by itself fiscal crises in cities (as the author himself acknowledges). However, it points to an important criticism in normative approaches to intergovernmental grants. Designing grants based on the perceived needs of the jurisdictions may have the perverse effect that jurisdictions “produce” their increased levels of needs. This can be interpreted as indicating that local governments do not have incentives to reduce X-inefficiency if they know that they will be compensated for the increased costs of providing local services.

The paper of Alperovich (1984) also approaches grants from a positive point of view, and considers economic and political factors that might explain grants in the case of Israel. He does not develop a formal model but postulates that the allocation of grants is a function of the objective needs of the population common to all jurisdictions, social and economic factors which are specific to each community and political factors related to the voting pattern of the residents.
More specifically, he considers that central government behaviour in allocating grants can be given by the following relationship:

\[ g' = f(N^i, P^i, Dep^i, Def^i, u^i) \]

where \( g' \) is grants per capita, \( N^i \) is population, \( P^i \) the percentage of voters in jurisdiction \( i \) who voted for parties which form the government, \( Dep \) is a dependence ratio (ratio of people aged over 65 and under 14 to those people aged 14 to 65), \( Def \) the per capita annual deficits and \( u^i \) the error term.

Since this relationship is ad hoc he estimates several specifications of the functional relationship (linear, logarithmic and exponential) for 1976 when Labour was in power and 1978 when Likud was in office. Alperovich’s results show that the estimated coefficient are almost all significant\(^1\) and have the expected signs. Intergovernmental grants (per capita) seem to increase with jurisdiction needs (\( Def \) and \( Dep \)) and also with the share of voters that support the incumbent government whether it is Labour or Likud. On the other hand grants per capita decrease with the population size of communities, which Alperovich interprets as indicating economies of scale.\(^2\)

Finally, Inman (1988) analyses what he considers to be the evolution of a new federalist fiscal order in the United States, the main characteristics of which are the historical increase in federal, state and local spending in real terms and in percentage of national income, the increase in the state/local share of nondefense expenditures and finally the centralization of revenues (increasing weight of receipts from grants) in both state and local governments.

Inman’s aim is to understand whether the traditional economic ratio-

\(^1\)Except for the political variable (\( P \)) in 1978 (logarithmic specification) and the population variable (\( N \)) also in 1976 (linear and exponential specifications).

\(^2\)However, Alperovich does not clarify whether he is referring to economies of scale in consumption or to economies of scale in production. It is one important argument of this thesis that the population coefficient can not be interpreted solely in terms of economies of scale in production but also in terms of decreasing quality of local services.
nates underly the evolution of this new fiscal order or if there are political and institutional factors which are more relevant.

On efficiency grounds, grants might be necessary to induce state/local governments to provide the efficient level of national public goods (e.g. public infrastructures), to increase the provision of goods which have a significant level of spillovers across jurisdictional boundaries and finally to achieve within community allocative efficiency. On the other hand a different rationale for grants is to achieve equity goals in particular to equalize the provision of meritorious local public goods (e.g. education) across states and local communities.

Using regression analysis Inman concludes that "if one is to find a compelling public purpose logic to the present structure of federal aid to state and local governments it would have to be on the grounds of economic equity not economic efficiency." Therefore, he turns to politico-economic rationales and suggests the hypothesis that the growth in federal grants might be better understood as an outcome of redistributive politics.

Inman's argument runs briefly as follows. After World War II there was a substantial increase in demand for local services so that the mayors, governors and politicians turned to the federal government aiming to receive increased grants. However, an important shift in the institutional structure was necessary to allow for an important increase in federal grants. This structural change was the decentralization of Congress decision-making.

Elected representatives of Congress are assumed to derive political benefits from the projects (measured in dollars) aimed at their constituents taking into account a set of exogenous characteristics of the residents.

On the other hand they have political costs which can be given by the district's share of total taxes needed to finance total federal grants in all jurisdictions. Thus, the political cost is a positive function of this “tax-

\[\text{In Inman's table 2.4 there are 31 (!) independent equations, so that a meticulous analysis of each would be tedious.}\]

\[\text{Inman (1988) pg. 51.}\]
share”.

Under a pure centralized Congress (what Inman calls a cooperative legislative regime) where there is a single political leader, there is full internalization of spillovers from centralized financing. In this case marginal political benefits equate with marginal political costs. As we move from this centralized regime to a majority-rule legislative regime and from here to a decentralized regime the political cost associated with each project decreases since the (perceived) tax-share is lower.

The main difference between the previous majority-rule regime and the present decentralized regime is not the actual rule of decision-making (which is the same) but the changes in political environment which Inman relates to the declining influence of political parties, the increasing sophistication of voters and congressional redistricting favouring urban areas. These changes contributed to increase logrolling or “pork-barrel” budgeting where each member of congress submits his/her preferred project and the others vote in favour so that their projects are also approved. As a consequence of redistributive politics it is expected that the aggregate level of spending increases, spending becomes more equalized across congressional districts and the number of projects and programmes should expand to incorporate the needs of each legislative district. Therefore, Inman concludes that the changes in the U.S.A. federalist fiscal order are consistent with redistributive politics rationale.

3.4 Conclusions

Empirical literature that has addressed the problem of economies of community size, has been mainly within the median voter framework. The initial results of this theory suggested that there were no economies of community size, and this result was obtained under the assumption of constant returns to scale in the production of local public goods and a decreasing marginal

---

On logrolling see Mueller (1989) pgs. 82-89.
congestion specification of the "crowding" parameter.

Developments of the literature addressed the issue that those assumptions would have to be carefully considered since they have implications for the results of the non existence of economies of community size and also the "privateness" of local public goods.

Moreover, other approaches to local decision-making emphasized supply side factors and the role of politicians and in particular bureaucrats. Under this framework the "privateness" result may have a different interpretation. Several authors argue that there are economies of city size but that they are appropriated by budget maximising local governments.

Therefore the "publicness" or "privateness" of local public goods which is a different way to look at the problem of economies of community size is still an open issue. We will dedicate special attention to it in the next chapter.

The literature on intergovernmental grants also has empirical evidence that contributes to the rejection of the general validity of the median voter model as a model of local political decision-making. Namely the flypaper effect is consistent with several bureaucratic models of political decision-making. In general, when institutional factors are taken into account the impact of intergovernmental grants is different from the median voter model predictions.

Positive approaches to intergovernmental grants considered in the previous section, suggest that underlying the normative approaches based on efficiency and/or equity rationales, there is a need to search for politico-economic rationales. Such an approach will be developed in chapter 4.

The analysis thus far was developed in the implicit context of decentralized governments. A departure from this institutional context is the existence of centralized governments the meaning of which is developed in the next chapter. It will be apparent then, that the flypaper effect is easily explained within the context of centralized governments.
Chapter 4

Economies of community size and the quality of services within centralized fiscal structures

In chapter 1 we introduced the need for a distinction between centralized and decentralized fiscal structures. This chapter explores the theoretical and empirical implications of approaching local public economics within the framework of a centralized government.

The decentralized framework has been implicitly assumed in all of the median voter literature and most of the bureaucratic literature, as we have shown in the last chapter. The centralized approach has not received much consideration in the literature. However, as shown in chapter 2, it is possible to consider the Tiebout model in this latter framework.

Section 4.1 clarifies the distinction between the two frameworks after clarifying descriptive concepts related to the fiscal position of communities. It also summarizes the argument to be developed in this chapter.

The critical role that congestion and the changing quality of local services play in the context of our argument requires a particular attention to the problem of the "publicness" of local public goods. A new approach to
crowding and congestion is developed in section 4.2.

Section 4.3 develops a simple model that enables an understanding of the economies of community size and the provision of local public goods. It also clarifies the importance of intergovernmental grants within centralized governments and introduces some hypotheses that can be subjected to empirical scrutiny. This empirical research will be developed in chapters 5 and 6. Section 4.4 summarizes the chapter's results.

4.1 Decentralized versus Centralized fiscal structures

When Oates (1972) addressed the issue of fiscal federalism he set out two opposing polar forms of government. At one end a unitary form of government where there is a unique (central) institution that performs the major functions of government, which according to Musgrave's taxonomy are: allocation, (re)distribution and stabilization. At the other end is a decentralized system where numerous local governments perform almost all of the economic functions making central government a minimal and in effect residual government. Of course, these are only ideal types in the Weberian sense of the concept. Real fiscal structures in different countries are in-between these polar forms.

Oates analysed empirically the degrees of fiscal decentralization across a set of countries and used four variables which are: the share of central government in general government current revenues, current expenditures, consumption and civil expenditures.

The fact that these measures give an idea of the degree of fiscal centralization of a country is not in question. However, they do not focus on a crucial aspect of the centralization issue which is essential to understand local governments' fiscal behaviour. This aspect is local governments' degree of discretion in setting local revenues and expenditures. In other words, how and to what extent are local governments able to operate marginal
changes in the level of expenditures and within what limits.

Therefore, we distinguish two polar cases. The *centralized* fiscal structure arises when local governments are unable, in the short run, to set their own level of expenditures, because their revenues are totally exogenous. Moreover, the share of central government expenditures in total government expenditures is very high.

If local revenues are to be exogenous this implies two things. Firstly, that intergovernmental grants are only of the lump-sum form. Secondly, local tax rates are set by central government. Therefore, assuming immobile tax bases in the short run, *resources are effectively exogenous even in the presence of local taxes.*

On the other hand, what characterizes *decentralized* fiscal structures is the ability of local governments to autonomously decide the level and composition of local expenditures. Usually decentralisation is also associated with a great degree of sub-national governments' share of total government expenditures.

This chapter analyses the theoretical implications of centralized fiscal structures. The first methodological implication is that there is no need for a model of local governments' fiscal behaviour to determine overall local expenditures, because they are set by central government. However, a model is needed to explain the composition of local expenditures.

The first corollary of the exogeneity of revenues is that allocative efficiency will not be reached unless we assume that reality is as in the Tiebout model where citizens are distributed among communities according to their preferences for local public goods. In chapter 2 we have already dismissed this approach when applied to a whole country.

The second corollary is that demand approaches are somewhat mean-

\[^{1}\text{Matching grants have to be excluded since, even if local governments were unable to change directly the size of the budget they could do it indirectly, through an alteration of the composition of the budget enlarging the share of the programmes subsidized under matching grants.}\]
ingless, because what happens in centralized fiscal structures is that local governments supply local services under an exogenous financial constraint. The emphasis, at least in theory, is on horizontal fiscal equity. Disregarding the (important) problem of asymmetry of information between central and local government, within centralized fiscal structures it is possible to equalize fiscal positions of local governments, whereas in decentralized systems it is not.

It is useful at this stage to clarify some concepts that will have a mere descriptive function. They are mainly developments of concepts introduced in Musgrave and Musgrave (1989). It is assumed that there is only one tax base \( B \), one tax rate and that local revenues are the sum of local taxes with intergovernmental lump-sum grants \( G \). All definitions will be related to community \( i \) which appears always as a superscript.

**Fiscal capacity** of community \( i \) within a decentralized government is given by:

\[ F_i = t_i B_i \]

where \( t_i \) is a standard tax rate. In centralized governments it is given by:

\[ F_i = t_f B_i \]

where \( t_f \) is the local tax rate set by central government. In the definitions below, to obtain the equation applied to a centralized government it is sufficient to introduce \( t_f \) instead of \( t_i \).

The *isoquality cost function* \( C^*(\cdot) \) is the cost function for a local public good \( X \) where the quality of services is constant at the level \( z \).\(^2\)

**Fiscal need** \(^3\) associated with the level of services \( x \) \((R^x)\) is the value of

\(^2\) An algebraic presentation of the function will be considered in section 4.3.

\(^3\) We deeply dislike the concept of "need" due to the ambiguity and relativity of the concept. Some "needs" for developed countries appear as "luxuries" for underdeveloped ones. All the discussion below is developed in terms of quality or standards of service provision. Nevertheless, we will keep the concept of "fiscal need" for the sake of consistency with some existing literature.
the isoquality cost function in community \(i\), i.e.:

\[
R^i = C^i(N_i) \quad \text{for} \quad N = N^i
\]

where \(N^i\) represents the population of community \(i\) and \(R^i\) the level of fiscal resources needed to provide the services at quality \(x\) to the residents of the community.

**Fiscal position** of a community is the relationship between fiscal capacity and fiscal "needs" at a standard level \(x\):

\[
F^i = \frac{t_i b^i}{R^i} = \frac{t_i b^i}{\nu^i}
\]

where \(b\) and \(r\) refer to per capita values. Therefore, potential fiscal position is a relative index that indicates local governments' ability to provide local public goods at a standard level \(x\) if it had set a standard tax rate \(t_s\). If instead of the standard tax rate \(t_s\) it is considered the effective tax rate \(t\) we may speak of real fiscal position.

Note that if \(R^i\) is a constant independent of the size of the jurisdiction, fiscal position is proportional to fiscal capacity.

Figure 4.1 clarifies the distinction between a centralized and a decentralized fiscal structure.

To interpret the figure it is convenient to assume that the standard local tax rate under the decentralized system \((t_s)\) is identical to the central government set local tax rate \((t_f)\). We will also assume for a start that \(r^i\) = constant and therefore that fiscal capacity is a proxy of fiscal position. This assumption will be relaxed later on.

Tax rates are measured on the horizontal axis and fiscal resources on the vertical axis. The lines 0A and 0B show the own fiscal resources of communities A and B respectively, for different values of the tax rate. In this situation there are no intergovernmental grants. Since the slope of 0A is higher than the slope of 0B it is assumed that community A has a larger tax base.

The first difference between systems of government is obvious. Within centralized governments the figure reduces to the vertical dotted line \(t_f t_f\).
Community A will have local taxes $A_1$ and community $B$, $B_1$. If central government gives respectively grants $G^a$ and $G^b$, it will put communities in the same real post-grant fiscal capacity. It will also equalize the real post-grant fiscal position of the two communities, due to the assumption of $r^z = constant$.

On the other hand, under a decentralized government, the same set of grants would equalize the post-grant fiscal capacity and fiscal positions of the two jurisdictions albeit the real situation of both might well differ.

In other words, under a centralized regime, intergovernmental grants can be designed to offset disparities in communities’ real fiscal positions, since tax rates are legally binding. Therefore, grants can achieve an equalization of fiscal positions across communities.

However, under a decentralized regime, central governments can achieve, at best, an equalization of potential fiscal positions.

If we relax the assumption that $r^z = constant$ the system of lump-sum
grants to communities $A$ and $B$ under the centralized system of government will equalize post-grant fiscal resources in both communities. However, fiscal positions would not be equalized.

There has been a continuing and extensive normative debate on horizontal fiscal equity which will not be addressed in this thesis. The approach here is positive in the sense given to this term in chapter 1. Basically this chapter addresses two problems in the context of fiscally centralized countries:

1- What is the relative fiscal position of communities with different population size?

2- In what sense can it be predicted that intergovernmental grants are designed to offset differences in the fiscal position of communities?

The answer to the first problem can be simplified under a centralized regime since fiscal capacity of jurisdictions is easily dealt with. There remains the problem of analysing the existence and magnitude of economies of community size, or whether bigger communities are (or are not) able to provide the same levels of services at a lower per capita cost.

The existence of economies of community size is essentially related to the existence (or not) of economies of scale in production and "publicness" in consumption of local public goods.

The approach developed here distinguishes two kinds of local services according to technology. Labour intensive services, where economies of scale are exhausted at a low level of output, and capital intensive services where the range of output having increasing returns is much broader.

It will be shown theoretically and analytically that under certain assumptions local public goods have horizontal "privateness" characteristics, i.e. that to provide the same level of services in different size communities

---


$^5$A better concept clarified below is constant returns to scale in consumption.
requires a provision of local public goods proportional to population.

From these two aspects it follows that there are no economies of community size in the provision of labour intensive local services but that there are economies in the provision of capital intensive services (e.g. public infrastructure). These economies are a result of economies of scale in production and not of economies of sharing in consumption.

Therefore, economies of community size are said to exist and be greater the higher the "preferences" for capital services in larger communities.

Within centralized fiscal structures the level and distribution of intergovernmental grants is crucial in determining the quality of local services provided by each jurisdiction. If grants are designed to offset differences both in fiscal capacity and economies of community size we would expect that the quality of services provided would be similar.

On the other hand, if intergovernmental grants depart significantly from the isoquality horizontal design, we would expect considerable variations in the quality of services provided among communities. In fact a distinct feature of centralized fiscal structures is that changes in quality are precisely the way of adjustment of exogenous revenues to particular local "needs".

This leads us to the second problem stated above. This chapter falls short of developing a positive theory of intergovernmental grants. Perhaps the development of such a theory is too ambitious a task. Specific peculiarities of constitutional, political and economic conditions in different countries, might dictate different systems of intergovernmental grants. Nevertheless, hypotheses can be formulated and empirically tested for specific countries.

In this chapter three hypotheses are confronted. One considers that

---

6 All the analysis excludes small communities (i.e. under 10000 inhabitants).
7 As we referred to in section 3.3 there are few positive approaches to intergovernmental grants' design. The majority of positive approaches address the impact of intergovernmental grants on local governments' behaviour.
intergovernmental grants are designed according to central government's macroeconomic objective of controlling public sector expenditure growth. Another is the more traditional hypothesis that revenue-sharing grants are designed to offset disparities of fiscal capacity across communities. Finally, the third hypothesis states that lump-sum grants aim at equalising communities' fiscal positions, with the meaning given to this concept in this chapter.

To address the issue of fiscal position and community size, we need firstly to consider the "privateness" or "publicness" of local public goods and also to face the issue of economies of scale in production. This is done in the next section.

4.2 Properties of the Crowding and Congestion Functions

Let us start with a conceptual clarification. Congestion will be used here as in club goods theory to indicate the relationship between the level of services subjectively perceived by an individual who shares the consumption of a given facility and the membership and capacity sizes of that facility. Congestion is mainly a facility related concept but might also be applied to a community whenever there is joint production and joint consumption of a unique local public good by all members of the jurisdiction. However, if the LPG is provided in different facilities and locations within the community congestion applies to each facility. In general we will assume that given a local public facility $X^i$ and the users of that facility $n^i$ there exists a congestion function $\phi$ such as:

$$ z = \phi(X^i, n^i) $$

with $\phi_x > 0$, $\phi_{n'} < 0$ and $\phi_{n'n'} < 0$.

Let us define, like Brueckner (1981), returns to scale in consumption as the services perceived by individuals when per capita provision is con-
stant but the scale of provision (and therefore population) increases. If, for instance, \( \phi \) is homogeneous of degree \( k \) there will be increasing, constant or decreasing returns to scale in consumption whenever \( k \) is greater than, equal to or less than zero. This concept is meaningful both within and across jurisdictions. Moreover, we will refer to the effect on the level of subjective consumption of an increase in the scale of LPG provision as the scale effect.

This clarifies another feature of the BG and BD formulation of the crowding function. Since with a scalar \( \lambda > 1: \)

\[
\frac{\lambda X}{(\lambda N)^{\gamma}} = \lambda^{(1-\gamma)} \frac{X}{N^{\gamma}} = \lambda^{(1-\gamma)} x
\]

services are homogeneous of degree \( k = 1 - \gamma \) regarding population and capacity sizes. If \( \gamma = 1 \) there are constant returns to scale in consumption, if \( 0 < \gamma < 1 \) there are increasing returns to scale in consumption since \( \lambda^{(1-\gamma)} > 1 \) and if \( \gamma > 1 \) there are decreasing returns to scale.

Finally, changes in the level of subjective services perceived by the individual when capacity (\( X \)) is fixed and the size of the sharing population changes will be called the density effect.

Crowding is the relationship (if it exists) between the level of services perceived by an individual and the level of aggregate provision and aggregate membership size of the consumption group in each community. Therefore crowding applies to the relationship widely used in the empirical median voter approaches. Under a given set of assumptions, to be clarified below, there is a crowding function with at least two arguments (\( X \) and \( N \)):

\[
\psi = \psi(X, N, ..)
\]

It will be argued below that for the crowding function to be meaningful it has to have two main characteristics: constant returns to scale and decreasing marginal crowding. This result is compatible with variable returns to scale of the congestion function.

The argument will be developed through a careful distinction between crowding and congestion and between scale and density effects. As will
became apparent, changes in crowding (assuming the function exists) are the sum of the scale and density effects.

4.2.1 Crowding and congestion with a unique facility per jurisdiction

Let us assume that (i) individuals are homogeneous across jurisdictions, (ii) jurisdictions have different population sizes, (iii) each jurisdiction provides one local public good jointly consumed by all members of the jurisdiction and that LPG’s are not pure local public goods ($\phi_N < 0$). In this particular case the crowding and congestion functions would be identical:

$$\psi(X, N) \equiv \phi(X, N)$$

The decomposition of the crowding effect into a scale effect and a density effect can be illustrated by Figure 4.2, where the domain of the $\psi$ function is drawn.

Community A has $N$ citizens sharing the consumption of a capacity size facility $X$ while community B has $N + dN$ residents consuming a facility with size $X + dX$.

The difference in population between the two communities ($dN$) can be given by a sum of a proportional change in population given the change in facility size $dN^* = \frac{N}{X}dX$ plus a residual change $dN - dN^*$. Therefore it is possible to write:

$$(dX, dN) = (dX, dX\frac{N}{X}) + (0, (dN - dX\frac{N}{X}))$$

The scale effect $(dX, dX\frac{N}{X})$ from A to A' (see figure 4.2) assumes that per capita provision remains constant while capacity adjusts to the level

---

8This is a critical assumption in the development of our argument. We will return to it below.

9In this section functions $\psi$ and $\phi$ could be used interchangeably since, by assumption, each jurisdiction provides the LPG using only one facility. However, this is a particular case of the function $\phi$ and this is the reason why we use the general notation for the crowding function.
effectively provided in community B. Therefore, A' is the position a community would be in if it had the same density of utilization of the local public good as community A and the same level of provision of community B.

The density effect \((\mu, dN - dX \frac{N}{X})\) considers that capacity is given and that the population changes from what the level of population would be, if per capita provision had remained constant, to the actual level existent in community B.

In general, to see how the level of subjective services changes with both population and capacity sizes we differentiate totally the crowding function:

\[
d\psi(X, N) = \psi_X.dX + \psi_N.dN = \psi_X.dX + \psi_N.dX \cdot \frac{N}{X} + \psi_N(dN - dX \frac{N}{X})
\]
Which after some rearrangements yields:

\[ d\psi(X, N) = (1 + \frac{\psi_{N,N}}{\psi_{X,X}})\psi_X dX + (1 - \frac{\psi_{X,N}}{N})\psi_N dN \quad (4.1) \]

The first term of the right hand side is the \textit{scale} effect on subjective or perceived consumption, while the second term is the \textit{density} effect.

The scale effect shows how services change as a result of an increased capacity \((\psi_X dX > 0)\) and a proportional increase in population \((\psi_N \frac{X}{N} dX < 0)\).

The second term captures the effect on congestion of changes of density. If provision per capita is constant \((\frac{\psi_{X,N}}{dX/N} = 1)\) when considering two communities with different population sizes, differences in the level of services can be due only to the existence of economies (or diseconomies) of scale of provision.

On the other hand if citizens are indifferent to the size of provision of the LPG, provided that per capita provision \((X/N)\) is constant, the only relevant effect is density. In this case \(\psi = \psi(X/N) \equiv \Psi(z)\), \(\psi_N = -\frac{\partial}{\partial z} \Psi \) and \(\psi_x = \frac{1}{N} \Psi_x\). Therefore,

\[ \frac{\psi_{N,N}}{\psi_{X,X}} = -1 \]

which equates the scale effect to zero.

Therefore, changes in the level of services as a result of crowding are a sum of scale and density effects as defined above.

It should be emphasized, that if there are constant returns to scale in consumption, differences in the quality of services arise only from differences in \textit{per capita} provision. The next section clarifies that the homogeneity of the \textit{crowding} function is the only realistic assumption.

\subsection*{4.2.2 Crowding and congestion when each community provides several facilities}

The approach introduced in the last section applies to some LPGs provided by communities but not to others. Many communities provide not one but several parks, public libraries, swimming-pools, etc. In general the bigger
is a community the greater the number of facilities providing the same service.

As a consequence a distinction must be made between $\phi$ and $\psi$ and the assumptions under which a crowding function is meaningful must be considered.

There are two alternative sets of assumptions whereby, given a congestion function $\phi = \phi(X',n')$, a crowding function $\psi(.)$ exists.

Firstly, there is the case when $\phi$ is homogeneous of degree zero and population is distributed across facilities proportional to their size. In this case:

$$x = \phi(X',n') = \Phi\left(\frac{X'}{n'}\right) = \psi\left(\frac{X}{N}\right) \equiv \Psi(x)$$

A simple example illustrates this aspect. Consider two communities A and B with the same population ($N$) and the same aggregate provision (area) of parks ($X$). The only difference is that community A has one park, while community B has two parks ($b_1$ and $b_2$) with dimension (area) ($X^{b_1} = X/3$) and ($X^{b_2} = 2X/3$) respectively. Assume for a start that the population distributes between the parks proportionally to their area. Therefore, the level of services received by each individual in community A is $x^A = \phi(X,N)$ and in community B is $x^{b_1} = \phi(X^{b_1},n^{b_1}) = \phi(X/3,N/3)$ for a user of park $b_1$ and $x^{b_2} = \phi(X^{b_2},n^{b_2}) = \phi(2X/3,2N/3)$ for a user of park $b_2$. It is easy to understand that $x^A = x^{b_1} = x^{b_2}$ if and only if $\phi$ is homogeneous of degree zero.

To clarify the critical role of the proportional distribution of population between parks consider that in community B the users were distributed between the less dense park ($b_1$) and the more dense park ($b_2$) in the following proportions $n^{b_1} = \frac{N}{4}$ and $n^{b_2} = \frac{3N}{4}$. Assuming now homogeneity of the congestion function ($\Phi = \Phi\left(\frac{X}{n'}\right)$) we would find that the services received by users of the more dense park would be: $x^{b_2} = \Phi\left(\frac{X^{b_2}}{n^{b_2}}\right) = \Phi\left(\frac{\frac{2X}{3}}{\frac{3N}{4}}\right)$ and those of the less congested park $x^{b_1} = \Phi\left(\frac{X^{b_1}}{n^{b_1}}\right) = \Phi\left(\frac{\frac{X}{3}}{\frac{N}{4}}\right)$. No crowding function ($\psi$) would exist because the same aggregate provision of parks ($X$) and the same population ($N$) in the two communities was associated with different
levels of services.

Therefore with \( \phi \) homogeneous and proportional distribution of population, \( \psi \) would also be homogeneous with only two arguments \((X, N)\).

An alternative set of assumptions considers that communities of different sizes differ by the number of identical clubs they have. Therefore, all facilities have identical membership and capacity sizes and \( \phi \) is not homogeneous of degree zero.

In this case the crowding function can be obtained as follows:

\[
z = \phi(X', n') = \psi(X/k, N/k) = \psi(X, N, k)
\]

where \( k \) is the number of identical clubs that the jurisdiction has.

The reason why \( X \) and \( N \) are not the only arguments in the function is that an individual is indifferent regarding the population size of the community where he lives and the aggregate provision of LPG, insofar as the characteristics of the particular club he uses are the same. This happens whenever \( X/k \) and \( N/k \) are the same, independently of the absolute size of the community.

To summarize, if a crowding function is to exist when communities provide LPGs in several facilities, the population has to divide between facilities proportionally to their size. If facilities have different sizes, it has to be assumed that the congestion function is homogeneous (degree zero) and the crowding function is represented by:

\[
z = \Psi(x)
\]

On the other hand if facilities have similar dimensions, it is possible to allow for non homogeneity of the congestion function, but in this case we would have:

\[
z = \psi(X, N, k)
\]

The assumption of identical dimension of LPGs does not appear to be a reasonable one. Moreover, the use of \( \psi(X, N, k) \) in empirical analyses would imply the introduction of the additional (discrete) variable \( k \).
If a crowding function is to be used, the constant returns to scale assumption seems the only realistic one, particularly if each community provides several facilities.

As an example consider local schools. Within a country the size of schools has a considerable range as can be seen when low density rural areas and high density urban areas are considered.\(^\text{10}\)

It is reasonable to assume that \textit{ceteris paribus} (i.e. keeping constant the skills and qualification of teachers, the characteristics of the student population, etc.) the quality of education services does not change considerably if the number of students per class and the size of the recreational area per pupil is similar. However, a significant change in these variables will most likely affect the quality of educational services received.

It might be argued, against the assumption of constant returns to scale in consumption, that too big a school may be subject to decreasing returns from educational services, even if per capita provision is constant.

This seems to be true. However, schools are replicated not only to minimize transportation costs but also to avoid diseconomies of scale in consumption. Therefore we can qualify the assumption of constant returns to scale, saying that facilities have a size within the range of constant returns to scale in consumption.

There is, however, an important case where the assumption of constant returns to scale in consumption does not hold. This is when the capacity of LPGs is partially idle. Pure local public goods can be seen as a particular case and services provided on a non continuous basis another.\(^\text{11}\)

\(^{10}\)This happens in Portugal and it is likely to happen in other countries. In fact, if the size of schools were similar this would increase dramatically the size of the catchment areas in rural communities and with it the transportation costs (including the costs in time to travel from home to school).

\(^{11}\)This is the case of emergency services, in particular fire services. We have discussed already that the findings in Brueckner (1981) of increasing returns to scale in fire services can be accepted, since it is a particular case where capacity is idle part of the time. However, this does not happen for the majority of local services, and therefore is not
The case of “pure local public goods” seems to be a possibility only in small communities. In those communities it is likely that the technological minimum level of LPGs capacity, is such that no rivalry in consumption exists. Therefore, at the margin, LPGs appear as pure public goods. Increasing population sharing the consumption of the good can either not affect or even increase (“camaraderie” effect) the level of services that each individual receives. However, in bigger communities the threshold is soon reached where negative externalities in consumption arise.

As a consequence, the assumption of homogeneity of the crowding function can be supported only when small communities are excluded.

4.2.3 Crowding function, increasing marginal congestion and “publicness”

In the preceding section it was concluded that whenever some community provides one service through different facilities, a distinction must be made between crowding and congestion. Moreover, it was clarified that the only realistic approach regarding the crowding function is that it is homogeneous of degree zero in population and capacity, and that the population divides between facilities in proportion to their size. Otherwise, much stronger restrictions would have to be made. This section explores more deeply the implications of the specification 

\[ x = \psi(X, N) = \Psi(X/N) = \Psi(x) \text{ with } x = X/N. \]

It can be shown that \( \psiX \leq 0 \) when \( \psiX \geq 0 \) and \( \psiXN \leq 0 \) (increasing marginal congestion with population) as club goods theory suggests.

Differentiating \( \psi \) with respect to \( N \) yields:

\[ \psiN = -\psiX \frac{X}{N^2} \Rightarrow \psiX = -\frac{N^2}{X} \psiN \]

and differentiating with respect to \( X \),

\[ \psiX = \PsiX \frac{1}{N} \]

discussed below.
Therefore,

\[ \psi_{NN} = \frac{X^2}{N^4} \psi_{ss} + 2 \frac{X}{N^3} \psi_s \]  

(4.2)

Substituting the expression \( \psi_s \) into the equation and solving in order to \( \psi_{ss} \) yields:

\[ \psi_{ss} = \frac{N^3}{X^2} [N \psi_{NN} + 2 \psi_N] \]  

(4.3)

Since \( \psi_N \leq 0 \) and assuming, as in club goods theory, increasing marginal congestion \((\psi_{NN} \leq 0)\) it can be concluded that

\[ \psi_{ss} \leq 0 \]

The only case where \( \psi_{ss} = 0 \) is when local public goods are pure public goods, i.e. \( \psi_N = 0 \) (and obviously \( \psi_{NN} = 0 \)).

As a general conclusion on the properties of the “crowding” function for communities with medium to large population size \(^{12}\) we have:

\[ x = \Psi(z) \]  

(4.4)

with \( \Psi_s > 0, \psi_{ss} < 0 \) and \( z = \frac{X}{N} \)

Figure 4.3 illustrates the shape of the crowding function. The horizontal axis shows the per capita provision of the local public good and the vertical axis the level of services that citizens receive.

It should become apparent that if communities with different populations, provide the same level of per capita provision (say \( z_1 \)), they will all locate in the same point in the crowding function (B).

If communities do not differ too much in their level of per capita provision, i.e if they locate within the arc AB, a linear approximation to the crowding function will fit the cross-section data well (see the dashed line). This might explain the “privateness” result. In fact this empirical result is illustrated by:

\[ x = \frac{X}{N} \]
Someone could argue, against this “privateness” result, that it assumes the narrow range of per capita provision across communities \((z_1, z_2)\).

If the range is not narrow, it means that there is a wide variation in the services provided to the citizens in different communities. This situation would not be the one considered in the “privateness” versus “publicness” debate. Here, it is assumed that the services are relatively homogeneous across jurisdictions which implies, under our analysis, a narrow range of per capita provision.

Therefore, the “privateness” characteristic of LPG might be obtained, assuming constant returns to scale in consumption, and that (within the sample of communities considered) the range of per capita provision of local public goods is not too wide. It was also assumed that local public goods

\[12\] For empirical purposes we will adopt the range more often considered in the literature between 10,000 and 250,000 inhabitants.
goods are not "pure public goods" and we introduced the more realistic assumption of increasing marginal congestion in consumption.

We reach the same conclusion as Wyckoff (1988b), albeit using a different path. There is no need to introduce bureaucratic explanations for the "privateness" result. This does not mean that bureaucratic forces do not operate at the local level.

Another conclusion is that not only the subjective quality of services is identical when per capita provision is constant but also utility levels will be the same if we assume two additional assumptions: (i) constant returns to scale in the production of public facilities and (ii) equal sharing of the costs of public provision.13

Nevertheless, the assumption of constant returns to scale is not always realistic as discussed below.

4.2.4 Economies of scale

A similar problem found when discussing the crowding function appears when economies of scale are considered. The argument can therefore be summarized.

Assuming a homogeneous production function, economies of scale mean that the scale of production increases with a less than proportional increase in inputs. This may mean two different things when local public goods are considered. The scale of aggregate production may increase because the size of a unique facility increases (widening roads, enlarging a swimming-pool, etc.) or because there is an increasing number of facilities (e.g. the replication of facilities of equal size)14.

Therefore we should distinguish between the production function $f(.)$ as referring to the production on-site of a given facility and the production

---

13See section 4.3.1.
14This distinction is necessary whenever local services are (or might be) provided in different facilities.
function $F(.)$ as referring to the aggregate production of multiple facilities.

If each and every jurisdiction has only one facility, then only the production function $F(.) \equiv f(.)$ needs to be considered, which might have increasing, constant or decreasing returns to scale for different levels of output.

On the other hand, in the presence of several facilities providing the same service, if $F(.)$ is to exist, either (i) all facilities have identical size and $f(.)$ might have increasing returns to scale or (ii) facilities have different sizes and $f(.)$ has to have constant returns to scale. In both cases $F(.)$ has constant returns to scale.

Case (i) was considered by Berglas (1981) and Berglas and Pines (1981, 1984) in the context of club goods theory. They consider a facility production function that might have increasing returns to scale:

$$\% = f(A_x, N_x)$$

where $A_x$ and $N_x$ are the land and labour input per facility $X$ respectively. Assuming a constant level of services rendered to each club member the resources necessary to produce $N/n$ clubs are $N/n$ times greater, so that if we measure aggregate output of $X$ as the number of identical clubs $N/n$, the "aggregate" production function can be obtained multiplying both terms of the equation above and rearranging so that:

$$X = f(A_x, N_x)$$

which means that $F(.)$ has constant returns to scale in the production of facilities with identical size $X$.

Some qualifications are necessary to the B-P approach. If $f$ has increasing returns to scale within a large scale of production then it results from the definition of economies of scale that ($k$ being an integer):

$$kX > f(kA_x, kN_x)$$

15 We keep our notation (which is different from B-P's notation) where $X$ refers to the local public good, in this context the club good.
which means that from an efficient production point of view, there is no case to replicate clubs. Therefore, the only plausible case for the replication of clubs, within the B-P model framework is to assume that congestion costs (with production costs) lead to the optimal size of the club $X^*$ being reached well before economies of scale are exhausted.

The competitive case of B-P is useful from a theoretical point of view but unrealistic from an empirical perspective. Most local public goods are not provided in the form of clubs (there is no excludability of consumption based on an entrance fee) and have different sizes. Therefore, case (ii) has to be accepted as the more realistic one. The necessity of the assumption of constant returns to scale, if facilities have different sizes, appears intuitive.

For the purpose of our analysis, it will be considered that communities provide current services ($X$) replicable in different facilities at constant average costs. However, capital related services ($V$) (infrastructures and so forth) are modelled as being a unique facility per jurisdiction, allowing for increasing returns to scale.

The simplest model considers that each community produces only one local public good under constant returns to scale:

$$C(X) = c(X)$$

where $c$ is the constant unit cost of provision.

A second case introduces two local public goods $X$ and $V$ with the respective long term aggregate\footnote{There is another powerful case for the replication of clubs which is the consideration of transport costs to access the facility, which should be introduced in the budget constraint of the individual. However, this is not the model we are dealing with.} cost functions:

$$C(X) = c_0X$$

and

$$C(V) = c_1V^\eta$$

where $c_0$ is the constant unit cost of $x$, $c_1$ is a constant (technological parameter) and $\eta$ is a parameter indicating the economies of scale.

\footnote{This applies only to the good $X$ as the discussion in this section clarifies.}
It is assumed also that there are no economies of scope or economies of joint production so that the joint cost function for the community in producing $X$ and $V$ is additive:

$$C(X, V) = c_0X + c_1V^n$$

The form of $C(V)$ is a constant elasticity specification and does not allow for a U-shaped long term average cost function. This would be possible with a more flexible functional form such as the translog cost function.\(^{18}\) The use of such a function would introduce such complexity in the approach, that would invalidate any relevant conclusions. Moreover, there are some arguments against the use of such flexible functional forms.\(^{19}\)

However, analysis of economies of scale in capital intensive goods using the translog cost function\(^{20}\) shows that, after an initial range of output where there are significative economies of scale, there follows a considerable range where the long term average cost function is almost flat. Within the specification considered in our approach, economies of scale are also decreasing at the margin, so that for high levels of output the long term average cost function becomes almost horizontal. Moreover, we exclude from our analysis small communities.

### 4.3 Community Size and Fiscal Position of Communities

At the beginning of this chapter fiscal position was defined as the relationship between the fiscal capacity of jurisdictions and the fiscal resources needed to provide a standard quality of local services.

There are several factors, other than population size, which might affect the fiscal "needs" of jurisdictions, most of them related to the production

---

\(^{18}\)See Christensen, Jorgensen and Lau (1975).

\(^{19}\)See Appelbaum (1979).

\(^{20}\)Christensen and Greene (1976).
function of local public services. As clarified already, it is possible to approach the “production function” of local services as a two stage process. In the “first” stage,\(^1\) local public goods are produced using public sector inputs and in the “second” they are combined with local characteristics to produce local public services. Within this framework there are at least three main factors which might account for differences in per capita fiscal “needs” between jurisdictions. Firstly, there might be differences in unit prices of local inputs purchased by the public sector. Secondly, there might be differences in local characteristics of the community (environment, population). Thirdly, communities have different population sizes and therefore there are differences in the number of citizens sharing the consumption of local public goods.

In the approach developed below it will be assumed that the first two conditions either do not hold or they are orthogonal (not correlated) with the population size effect. Therefore, we can isolate the consequences of this latter effect.

4.3.1 Local public good produced under constant returns to scale

The model now introduced is a very simple model of centralized fiscal structures. Local governments’ revenues are completely exogenous and come from two sources: intergovernmental lump-sum grants \((G)\) and local taxes \((T)\). Central government sets a local tax rate \((t_f)\) uniform across communities so that local governments are unable to raise autonomous revenues, at least in the short run, when the tax base \(B\) is given. They provide only one local public good \((X)\), have a balanced budget and do not “refund” any part of the grants received. Moreover, there is no \(X\)-inefficiency in the production of the local public good.

Citizens are assumed to consume a private numeraire good \(y\) and cap-

\(^1\)The distinction between first and second stages is an heuristic distinction and does not have any real temporal dimension.
Figure 4.4: Intergovernmental Grants, Crowding and the Quality of Services
ture services from the local public good on the basis of the crowding function 4.4:

\[ x = \Psi(x) \equiv \Psi(X/N) \]

Therefore, they have the utility function:

\[ U = U(y, \Psi(X/N)) \quad (4.5) \]

The local public good is produced under constant returns to scale, at a unit cost \( c \). Therefore, jurisdictions will have the cost function,

\[ C(X) = cX \quad (4.6) \]

own fiscal resources,

\[ T = t_fB \]

and a resource constraint,

\[ C(X) = t_fB + G \]

which can be given in per capita terms (after introducing equation 4.6) by:

\[ cz = \frac{t_fB}{N} + \frac{G}{N} = t_fb + g \quad (4.7) \]

Solving for \( z \) yields,

\[ z \equiv \frac{X}{N} = \frac{t_fb + g}{c} \quad (4.8) \]

where \( b \) and \( g \) are variables indicating respectively the per capita tax base and per capita grants. If we consider communities \( i \) and \( j \) with per capita tax bases \( b' \) and \( b'' \) respectively, the two following equations can be written:

\[ z_i' = \frac{t_fb'}{c} + \frac{1}{c}g = z_i' + \frac{1}{c}g \quad (4.9) \]

and

\[ z_j' = \frac{t_fb''}{c} + \frac{1}{c}g = z_j' + \frac{1}{c}g \quad (4.10) \]

with \( z_i' = \frac{t_fb'}{c} \) and \( z_j' = \frac{t_fb''}{c} \) the levels of per capita provision of \( X \) in communities \( i \) and \( j \) when they do not receive any amount of intergovernmental
grants. Since per capita provision in each community is only a function of intergovernmental grants we may write:

\[ z' = z'(g) \]  

(4.11)

and,

\[ z^f = z^f(g) \]  

(4.12)

We can introduce these equations into the crowding function and obtain,

\[ z' = \psi(z') = \psi(z'(g)) = \Gamma'(g) \]  

(4.13)

and

\[ z^f = \psi(z^f) = \psi(z^f(g)) = \Gamma^f(g) \]  

(4.14)

where \( \Gamma \) is the function that relates local public services to intergovernmental grants in centralized countries.

These relationships can be better understood with the help of figure 4.4.

In the first quadrant are plotted equations 4.9 and 4.10.

The per capita provision of \( X \) changes when either the per capita tax base changes (parallel shift of the lines) or intergovernmental grants change (movements along the lines). Jurisdiction \( J \) has a higher tax base and can provide \( z^J \) without any intergovernmental grants while community \( i \) can only provide \( z^i \). As per capita intergovernmental grants increase, jurisdictions are able to increase per capita provision of the local public good.

In the second quadrant is plotted the crowding function with the shape analysed in section 4.2 and given by equation 4.4. Therefore the quality of services is a function of tax base and grants both in per capita terms.

The fourth quadrant gives local public services \( z \) as a function of per capita intergovernmental grants \( g \) which is given by the \( \Gamma \) function (vide equations 4.13 and 4.14).

If jurisdictions have the same per capita tax base they have to receive the same amount of per capita grants to be able to provide the same quality of services to their citizens. On the other hand, with different per capita tax bases central government would have to offset these fiscal differences to put
jurisdictions in the same (post-grant) fiscal position. To provide identical standards of services \( x^* \) it should transfer \( g_i \) and \( g_j \) to communities \( i \) and \( j \) respectively.

In general, if \( x^* \) is such that \( x^* = \Gamma(x^*) \), intergovernmental grants to equalize provision of local services at a standard \( x^* \) are given by:

\[
g = c x^* - t f b
\]

(4.15)

Furthermore, the marginal effect on quality of an increase in per capita grants decreases with the size of the per capita grant. As a consequence, in communities where the quality of services is poor, a marginal increase in per capita grants will have a much greater effect in raising standards of service, than in communities where that standard is already high.

Finally, it should become clear that under the assumptions of this simple model there are no economies of community size. In other words, more populated communities are unable to provide the same quality of local services at a lower per capita cost than smaller size communities do.

The homogeneity of the crowding function implies that if the quality of services is to remain constant (\( x^* \)) the provision per capita must remain the same (\( x^* = \Psi(X^*/N) = \Psi(\lambda X^*/\lambda N) \) with \( \lambda > 0 \)).

On the other hand, considering costs per capita given by \( C_p(X, N) \), the assumption of constant returns to scale implies that:

\[
C_p(X, N) = C(X)/N = \frac{\lambda C(X)}{\lambda N} = \frac{C(\lambda X)}{\lambda N} = C_p(\lambda X, \lambda N)
\]

To see the implications of the analysis in terms of individual "welfare", some additional assumptions will be considered. Individuals are homogeneous regarding incomes (exogenous) and preferences and the local tax is a local income tax.

We can write the citizen resource constraint as:\(^{22}\)

\[
y + C(X)/N = I
\]

\(^{22}\) is a composite private good with unitary price. Note that \( N \) is not introduced in the cost function, as club good models sometimes do, since we want to isolate the effect of population through subjective congestion and not through the production function.
which using the assumption of constant returns to scale becomes

\[ y + c \cdot z = I \]

Therefore, the indirect utility function will be dependent only on the variable \( z \), the parameter \( c \) (unitary cost of the local public good) and the exogenous variable \( I \) (income):

\[ S = S(I - cz, \Psi(z)) = S(z, I) \quad (4.16) \]

If provision per capita is identical in different size communities a citizen is indifferent in living in any of them since he could not increase his "welfare" by doing so. It is also obvious that the bundle of private and local public goods consumed is not the result of a utility maximizing process, since individuals cannot choose their budget shares. Therefore, efficiency in consumption cannot be achieved.

Below, we are going to introduce a second local public good and relax the constant returns to scale assumption.

### 4.3.2 Extending the model to two local public goods

A second local public good \( V \) is introduced which is produced under increasing returns to scale in a unique facility per jurisdiction. For simplicity, and without loss of generality it will be assumed that the crowding function \( \Psi \) also applies to \( V \). Therefore the utility function is now:

\[ U = U(y, \Psi(X/N), \Psi(V/N)) = U(y, \Psi(z), \Psi(s)) \quad (4.17) \]

where \( s = \frac{V}{N} \).

Assuming the absence of economies of scope in production we have the following cost function:

\[ C(X, V) = c_0 X + c_1 V^n \quad (4.18) \]

The budget constraint of the community is now:

\[ c_0 X + c_1 V^n = tB + G \]
The left hand side of this equation are total costs and the right hand side total revenues including grants.

The budget constraint in per capita terms is,

\[ \frac{c_0 X}{N} + \frac{c_1 V^n}{N} = t b + g \quad (4.19) \]

Per capita costs \( C_p \) are given by:

\[ C_p(X, V, W) = \frac{c_0 X}{N} + \frac{c_1 V^n}{N} = c_0 x + c_1 s^n N^{n-1} \equiv C_p(x, s, N) \]

This means that per capita costs are a function of per capita provision of local goods \( X \) and \( V \) and also of the population size of communities. If we introduce the inverse of the crowding function \( (\Psi^{-1}) \) it is possible to write the per capita cost function as a function of the quality of services \( (x \) and \( v) \) provided:\[^2\]

\[ C_p(x, v, N) = c_0 \Psi^{-1}(x) + c_1 (\Psi^{-1}(v))^n N^{n-1} \quad (4.20) \]

To analyse economies of community size, firstly we will calculate the differential of the per capita cost function and then the elasticity of the per capita cost with respect to population. Finally, it will be assumed that the quality of services is constant and therefore we will obtain the *isoquality population elasticity of the per capita cost*.

The differential of the per capita cost function is given by:

\[ dC_p = \frac{\partial C_p}{\partial X} dX + \frac{\partial C_p}{\partial V} dV + \frac{\partial C_p}{\partial N} dN \quad (4.21) \]

and therefore,

\[ dC_p = \frac{c_0}{N} dX + \eta c_1 \frac{V^{n-1}}{N} dV - c_0 \frac{X}{N^2} dN - c_1 \frac{V^n}{N^2} dN \]

\[^2\]Note that if in the equation below we make \( z = x_0 \) and \( v = v_0 \) we will obtain

\[ C_{p^{stx}} = c_0 \Psi^{-1}(x_0) + c_1 (\Psi^{-1}(v_0))^n N^{n-1} \]

We label this function as *isoquality per capita cost function* and it is only a function of population and three parameters: the quality of local services \( (x_0 \) and \( v_0) \) and the technology of production parameter \( (\eta) \). To obtain the *isoquality cost function* \( C_{p^{stx}} \) that we referred to in the beginning of this chapter we just multiply \( C_{p^{stx}} \) by \( N \).
which after rearrangements yields:

\[ \frac{dC_{p}}{dN} = -\frac{1}{N^2}[c_0X(1 - \frac{dX}{N}) + c_1V^n(1 - \frac{dV}{N})] \]  

(4.22)

This expression can be simplified using the following expressions for the elasticities:

\[ \epsilon_{x,n} = \frac{dX}{X} \]

and

\[ \epsilon_{v,n} = \frac{dV}{V} \]

By definition the budget share for \( X \) is:

\[ \alpha = \frac{c_0X}{C} \]

and the budget share for \( V \) is:

\[ (1 - \alpha) = \frac{c_1V^n}{C} \]

Therefore, the expressions for the elasticities can be immediately introduced in equation 4.22 and those for the budget shares can also be put in the same equation after dividing the expression within brackets by \( C \) and pre-multiplying by \( C \).

Thus, the intermediate result is,

\[ \frac{dC_{p}}{dN} = -\frac{C}{N^2}[\alpha(1 - \epsilon_{x}) + (1 - \alpha)[(1 - \eta \epsilon_{v})]] \]

(4.23)

Since \( \frac{X}{N} = \frac{V}{N} \) we can solve for the elasticity of per capita costs with respect to population and obtain the fundamental result:

\[ \epsilon_{x,n} = -\alpha(1 - \epsilon_{x,n}) - (1 - \alpha)(1 - \eta \epsilon_{v,n}) \]

(4.24)

This result, is the essential result regarding economies of community size in the model we are considering here. The change in per capita costs (in percentage terms) with respect to changes in the population size of communities is dependent on (i) the economies of scale in producing the capital intensive good \( V (\eta) \), (ii) the structure of the local government budget \( (\alpha, 1 - \alpha) \) and (iii) changes in local public goods crowding \( (\epsilon_{x,n}, \epsilon_{v,n}) \).
In respect to (iii) it is worth clarifying the meaning of $\epsilon_{z,n}$ and $\epsilon_{v,n}$.

If these elasticities are identical to one this means that per capita provision is constant. Thus, the density of utilization will remain constant and the standards of provision will be similar across communities.

In fact, keeping standards of quality constant ($z_0 = \Psi(z_0)$ and $v_0 = \Psi(s)$) is tantamount to keeping $z_0$ and $s_0$ constant since $\Psi_s > 0$.\footnote{When $x$, $s$ and $z$ have subscripts this denotes that they are no longer variables and become parameters.}

This leads us to the second result which can be obtained by substituting in equation 4.24 $\epsilon_{z,n}$ and $\epsilon_{v,n}$ both equal to one:

$$\epsilon_{z,n}^{z_0} = -(1 - \alpha)(1 - \eta) \quad (4.25)$$

The superscript $^{z_0}$ indicates that this is the elasticity of per capita cost regarding population size when the quality of services do not change.

Therefore, this equation shows that the elasticity of the isoquality per capita costs with respect to the population size of communities is negative under increasing returns to scale ($1 - \eta > 0$ and $1 - \alpha > 0$) and positive for decreasing returns to scale ($1 - \eta < 0$). Moreover, the greater the share of labour intensive local services within local budgets ($\alpha$) the smaller the effect of changes in population size on per capita costs.

If we assume, as is usual in the literature, that per capita expenditure is a proxy of per capita cost, the elasticity in equation 4.25 can be considered as the population elasticity of per capita expenditure controlling for the quality of services.

For purposes of empirical analysis developed in chapter 5 it is useful to have a range of plausible values for this elasticity. The $\alpha$ parameter can be obtained from available data. However, the same does not apply for $\eta$ which is a parameter indicating returns to scale in the production of capital intensive local public goods. The discussion in chapter 3 suggested that economies of scale are likely in the production of these goods although a precise quantification is difficult (if not impossible).

Therefore, the maximum value for $\eta$ should be considered as $\eta = 1$
which is the case of constant returns to scale in production. A reasonable minimum for this parameter is \( \eta = 0.5 \). In this case the cost function for \( V \) would be given by:

\[
C(V) = c_1 \sqrt{V}
\]

This would mean that total costs of \( V \) would increase proportionally to the square root of output. This would represent a very considerable degree of economies of scale and values below this are not plausible.\(^25\)

Under these assumptions we would expect that the population elasticity of per capita expenditures (with output homogeneous with respect to quality) be given by:

\[
-0.5 (1 - \alpha) < \epsilon_{p,n}^{\text{eq}} < 0
\]

It is possible to compute the effective value of the elasticity \( \epsilon_{p,n} \) from available data and to derive conclusions regarding the relative (post-grant) fiscal position of communities.

If,

\[
\epsilon_{p,n} < -0.5 (1 - \alpha)
\]

more populated communities are in a worse fiscal position than small/medium size communities, and the reverse happens when

\[
\epsilon_{p,n} > 0.
\]

When \( \epsilon_{p,n} \) falls into the plausible range for \( \epsilon_{p,n}^{\text{eq}} \), i.e.:

\[
-0.5 (1 - \alpha) < \epsilon_{p,n}^{\text{eq}} < 0
\]

no serious discrimination exists between communities with respect to their (post-grant) fiscal position.

Within centralized systems of government, central governments can define the real (post-grant) fiscal position of communities. Therefore we can

\(^{25}\)Note that these are maximum and minimum values for the parameter. We would not expect that the true value of the parameter lies in their neighbourhoods.
also approach the issue just discussed above in terms of intergovernmental grants.

The per capita grants necessary to enable jurisdictions to potentially provide the same quality of services can be found by rearranging the community budget constraint equation in order to obtain $g$:

$$ g = c_0 z + c_1 s^n N^{n-1} - tb $$  

Assuming that the per capita tax base $(b)$ is independent of population leads to the conclusion that:

$$ \frac{\partial g}{\partial N} = (\eta - 1)c_1 s^n N^{n-2} $$  

This derivative is less than zero $(\frac{\partial g}{\partial N} < 0)$ since $\eta - 1 < 0$ and the second derivative regarding population is positive:

$$ \frac{\partial^2 g}{\partial N^2} = (\eta - 2)(\eta - 1)c_1 s^n N^{n-3} $$  

Therefore, to keep constant the quality of services provided by communities, per capita grants should decrease with the size of communities. This decrease is only needed to offset the economies of scale in production of the LPG $V$. The higher is the equal standard of quality (of $V$) the greater is the decrease in per capita grants (when population increases).

A comparison of lump-sum grants necessary to put communities in a similar (post-grant) fiscal position with the actual intergovernmental grants' design will enable an understanding of communities' ability to provide local services within centralized systems of government.

In this context it is useful to have a positive approach to intergovernmental lump-sum grants.

### 4.3.3 Towards a positive approach of lump-sum grants

The intergovernmental lump-sum grants will be approached as the result of an implicit or explicit bargaining process between central and local governments.
There are two main issues at stake: the determination of the size of the "cake", i.e. total grants to lower level jurisdictions and the distribution of total grants between jurisdictions.

The total amount of intergovernmental grants may increase due to an increase in general taxation or due to a decrease in central governments' (post-grants) resources, other things being equal. In the former case there is an increase in the overall size of the public sector while in the latter it remains constant.

The general problem of the distribution of resources between tiers of government will be labelled the decentralization issue. We will reserve the expression pure decentralization for the particular case where overall taxation remains constant.

Finally when the issue is the distribution of grants across communities, keeping the overall amount of grants constant we will refer to the pure distribution problem.

In the bargaining process on the decentralization issue local governments are usually starving for extra funds and lobby for the increase in total grants (G) but seldom, if ever, is there a proposal to change the distributional shares between communities.\(^\text{36}\)

The reason for this is simple. The first "game" is a positive sum game for local governments while the second one is a zero sum game, since in the latter what some jurisdictions win is just offset by the losses of the others. Therefore, unanimity is possible and probable in the first case\(^\text{37}\) while it is most unlikely in the second.

From the point of view of central government's macroeconomic objectives, the redistribution game is almost innocuous if it does not change the total amount of grants.

\(^\text{36}\)This is based on Portuguese experience. We believe, however, that it is a more general situation since it has a game theory rationale behind it (see below).

\(^\text{37}\)It is not guaranteed that unanimity will prevail for the same reason as in an Edgeworth box type situation the contract curve may not be reached because there might not be an agreement on the shares of the gains from trade.
On the other hand restraining overall public expenditure is on the agenda of most developed and developing countries. Therefore, containing the amount of intergovernmental grants is clearly an objective of central government's policies while the way these grants are distributed seems a second priority.

In periods of economic growth, and due to the relatively elastic nature of fiscal revenues in relation to GDP, total grants have a tendency to increase not only in real terms but also as a proportion of GDP. In recession, faced with shortening resources, central governments will try to reduce grants.

As a conclusion it is predictable that central government is willing to satisfy local governments' preferences regarding the distribution game (even if not revealed for the reasons stated above) but unwilling to make much concessions regarding the decentralization issue.

The nature of the redistributive game has to be clarified and also the meaning of a "self-interest" community. Each community is assumed to want to maximize the amount of grants received and therefore the share it has in total grants.

However, it is reasonable to assume that similar communities are treated alike. Therefore redistribution does not go to a particular community but to communities with similar characteristics.

One of the more important characteristics that discriminates communities is community size. Populous jurisdictions are urban or suburban while low populated communities are usually rural, with distinct socio-economic and productive patterns.

A way of formalizing the distribution problem is to consider that total grants for \( k \) communities are given by:

\[
\bar{G} = \sum_{i=1}^{k} G^i
\]  

(4.30)

and grants for each jurisdiction are:

\[
G^i = AN^{ii \sigma + 1}
\]  

(4.31)
which in per capita terms is:

\[ g^* = AN^{\mu} \]

(4.32)

where \( \mu \) is the distribution parameter. If \( \mu = 0 \) all jurisdictions receive the same amount of per capita grants \( A \).

Therefore, each lump-sum intergovernmental grants' scheme can be determined by the total amount of grants \( G \) and the distributional parameter \( \mu \) when the size distribution of the communities is given.

Empirical analysis in several countries shows that in general the population hierarchy of cities follows a "Pareto" distribution given by:

\[ N^* = \left( \frac{1}{D} \right)^{\beta} \]

(4.33)

where \( i \) is the rank of the community when communities are ordered by decreasing population size and \( D \) and \( \beta \) are parameters to be estimated.

In almost every study of nontruncated hierarchies of communities \( \beta \) is close to minus one. Some studies show that \( \beta \) is significantly different from minus one and others that it is not.

For purposes of the development of the theory it is convenient to consider that \( \beta = -1 \), so that we have got the rank size rule:

\[ N^i = \frac{D'}{i} \]

(4.34)

where \( D' = N^1 \) i.e. the population of the largest community.

Under the rank size rule, the size of each community is given by the ratio of the population of the largest community divided by the rank of the community.

This enables us to calculate the size of the median rank community \( m \) which is defined as being the one that occupies the median position within the population hierarchy.

Assuming for simplicity an odd number of communities \( k \) we have:

\[ m = \frac{k + 1}{2} \]
Introducing this expression in equation 4.34 we obtain:

\[ N^m = \frac{2}{k+1} N^1 \]  \hspace{1cm} (4.35)

This result is interesting and shows that the median rank size community is a relatively small community. If for example a country has 199 communities the medium rank community will have \( \frac{1}{100} \) of the largest community population.

Moreover, we can obtain an expression for the total amount of grants introducing 4.34 in 4.31 and the result in 4.30, so that:

\[ \bar{G} = AN^1 \mu+1 \sum_{i=1}^{k} \frac{1}{i^{\mu+1}} \]  \hspace{1cm} (4.36)

Therefore total grants change with the distribution parameter \( \mu \), given \( A \) and \( N^1 \).

It is interesting to analyse what happens to total grants when \( \mu \) changes and particularly when \( \mu \) changes and the grants received by the median rank community are constant.

Keeping constant the per capita grants received by the median size community, if \( \mu < 0 \) smaller jurisdictions will be better off and bigger jurisdictions worse off.

This can be illustrated in figure 4.5 where total grants are measured on the Y axis and the k communities of a country are ranked according to (decreasing) population size on the X axis.

Since with \( \mu = 0 \) total grants per jurisdiction are proportional to population, the curve 1 (with \( \mu = 0 \)) can measure both the rank-size distribution of communities and total grants received by communities when per capita grants across jurisdictions are the same. Total grants (\( \bar{G}_1 \)) is the area under
Consider now the curve 2 (with \( \mu = a < 0 \)) drawn so that the median rank community receives the same amount of grants. It is clear that communities bigger than the median (at the left of \( m \)) are worse off and smaller than the median (at the right of \( m \)) are better off. This area can be approximated by \( \int_{x=1}^{k} \gamma x + \gamma = \ln k + \gamma \), where \( \gamma \) is the Euler's constant. An equation can be written using the "big oh" notation. (See Apostol (1974) section 8.13 on "The Big Oh and Little Oh Notation" example 1.) Using this notation we can write: \( \tilde{G}_1 = A_1 N^1 \left[ \int_{x=1}^{k} \gamma x + \gamma + O(1/n) \right] \), where \( O(1/n) \) is "the big Oh" of \( 1/n \).
communities smaller are slightly better off. Now total grants are given by:\footnote{Again we may use the big oh notation to pass from the summation to the integral. However, in this case the expression would be different. (see Apostol op. cit. example 2).}

\[ G_T = A_T N^{1+\mu} \sum_{i=1}^{k} \frac{1}{i^{\mu+1}} \]  

(4.38)

Total grants \( G_T \) are smaller in this case. The difference \( G_1 - G_2 \) is the difference between the two shaded areas at the left and right of the median community respectively.

To put the issue in pure redistributive form, the surplus \( G_1 - G_2 \) must be allocated uniformly across jurisdictions resulting in an upwards parallel shift of the curve 2. This is illustrated by curve 3. Under the new scheme \( (C_3 = G_1, \mu = a < 0) \) the majority of communities that will be better off under the new scheme is enlarged.

The case of \( \mu > 0 \) is not drawn in order not to overburden the figure. Nevertheless, it can easily be seen that in this case there would be a clockwise rotation of the curve 1, so that all the smaller communities would be worst off. Moreover, the total amount of grants \( G_4 \) would be considerably larger so that a downwards shift of this curve would be necessary to keep total grants \( G \) constant.

Therefore, some conclusions can be drawn. In a pure distribution issue the large majority of (small) communities will support a scheme of distribution of per capita lump-sum grants that is regressive towards the population size of communities (i.e. \( \mu < 0 \)). This scheme is supported by all small communities and a considerable range of medium size communities.

On the other hand, central government can decrease the overall amount of grants keeping constant the per capita grants of the median size community and improving the situation of all the smaller communities.
4.3.4 Equalization versus Political Hypotheses on Intergovernmental Lump-Sum Grants

The last two sections introduced two distinct rationales for lump-sum intergovernmental grants. The first, considers a benevolent central government which wants to equalize the provision of local services across jurisdictions of a country and uses intergovernmental grants to offset differentials in fiscal capacity and/or fiscal position across jurisdictions.

The second approach envisages central government wanting mainly to keep overall public expenditure under control and therefore managing transfers of resources to lower level jurisdictions in order to enlarge a qualified majority of (small) communities supporting its proposals.

Three hypotheses were considered.

The first (equalization) hypothesis assumes that central government wants to equalize the fiscal capacity of jurisdictions and can be given by equation 4.15. If we add an error term we have:

\[ g^i = \gamma_0 + \gamma_1 b^i + \varepsilon^i \]  

and predict that \( \gamma_0 > 0 \) and \( \gamma_1 < 0 \).

The second (equalization) hypothesis can be tested with the help of equation 4.27. Therefore, the following relationship can be estimated:

\[ g^i = \alpha_0 + \alpha_1 N^{\alpha_2} + \alpha_3 b^i + \varepsilon^i \]  

where it is predicted that \( \alpha_0 > 0, \alpha_1 > 0, -0.5 < \alpha_2 < 0 \) and finally \( \alpha_3 < 0 \).

The third (politico-economic) hypothesis can also be tested using a generalization of equation 4.32:

\[ g^i = \beta_0 b^{\beta_1} N^{\beta_2} \varepsilon^i \]  

30 As clarified above under the (median voter model) assumption of constant returns to scale in production, this hypothesis is equivalent to equalize the fiscal position of communities.
Here, it will be predicted that $\beta_0 > 0$ and $-1 < \beta_1 < 0$. These predictions summarize the rationale of constraining overall intergovernmental lump sum grants.

There are no restrictions for $\beta_1$. If $\beta_1 = 0$ (not significantly different from zero), we have again equation 4.41. If $\beta_1 > 0$ there is also a revenue sharing aim associated with intergovernmental grants. On the other hand if $\beta_1 < 0$ there is also a fiscal equalization objective associated with grants' design.

The fact that both population parameters ($\alpha_2, \alpha_3$) are predicted to be negative may introduce some problems in discriminating between the hypotheses 2 and 3. It is not possible to test the identity of the two parameters since equation (4.41) is estimated using ordinary least squares after applying logarithms to both members of the equation while equation (4.40) is estimated using non linear least squares.\(^{31}\)

If the equalization hypothesis is not rejected, and the political hypothesis is, the comparison between the estimates is not particularly relevant. However, in the contrary case, the test would be very useful since if $\beta_2 < \alpha_2$ it would mean that when communities increase in size, what a jurisdiction gets, in per capita terms, is lower than necessary to keep similar standards of service.

A corollary of (i) is that quality of services will deteriorate in more populated communities, and a manifestation of this is increased congestion of local public goods provision.

### 4.4 Conclusion

This chapter introduced a positive approach to economies of community size within centralized governments. The distinction between centralized and decentralized governments and the concept of economies of community size were clarified.

\(^{31}\)However, the expenditure analysis will enable us to discriminate the two hypothesis.
Centralized governments have been defined in this thesis as fiscal structures where local governments' revenues are mainly exogenous. This is due to the joint effect of intergovernmental grants and local tax rates set by central government. Thus, central government has a broader influence on the quality of local services provided than would be the case within a decentralized fiscal structure.

In opposition to the median voter model, under a centralized regime, citizens pay taxes to finance local goods the level and quality of which is determined elsewhere. It is not the maximization of the utility function of a representative voter that determines levels of public sector output. On the contrary, the consumer budget shares between private and local public consumption are set by central government. There is no room for choice within the jurisdiction. However, as pointed out in section 2.5 the centralized model has some affinity with the Tiebout model since citizens may reveal their preferences by “voting with their feet”. Yet, this is likely to be associated (as in the Tiebout model) with distortions since the preferences for a community on the grounds of the fiscal “package” it offers may not coincide with job preferences.

Turning now to the economies of community size, we saw that they are mainly dependent on the existence of economies of scale in the provision of local public goods and on economies of “crowding”.

A theoretical approach to crowding developed the argument that (apart from small communities) there are no economies of “crowding”. In other words, local public goods should be considered as having “privateness” characteristics.

The concept of crowding was distinguished from the concept of congestion because the latter applies to joint consumption while the former applies to the relationship between aggregate capacity of LPG and aggregate consumption.

The problem of economies of scale was not developed any further than
the existing few approaches in the literature suggest. It was assumed that local governments provide labour intensive services at constant returns to scale and capital intensive services at increasing returns.

The combined effect of the existence of no economies of "crowding" and economies of scale in capital intensive services is that there are modest economies of community size.

On the other hand a positive approach to intergovernmental grants suggests that per capita lump-sum grants decrease with the population size of communities independently of any fiscal equalizing factor.

Since local tax revenues are assumed mainly exogenous to local municipalities within centralized systems of government, we are able to predict how local revenues will change with the size of the jurisdictions.

The analysis of economies of community size and revenues enable us to derive some hypotheses that will be the subject of empirical analysis in the chapters 5 and 6 below.
Chapter 5

Anatomy of a Fiscally Centralized Country: Portugal

This chapter analyses the case of Portugal, and tests the hypotheses developed in the last chapter, which apply to fiscally centralized countries.

Fiscally centralized countries were defined by the essential characteristic of local governments being unable or severely limited in their ability of raising their revenues at their own discretion.

Therefore, an inquiry into the degree of centralization has to consider the institutional and legal frameworks that shape local governments' spending and revenue responsibilities and on the other hand an analysis into the actual structure of revenues of local governments. These two aspects will be considered in section 5.2 below.

Beforehand, the Portuguese case will be contextualized within European countries members of OECD (section 5.1), using more traditional measures of fiscal centralization. Using these traditional indicators Portugal appears clearly as one of the more (if not the most) centralized of European countries.

The analysis of local governments' revenues will be addressed in sections 5.2 and 5.3. The former clarifies the main sources of Portuguese local governments' revenues and why the Portuguese legal framework implies that local tax yields are mainly exogenous to local governments in
the short run. The latter analyses empirically the hypotheses developed in
the last chapter regarding intergovernment grants, fiscal equalization and
community size.

Section 5.4 puts together the conclusions of the past two sections on
the exogeneity of revenues with the conclusions of chapter 4 regarding the
economies of community size. The main implications will be addressed as
will the reasons why urban communities will tend to undersupply capital
stock under the present legal framework and therefore to provide local
public goods with increased congestion. Finally, section 5.5 concludes.

5.1 Fiscal Centralization in European (OECD)
Countries

When dealing with cross-country comparisons of the degree of fiscal cen­
tralization, economists usually use some ratio of central government ex­
penditures (and/or revenues) to overall government expenditures (and/or
revenues) . This is a straightforward way of dealing with the problem but
faces several methodological problems, some of them are highlighted below.

Firstly, there is a problem of scale. European countries differ largely on
dimension and population size. To put this diversity into perspective, the
average area of a Lande in Germany (West) is roughly seven times the size
of Luxembourg and half the size of Netherlands and of Denmark. On the
other hand the average population size of a Lande (5605) is larger than the
population of Denmark (5124), Finland (4788), Ireland (3440), Luxembourg
(364) or Norway (4092). Therefore, centralization means different things
for different countries.

Secondly, there are differences in the administrative and political struc­
ture of the different states. The majority of countries are unitary countries

1See for instance Oates (1972).
2Figures cited (in thousands) come from Karran, T. "Local Taxing and Local Spending:
with two relevant decision-making tiers of government (central and local), having also most of them an intermediate level of government with relatively minor functions and responsibilities. However, others have a federal or confederal structure (Germany, Austria and Switzerland) or strong regional governments (Spain with Catalonia and the Basque country).

Thirdly, there are accountability problems. Within the countries that have intermediate levels of government, only Austria, Germany and Switzerland publish provincial accounts. In the other countries where that tier of government exists, provincial or regional accounts are consolidated into local governments accounts. Therefore, figures are biased towards decentralization, since the concept of local governments is much broader. Moreover, national accounts published in each country produce social security accounts consolidated into central government accounts. Therefore the way OECD constructs the social security account is not clear from their published data and may also introduce distortions when interpreting the ratio as we will see shortly.

Finally, there are problems related to the degree of central governments' control over local governments' "own" expenditures and revenues.

Local governments' expenditures on behalf of central government mislead the interpretation of data on expenditures by level of government. This is particularly important when considering which level of government is responsible for payments to employees in the educational sector. As shown below this has a very important impact on the centralization ratios.

In what concerns revenues, only a careful analysis of local governments' legal and institutional frameworks would allow a better understanding of the real degrees of autonomy in raising own revenues.

Having these provisos in mind it is possible to turn now to the national accounts data, for European-OECD countries.

---

3The U.K. is an exception with only two tiers of government, and Portugal only recently created the Lisbon and Porto Metropolitan Areas.
4Turkey is not included since the OECD does not publish their national accounts. Luxembourg and Iceland should both be omitted according to the scale problem referred
Table 5.1: Final Consumption Expenditures by levels of Government in OECD-European countries (1985)

Table 5.1 shows the ratio of final consumption expenditures by levels of government to general government expenditures. Social security expenditures are considered separately.

Final consumption expenditures seems a good indicator to analyse centralization since it excludes intergovernmental transfers and interest debt. Taking into consideration the high weight of social security expenditures in countries such as Austria, Germany, Spain and France (which might reflect an accountability problem) it is worth considering central government expenditures without (column 2) and with (column 6) social security expenditures.

According to the final consumption ratio (column 2) Portugal is the most centralized of European countries (OECD members) with 88% of government expenditures. We keep data on Luxembourg (given EEC membership) but we will not discuss it.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>17.07%</td>
<td>66.66%</td>
<td>27.01%</td>
<td>6.33%</td>
<td>72.99%</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>25.28%</td>
<td>31.14%</td>
<td>68.15%</td>
<td>0.71%</td>
<td>31.85%</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>19.37%</td>
<td>57.95%</td>
<td>21.96%</td>
<td>20.10%</td>
<td>78.04%</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>20.06%</td>
<td>18.87%</td>
<td>17.81%</td>
<td>49.00%</td>
<td>51.06%</td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td>20.40%</td>
<td>70.56%</td>
<td>19.90%</td>
<td>9.51%</td>
<td>80.07%</td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>18.56%</td>
<td>49.65%</td>
<td>48.62%</td>
<td>1.73%</td>
<td>51.38%</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>16.44%</td>
<td>54.37%</td>
<td>42.95%</td>
<td>2.68%</td>
<td>57.05%</td>
<td></td>
</tr>
<tr>
<td>Luxembourg</td>
<td>15.74%</td>
<td>66.10%</td>
<td>28.14%</td>
<td>5.76%</td>
<td>71.86%</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>18.18%</td>
<td>46.19%</td>
<td>49.19%</td>
<td>4.61%</td>
<td>50.81%</td>
<td></td>
</tr>
<tr>
<td>Portugal</td>
<td>15.52%</td>
<td>87.62%</td>
<td>8.80%</td>
<td>3.57%</td>
<td>91.20%</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>14.72%</td>
<td>46.97%</td>
<td>32.00%</td>
<td>21.03%</td>
<td>67.99%</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>20.80%</td>
<td>60.92%</td>
<td>37.82%</td>
<td>1.26%</td>
<td>62.18%</td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>18.91%</td>
<td>37.48%</td>
<td>19.36%</td>
<td>41.36%</td>
<td>21.16%</td>
<td>58.64%</td>
</tr>
<tr>
<td>Finland</td>
<td>20.36%</td>
<td>30.30%</td>
<td>66.43%</td>
<td>3.27%</td>
<td>33.57%</td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>43.48%</td>
<td>37.26%</td>
<td>61.32%</td>
<td>1.41%</td>
<td>38.68%</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>27.62%</td>
<td>27.17%</td>
<td>71.50%</td>
<td>1.33%</td>
<td>28.50%</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>13.35%</td>
<td>24.23%</td>
<td>30.85%</td>
<td>71.98%</td>
<td>28.02%</td>
<td></td>
</tr>
</tbody>
</table>

Source: Calculations based on OECD's National Accounts 1978-1990, Paris
Note: GG, CG, LG, PG mean general, central, local and provincial governments respectively.
S.S. stands for social security which is a separate account.
ernment final consumption expenditures made by central government. It is followed by Greece (71%), Belgium (67%) and the United Kingdom (61%), all of them with more than 60% of government expenditures undertaken by central government. If social security contributions are included France and Spain would have to be added to the list. On the other hand Scandinavian countries, with Germany and Austria form the group of the more decentralized ones.

Some corrections can be made to this data and one important one is expenditures on education, since it is often the major single item of local governments' expenditure. In most countries the wage rates of educational employees (basically teachers, administrative staff and other school employees) are set by central government. However, the authority which actually pays the bill is in some cases the central government and in others the local government. Whenever rates are set by central government, these expenditures should be accounted for at central government level, as far as centralization is concerned. Even if local governments have the ability to set levels of employment and qualifications of staff there are many regulations that constraint this autonomy.

To understand better the importance of this correction consider the U.K. case. This correction being made for the U.K. would increase the weight of central government from 60.9% to 77.9% and decrease the share of local governments from 37.8% to 20.2%.

Considering gross investment expenditures Table 5.2 shows the responsibilities of the different tiers of government.

Portugal and Belgium are the only countries where central government outweighs (and almost double) local governments' role in investment. The United Kingdom occupies an intermediate position where the share of investment responsibilities between central and local governments is almost

---

5The correction was made taking local educational wages out of local government current consumption expenditure and including them in central government expenditures.
Table 5.2: Gross Fixed Capital Formation by levels of Government (1985)

equal, while the majority of countries have greater responsibilities at a "local" level.

It should be referred again that some regional responsibilities on investment (e.g., highways), in the absence of an intermediate level of government, are accounted for at the central government level. On the other hand, where provincial governments exist, they are accounted for at a local level.

Turning now to the revenue side, Figure 5.1 is based on panel data indicating the ratio of current intergovernmental transfers to local governments' current receipts. *Ceteris paribus*, which in this case means essentially keeping constant the share of local government expenditures, the greater the share of intergovernment grants the greater the degree of centralization.

---

*See Appendix A.1.*
Figure 5.1: Ratio of Intergovernment Grants on Local Governments' Receipts in OECD-European countries
Cross-section analysis using figures for 1989\(^7\) shows that there are essentially three groups of countries. Those where the centralization ratio is higher than 70\% (Netherlands, Ireland, Italy, Greece). Those where intergovernmental grants amount to 40-70\% of local current revenues (Belgium, Denmark, Finland, Norway and Spain). Finally, countries where the proportion of grants is lower than 40\% (the other countries).

Federal countries (Germany, Austria and Switzerland) are among the less centralised countries with only a third or less of local revenues financed through higher levels of government.

From a time-series perspective different time patterns emerge. Countries with a very low share of intergovernmental grants (Austria and Switzerland) have been increasing this share during the 1978-1990 period. Countries with moderately high shares in the beginning of the period (Belgium, Denmark, Greece and the United Kingdom) have been decreasing or stable in the period, with the exception of Greece.\(^8\) Finally countries with high shares in the beginning of the period keep high levels of intergovernment grants (Ireland, Italy and the Netherlands).

These figures show the diversity of situations across European countries. The degree of (de)centralization of a country, however measured, is a consequence of a long historic process peculiar to each country.

Yet, the particular choice of the kinds of taxes and the size and sort

\(^7\)It is the last year where there is information available for all countries except Portugal. In this case we used local governments’ budgets to calculate the ratios for 1987 until 1990. These ratios are not directly comparable with the rest of the series since the information source is different. We believe that estimations based on budget accounts under estimate the ratio.

\(^8\)Greece does not have any major tax such has income or property tax the revenues of which are directed to local governments. Therefore an increase on local expenditure has been financed through increased intergovernment grants. The U.K. faces a displacement effect in 1990 due to the introduction of the poll tax in England and Wales in that year. Community charge yields were significatively lower than previous rate yields and this lead to the introduction of new grants (“national non-domestic rates distribution”) which have the result of increasing the centralisation of revenues.
of tax bases that are used to finance local governments are critical in explaining the moves towards (de)centralization. In particular it is the income elasticity of local taxes in relation to national taxes that commands changes in centralization.

Countries that base their local revenues on property taxes (or other sources of revenue related to the property tax base) will tend to increase their degree of revenue centralization over time since they are less elastic than income taxes or VAT used to finance central governments.

Revenue centralization, measured by the ratio of intergovernmental grants to local revenues, is a weak indicator of the centralization of systems of government for two main reasons. Firstly, because it does not take into account the relative value of local governments' revenues on overall public sector revenues. Secondly, it does not clarify whether local governments are able (or not) to change local tax rates at their own discretion.

The Portuguese situation is defined by a very low ratio of local revenues on overall public sector revenues and also by a great control of central government on local governments' tax rates as we will see below.

5.2 Fiscal centralization in Portugal: the revenue side

5.2.1 Introduction

The last section highlighted the fact that Portugal is one of the more (if not the most) centralized of European-OECD countries. This section will introduce a better understanding of the Portuguese case looking mainly to the revenue side of local budgets and also to the legal framework that shapes local governments' behaviour.

After the 1974 revolution, which ended almost fifty years of dictatorship and political centralization, three main bills were enacted to set the legal
framework of local governments' revenues. Their aim was to decentralize government fiscal structure, as stated in the preamble of Bill 98/84 which is said to "keep the deeply decentralized aim of the Local Finance Bill [1/79]."

This is not the place to discuss in detail the major changes introduced with these bills.10

<table>
<thead>
<tr>
<th>Year</th>
<th>Taxes</th>
<th>Property Tax</th>
<th>Siza</th>
<th>Other T.</th>
<th>Interg. Trf.</th>
<th>FEF</th>
<th>Other trf</th>
<th>Borrowing</th>
<th>Other Rev.</th>
<th>Tot. Rev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td>14.05%</td>
<td>8.37%</td>
<td>4.14%</td>
<td>5.70%</td>
<td>59.53%</td>
<td>50.49%</td>
<td>9.04%</td>
<td>11.14%</td>
<td>15.28%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1982</td>
<td>13.33%</td>
<td>8.52%</td>
<td>5.11%</td>
<td>6.21%</td>
<td>57.16%</td>
<td>51.71%</td>
<td>5.44%</td>
<td>11.98%</td>
<td>17.53%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1986</td>
<td>17.66%</td>
<td>9.99%</td>
<td>4.04%</td>
<td>5.65%</td>
<td>62.94%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>3.60%</td>
<td>15.60%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1987</td>
<td>25.09%</td>
<td>5.11%</td>
<td>4.04%</td>
<td>5.65%</td>
<td>56.26%</td>
<td>47.22%</td>
<td>9.04%</td>
<td>2.90%</td>
<td>15.76%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1988</td>
<td>29.82%</td>
<td>6.21%</td>
<td>4.04%</td>
<td>5.65%</td>
<td>49.23%</td>
<td>38.62%</td>
<td>10.61%</td>
<td>4.17%</td>
<td>16.78%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1989</td>
<td>28.81%</td>
<td>6.21%</td>
<td>4.04%</td>
<td>5.65%</td>
<td>48.25%</td>
<td>36.98%</td>
<td>11.27%</td>
<td>5.73%</td>
<td>17.21%</td>
<td>100.00%</td>
</tr>
<tr>
<td>1990</td>
<td>30.15%</td>
<td>5.67%</td>
<td>5.67%</td>
<td>5.67%</td>
<td>47.75%</td>
<td>36.11%</td>
<td>11.64%</td>
<td>5.52%</td>
<td>16.59%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>


* Tax on the purchase of housing **Surcharges on the property tax for capital improvements

Table 5.3: The Structure of Revenues of Portuguese Local Governments

Figure 5.3 shows the evolution of the structure of revenues of Portuguese local governments in recent years.

After 1986 the role of grants has been decreasing. This is explained by the allocation to local governments of an already existent excise tax on the sale of housing11 and an increase in borrowing and other revenues.

For 1990 12 the great part of local revenues comes predominantly from

9 The bills which will be labelled Bill 1/79, Bill 98/84 and Bill 1/87 are: Lei 1/79 de 2 de Janeiro; Decreto-Lei 98/84 de 29 de Marco; Lei 1/87 de 6 de Janeiro. Also relevant is the Bill 442-C/88 (Decreto-Lei 442-C/88; Código da Contribuição Autárquica) which sets the property tax framework.
10 For more details see in particular bill 1/87 and also Mozzicafredo et al. (1988), Porto (1990) and Santos (1989).
11 The Siza became a local revenue after 1987 (inclusive). See Bill 1/87.
12 At a final stage of the elaboration of this dissertation data for 1990 was published. It is considered here, but cross-section analysis will use figures for 1989. However, the legal framework is still the same and therefore there are no significant changes either on the
two main sources: local taxes and intergovernment (and other) grants.

Tax yields come essentially from three sources: a property tax (26.2%) an excise tax on housing (the Siza) (38.2%) and discretionary surcharges on the property tax (the derramas) levied for purposes of financing capital projects (16.8%). Together, these three taxes represent 81% of total taxes.

It will be shown below why local tax revenues are mainly exogenous to local governments in the short run when the tax base is assumed to be given.

5.2.2 Local Tax Revenues within Centralized Governments

It is worth being more precise about the concept of fiscal centralization. A necessary condition for a centralized government is that each local tax revenue is a linear function of the respective tax base. Moreover, the tax base should be the essential explanatory variable of the level of local taxes. In fact, this means that implicit tax rates are uniform across communities.

This should be contrasted with the situation in decentralized governments where local governments are able to manipulate their tax rates at their own discretion, and therefore demand factors might be in place in explaining the level of local taxes.

Property tax rates up to 1989 were uniform across communities. From 1989 onwards very limited discretion was introduced on these rates (Bill 442-C/88). Local governments were given the choice to set tax rates between 1.1 and 1.3% of the assessed values of urban property. A uniform rate of 0.8% was set for rural property. However, the rates of the excise tax on housing (Siza) are defined by central government. Finally, concerning the derramas, they are also related (although less stringently) with the property tax base since the derramas can not exceed 10% of property tax yields.

distribution of grants or the structure of tax revenues.
Under this legal framework it is not surprising that local taxes are almost entirely linearly related to the property tax base. Local governments have very limited discretion on their ability to raise revenues from taxes other than those allowed by their tax bases.

To clarify this point per capita local tax revenues \( T \) were regressed against the per capita tax base \( b \) for all Portuguese municipalities. The variables were considered in per capita terms to avoid heteroscedasticity since the variance of the error terms usually increases with the size of the community.

Per capita tax revenues include all taxes (direct and indirect) collected by local governments.

Since there is no direct measure of the property tax base it was calculated through the property tax yields, according to the expression:

\[
\text{TAXBASE} = \frac{\text{Rural Property Tax}}{0.008} + \frac{\text{Urban Property Tax}}{0.012}
\]

This process is relatively rigorous since the rural property tax rate is fixed at .008 and the urban property tax rate can change just between 1.1\% and 1.3\% , being the average value adopted.\(^\text{13}\)

Therefore the following equation was estimated using OLS\(^\text{14}\):

\[
T_i = \alpha_1 + \alpha_2 b_i + \epsilon_i \tag{5.1}
\]

With \( T_i \) and \( b_i \) defined above, and \( \epsilon_i \) a random error term.

Results are the following (t-statistics in parenthesis)\(^\text{15}\)

\[
\hat{T}_i = -1.1546 + 0.048373 b_i \quad R^2 = .86 \quad N = 272
\]

\[\begin{array}{cc}
(-5.4082) & (41.4998)
\end{array}\]

\(^\text{13}\)There remains the problem of the reliability of the assessment of the property values. This problem would still exist if there were direct observations on property values. Note, however, that if the property values were uniformly underestimated this would have an implication for the intercept estimator, but not in the slope estimator.

\(^\text{14}\)Ordinary Least Squares

\(^\text{15}\)Full econometric results can be found in Appendix A.2 . Herein after, figures in parentheses below the coefficients are always the t-statistics.
This result shows the importance of the property tax base as a major factor explaining municipal tax yields. Since communities are constrained in their ability to manipulate tax rates, their degree of discretion in setting their own tax revenues is rather limited.

It might be argued that the explanatory power of the per capita tax base, is a consequence of the way the proxy to the community tax base was constructed. Property taxes, as clarified above, only account for a quarter of local taxes in Portugal. Anyway, to confirm that the relationship holds, independently of the property tax, we regressed per capita local taxes excluding the property tax rate \( T_{h} \) on the per capita tax base.

\[
T_{hl} = \alpha_1 + \alpha_2 b_i + \varepsilon_i \quad (5.2)
\]

The results are very similar:\(^{16}\)

\[
T_{hl} = -1.0586 + 0.036577 b_i \quad R^2 = .79 \quad N = 272
\]

\[
(-5.0795) \quad (32.1443)
\]

The next section addresses the issue of intergovernmental grants, namely whether these grants are aimed at equalizing fiscal capacity, fiscal position or are explained by other politico-economic factors as developed in chapter 4.

5.3 Intergovernmental Grants, Fiscal Equalization and Community Size

After clarifying the importance of intergovernmental grants on local governments' revenues it is now time to analyse empirically the propositions stated in section 4.3.4 above.

A positive analysis of intergovernmental grants should not accept at face value the normative criteria used by decision-makers (in this case central

\(^{16}\)See full results in Appendix A.2.
government) to design a particular distribution of grants. On the contrary, it is the purpose of a positive analysis to submit to empirical scrutiny central governments' stated rationales and go behind them searching for alternative explanations for a specific grant design.

Section 4.3.4 introduced three rationales for intergovernmental grants. Equalization of fiscal capacity of jurisdictions, equalization of fiscal position and maximization of political support from communities of a redistributio­

Before testing these hypotheses, a short clarification of the Portuguese intergovernmental grants scheme is in order.

After 1984, central government's grants were consolidated in a Fund for Financial Balance (FEP$^{17}$). The stated objectives of this Fund were to contribute to offset fiscal disparities across communities, both in terms of the fiscal capacity and the fiscal position ("needs") of communities. The size of the Fund was established as a percentage of the general government expenditures (Bill 84) set annually in the public sector budget. Later on (Bill 1/87) it was a percentage of the VAT yields. The distribution of the Fund across communities was set according to some criteria which supposedly are indicators of needs and fiscal capacity of communities. The Fund is also split into current and capital components, so that it should be allocated to current and capital expenditures.

The distributional criteria are rather subjective and have been subject to change in each new legislation without a clear justification of the rationale underlying these changes.

As stated above we will not discuss these indicators, nor the consistency between them and the objectives they are supposed to be related to. This would be the task of a normative approach to fiscal equalization. Our task is to address directly the issue of whether the actual scheme of grants operates to equalize fiscal capacity, fiscal position across jurisdictions (as stated in the legislation) or can be explained by a different rationale.

$^{17}$Fundo de Equilíbrio Financeiro
The empirical application will start with the equalization hypotheses introduced in section 4.3 and 4.3.4 above.

5.3.1 The fiscal equalization hypotheses

The first hypothesis states that central government wants to equalize local governments' per capita revenues. This means to equalize fiscal capacity of jurisdictions and would amount to equalizing fiscal position only under the assumption of constant returns to scale in the production of local public goods.

Let us recall that in centralized governments, real equalization of fiscal revenues of local governments is possible if central governments do aim at this objective. However, this is not the case in decentralized countries where equalization is always potential.

Lump-sum per capita grants to equalize per capita fiscal capacity across jurisdictions can be given by equation 4.15:

\[ g = cz^* - t/b \]

Therefore, the following equation was estimated using OLS:

\[ g^i = \alpha_1 + \alpha_2 b^i + \epsilon_i \]

and the estimated equation is:

\[ \hat{g}^i = 20.2341 - 0.013442b^i \]

\[ R^2 = 0.03 \quad N = 272 \]

\[ (25.3948) \quad (-3.0899) \]

It is sufficient to look at the adjusted \( R^2 \) to see that this relationship is meaningless. In fact, since we are dealing with a two-variable regression model, the \( R^2 \) indicates the proportion of the total variation in per capita grants explained by per capita tax base. A \( R^2 \) close to zero indicates in this case that the regression line is almost horizontal and thus per capita tax base do almost nothing to help explaining per capita grants.\(^{18}\)

\(^{18}\)Note that the intercept is 20.23 and the mean value of the dependent variable is 18.48 (see Appendix A.2).
Therefore, we can easily refute the hypothesis that intergovernmental lump-sum grants are only aimed at equalizing the fiscal capacity of Portuguese jurisdictions.

We can turn now to the second hypothesis that lump-sum grants are aimed at equalizing the fiscal position of communities as defined in chapter 4, i.e. equalizing the ability of communities providing the same quality of local public goods, taking into account their fiscal capacity and the economies of community size.

Equalization of fiscal position is given by equation 4.27:

\[ g = c_0 z + c_1 s^q N^{p_1} - t b \]

Therefore, the following equation was estimated using nonlinear least squares estimation:

\[ g' = \beta_1 + \beta_2 N^{\beta_3} + \beta_4 b' + c' \]  \hspace{1cm} (5.4)

The predicted values for the parameters according to the equalization of fiscal position hypothesis are the following: \( \beta_1 > 0 \), \( \beta_2 > 0 \), and \( \beta_4 < 0 \).

Let us recall from chapter 4 that \( \beta_3 = \eta - 1 \) is a parameter indicating the existence of economies of scale in production of capital intensive goods. Plausible values for \( \beta_3 \) as indicating only economies of scale are in the range \( -5, 0 \).

The results for all mainland Portuguese communities with less than 250000 inhabitants are:

\[ g' = 3.5261 + 88.0749 N^{-.69401} + .003069 b' \ R^2 = .83 \quad N = 272 \]

(2.9093) \hspace{1cm} (17.546) \hspace{1cm} (-13.4188) \hspace{1cm} (1.6339)

Reducing the sample to the communities with more than 10000 inhabitants produces the following estimates:

\[ g' = 1.9366 + 76.924 N^{-.95789} + .003089 b' \ R^2 = .67 \quad N = 186 \]

(.79147) \hspace{1cm} (4.1817) \hspace{1cm} (-4.1603) \hspace{1cm} (1.9627)
These results show that the equation has better explanatory power when all communities are considered, which decreases when the smaller communities (under 10000 inh.) are excluded. The intercept is positive (as predicted) and significantly different from zero in the larger sample but not in the smaller sample.

It is clear that per capita intergovernmental grants have a nonlinear relationship with the population size of communities. $\beta_2$ is positive and $\beta_3$ negative and they are significantly different from zero (at a 99% degree of confidence) in both samples, as predicted.

However, the value of $\beta_3$, is lower than mere economies of scale would suggest, indicating that more populated communities seem to be in a worse position to provide a similar quality of local public goods.

We fail to reject the null hypothesis that $\beta_4 = 0$, and thus conclude that per capita intergovernmental grants do not seem to be a linear function of the per capita tax base.

### 5.3.2 A politico-economic hypothesis on intergovernmental grants

This hypothesis was formulated in section 4.3.4 on the assumption that the non-truncated hierarchy of communities in a country follow the "Pareto" distribution and the rank-size rule is a reasonable approximation to the actual hierarchy of communities.

Therefore, firstly the actual population distribution of Portuguese communities will be considered. Secondly, the actual specification that explains intergovernmental grants will be tested.

In short if $D$ is a parameter, $i$ is community's rank when they are ordered by decreasing population size, $N^i$ is the population size of community (ranked) $i$ and $\hat{\beta}$ is an estimator, the hierarchy of communities follow a "Pareto" distribution if:

$$N^i = \left[ \frac{i}{D} \right]^\frac{1}{\beta}$$

(5.5)

Taking logarithms of both sides of this equation, rearranging and adding
an error term yields:

\[ \ln i = \ln D + \beta \ln N^i + \epsilon^i \]  

(5.6)

Estimating using OLS for all Portuguese communities (population data for 1991) yields:

\[ \hat{\ln i} = 13.6943 - .9229 \ln N^i \quad R^2 = .93 \quad N = 275 \]

(93.6477) (-62.3086)

This result shows that the distribution of population communities follow a Pareto distribution, but that the rank-size rule does not apply strictly. In fact \( \hat{\beta} \) is significantly different from minus one even at 90% degree of confidence. \(^{19}\) However, the argument in section 4.3 was developed on the basis of \( \beta = -1 \) for purposes of analytical tractability of the problem. The actual value of -.92 is close enough to one to keep the argument valid.

Therefore, we turn now to the analysis of the politico-economic hypothesis on intergovernmental grants.

As introduced earlier it can be tested using the following specification:

\[ g^i = \alpha_1 N_i^{\alpha_2} \beta_i^{\alpha_3} \]

which after taking logarithms and adding an error term yields:

\[ \ln g^i = \ln \alpha_1 + \alpha_2 \ln N_i^i + \alpha_3 \ln b^i + \epsilon^i \]  

(5.7)

The main prediction is that \( \alpha_3 < 0. \)\(^{20}\) If additionally there is a revenue-sharing objective \( \alpha_2 > 0. \) Conversely, if there are also objectives of equalization of fiscal capacity, we would expect \( \alpha_3 < 0. \)

The actual estimate of equation 5.7 is:

\[^{19} t_{\beta} = - .9229 - (1) = 5.205 \]

\[^{20} \text{It is necessary to make clear the distinction between the population parameter in equation 5.4 (} \beta_3 \text{), and in here (} \alpha_3 \text{). The former is a parameter related to economies of scale in the production of capital intensive local goods and the latter is not. } \alpha_3 \text{ is a (constant) population elasticity and } \beta_3 \text{ is not.} \]
\[ \ln g^i = 7.4636 - 0.49912 \ln N^i + 0.046149 n_6^i \quad R^2 = 0.86 \quad N = 272 \]

\begin{align*}
(59.6459) & \quad (-37.6148) & \quad (2.2965)
\end{align*}

This seems the best specification of intergovernmental grants and therefore supports the politico-economic hypothesis. Every estimate is statistically significant at a 99% degree of confidence, \( \alpha_2 \) is negative as predicted and the positive value for \( \alpha_3 \) confirms the revenue sharing and not equalization purpose of intergovernmental grants.\(^{21}\)

The elasticity of per capita grants regarding population shows that an increase of 1% in community size decreases per capita grants nearly 0.5%. This means that grants are quite rigid regarding population size. In percentage, the increases in total grants are almost half the increases in population.\(^{22}\)

This section and the last one reviewed empirical evidence that Portugal has a centralized government since local revenues are mainly determined by central government directly through lump-sum intergovernmental grants and indirectly through the legal framework that sets the main tax rates.

There is some degree of discretion but it is rather limited. Local governments can introduce user charges on some services, set the level of the derramas within the limits imposed by law and also make use of borrowing.

A consequence of the degree of discretion over local resources being limited is that local expenditures should be mainly explained by the budget constraint itself.

The main difference with decentralized countries is that in these coun-

\(^{21}\)It is important to note that the per capita tax base coefficient (which is an elasticity) is now statistically significant. It was not under the equalisation of fiscal position specification.

Moreover, the per capita tax base is not correlated with population size, and therefore there is no equalisation through the population size of communities. In fact if they were correlated there could be a redistribution of income to lower size communities through the regressivity of grants regarding population. But in this case there would be multicollinearity in the model.

\(^{22}\)If \( g' = \alpha_1 N^{a_1} y^{a_2} \Rightarrow \sigma' = \alpha_1 N^{a_1+1} y^{a_2} \).
tries the direction of the causality might be reversed since local revenues are (assumed to be) endogenous to local governments. Either expenditures increase because revenues did so or revenues rise to cope with increasing demand.

5.4 Local Government Expenditures

This section will discuss results regarding local government expenditures. As clarified in chapter 4 the determination of the level of local expenditures is different from the issue of the allocation of expenditures between its current and capital components. In the former case, there is not much scope for local autonomy in centralized systems of government while in the latter case there is a greater degree of discretion.

Starting with per capita expenditures the fact that grants are a nonlinear function of population and taxes a linear function of tax base suggests the following nonlinear specification:

\[ c' = a_1 N_i^{a_2} + a_3 b' + e' \]  

(5.8)

where \( c' \) stands for per capita expenditures and \( N_i, b' \) and \( e' \) have the usual meaning.

The estimation of this equation yields the following results:

\[ \hat{c} = 140.4684 \quad N^{-0.61894} + 0.089356 b \quad R^2 = .71 \quad N = 186 \]

\( (4.9559) \quad (-8.3571) \quad (19.4571) \)

The fit of the equation is good particularly taking into account the high degrees of freedom. This seems to give additional support that local expenditures are mainly exogenously determined.

Figure 5.2 shows the actual values of expenditures and grants per capita for Portuguese local governments and clarifies the negative relationship with population size. 23

23 Communities have 10 to 250 thousand inhabitants and are ranked by increasing population size. Algarve's municipalities are special cases as a consequence of being a highly
Figure 5.2: Per Capita Intergovernmental Grants and Local Governments' Expenditures when Communities are Ranked by Increasing Population Size
From equation 5.8 it is possible to obtain the expression for the population elasticity of per capita expenditures, which is given by:

$$ \epsilon_{e,N} = \frac{\delta e}{\delta N} = \alpha_2 \alpha_1 N^{a_2 - 1} \frac{N}{e} = \alpha_2 \alpha_1 N^{a_2 - 1} \frac{1}{e} \quad (5.9) $$

The value of the population elasticity of per capita expenditures was calculated for the median rank community (m=93) considering its population and per capita expenditure:

$$ \epsilon_{e,N_{93}} = -0.49924 $$

This indicates that in the neighbourhood of the median community a 1% increase in population means a -0.5% decrease in per capita expenditure.

Is this the consequence of economies of community size only or is it also a manifestation that more populated communities are unable to provide local services with a similar quality?

To answer this problem it is necessary to calculate the value for the population elasticity of per capita expenditures when output is homogeneous regarding quality. In chapter 4 it was shown that economies of community size are dependent on the structure of the local budgets and that a plausible range for this elasticity was:

$$ -0.5(1 - \alpha) < \epsilon_{e,0}^{e_{90}} < 0 $$

where $1 - \alpha$ is the local governments' budget share of capital intensive goods $z_0$, $s_0$ are fixed levels of quality of local services and indicate precisely that this elasticity assumes that levels of provision are similar across jurisdictions.

Since the budget share of capital expenditures lies in the range 40%-60% for the great majority of Portuguese local governments, using the highest of these values (60%) we obtain the lowest minimum value for the elasticity stated above:

$$ -0.3 < \epsilon_{e,0}^{e_{90}} < 0 $$

exclusively those municipalities.

24It is assumed that per capita costs are identical to per capita expenditures.
Therefore, we have greater confidence in the result obtained that the effective population elasticity of per capita expenditure (-0.49924) lies outside the range accepted as indicating economies of community size.

This is another critical reason to support our argument that per capita expenditures decrease more with the population size of jurisdictions than economies of community size would suggest. As we have seen above this is mainly a consequence of intergovernmental lump-sum grants' design.

It is interesting to confront these supply side conclusions with the demand results from the median voter literature.

To do this, let us postulate that the relationship between per capita expenditures, tax base and population size assume the following multiplicative form:

\[
e = AN^a b^\phi
\]  

(5.10)

It is important to clarify that this functional specification is wrong according to the approach being developed for two main reasons. Firstly, because per capita expenditures do not have a constant elasticity in relation to population. Secondly because the relationship with the per capita tax base is linear which is not the case in this specification.

However, it might fit the data reasonably well because this specification shares with our specification the fact that the logarithm of per capita expenditures is a linear function of the logarithm of population.\(^{25}\)

To put our results in perspective with the median voter results let us introduce the following restrictive and unrealistic assumptions:

i) There is only domestic property tax base or the relationship between domestic and nondomestic property is constant across jurisdictions.

ii) Personal income is normally distributed within each jurisdiction so that the median voter is the average voter.

iii) Domestic (assessed) property values are proportional to incomes.

\(^{25}\)Note however, that ours is not a constant elasticity with regard to population.
These three assumptions imply that the per capita tax base is proportional to the income of the median voter of the jurisdiction and therefore can be used as a proxy to his income.

Let us recall that according to the crowding and congestion approach developed in section 4.2 there are constant returns to scale in consumption (except for small communities). If there are constant returns to scale in production as well, the price for the community of the local public good does not change with community size. However, with increasing returns to scale the price will decrease with the size of the jurisdiction.

With the assumptions i), ii) and iii) it can be seen that the "tax-price" for the median voter will be the same or decrease with jurisdiction size when there are constant or increasing returns to scale in production respectively.\footnote{With increasing returns to scale it is assumed that jurisdictions adopt an average cost pricing rule. Otherwise they would be unable to finance the production of local public goods.}

The above equation can be expressed as:

\[ e = AN^{a_1}N^{a_2}b^\phi = AN^{a_1+a_2}b^\phi \]

and interpreted as follows: \(N^{a_1}\) accounts for the "tax-price" effect, \(N^{a_2}\) shows the quality effect and \(b^\phi\) is the income effect.

Taking logarithms of both members of the equation and adding an error term yields:

\[ \ln e^i = \ln A + \alpha \ln N^i + \beta \ln b^i + \epsilon^i \]

(5.11)

where \(\alpha = \alpha_1 + \alpha_2\). Results of the estimation are the following:

\[
\begin{align*}
\ln e^i &= 5.1768 - 0.38812 \ln N^i + 0.46272 \ln b^i + \epsilon^i \\
& (23.3570) (-16.6424) (16.4864) \\
R^2 &= 0.68 \quad N = 186 
\end{align*}
\]

As expected, although it is a wrong specification, this functional form fits the data relatively well.

Interpreting the taxbase elasticity as an "income" elasticity and confronting the estimate with the ones found in the literature, it is clear the
Turning to the population elasticity, median voter theory as interpreted it as indicating economies of "crowding". Inman (1979) considers it as being economies of crowding and economies of scale in production. The approach developed in chapter 4 shows that it is essentially changes in quality and economies of scale in production. The latter originates a price effect.

5.5 Conclusions

This chapter analysed the case of local government revenues and expenditures in Portugal.

It is clear that in comparison with other European countries members of the OECD, Portugal is one of the most centralized European countries. This conclusion was not based only on the revenues and expenditures by tiers of government but also in the legal framework which constrains the municipalities' ability to manipulate the main local tax rates.

Within a centralized country revenues are mainly exogenous to local governments. The two main sources of revenue (intergovernmental grants and local taxes) are in fact directly and indirectly controlled by central government.

Empirical evidence shows that local taxes are a linear function of the tax base which implies that the implicit tax rate is similar for all communities.

On the other hand, intergovernmental grants they are not designed towards the equalization of the fiscal capacity of jurisdictions. On the contrary communities with higher per capita tax bases receive higher per capita grants. Regarding equalization of the fiscal position of jurisdictions, as given by the economies of community size, empirical evidence suggests that per capita grants are more regressive towards population size than these economies indicate.

This result is consistent with the hypothesis developed in section 4.3.4

in which this design of intergovernmental grants may maximize political support from the jurisdictions while keeping relatively low the level of intergovernmental grants.

However, we have to be cautious in these conclusions. As happens frequently in economics, empirical evidence may be consistent with alternative (and even conflicting) hypotheses.

The exogeneity of revenues implies that receipts command levels of expenditure. Since per capita revenues and expenditures are lower than economies of community size would suggest, this indicates that a decreased quality in local public goods provision should be expected in urban areas.

The next chapter addresses this issue by considering the case of local provision of primary school facilities. A direct look at output indicators will allow a better understanding of the problem than the simple use of expenditures data enable.
Chapter 6

Congestion and the Provision of Primary School Facilities

In chapter 5 we showed that within the context of a centralized country (Portugal) the level of local expenditures is mainly exogenous to local governments.

Per capita expenditures decrease with the population size and increase with the per capita tax base of jurisdictions. The former effect is a consequence of intergovernmental grants design, a rationale for which was introduced in chapter 4. The latter effect is a mere consequence of the exogeneity of local tax rates.

Moreover, empirical evidence suggests that the decrease in per capita expenditures is greater than economies of community size would indicate. Therefore, a lower level of quality and/or a higher degree of local public goods' congestion should be expected in more populated areas.

This chapter takes a bit further the inquiry into this issue, analyzing the case of primary school facilities in Portugal.

The choice of local schools provision was based on four main aspects. Firstly, the importance of education per se. Secondly, because school buildings are facilities that can be replicated at any size or enlarged, at the will of local governments and therefore, differentials in congestion or quality can not be explained by community size. In short, schools are not capacity
constrained. For those goods that are capacity constrained (e.g. the network of roads in a given community assuming that it can not be enlarged) increasing population would mean increasing congestion. However, this is not the case for goods that might be replicated.\(^1\) Thirdly, local schools are less subject to negative externalities than other local public goods as will become apparent below. Finally, because there is some available data on outputs which enables us to directly address the issue of congestion, instead of using imperfect proxies such as expenditures.

It will be shown that urban communities under provide educational capital stock and low populated communities have an excess capacity. The latter effect, is mainly due to demographic changes and political pressures. However, the former effect can not be fully explained by demographic trends but relies on fiscal constraints within urban communities.

Section 6.1 addresses the main hypotheses suggested in the literature on under provision of capital stock in urban communities, as well as the difficulties in dealing with the problem empirically. Section 6.2 introduces the general framework of responsibilities regarding educational building management by local authorities in Portugal and highlights the main demographic trends in the past decade. Section 6.3 considers the issue of whether there are economies of scale in the production of school buildings. Section 6.4 analyses the relationship between primary school congestion and the population size of communities. Finally, section 6.5 concludes.

### 6.1 Community Size, Fiscal "Stress" and Municipal Capital Spending

Capital expenditures have not been the focus of much empirical research by economists for understandable reasons. With few exceptions,\(^2\) the empir-

\(^1\)See discussion supra on section 3.1.4.

\(^2\)One is Wyckoff (1984,1989). Inman considers that only current expenditures should be considered in a median voter framework.
ical median voter literature has considered only current expenditures, the benefits of which are fully "internalized" in the fiscal year to which they relate. On the other hand, capital expenditures generate flows of services which extend for many years.

There has been a substantial literature on urban fiscal "crisis" or fiscal "stress" and the important issue of financing urban infrastructure. The purpose of this section is not to review this literature but only to introduce some landmark approaches and to integrate our approach into the ongoing debate.

Considering capital expenditures in general is a good starting point for discussing the specific case of the capital stock in primary education.

Baumol wrote an influential paper (Baumol 1967) with the subtitle: "the anatomy of the urban crisis". In short, he argued that productivity increases in the public sector are much harder to obtain than in the private sector since technological "rigidities" in the production functions of public sector services (in particular local services) prevent the substitution of capital for labour when wages rise and also limit the scope of technological innovations that can be introduced. This is the case, he argues, of education, police, hospitals, social services and inspection services.

This argument gives a rationale for the increase in the share of the public sector (in the GDP, for instance) but does not explain the particular problems of urban areas. This, Baumol argues, can be explained by negative externalities such as pollution and traffic congestion. Since these externalities are public goods (bads) each home will "consume" the total amount of the public good and therefore the "cost of externalities" to the community rises with the square of population rather than proportional to population.\footnote{\textsuperscript{5}See for instance Kirwan (1989), Peterson (1984), Bailey (1991) and references therein. \textsuperscript{4}The example in Baumol (1967) is simple and elucidating. If pollution is proportional to the population size of the community (say $kn$), each home will "consume" that amount of the public bad. Since houses are proportional to population (on) the total amount of...}
This introduces a clear rationale for the decreasing quality of local services in urban areas. Costs increase more than proportional to population while resources do not. In particular the flight to the suburbs of those who can pay erodes the tax base of the city. At the same time it increases the per capita costs of providing the same level of services, since they are assumed to generate less negative externalities.

Bradford, Malt and Oates (1969) built on Baumol’s paper and made an important distinction between two notions of output already referred to. The output directly produced by local governments is an input, together with community’s environment (or characteristics), in the production of the “final” output that appears in consumers’ utility functions.

Most of the contemporary approaches to the production function for local services uses a variant of the “characteristics” approach. Income, was considered by Baumol and more recently by Hamilton (1983a) and others as a good proxy for the population characteristics of a community. Wealthier families are less likely to commit violent crimes, more easily maintain their houses and more likely to contribute to the education of their children.

In particular Oates has been arguing that the population size of a community is likely to be negatively correlated with the characteristics that have a positive role in the “production function” of local services.

This approach enables us to extend Baumol’s approach from pure local public goods to club goods. Even when the negative externalities are not pure public goods (bads), but are confined to the individuals sharing the consumption of a shared good (school, swimming-pool,...).6

To understand this, consider the case where the population of every community is partitioned into identical capacity and membership clubs and where there is some degree of congestion (negative consumption externalities) within each. It is clear that each individual faces the same conditions

domestic sootfall will amount to $abn^2$.

5See subsection 3.1.3.
6Actually, Bradford, Malt and Oates do not mention or refer club goods theory which was being developed at that time after Buchanan’s seminal paper (1965).
independently of the size of community where he lives. Each club has the same amount of "cost of externalities" and therefore total "costs" will be proportional to community size.

However, this would not be the case if the characteristics of the individuals sharing the club change systematically with community size, which is precisely the (not tested) argument of Oates.

Another interesting hypothesis found in the literature was introduced by Peterson (1980) and refers specifically to the maintenance of the capital stock. Peterson suggests that governments under fiscal distress decrease maintenance expenditures and new capital formation and shift expenditures to other more visible services. The idea is further developed in Hulten and Peterson (1984) who consider that maintenance expenditures can be deferred in time without much damaging the quality of the capital stock. However, at some point in time further deferral of maintenance will have a dramatic effect on the quality of the capital stock.

As they say, referring to the USA situation, "the 1970's saw a significant increase in the share of the budget devoted to social welfare programs. Given limits, implicit and explicit, on overall spending levels, the crowding-out of other state-local programs was inevitable. Capital spending was a natural candidate for the crowding-out process since (...) it can be deferred for some time before adverse consequences are created (and even longer before these are noticeable). This crowding-out effect was most pronounced in cities experiencing fiscal distress".7

The problem with the concept of fiscal distress is its ambiguity and difficulties in measuring it.

An empirical test of the capital deferral hypothesis was undertaken by Bumgarner et al. (1991). It is a good paper to understand the difficulties inherent in analyzing capital expenditures and specifically the hypothesis stated above.

One of the obvious problems is the indicator for fiscal stress. Should it be a relative concept for a given point in time comparing different jurisdictions or should it take into account for each jurisdiction the changes in fiscal health over a given period of time?

Bumgarner et al. chose the latter option and used an index developed by Bradbury, Downs and Small (1982) which is an index of urban decline. It is a composite index taking into account for a given period (1970/1975) the negative changes in the unemployment rate, in the violent crime rate in the city debt burden and the percentage change in per capita income (1969/1974).

They acknowledge the existence of other indexes such as those developed by Ladd and Yinger (1989) and Fossett and Nathan (1981). However, they consider them to be measures of relative fiscal capacity and health among cities and not of change in fiscal health.®

Their approach runs three regressions corresponding to three different ratios: (i) maintenance expenditures on total services expenditures (ii) maintenance expenditures on capital expenditures and (iii) capital expenditures on service expenditures.

Total expenditures are the sum of non capital inputs, maintenance of capital stock and gross investment. Each of these components of expenditure is assumed to be a function of: the interest rate, the vintage of the capital stock, the depreciation rate, service delivery conditions, intergovernmental grants and total expenditure. Apart specification issues® there is some difficulty in getting good proxies for several variables.¹⁰

The authors conclude with caution, given data constraints, that fiscal

---

®It is important to note that the estimate of the “fiscal distress” variable using the Fossett-Nathan index and the Ladd-Yinger index were not statistically significant.

®We suspect the existence of multicollinearity between intergovernmental grants and total expenditure.

¹⁰For instance the proxy for the depreciation rate is the percentage of the labour force in the jurisdiction employed in manufacturing. The indicator for the vintage of the capital stock is the years elapsed since the housing stock in the jurisdiction was half its actual size. The proxy for service of delivery conditions is the area of the jurisdictions.
distress does in fact contribute to a decrease in the share of maintenance and investment expenditures and an increase in other expenditures.

Holtz-Eakin and Rosen (1989) also address the issue of municipal capital spending in a quite different context. Their aim is to analyze if capital decision-making is a response to changing economic and demographic conditions or if it is a myopic decision by politically motivated individuals. The idea is an interesting one, but again empirical testing is rather problematic. The authors use panel data, but are forced to pool data according to a stratification of the sample in categories according to which communities are or are not suburbs, and have high or low population.

The problem with using expenditures as a proxy for output has been already referred to in this thesis. It is not a good proxy particularly when investment is concerned, due to big fluctuations in the capital formation from year to year. Therefore, we will analyze directly measures of capital stock output and whether differentials in congestion are explained by demographic conditions and/or fiscal constraints.

In so far as the hypotheses here introduced are concerned, our approach in chapters 4 and 5 gives a different rationale to fiscal "stress" in more populated communities. The combined effect of (i) regressive intergovernmental grants regarding population size and (ii) tight fiscal constraints on the ability to raise taxes, put urban communities in a worse (post-grant) fiscal position. Therefore, they will have to cut back on current and capital expenditures. Increased congestion is just the result of underprovision\(^\text{11}\) of capital stock.

\(^{11}\)The concept of underprovision we are using here has not the usual meaning related to allocative efficiency but to the norm of equalization of provision according to a (central government set) standards of quality.
6.2 Education Responsibilities and Demographic Changes in Portugal

In Portugal the Ministry of Education assumes almost total responsibility regarding curriculum, maintenance and construction of new schools and payments to school employees (teachers and other civil servants). An exception has been the maintenance and construction of primary schools which have been for a long time the responsibility of local governments. Therefore, we may say that the capital stock of primary schools as it exists today is a main consequence of past and present local governments' investments in the sector.

Only recently, have local governments begun to share some responsibilities in building construction for other levels of education within the compulsory school system (or the remainder of basic education). At this educational level there are cooperation agreements between the local authorities and the Ministry of Education in a joint provision. Local authorities are mainly responsible for providing the site and preparing it for construction which in terms of total costs represents between 10 and 30% of total costs assigned to local authorities.

This chapter will concentrate on primary schools since they have been a long standing responsibility from local governments.\footnote{Only in section 6.3 will economies of scale being discussed for non primary schools which are the only ones where data is available.}

The provision of primary school education is mainly done through the public sector that shares 95% of student enrolments. Private provision is concentrated in a few urban communities which have the majority of Portuguese private schools.\footnote{Lisbon and Oporto are two of them and they are excluded from the analysis for the reasons stated in chapter 5.}

Figure 6.1 has on the horizontal axis Portuguese municipalities with more than 10000 and less than 250000 inhabitants in 1991 ranked by increasing population size of the jurisdiction in that year. On the vertical
axis is the average annual growth rates of population during that period.

Figure 6.1: Average Annual Growth Rates of Population (1980/91) for Portuguese Municipalities (ranked by increasing population size (1991))

It emerges from the figure that smaller communities have been loosing population during the decade, the more populated communities are getting still more populated and the medium range ones have mixed situations. The urbanization process has been from low populated communities to coastal communities and the suburbs of the two metropolitan areas of Lisbon and Oporto.\(^1\)

The population change variable is an indicator of the general fiscal pressures that jurisdictions are facing and will be used in section 6.4. However, to understand the changes in the "demand" for local public schools we

\(^1\)Note that there are small communities (under 10000 inhabitants) which are not considered in the analysis and which are also loosing population.

\(^2\)Having faith in the latest statistics (INE 1993) Lisbon lost more than 18% of population and the Oporto lost 9%, confirming the tendency of the flight to the suburbs that is typical in most countries. They are not contemplated in the figure since they are above the ceiling of 250000 inhabitants.

\(^3\)The quotation marks clarify that we are not speaking of the usual economic concept of demand.
should look at changes in the enrolments in public primary schools.

Figure 6.2 gives this information. In the horozontal axis we have again the communities ranked by increasing population size and in the vertical axis we have now the average annual growth rates of the enrollments in primary public schools.

![Graph showing average annual growth rates of primary school enrollments.](image)

**Figure 6.2: Average Annual Growth Rates (1980/91) of Primary Schools’ Enrollments for Portuguese Municipalities (ranked by increasing population size (1991)**

There was a substantial decline on the total number of enrolments during the eighties which amounted to -24% in the decade. This is mainly explained by the decline in birth rates from the seventies onwards and a slight decrease in the rate of failure.\(^{17}\)

However, it is important to state that almost all communities had a

---

\(^{17}\) The primary school in Portugal has four years divided in two “phases” of two years each. Around 30%(!) of children fail progression in studies either in the first phase or in the second phase of primary school. The situation is slightly better today when compared with 1980, however, it is still a dramatic one.

Of course statistics are what they are, and a way to change the figures was simply to legally abolish the possibility of failing. This sort of move would have its benefits and also its natural costs. A new generation of adults without some essential basic skills would
sharp decrease in the number of primary school enrolments during the
eighties, although a significant number of them had an increase in total
population.

Therefore, for the decade as a whole, actual demand for local public
schools has decreased particularly in rural areas, but also in urban commu­
nities. This introduces the general framework for the discussion of conges­
tion in primary schools in section 6.4.

Beforehand, the issue of the existence of economies of scale in the build­
ing of schools is addressed.

6.3 Economies of scale in the production of
educational facilities

There is no data available to analyse economies of scale in primary schools.
However, there is a data set (GEP (1990)) for investments on schools at a
higher degree of education built during the period 1985-1988. These schools
are considerably bigger than primary schools, but the fact that there is a
relatively wide range of sizes, allows us to make some inductive reasoning
on the existence of economies of scale in primary schools.

Apart from certain limitations the data set is very comprehensive and
has individualized information on contract costs for each of the 149 schools
in the sample. There is no data that enables us to develop a fully specified
production function which would require information on the quantity and
quality of factors of production, factor prices and output.

However, there is information about total costs and two variables that
may serve as indicators of output: area of the school and number of class­
rooms. We choose the latter variable because it seems a better indicator of
output.

Therefore a preliminary approach to the data can be done by inspecting
arise. Some commentators argue that this is what happens to some extent in the U.K.
nowadays as a result of the abolition of primary school tests several years ago.
figure 6.3 where average costs (per classroom) are plotted against the size of schools as measured by the number of classrooms.

![Figure 6.3: School Investment Costs per Classroom](image)

It is clear that output is here a discrete variable being schools of the following sizes: 11, 12, 18, 20, 24, 30 and 42. The arrows indicate the average cost within each size.

Apart from factors that might partially explain the variations on average costs, it seems that either the market is not very competitive or the contracting out procedures were not very efficient. These schools were contracted out to private firms, but it is not clear which was the process of tendering.

Total costs are split into the initial contract cost, price revisions due to the length of time between the date of the contract and the date in which building works started, additional costs due to imponderable situations, and premia for the anticipation of building works. Initial contract costs represent around 70% of total costs and there is a considerable variety in the way entrepreneurs use their possibilities to increase costs after the contract was actually signed.

There is no evidence however that the variety of average costs has any
relationship with differences on quality of schools. On the contrary it seems that these differences show a non competitive market and some loose regulations regarding the ability to change \textit{ex post} the values established in the initial contract.

The average cost values within each size of schools suggest that some economies of scale are in place although there is a great variation in average costs.

To analyse this, average costs ($\bar{C} \equiv \text{ACOST}$)\textsuperscript{18} were regressed against the size of the school the length of time that took the construction process and a variable indicating the location of the school.

The size of the school was measured by the number of classrooms ($S \equiv \text{SIZE}$), the time variable is measured in days ($T \equiv \text{DURATION}$) and the location variable is a dummy variable ($U \equiv \text{URBAN}$) which assumes a value one if the school belongs to the Lisbon and Oporto metropolitan areas\textsuperscript{19} and a not value otherwise.

The output variable is a discrete variable clustered in seven categories and it is clear from figure 6.3 that the "explanatory" capacity of such a model would be rather poor. Anyway, it would help understanding if the exogenous variables are significant in the regression.

The postulated relationship is the following:

$$\ln \bar{C}^i = \alpha_1 + \alpha_2 \ln T^i + \alpha_3 \ln S^i + \alpha_4 U^i + \epsilon^i$$

(6.1)

where $\epsilon^i$ is the error term.

A double-logarithmic model was specified and the results are the following:

$$\ln \bar{C}^i = 11.0049 - .1207 \ln T^i - .3925 \ln S^i - .6783 U^i \quad R^2 = .38, N = 149$$

(48.1245) (-3.5065) (-6.6818) (-2.2356)

\textsuperscript{18}The first letter is the short reference to the variable which appears in the regression expression in the main text. The second expression is the reference in the output results in appendix A.2.

\textsuperscript{19}More precisely if they belong to the Distritos of Lisbon, Setubal and Oporto which almost overlap these metropolitan areas.
Although the variation in average costs is only partially explained by these variables, all three are statistically different from zero.

The negative sign on the size variable confirms the existence of some economies of scale in the production of schools. The negative sign on the duration variable is consistent with the existence of premia that the entrepreneurs receive when they are able to anticipate the construction schedule. The negative sign for the urban variable is also in conformity with the expectations that in urban areas the market is likely to be more competitive and raw materials more cheap.

As indicated in the beginning of this section the great majority of primary schools are much smaller than those considered here. If we assume a monotonically decreasing average cost function this suggests that average costs in primary schools would be even higher than those in secondary schools and economies of scale higher than those obtained for secondary schools.

6.4 Urban Size and Congestion of Local Schools

This section analyses empirical results on congestion in local primary schools. In chapter 4 we analysed the crowding function as a nonlinear increasing function of per capita provision of local public goods and clarified that the density of utilization and clarified that it is possible to shift from the former to the latter and vice-versa.

The endogenous variable that will be used in this section \( k \equiv UTILCAP \) is the density of utilization of primary schools indicating the relationship between aggregate levels of utilization \( (N^i) \) and aggregate capacity of schools.
Since there is no precise data on capacity of schools, but there is available data on number of classrooms the process to construct the capacity variable is the following:

Capacity = 20 \times \text{Number of classrooms, where 20 is the average of the minimum and maximum recommended values of class sizes in primary schools.}

Utilization of capacity is therefore the ratio of total students' enrolments on capacity of schools (as defined above).

Figure 6.4 shows the values of this ratio (minus one) for medium/large Portuguese communities ranked by increasing population size. The line (0,0) indicates situations where the number of enrollments is identical to schools' capacity in that jurisdiction. If a community is on the line it means that (on the average of its schools) there is neither over utilization nor under utilization of schools' capacity. Columns above the line indicate over utilization and below the line under utilization. A first inspection of the figure indicates two things. The level of utilization increases with the population size of communities and other factors should be in place to

\( k^t = \frac{N^t}{X^t} \)

20Note that \( k^t = 1/z^t \), being \( z^t \) (as defined in chapter 4) the per capita provision of local public goods in community \( i \).

21The study is GEP (1992) and the values are respectively 15 and 24 (the average (19.5) was rounded). In a final stage of writing this thesis we had access to another publication (GEP (forthcoming)) which has important detailed information on utilisation rates in primary schools and could not be incorporated here. In this study, the capacity variable was measured in the same way but it was used the maximum value of the range (24). This is also a valid approach which assumes as a norm for capacity the maximum value whether in our approach we are more restrictive. Nevertheless, the indicator of capacity, should use the value that those engaged in pedagogical approaches to primary school education consider as the ideal ratio students per teacher.

22We have subtracted one from the UTILITY variable only for purposes of graphical presentation and not in the regression analysis below.

23Note that it might be erroneously suggested from the figure that the relationship could be linear. It can not because population size increases nonlinearly with the rank of
Figure 6.4: The Utilization of Capacity of Primary Schools and Community Size

explain levels of utilization of capacity.

We investigate this issue through regression analysis where the endogenous variable is $k \equiv UTILCAP$ and the exogenous variables considered are: the variation of population during the decade ($V \equiv VAR8191$), the change on the number of enrolments ($E \equiv AL9092$) and a dummy urban variable considered in terms of population density ($D \equiv DUMURB$).\(^{24}\)

The variable $V \equiv VAR8191$ indicates the expected effect of population growth (decline) as a fiscal "pressure" ("release") on local governments. Therefore, it is expected to be positively related to the utilization of capacity variable.

$AL9092$ is a precise indicator of how changed the level of enrolments and therefore incorporates the "demand" factors in explaining utilization communities, as we have seen when discussing and empirically analysing the population hierarchy of Portuguese communities.

\(^{24}\) $D$ assumes the value one when population density is higher than 200 inhabitants per squared kilometre and zero otherwise. Data for $UTILCAP$ (reported to 1992) and enrolments 1990/1992 were obtained from GEP (unpublished data). Statistics of population come from INE.
of capacity in public primary schools.

Increased enrollments should be associated with increased density of utilization of primary schools.

Finally, the variable which accounts for the effect of population size is a dummy variable on population density. There are two main reasons why we had to use this dummy variable and not population size ($PS$) itself. The first is the high correlation between $PS$ and $VAR8191$ which would introduce problems of multicollinearity if we wanted to keep both variables. However, the two variables account for different factors: population size and growth respectively and we would like to analyze separately if both were relevant. On the other hand, due to the (approximate) rank-size rule on the distribution of communities' population, the introduction of the $PS$ variable would have the additional consequence of avoiding a linear specification.

This was in fact the specification chosen:

$$k^i = \beta_1 + \beta_2 E^i + \beta_3 V^i + \beta_4 D^i + \epsilon^i$$

and the results are the following:

$$\hat{k}^i = 1.2594 + 0.0058759 E^i + 0.016 V^i + 0.1966 D^i$$

$$\hat{R}^2 = .61 N = 179$$

All parameters are statistically significantly different from zero at a 99\% degree of confidence and have the expected positive signs.

It is important to recall here that population increased in urban areas and decreased in rural areas but enrollments decreased almost everywhere.

The results show that the variable indicating changes in “demand” for public primary schools are relevant to understand the rates of utilization of capacity and therefore congestion. The fact that changes in enrollments are positively related with congestion might indicate that there are lags in

\footnote{The choice was ad hoc, perhaps influenced by some "econometric verificationist" bias.}
the process of adjustment of the capital stock to variations in demand for public primary schools.

However, they do not tell the whole story. Factors affecting the fiscal position of communities, in particular population growth and population size have also an important explanatory role. This suggests that the Hulten and Peterson's capital deferral hypothesis might be relevant in understanding investments in local schools. Communities under fiscal "stress" will defer capital formation from services subject to increased congestion and will shift resources to others more "visible" where monitoring quality can be done at a lower cost.

The quality indicator used in this section is associated essentially with congestion in local schools. There is no available data to analyse Peterson's hypothesis regarding maintenance of local schools either in terms of expenditures or output. However, there is another dimension of the problem of quality in education provision that we would like to consider.

6.4.1 Schools' Capacity, Congestion and another Dimension of the Quality of Educational Services

The questions we would like to address here are the following: are primary schools' students attending normal classes, with normal timetables where classrooms are assigned to classes on a one to one basis? If the answer is no, are schools' capacity constraints responsible for this situation?

The normal *regimen* in which a primary school should function is that each classroom should be assigned to only one class. Therefore, if a school has three classrooms it should have the same number of classes.

26It was our purpose to analyse here output data on the quality of local primary schools. There was a recent survey made by GEP on the characteristics of all local primary schools in Portugal. Unfortunately this data is still not available.

27We are aware that this question might seem peculiar to some readers. The clarification of the problem follows below.
However, under Portuguese law schools may have two or even three\textsuperscript{28} classes which share the same classroom during different periods of the day. This will obviously have negative implications for the quality of educational services in schools that work in this way. The first group of children have to come earlier than children in “normal” schools and leave earlier. The second group of children will attend classes in the afternoon. The classroom can not function as a space which children feel is theirs, where they can display and decorate according to their specific needs and wishes because it is a shared space. This sort of rota system was intended to be a temporary solution but it has remained as a more or less permanent reality. It enables schools to have much more students than their real capacity allowed for. The second question addressed above can be restated as follows: is the use of the “double” scheme explained by \textit{present} capacity constraints or not?

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.5.png}
\caption{Proportion of Primary Schools' Classes with Normal Timetables}
\end{figure}

Figure 6.5 shows for each municipality\textsuperscript{29} the ratio (in percentage) of pri-

\textsuperscript{28} The “triple” regimen was used in the past (but not today) when the pressures on the educational system were greater.

\textsuperscript{29} Again only municipalities with population size within the range 10000/250000, ranked
mary schools' classes that have normal timetables to the total number of existing classes in the jurisdiction. For the considered range of municipalities, provision is always done with some classes in the "double" regimen. This situation is more frequent in more populated communities, although the observed variation indicates that factors other than community size should be considered.

The important aspect that we would like to stress here is that it is not present schools' capacity which is the main factor underlying this situation. In fact, if we look at Figure 6.4 only those communities which have an over utilization of capacity (where the column is above the 0.00 line) would have a rationale (based on capacity constraints) for using two classes in the same classroom. All the others could function in the normal regimen.®°

In the past, when enrolments in primary schools were higher, capacity constraints might have been the essential rationale for introducing the rota system in school classrooms. Nowadays, the rationale for the permanence of this system should be searched elsewhere.®¹

6.5 Conclusions

This chapter analysed the specific case of the provision of primary education facilities in Portugal, addressing the issues of economies of scale in production and particularly the level of utilization of local schools.

The demographic changes occurred in the past decade show that population has been ageing as a result of decreased birth rates following the 1970's and therefore the total number of enrolments in primary schools has by increased population size in 1991.

®°It is important to note here that if we had used a less restrictive indicator of capacity (e.g. the one used in GEP (forthcoming) considering 24 students per classroom) the line (0.00) would shift upwards and the number of communities having an over utilisation of schools' capacity would be smaller.

®¹In particular, we suspect that there might be teachers' interests in maintaining the current situation since working hours might be reduced when classes share the same classroom. Working GEP's new data on the issue might highlight the current problem.
been decreasing sharply during the eighties. At the same time there was a process of urbanization so that this decrease was more accentuated in rural areas and less in urban areas.

In this context the high degree of congestion in local schools in urban areas can not be explained by increased demand and a lag in the adjustment of the capital stock to changes in demand. If anything, changes in demand with a lagged process of adjustment would indicate that there would be an under utilization of capacity everywhere, more pronounced in the rural areas.

Moreover, it can not be explained by per pupil costs since empirical evidence suggests the existence of economies of scale in production and therefore these costs are lower in urban areas.

What really seems to affect provision of local schools is conditions affecting supply and financial constraints that might be associated with population growth during the decade. It is the main argument of this thesis that the combined effect of intergovernmental grants' design and the constraints imposed by the legal framework, put urban communities in a relatively worse fiscal position, what some authors denote as fiscal "stress".
Chapter 7

Conclusions

Three types of considerations will be analysed here. Firstly, more general theoretical conclusions and implications of the analysis will be developed in section 7.1. Specific issues regarding the Portuguese case will be introduced in section 7.2. The case of education provision in the context of institutional design will be approached in section 7.3. Finally, section 7.4 concludes with the limitations and possible developments of this research.

7.1 On the economies of community size

The theory developed in this thesis supports the statement that, excepting small communities, there are rather limited economies of community size. They do not seem to exist for labour intensive local services and there are economies in the provision of capital intensive services due mainly to economies of scale in production.

However, some authors suggest that the characteristics of the population in more urban areas require additional expenditure to obtain the same quality of local services. If this is the case, economies of community size may be partially or totally offset by the diseconomies associated with changing characteristics of the population when communities get larger in size.

It was also clarified that central governments aiming at controlling public expenditure will, *ceteris paribus*, prefer a distribution of grants to local
authorities less favourable to urban communities than economies of city size would suggest.

In countries with centralized governments these factors will mean that urban communities will be in a situation of fiscal stress when compared with their smaller counterparts.

Local public goods were considered to have "privateness" characteristics rather than showing "publicness". This was argued in pure theoretical and analytical terms and it was also suggested that the dispute over the characteristics of local goods can not be settled on the basis of empirical evidence.

This result only shows that technologically speaking, local public goods can be provided publicly or privately the decision being dependent on an appropriate assignment and transfer of property rights.

The analysis in this thesis was made within a positive framework. However some normative implications can be addressed at this stage.

If efficiency was the unique criterion, provision of all local "public" goods that can be provided competitively through the market should be provided privately. In what concerns "natural monopolies" the issue would have to be considered case to case since it is not clear a priori that private provision means more efficiency in the long run.

However, if equity issues are considered, a case for public provision may exist. In particular if central government wants to equalize the provision of meritorious local goods, the "privateness" conclusion implies that intergovernmental grants should equalize the fiscal position of communities trying to offset only the economies of community size originated by economies of scale in the production of capital intensive goods.

The main problem in using intergovernmental grants with equalization purposes seems the fungibility of resources even when matching grants are in place. Assuming plurifunctional local governments providing both merit and non merit local public goods⁴ there is no guarantee that grants will

⁴The merit characteristic should be seen from a Constitutional point of view and not from the points of view of the economist or the policy maker.
equalize the provision of the former ones.

### 7.2 On the Portuguese Case

It was clear from the empirical work developed earlier that local revenues (and expenditures) in Portugal are mainly exogenously defined in spite of the great variety of fiscal instruments and other revenue resources at the discretion of local governments.

Central government has a great degree of control of local expenditures both directly through intergovernmental grants and indirectly through the legal framework that assigns to the central government the responsibility of setting local tax rates and the limits of borrowing. Only in the reduced field of user charges and selling of services do local governments have real fiscal autonomy.

Moreover, Portugal was found to be one of the most centralized of European (OECD) countries both in terms of traditional measures of fiscal centralization (ratios of current consumption expenditures and capital formation by tiers of government) and also according to the structure of local tax rates.

A consequence of the main local taxes being related to the property tax base is that as the economy grows local revenues will rise less quickly than central revenues and hence, centralization of revenues will follow naturally. This occurs, because income tax and VAT are more elastic regarding GDP than taxes related to the property tax base.

This suggests that moves towards greater decentralization will imply a considerable reform of local government finance.

Experience of other countries indicates that where local governments are financed through a mix of income and property taxes the degree of decentralization is greater and so is the stability of the share of responsibilities among tiers of government. The implementation of a local income
tax together with a liberalization of property tax rates would enhance the flexibility in raising autonomous revenues. Competition between local governments would be a good instrument for keeping rates from increasing dismeasurably.3

In what concerns the distribution of local grants across communities, the analysis of economies of community size indicates that population is a rough but perhaps the best overall indicator of fiscal "needs" with mainly two exceptions. Small communities were excluded from the analysis since in these communities the optimal level of local service provision may exceed the population of the community. Also the biggest communities 4 should be considered as having net "spillovers" going to non-residents that live outside these jurisdictions. Therefore, the resident population indicator would underestimate the "needs" for fiscal resources in these jurisdictions. If the design of intergovernmental lump-sum grants aims at equalizing the fiscal position of communities, it should give a greater weight to the population size of communities.5

It should also take into account the existing differentials in per capita tax base in order to partially offset them instead of reinforcing them as presently is the case. Therefore, it seems that a revaluation of the criteria for distributing these grants should be considered if that normative criterion is to be accepted.

It should be pointed out that the suggestion being made for further

---

3Which most likely should be related (as a surcharge) to the existing income taxes.
4Other processes for expenditure control will be considered below.
5We have excluded from the analysis communities with more than 250,000 inhabitants to be consistent with other empirical literature. In practice this amounts to excluding Lisbon, Oporto and Loures (a suburban community in the Lisbon metropolitan area).
6This means that the benefits to non-residents from using local services within these jurisdictions ("spillovers") exceed the benefits that residents receive from using services in neighbour communities ("spillins"). This has been accentuated in the last decade with the process of "flying to the suburbs" typical in metropolitan areas of several countries. Lisbon has lost 18% of population and Oporto 9% (data from INE (1993)) in the eighties.
7We are referring to the formulae which determines the distribution of the Fund for financial balance that consolidates most of intergovernmental grants to local governments.
decentralization would imply, *ceteris paribus*, a larger overall public sector.

Portugal has a smaller ratio of public sector expenditure to GDP than other European countries do, but this does not suggest that it should become much larger. On the contrary, moves towards greater decentralization are envisaged mainly as a reallocation of functions and responsibilities between levels of government and therefore should be performed simultaneously with a contraction of central government expenditures.

There is currently a process of reviewing the size and structure of public administration in Portugal, and the status of public employees. Whether the present changes will lead only to a reorganization of central government administration with some Ministries being enlarged at the expense of others (in which case no reduction in central government expenditures is to be expected) or whether this will lead to further decentralization is yet unknown.

The implementation of the above mentioned changes on the *distribution* of grants would make urban communities better off and rural communities relatively worse off. However, total grants may increase in size and therefore make rural communities in the same (or better) fiscal position.

It is important to note here that scarcity of local revenues should not constitute a panacea for local governments' difficulties. The fact that incentives for allocative and production efficiency are lower in the public sector than the private sector is commonly acknowledged by economists nowadays.

Therefore, improving the management of local services, increasing the tendering of services and enhancing local accountability should certainly also be on the agenda for reforming local governments' institutional context.

The principal agent problem resulting from asymmetric information is usually stated in terms of central and local governments respectively. However, the principal agent relationship should be seen between the citizens tax-payers (the principals) and their local government (the agent).

---

1 Economists have written very much about the optimal size of government. However, it seems that there is no definitive conclusion about optimality in this context and particular views (like ours) are more idiosyncratic than anything else.
Fears for political and economic decentralization are often associated with the threat that overall public expenditure gets out of control. However, it is not only the central government that wants to contain public expenditure. Those who ultimately share the burden of the taxes - the citizens - also want to keep the local and central government expenditures under control and will do so if they have access to the appropriate information and the mechanisms to make their fiscal choices operative.

This would happen if citizens were more aware of their fiscal burden and also the benefits associated with the provision of public goods.

Information on these issues is enhanced by effective and more transparent political competition which would improve if current obstacles in the "local political market" were removed. Within more competitive local political markets the median voters' voice would become louder even in the likely presence of some voters' "rational ignorance".

7.3 Local Provision of Education

If primary education is to be considered a meritorious local public service, the current system of educational finance should also be under scrutiny.

When the emphasis of the analysis shifts from general local government expenditures to specific expenditures in particular services (such as education) several degrees of freedom are introduced. The present local governments' finance bill rules out the earmarking of revenues to specific expenditures. Therefore, given that revenues are exogenous, it is natural that where local governments have a relatively weak fiscal position they will cut down several services, hence, decreasing their quality.

The analysis developed in this thesis shows that there is in fact a high

---

8The main one is perhaps the monopoly of local political competition by national political parties. Under Portuguese Law, there can be no independent lists in the elections at municipality level. Only at the lower level of freguesias, which have very small responsibilities, it is possible for groups of citizens to run for local elections.

9See Downs (1957).
degree of congestion in urban primary schools in spite of a decrease in enrolments in almost every community in the country during the eighties.

Urban communities have in general increased population during the decade and this must have had an impact on the fiscal needs of these communities. However, in primary education there was a decrease in demand, so that congestion cannot be explained directly through the demographic effect.

With respect to the capital costs of investment, either the market is not competitive or the tendering procedures for the construction of schools are not efficient. Only one of these reasons seems to be capable of explaining the great variety in average construction costs for schools with similar dimension.

Urban communities face lower costs per pupil both because they can appropriate the economies of scale in the production of local schools and also because spatial location (belonging to the Lisbon and Oporto metropolitan areas) lowers average costs. Again this is an additional factor that does not explain why there is increased congestion in urban schools.

Extra resources to local governments would not increase ipso facto the expenditures on education. This is due not only to the potential X-inefficiency reasons but also because more "visible" services can receive the additional revenues. Even specific matching grants to education instead of the current lump-sum grants might not improve the situation, since the literature has found a considerable degree of fungibility of these type of grants.10

If there is a strong will of equalizing standards of quality the creation of local education authorities should be considered as a strong candidate for institutional reform.

Unifunctional local authorities is an institutional solution to circumvent the problems of general fund financing stated above and the deficiencies of specific matching grants. It is a way of earmarking revenues to expenditures

10On the fungibility of grants see section 3.3.1 above.
and is the proper institutional context where the median voter may have some decisive importance on local decision making. Maybe because of this, it has been the major "peaceful revolution" in local finance in the USA during the eighties and early nineties which saw a proliferation of special (unifunctional) districts. Although there are some potential pitfalls in its application, there is no doubt that it has great potential for improving efficiency and equity and it should be considered seriously.

Under the present system of local accountability it is even difficult to quantify local expenditures on education since they are mixed with other items in the local budget. Therefore, it is not surprising in the current debate that central government wants to shift some responsibilities on education to municipalities which are reluctant to accept them because they believe they will not have the extra revenues necessary to cope with the increased responsibilities. It seems that neither one nor the other knows much about what is and what should be spent on education and both play the same strategy: try to maximize (minimize) intergovernmental transfers in the context of imperfect and asymmetric information.11

There is no doubt, however, that the present fragmentation of responsibilities between local and central governments does not seem the best solution.

The possibility of the creation of local education authorities should represent an effective decentralization not only in terms of expenditures but also in terms of curriculums and pedagogical experiences. It would not be a great improvement if they become only authorities on behalf of the Ministry of Education and if central government keeps a tight control on curriculum issues.

11 The relation between central and local government can be considered as an agency relationship where the agent (local government) has more information (although imperfect) than the principal (central government) on the production function of local services. Therefore, local governments lobby for increasing intergovernmental transfers while central government try to constrain these transfers.
7.4 Final Comments

At this stage we would like to point to some possible research directions that go beyond the limits of our approach.

In chapter 2 we argued that provision of local public goods is likely to be inefficient since communities can not be modelled as clubs. A Tiebout-like model could be developed to address the issue of the equilibrium resulting from the sorting out process of citizens among jurisdictions offering local public goods with different levels of congestion and quality.

The important prediction of such a model would be related to the migratory response from citizens with different levels of income and heterogeneous preferences regarding the quality of local services.

Another important development is to address the issue of the allocation of local governments' expenditures among different local public goods. As we clarified earlier, within centralised governments it is not necessary to have a model to explain and predict overall local expenditures. However, such a model is needed if we want to predict the equilibrium level of expenditures among alternative uses.

We have shown the limitations of the median voter approach and therefore other approaches should be considered. In this context special attention should be given to the role of interest groups and of the media in setting local governments' agendas.

Another issue which deserves further attention is the positive analysis of intergovernmental grants' design. What is the real rationale behind grants' design? Further empirical research on this issue should highlight the comparative weights of economic and political rationales underlying grants' design. This would enable a better understanding of the real scope (effectiveness) of the normative analysis on intergovernmental grants.\footnote{What is the relevance of analysing "optimal" grants' design i f we conclude that economic rationales occupy a secondary importance in central governments' decision making?}

Other developments of research are more related to normative analysis, in particular comparative fiscal and institutional design. We referred to the
issue of creating special districts to provide particular local merit goods. Theoretical and empirical analysis should clarify the allocative efficiency and potential equity advantages of such a move as well as clarifying the increasing administrative costs likely to be associated with this institutional reform.

A different issue is to compare benefits and costs of alternative systems of local public goods' finance. Local governments may use general funds or earmark taxes for particular expenditures. They can apply user charges and fees to the consumption of certain local public services. Again the choice among these fiscal instruments will have implications for allocative efficiency, X-efficiency and equity. Further theoretical and empirical approaches should clarify the main existing trade-offs and the institutions that should be set up to decide upon them.13

It is worth noting that tax reform is usually cheaper and more flexible than institutional reform. On the other hand, it seems that institutional reform can easily handle problems of temporary chronic under supply of local public goods.

These are but a few possible developments of the analysis. It is our firm intention to pursue this line of research and sincerely hope that other academics, with more knowledge and experience of this vast field than ourselves, will share in this challenging task.

---

13 Our approach to the problem is essentially procedural. The role of the political economist here is not to give his/her view on the efficiency/equity trade-offs and the "right" social welfare function, but to devise the procedures (institutions, rules of decision-making) so that citizens' preferences on values are channelled through the political process into actual decision-making.
Appendixes
## Local Governments' Current Receipts and Current Inter. Transf. (current prices in national currencies)

### Appendixes

**Tables not included in the main part of the thesis**

<table>
<thead>
<tr>
<th>Year</th>
<th>Portugal*</th>
<th>Belgium</th>
<th>Denmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>10524</td>
<td>20149</td>
<td>94382</td>
</tr>
<tr>
<td>1979</td>
<td>18963</td>
<td>220141</td>
<td>107261</td>
</tr>
<tr>
<td>1980</td>
<td>23258</td>
<td>229599</td>
<td>122994</td>
</tr>
<tr>
<td>1981</td>
<td>27518</td>
<td>257840</td>
<td>135201</td>
</tr>
<tr>
<td>1982</td>
<td>37909</td>
<td>288068</td>
<td>152479</td>
</tr>
<tr>
<td>1983</td>
<td>45466</td>
<td>302199</td>
<td>168928</td>
</tr>
<tr>
<td>1984</td>
<td>56124</td>
<td>341643</td>
<td>178686</td>
</tr>
<tr>
<td>1985</td>
<td>69850</td>
<td>358309</td>
<td>188633</td>
</tr>
<tr>
<td>1986</td>
<td>83996</td>
<td>375308</td>
<td>191327</td>
</tr>
<tr>
<td>1987</td>
<td>121840</td>
<td>374091</td>
<td>204940</td>
</tr>
<tr>
<td>1988</td>
<td>150788</td>
<td>387358</td>
<td>219066</td>
</tr>
<tr>
<td>1989</td>
<td>176721</td>
<td>374237</td>
<td>230663</td>
</tr>
<tr>
<td>1990</td>
<td>213601</td>
<td>385131</td>
<td>239512</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Austria</th>
<th>Finland</th>
<th>Norway</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>61555</td>
<td>22793</td>
<td>35778</td>
</tr>
<tr>
<td>1979</td>
<td>65408</td>
<td>25906</td>
<td>37333</td>
</tr>
<tr>
<td>1980</td>
<td>73123</td>
<td>29033</td>
<td>44432</td>
</tr>
<tr>
<td>1981</td>
<td>78384</td>
<td>34707</td>
<td>50458</td>
</tr>
<tr>
<td>1982</td>
<td>81813</td>
<td>39135</td>
<td>57027</td>
</tr>
<tr>
<td>1983</td>
<td>65290</td>
<td>44257</td>
<td>62573</td>
</tr>
<tr>
<td>1984</td>
<td>51473</td>
<td>51195</td>
<td>70510</td>
</tr>
<tr>
<td>1985</td>
<td>99540</td>
<td>57644</td>
<td>80552</td>
</tr>
<tr>
<td>1986</td>
<td>104830</td>
<td>63108</td>
<td>89484</td>
</tr>
<tr>
<td>1987</td>
<td>107932</td>
<td>67581</td>
<td>99500</td>
</tr>
<tr>
<td>1988</td>
<td>107946</td>
<td>71825</td>
<td>105997</td>
</tr>
<tr>
<td>1989</td>
<td>114308</td>
<td>79389</td>
<td>111968</td>
</tr>
<tr>
<td>1990</td>
<td>123883</td>
<td>91334</td>
<td>120627</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>FRANCE</th>
<th>GERMANY (West)</th>
<th>GREECE</th>
<th>IRELAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>121096</td>
<td>52471</td>
<td>43.33%</td>
<td>7200</td>
</tr>
<tr>
<td>1979</td>
<td>141130</td>
<td>61289</td>
<td>43.43%</td>
<td>7780</td>
</tr>
<tr>
<td>1980</td>
<td>164309</td>
<td>71103</td>
<td>43.27%</td>
<td>87500</td>
</tr>
<tr>
<td>1981</td>
<td>190119</td>
<td>82574</td>
<td>43.43%</td>
<td>88020</td>
</tr>
<tr>
<td>1982</td>
<td>224312</td>
<td>99874</td>
<td>44.52%</td>
<td>89940</td>
</tr>
<tr>
<td>1983</td>
<td>260536</td>
<td>113336</td>
<td>43.46%</td>
<td>92950</td>
</tr>
<tr>
<td>1984</td>
<td>286771</td>
<td>102732</td>
<td>35.82%</td>
<td>97950</td>
</tr>
<tr>
<td>1985</td>
<td>318060</td>
<td>113033</td>
<td>35.76%</td>
<td>104120</td>
</tr>
<tr>
<td>1986</td>
<td>348826</td>
<td>130196</td>
<td>37.32%</td>
<td>108490</td>
</tr>
<tr>
<td>1987</td>
<td>370924</td>
<td>139295</td>
<td>37.58%</td>
<td>112000</td>
</tr>
<tr>
<td>1988</td>
<td>399570</td>
<td>148374</td>
<td>37.13%</td>
<td>117730</td>
</tr>
<tr>
<td>1989</td>
<td>429389</td>
<td>163018</td>
<td>37.97%</td>
<td>124770</td>
</tr>
<tr>
<td>1990</td>
<td>460732</td>
<td>171375</td>
<td>37.20%</td>
<td>129700</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>SWEDEN</th>
<th>SWITZERLAND</th>
<th>UNITED KINGDOM</th>
<th>SPAIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>126916</td>
<td>43640</td>
<td>32.26%</td>
<td>11645</td>
</tr>
<tr>
<td>1979</td>
<td>12085</td>
<td>1030</td>
<td>8.52%</td>
<td>12695</td>
</tr>
<tr>
<td>1980</td>
<td>13648</td>
<td>48677</td>
<td>32.14%</td>
<td>13755</td>
</tr>
<tr>
<td>1981</td>
<td>167225</td>
<td>50564</td>
<td>30.24%</td>
<td>14650</td>
</tr>
<tr>
<td>1982</td>
<td>184487</td>
<td>55420</td>
<td>29.98%</td>
<td>15650</td>
</tr>
<tr>
<td>1983</td>
<td>200356</td>
<td>60526</td>
<td>30.21%</td>
<td>16475</td>
</tr>
<tr>
<td>1984</td>
<td>219859</td>
<td>65339</td>
<td>30.54%</td>
<td>17645</td>
</tr>
<tr>
<td>1985</td>
<td>225997</td>
<td>66182</td>
<td>29.28%</td>
<td>18795</td>
</tr>
<tr>
<td>1986</td>
<td>241417</td>
<td>70261</td>
<td>29.10%</td>
<td>19570</td>
</tr>
<tr>
<td>1987</td>
<td>253566</td>
<td>70138</td>
<td>27.66%</td>
<td>20375</td>
</tr>
<tr>
<td>1988</td>
<td>279236</td>
<td>77288</td>
<td>27.68%</td>
<td>21715</td>
</tr>
<tr>
<td>1989</td>
<td>311810</td>
<td>84610</td>
<td>27.20%</td>
<td>22885</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>8274 277 3.35%</td>
<td>8453 432 5.11%</td>
</tr>
<tr>
<td>1979</td>
<td>40722 33649 82.63%</td>
<td>9255 539 5.62%</td>
</tr>
<tr>
<td>1980</td>
<td>49232 39819 80.88%</td>
<td>9596 572 5.96%</td>
</tr>
<tr>
<td>1981</td>
<td>59337 48587 81.88%</td>
<td>10778 690 6.40%</td>
</tr>
<tr>
<td>1982</td>
<td>79656 63524 79.75%</td>
<td>12672 690 5.45%</td>
</tr>
<tr>
<td>1983</td>
<td>86377 68279 79.05%</td>
<td>13057 733 5.61%</td>
</tr>
<tr>
<td>1984</td>
<td>96425 76466 79.32%</td>
<td>14194 812 5.72%</td>
</tr>
<tr>
<td>1985</td>
<td>108008 83132 78.42%</td>
<td>14856 1026 6.91%</td>
</tr>
<tr>
<td>1986</td>
<td>120080 95555 79.86%</td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td>131074 104153 79.46%</td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>142963 109538 76.62%</td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>155724 115502 74.17%</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>NETHERLANDS Tot. Rec. Interg. Trans Gr./Rec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1978</td>
<td>46920 39140 80.01%</td>
</tr>
<tr>
<td>1979</td>
<td>51880 41390 79.81%</td>
</tr>
<tr>
<td>1980</td>
<td>54580 43790 80.20%</td>
</tr>
<tr>
<td>1981</td>
<td>59510 47510 79.84%</td>
</tr>
<tr>
<td>1982</td>
<td>68770 55920 81.31%</td>
</tr>
<tr>
<td>1983</td>
<td>72970 59480 81.51%</td>
</tr>
<tr>
<td>1984</td>
<td>72770 56580 80.50%</td>
</tr>
<tr>
<td>1985</td>
<td>75450 60220 79.81%</td>
</tr>
<tr>
<td>1986</td>
<td>75600 60340 79.81%</td>
</tr>
<tr>
<td>1987</td>
<td>72660 57110 79.03%</td>
</tr>
<tr>
<td>1988</td>
<td>71880 57220 79.83%</td>
</tr>
<tr>
<td>1989</td>
<td>70550 55580 78.78%</td>
</tr>
<tr>
<td>1990</td>
<td>75020 59320 79.07%</td>
</tr>
</tbody>
</table>

### A.2 Regression Results

186 observations = mainland Portuguese communities (concelhos) with population between 10,000 and 250,000 inhabitants in 1991

272 observations = mainland Portuguese communities (concelhos) with population under 250,000 inhabitants

275 observations = all mainland Portuguese communities (concelhos)

Estimation results for equation (5.1), (N=272).

<table>
<thead>
<tr>
<th>Dependent variable is TAXPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>272 observations used for estimation from 1 to 272</td>
</tr>
<tr>
<td>Regressor</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>TXBASEPC</td>
</tr>
<tr>
<td>R-Squared</td>
</tr>
<tr>
<td>R-Bar-Squared</td>
</tr>
<tr>
<td>Residual Sum of Squares</td>
</tr>
<tr>
<td>S.D. of Dependent Variable</td>
</tr>
<tr>
<td>DW-statistic</td>
</tr>
</tbody>
</table>

Estimation results for equation (5.2), (N=186).

<table>
<thead>
<tr>
<th>Dependent variable is TAXLPTPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>186 observations used for estimation from 87 to 272</td>
</tr>
<tr>
<td>Regressor</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>TXBASEPC</td>
</tr>
<tr>
<td>R-Squared</td>
</tr>
<tr>
<td>R-Bar-Squared</td>
</tr>
<tr>
<td>Residual Sum of Squares</td>
</tr>
<tr>
<td>S.D. of Dependent Variable</td>
</tr>
<tr>
<td>DW-statistic</td>
</tr>
</tbody>
</table>
### Estimation results for equation (5.2), (N=272)

**Ordinary Least Squares Estimation**

<table>
<thead>
<tr>
<th>Dependent variable is</th>
<th>TAXLPTPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>272 observations used for estimation from 1 to 272</td>
<td></td>
</tr>
</tbody>
</table>

| Regressor | Coefficient | Standard Error | T-Ratio | Prot |
|-----------|-------------|----------------|----------|
| A         | -1.0586     | .20842         | -5.0795 | .00C |
| TXBASEPC  | .036977     | .0011379       | 32.1443 | .00C |

| R-Squared | .79283 | F-statistic F(1, 270) | 1033.3 | .00C |
| R-Bar-Squared | .79206 | S.E. of Regression | 2.41 |
| Residual Sum of Squares | 1572.0 | Mean of Dependent Variable | 3.71 |
| S.D. of Dependent Variable | 5.2915 | Maximum of Log-likelihood | -624.53 |
| DW-statistic | 2.1885 |  |

### Estimation results for equation (5.3), (N=272)

**Ordinary Least Squares Estimation**

<table>
<thead>
<tr>
<th>Dependent variable is</th>
<th>FEFPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>272 observations used for estimation from 1 to 272</td>
<td></td>
</tr>
</tbody>
</table>

| Regressor | Coefficient | Standard Error | T-Ratio | Prot |
|-----------|-------------|----------------|----------|
| A         | 20.2341     | .79678         | 25.3946 | .00C |
| TXBASEPC  | -.013442    | .0043903       | -3.0399 | .00C |

| R-Squared | .034152 | F-statistic F(1, 270) | 9.5472 | .00C |
| R-Bar-Squared | .030575 | S.E. of Regression | 9.2 |
| Residual Sum of Squares | 22976.4 | Mean of Dependent Variable | 18.4 |
| S.D. of Dependent Variable | 9.2692 | Maximum of Log-likelihood | -989.3 |
| DW-statistic | .45391 |  |
Estimation results for equation (5.4), (N=272)

**Non-Linear Least Squares Estimation**
The estimation procedure converged after 6 iterations

Non-linear regression formula:
\( \text{fefpc} = a_1 + a_2 * \text{popm191} + a_3 + a_4 * \text{tbasepc} \)
272 observations used for estimation from 1 to 272

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>T-Ratio</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>3.8261</td>
<td>1.2120</td>
<td>2.9093</td>
<td>.004</td>
</tr>
<tr>
<td>A2</td>
<td>88.0749</td>
<td>5.0197</td>
<td>17.5460</td>
<td>.000</td>
</tr>
<tr>
<td>A3</td>
<td>-.69301</td>
<td>.051644</td>
<td>-13.4188</td>
<td>.000</td>
</tr>
<tr>
<td>A4</td>
<td>.0030578</td>
<td>.0018715</td>
<td>1.6339</td>
<td>.103</td>
</tr>
</tbody>
</table>

R-Squared .83598 F-statistic F(3, 268) 455.3242 .000
Residual Sum of Squares 3901.8 Mean of Dependent Variable 18.480
S.D. of Dependent Variable 9.3692 Maximum of Log-likelihood -748.172
D-W-statistic 2.0965

Estimation results for equation (5.4), (N=186).

**Non-Linear Least Squares Estimation**
The estimation procedure converged after 4 iterations

Non-linear regression formula:
\( \text{fefpc} = a_1 + a_2 * \text{popm191} + a_3 + a_4 * \text{tbasepc} \)
186 observations used for estimation from 87 to 272

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>T-Ratio</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1.9356</td>
<td>2.4468</td>
<td>.791471</td>
<td>.430</td>
</tr>
<tr>
<td>A2</td>
<td>76.9240</td>
<td>18.3953</td>
<td>4.1817</td>
<td>.000</td>
</tr>
<tr>
<td>A3</td>
<td>-.39739</td>
<td>.14305</td>
<td>-2.764</td>
<td>.006</td>
</tr>
<tr>
<td>A4</td>
<td>.0030824</td>
<td>.0015705</td>
<td>1.9646</td>
<td>.053</td>
</tr>
</tbody>
</table>

R-Squared .67341 F-statistic F(3, 182) 125.0893 .000
Residual Sum of Squares 1767.3 Mean of Dependent Variable 13.83
S.D. of Dependent Variable 5.4084 Maximum of Log-likelihood -473.300
D-W-statistic 2.0326
### Estimation results for equation (5.6), (N=275)

<table>
<thead>
<tr>
<th>Ordinary Least Squares Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable is LNRANK1</strong></td>
</tr>
<tr>
<td>275 observations used for estimation from 1 to 275</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>T-Ratio [Prob]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>13.6943</td>
<td>.14623</td>
<td>93.6477 [.000]</td>
</tr>
<tr>
<td>LNPOP91</td>
<td>-.92290</td>
<td>.014812</td>
<td>-62.3086 [.000]</td>
</tr>
</tbody>
</table>

- **R-Squared**: .93430
- **R-Bar-Squared**: .93406
- **Residual Sum of Squares**: 16.7072
- **S.D. of Dependent Variable**: .96338
- **D.W.-statistic**: .024080
- **F-statistic F (1, 273)**: 3882.4 [.000]
- **S.E. of Regression**: .2473E
- **Mean of Dependent Variable**: 4.630
- **Maximum of Log-likelihood**: -5.075

### Estimation results for equation (5.7), (N=272)

<table>
<thead>
<tr>
<th>Ordinary Least Squares Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable is LNFEFPC</strong></td>
</tr>
<tr>
<td>272 observations used for estimation from 1 to 272</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>T-Ratio [Prob]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>7.4636</td>
<td>.12513</td>
<td>59.6459 [.001]</td>
</tr>
<tr>
<td>LNTXBPC</td>
<td>.046149</td>
<td>.020096</td>
<td>2.2965 [.02]</td>
</tr>
<tr>
<td>LNPOP91</td>
<td>-.49912</td>
<td>.013269</td>
<td>-37.6148 [.001]</td>
</tr>
</tbody>
</table>

- **R-Squared**: .85821
- **R-Bar-Squared**: .85716
- **Residual Sum of Squares**: 9.8027
- **S.D. of Dependent Variable**: .50509
- **D.W.-statistic**: 1.9728
- **F-statistic F (2, 269)**: 814.1061 [.00]
- **S.E. of Regression**: .190
- **Mean of Dependent Variable**: 2.79
- **Maximum of Log-likelihood**: 65.99

### Estimation results for equation (5.8), (N=186)

<table>
<thead>
<tr>
<th>Non-Linear Least Squares Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>The estimation procedure converged after 10 iterations</td>
</tr>
</tbody>
</table>

### Non-linear regression formula:

\[ \exp(pct) = a_1 + poptm1 + 191 + a_2 + a_3 + txbasep \]

| 186 observations used for estimation from 87 to 272 |

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>T-Ratio [Prob]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>140.4684</td>
<td>28.3437</td>
<td>4.9559 [.00]</td>
</tr>
<tr>
<td>A2</td>
<td>-.61894</td>
<td>.074062</td>
<td>-8.3571 [.00]</td>
</tr>
<tr>
<td>A3</td>
<td>.089356</td>
<td>.0045924</td>
<td>19.4571 [.00]</td>
</tr>
</tbody>
</table>

- **R-Squared**: .71050
- **R-Bar-Squared**: .70734
- **Residual Sum of Squares**: 17.0548
- **S.D. of Dependent Variable**: 2.1891
- **Maximum of Log-likelihood**: -675.71

---

223
Estimation results for equation (5.11), (N=186).

**Ordinary Least Squares Estimation**

<table>
<thead>
<tr>
<th>Dependent variable is LNEXPPC</th>
<th>186 observations used for estimation from 87 to 272</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regressor Coefficient</td>
<td>Standard Error</td>
</tr>
<tr>
<td>A</td>
<td>5.1768</td>
</tr>
<tr>
<td>LNP0P91</td>
<td>-.38812</td>
</tr>
<tr>
<td>LNTXBPC</td>
<td>.46272</td>
</tr>
<tr>
<td>R-Squared</td>
<td>.68021</td>
</tr>
<tr>
<td>R-Bar-Squared</td>
<td>.67672</td>
</tr>
<tr>
<td>Residual Sum of Squares</td>
<td>9.2464</td>
</tr>
<tr>
<td>S.D. of Dependent Variable</td>
<td>.39534</td>
</tr>
<tr>
<td>DW-statistic</td>
<td>2.3140</td>
</tr>
</tbody>
</table>

Estimation results for equation (6.1).

**Ordinary Least Squares Estimation**

<table>
<thead>
<tr>
<th>Dependent variable is LNACOST</th>
<th>149 observations used for estimation from 1 to 149</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regressor Coefficient</td>
<td>Standard Error</td>
</tr>
<tr>
<td>A</td>
<td>11.0049</td>
</tr>
<tr>
<td>LNDUR</td>
<td>-.1207</td>
</tr>
<tr>
<td>LNSIZE</td>
<td>-.3925</td>
</tr>
<tr>
<td>URBAN</td>
<td>-.0783</td>
</tr>
<tr>
<td>R-Squared</td>
<td>.3961</td>
</tr>
<tr>
<td>R-Bar-Squared</td>
<td>.3836</td>
</tr>
<tr>
<td>Residual Sum of Squares</td>
<td>6.1535</td>
</tr>
<tr>
<td>S.D. of Dependent Variable</td>
<td>.2624</td>
</tr>
<tr>
<td>DW-statistic</td>
<td>1.4218</td>
</tr>
</tbody>
</table>
Ordinary Least Squares Estimation

Dependent variable is UTILCAP
179 observations used for estimation from 1 to 179

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>T-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.2594</td>
<td>.0202</td>
<td>62.3983</td>
</tr>
<tr>
<td>AL9092</td>
<td>.0058759</td>
<td>.0009130</td>
<td>6.3978</td>
</tr>
<tr>
<td>VAR8191</td>
<td>.0160</td>
<td>.0018856</td>
<td>8.4964</td>
</tr>
<tr>
<td>DUMURB</td>
<td>.1966</td>
<td>.0391</td>
<td>5.0334</td>
</tr>
</tbody>
</table>

R-Squared  | .6138       |
F-statistic| F(3, 175)   | 92.7173  |
S.E. of Regression | .1928 |
Residual Sum of Squares | 6.5065 |
Mean of Dependent Variable | 1.2506 |
S.D. of Dependent Variable | .3077  |
Maximum of Log-likelihood | 42.6655 |

DW-statistic | 1.7269 |
A.3 Economies of scale and congestion within the median voter model.

Inman assumes a multiplicative demand function:

\[ x = AI^i p_x^d \]  \hspace{1cm} (A.1)

the cost function,

\[ C(X) = cX^n \]  \hspace{1cm} (A.2)

the crowding function,

\[ z = \frac{X}{N^a} \]  \hspace{1cm} (A.3)

and the tax-price variable (in units of consumption),

\[ p_x = \frac{C(x)}{N} \]  \hspace{1cm} (A.4)

Introducing equations A.2 and A.4 in equation A.1 yields:

\[ x = AI^i [ rc^{\eta-1} N^\frac{\eta}{\delta} ]^d \]

Introducing equation A.3 and rearranging we obtain:

\[ X^{1-(\eta-1)\delta} = AI^i (rc)^{\delta} N^{\frac{\eta-\delta}{\delta}} \]

Let

\[ 1 - (\eta - 1)\delta = q \]

then,

\[ X = A^i I^i (rc)^{\frac{\delta}{\eta}} N^{\frac{\eta - (\eta - 1)\delta}{\delta}} \]

Since from equation A.2

\[ X = \left[ \frac{C(X)}{c} \right]^\frac{i}{d} \]

we have:

\[ C(X) = c A^i I^i (rc)^{\frac{i}{d}} N^{\frac{\eta - \delta}{\delta}} \]  \hspace{1cm} (A.5)

The meaning of the notation can be found in section 3.1.1.
Let:

\[ \hat{\epsilon} = \frac{\eta}{q}, \]

\[ \hat{\delta} = \frac{\eta_{\alpha}}{q}, \]

\[ \hat{\alpha} = \frac{\eta (\gamma - \delta)}{q} \]

Therefore it is possible to rewrite equation A.5 as:

\[ C(X) = cA^2 I^2 (\tau c)^2 N^X \quad (A.6) \]

Adding an error term and taking logarithms it is possible to estimate this equation.

Note that:

\[ \alpha = \frac{\eta (\gamma - \delta)}{q} \Rightarrow \hat{\alpha} = \eta \gamma - \frac{\eta \delta}{q} \]

Since, \( \delta = \frac{\eta_{\alpha}}{q} \) it is possible to obtain:

\[ \eta \gamma = \hat{\alpha} + \hat{\delta} \quad (A.7) \]

and it is also obvious that

\[ \frac{\hat{\epsilon}}{\hat{\delta}} = \frac{\epsilon}{\delta} \quad (A.8) \]

Inman's equation A.8 is identical to our result but equation A.7 is different from the one we obtained.

Expenditures and Grants per Capita (excluding Algarve Municipalities)
Bibliography
Bibliography


Balsas, A., 1993, A Dimensão da Administração Local e Alguns Comportamentos Financeiros, mimeo, paper presented at the Seminário Internacional sobre Poder Local, Lisbon


230


Bramley, G., 1990, Equalization Grants and Local Expenditures Needs, Avebury, Aldershot

Brazer, H. and McCarty, T., 1987, "Interaction between demand for education and for municipal services", National Tax Journal, 40, p. 555-566


De Viti De Marco, A., 1936, First Principles of Public Finance, London, Jonathan Cape


Departamento de Programação e Gestão Financeira (DPGF), 1993, Taxas de Ocupação das Escolas 1991/92 (versão provisória), Lisboa, DPGF (Ministério da Educação)


Direcção-Geral da Administração Autárquica (DGAA), Finanças Municipais 1981/82, Lisboa, DGAA (Ministério do Planeamento e Ordenamento do Território),

Direcção-Geral da Administração Autárquica (DGAA), Finanças Municipais 1986, Lisboa, DGAA (Ministério do Planeamento e Ordenamento do Território),
Direcção-Geral da Administração Autárquica (DGAA), Finanças Municipais 1987, Lisboa, DGAA (Ministério do Planeamento e Ordenamento do Território)

Direcção-Geral da Administração Autárquica (DGAA), Finanças Municipais 1988, Lisboa, DGAA (Ministério do Planeamento e Ordenamento do Território)

Direcção-Geral da Administração Autárquica (DGAA), Finanças Municipais 1989, Lisboa, DGAA (Ministério do Planeamento e Ordenamento do Território)

Direcção-Geral da Administração Autárquica (DGAA), Finanças Municipais 1990, Lisboa, DGAA (Ministério do Planeamento e Ordenamento do Território)

Direcção-Geral da Administração Autárquica (DGAA), Finanças Municipais 1991, Lisboa, DGAA (Ministério do Planeamento e Ordenamento do Território)


Fischer, R., 1982, "Income and Grant effects on Local Expenditure: The Flypaper Effect and Other Difficulties", Journal of Urban Economics, 12, p.324-345


Friedman, M., 1953, Essays in Positive Economics, Chicago, University of Chicago Press


Gabinete de Estudos e Planeamento (GEP), 1990, Indicadores de Custos das Infraestruturas para os Ensino Básico e Secundário, Lisboa, GEP (Ministério da Educação)


Hamilton, B., 1975, "Zoning and property taxation in a system of local governments", Urban Studies, 12, p.205-211


Hamlin, A., 1985, "Federalism, Horizontal Equity and the Optimal Grant", Public Finance Quarterly, 13, p.115-131


Hirsch, W., 1984, Urban Economics, New York, MacMillan


Instituto Nacional de Estatística (INE), 1993, Censos 91: Resultados Pré-definitivos, Lisboa, INE


240


King, D., 1990, "Gains and losses from the changes in the system of grants paid to English local authorities", Environment and Planning C: Government and Policy, 8, p.69-86


241


Mozzicafredo, J., Guerra, I., Fernandes, M., Quintela, J., 1988, "Poder Autárquico e Desenvolvimento Local", Revista Crítica de Ciências Sociais, 25/26, p.79-113


Niskanen, W., 1971, Bureaucracy and Representative Government, Chicago, Addine-Atherton


243


Organization for Economic Cooperation and Development (OECD), 1992, Decentralization and Educational Building Management: The Impact of Recent Reforms, Paris, OECD


Peterson, G., 1980, "Cleveland and Cincinnati: A Tale of Two Cities," The Urban Institute Policy and Research Report, 10, p. 3-8


Popper, K., 1983, Realism and the Aim of Science, (vol 1 of the Postscript of the Logic of Scientific Discovery), London, Hutchinson (1st ed. 1956)


Rubinfeld, D., 1987, "The Economics of the Local Public Sector", in Auerbach, A. J. and Feldstein, M. (eds), Handbook of Public Economics, Amsterdam, North-Holland


Schumpeter, J., 1943, Capitalism, Socialism and Democracy, New York, Harper & Row


Wildasin, D., 1987, "Theoretical Analysis of Local Public Economics", in E.S.Mills (ed.), Handbook of Urban and Regional Economics (vol. 2), Amsterdam, North-Holland

Wildasin, D., 1989, "Demand Estimation for Public Goods (Distortionary Taxation and Other Sources of Bias)", Regional Science and Urban Economics, 19, p.353-379


