University College London
Department of Economics

Local government expenditure decisions: empirical models for the analysis of spatial data

Federico Giovanni Revelli

A thesis submitted for the degree of
Doctor of Philosophy
University of London

May, 1998

Abstract

This thesis is concerned with the English local authority expenditure decisions in the 1980s. I focus on two features of local governments that might affect their levels of taxation and spending. The first is the presence of a non-linear grant distribution system (the Block Grant). I estimate a demand function for local public spending in the presence of the piecewise-linear non-convex budget constraint created by the Block Grant. Since Ordinary Least Squares yield biased estimates - due to the endogeneity of the choice of segment on the kinked budget constraint - I use a two-error Maximum Likelihood procedure.

The second is the spatial character of the data collected at the subcentral government level. Firstly, I explore whether local public expenditure exhibits a spatial pattern. I compute spatial statistics which suggest that local governments' expenditures are positively spatially autocorrelated. Application of spatial econometric models shows that, when allowing both for a spatially lagged dependent variable and for spatial correlation in the error term, most of the spatial correlation in the data is captured by correlated shocks.

Secondly, I look at the implications of spatial dependence for local taxation and its effects on local election results. I model the relationship between voters and incumbent governments as a principal agent one, where local jurisdictions are subject to spatially correlated cost shocks and imperfectly informed voters decide whether to re-elect the incumbent after looking at their jurisdiction's relative performance. The empirical evidence on a panel data set of the English districts' elections shows that the own tax has a negative impact on the incumbent's re-election chances, which is only partially offset by the positive impact of neighbouring local governments' taxes. However, local election outcomes also appear to be largely affected by national politics issues.

To Margherita

Contents

Chapter Three

Spatial patterns in local public expenditure

Chapter Four

Local taxes, spatial effects and election results

List of tables

List of figures

Acknowledgements

In doing my PhD and writing this thesis, I have run up a long list of acknowledgements. First, I would like to take this opportunity to thank my professors at the University of Torino, who enthusiastically encouraged me to do a PhD. I owe a great debt of gratitude to Giorgio Brosio, Piervincenzo Bondonio, Ugo Colombino, Alberto Cassone, Enrico Luzzati, Terenzio Cozzi and Vittorio Valli.

During my four years at UCL, I learnt a great deal from my supervisor, Ian Preston. I am immensely grateful to him for his help and patience.

The Economics Department at UCL - as well as the Institute for Fiscal Studies - provided an excellent environment for my work. I benefited from discussions with Orazio Attanasio, Richard Blundell, Ian Crawford, John Hall, Hamish Low, Costas Meghir, Stephen Smith, and Donald Verry.

I am grateful to Jonathan Thomas and Dennis Epple for useful suggestions on chapters three and four respectively, and to Robert Bennett and two anonymous referees for very helpful comments on chapter two. I would also like to thank Douglas Hibbs, Jr., Dietmar Wellisch, and seminar participants at the Universities of Torino, London (UCL), Toulouse, Essex and Prague.

I greatly acknowledge the financial support from the University of Torino, the "Luigi Einaudi" Foundation, Torino, and the "Compagnia di San Paolo", Torino, without which this research would not have been possible. I would also like to thank the UCL Department of Economics for offering me the post of Teaching Assistant in the third and fourth year of my PhD. I learnt a great deal from that experience.

Finally, I owe special thanks to my PhD colleagues and friends at UCL, to my parents, and to my angel, Margherita.

Chapter One

Introduction

1.1 Local public expenditure determination

This thesis is concerned with the English local authority taxation and expenditure decisions in the 1980s. There are at least two reasons why local government seems a particularly interesting subject for empirical study. The first is the presence of a non-linear grant distribution system by which local governments are subsidised by central government. The second is the spatial character of the data collected at the subcentral government level. As a result, each local authority interacts both vertically with the higher tier authority (authorities) and horizontally with the communities that surround it.

I start by presenting for reference a popular, simple, and, perhaps, naive model of local public expenditure determination, where the above mentioned interactions are ruled out. In particular, I stress the simplifying assumptions it is based upon, and look at the consequences of relaxing them on empirical implementation. The model of local public expenditure determination that I take as a starting point is based on the following hypotheses:[1]

(*a*) there is no grant from other levels of government. The local public services are financed entirely by local taxes;

(*b*) a jurisdiction's taxation and expenditure decisions do not lead to spillovers into other jurisdictions, the tax base is immobile, and the jurisdictional boundaries are fixed;

---

[1] Clearly, the modelling of local public expenditure determination is only a special case of the more general problem of modelling how local jurisdictions choose a whole range of policies. They include, among the others, the levels of expenditures on various public services, and the combination and levels of taxes to be used to finance those expenditures.

(*c*) each local government is welfare-maximising, bureaucracy is no impediment to efficient production, and each output level is produced at least cost. Furthermore, all individuals in a jurisdiction are identical. This implies that each jurisdiction can be treated like a single household, and the standard microeconomic model of utility maximisation can be applied.

This approach to local government expenditure determination is discussed at length in Wildasin (1986) - who refers to it as the "community preference" model - as well as in Rubinfeld (1987) and Mueller (1989), who present extensive reviews of theoretical analyses that rely on this framework, and empirical applications that have tested its implications. I take it as a benchmark, and relax its assumptions in turn.

First, the model is based on the hypothesis that a local jurisdiction can be treated like a single household, or like a group of identical households. The preferences of the local community can be represented by a local social welfare function - which equals the utility function of the representative individual - defined over consumption of private goods (*c*) and consumption of public services (*x*) in jurisdiction *i*, as in equation (1.1):

$$\omega_i = \omega(c_i, x_i) \qquad (1.1)$$

If all inhabitants of the jurisdiction are identical and the publicly provided good is private, all variables can be expressed in per capita terms.[2] The jurisdiction's problem is to maximise the local social welfare function, subject to the budget constraint.[3] In the absence of a grant distribution system, and if units of *c* and *x* are chosen to have a price of £1, the budget constraint is simply:

[2] If the good provided by the local government is purely public in the Samuelsonian sense, total public expenditure X would enter the utility function. On the other hand, public expenditure per head ($x=X/n$) enters the utility function in the case of a publicly provided private good. Borcherding and Deacon (1972) and Bergstrom and Goodman (1973) propose a specification (X/n^{α}) that permits the estimation of a parameter (α) expressing the degree of publicness of the good. If $\alpha=0$, the good is a Samuelsonian pure public good, while if $\alpha=1$, it is a private good.

[3] I abstract here from problems related to the manner in which governmental goods and services are produced.

$$c_i + x_i = y_i \tag{1.2}$$

where y_i is given income (per head) in jurisdiction i. As for the tax instrument, jurisdiction i can finance its expenditures either with a head tax ($x_i=T_i$), or with a proportional income tax ($x_i=\tau_i y_i$), or with a property tax ($x_i=t_i B_i$). B_i is the average value of the property, and it is assumed that income and the average value of the property are unaffected by local spending and taxation.

The equation of local public expenditure determination deriving from this model has often been used in empirical applications (Oates, 1972; Gramlich, 1977; Craig and Inman, 1982). Basically, the level of public expenditure per head in a locality is related to income per head, urbanisation rate, population size and density, and other socio-economic, demographic and political characteristics (Mueller, 1989).

The fact that individuals within a jurisdiction are heterogeneous, however, has led many economists to give more explicit attention to the mechanism of collective action at the local level, and to attempt to model the local political process. The median voter model (Downs, 1957) has therefore been applied to the analysis of local public expenditure determination.

If the single-peakedness restriction on voter preferences is satisfied, a majority voting equilibrium exists, and this equilibrium is the preferred outcome of the voter who has the median ideal point. Several empirical studies have tested whether the level of public expenditure and taxation in a local jurisdiction is the one desired by the median voter, where the median voter is assumed to be the one with the median level of income.[4] As a result, the level of expenditure in a locality is related to the marginal tax-price and to the income of the median voter, as well as to population and a vector of other relevant local characteristics. Early examples of empirical analyses of this

---

[4] Bergstrom and Goodman (1973) analyse the expenditures of a number of US municipalities that mainly rely on the property tax as a source of revenue. They assume that the citizen with the median income owns the property with median value. The tax share variable in the regression equation is obtained by dividing the median value of property by total property tax revenue for the municipality.

sort are Barr and Davis (1966), Borcherding and Deacon (1972), Bergstrom and Goodman (1973), and Pommerehne (1978).[5]

In the following, I adopt the above model of local public expenditure determination, and look at the consequences on empirical analysis of relaxing the assumptions on which it is based.[6]

The presence of grants from central government - ruled out by hypothesis (*a*) - is considered in general terms in section **1.2**, and analysed in much greater detail in chapter two, with an application to the grant distribution system that was in place in Britain until 1990 - the Block Grant.

Section **1.3** extends the model of local public expenditure determination by allowing for spillovers between neighbouring jurisdictions. I derive an empirical equation of local public expenditure determination where spatial interactions are explicitly allowed for. The econometric issues involved in the estimation are treated extensively in chapter three, with an application to the English local governments' expenditures.

Finally, section **1.4** relaxes the assumption of welfare-maximising governments. I model the relationship between voters and representatives as a principal agent one, and I permit the objective function of the policymakers not to coincide with social welfare. The implications in the tax-setting and voting decisions are analysed in chapter four, where spatial interactions among local governments are explicitly taken into account.

---

[5] Even though the median voter model has become the most popular approach to the empirical analysis of local government spending, an alternative method for estimating the demand for local public goods uses micro (individual) survey data as opposed to macro or aggregated data (Bergstrom *et al.*, 1982; Gramlich and Rubinfeld, 1982; Preston and Ridge, 1995). This approach is reviewed in Blundell (1988) and Rubinfeld (1987). A great advantage of survey data is that since the individual (rather than the locality) is the unit of observation, it is possible to estimate the effect of individual characteristics - e.g., race, sex, age, number of children - on demand for local public services. On the other hand, that approach has two weaknesses. First, in most instances the desired level of public spending is not observed - but only whether an individual would like more or less of it - and qualitative choice techniques must therefore be used. Second, since survey respondents do not have to act on the basis of their answers, the results are hypothetical.

[6] In the empirical analyses in chapters two to four, median levels of income and tax base are not available. Mean community income and mean community tax price (obtained by using the mean value of the property) are used instead in the local public expenditure determination equation. As stressed by Mueller (1989), however, this model yields different predictions than a median voter model, only if the ratio of median to mean incomes differs significantly across communities, and these differences are important in determining the demand for public goods.

1.2 Central government grants and local government spending

In a system of local government, it is usually the case that central government subsidises local authorities' expenditures. I am mainly interested here in the consequences on the estimation of an equation of local public expenditure determination of a grant distribution system that alters the community's budget constraint.

The two most common models of the effects of grants on local expenditures are those which consider either an open-end matching grant - under which expenditures by the local government are subsidised at a constant rate - or a lump-sum grant - under which the local government is given a fixed amount of money, independent of its own level of spending. While in the former model the slope of the local community's budget constraint changes, in the latter model the budget constraint shifts outward. However, grants are often distributed according to much more complicated formulae, which can create a series of kinks - some convex and some non-convex. For instance, suppose that grants from central government to each local community (i=1,...,n) take the form of a lump-sum grant of the amount L_i and of matching grants at rate m_i (Wildasin, 1986). Assume also that both can depend on the level of per capita expenditure x_i. The local authority's budget constraint (1.2) is therefore transformed in the following way:

$$c_i + (1 - m_i(x_i))x_i = y_i + L_i(x_i) \qquad (1.3)$$

If the matching rate of the grant decreases if expenditure exceeds a threshold (x^T), the budget constraint is a piecewise-linear one:

$$c_i + (1 - m_i^1)x_i = y_i + L_i^1 \quad \text{if} \quad x_i \le x^T \qquad (1.4)$$

$$c_i + (1 - m_i^2)x_i = y_i + L_i^2 \quad \text{if} \quad x_i > x^T \qquad (1.5)$$

where: $L_i^2 = L_i^1 + (m_i^1 - m_i^2)x^T$. Any further reduction in the matching rate of the grant increases the price of public services to local residents, and creates additional convex kinks in the authority's budget constraint. However, the budget constraint can in some instances become non-convex. This occurs, for instance, when the matching grant rate is negative, local expenditure is so high as to lead to the loss of all entitlement to grant, and the local community has to rely entirely on local resources (see chapter two).

The presence of a piecewise-linear constraint creates three main problems for the estimation of an equation of local public expenditure determination. First, there might be non-standard price and income effects, because an authority can jump from one segment to another in response to changes in the parameters of the grant distribution formula, or might stick to a convex kink. Second, observationally equivalent authorities facing an identical budget constraint tend to spread themselves over different points along that constraint. Since, as a result, they respond in different ways to changes in the parameters of the grant distribution formula, the stochastic specification must allow for unobserved heterogeneity of preferences. Third, a demand function for local public expenditure in the presence of a kinked constraint cannot be estimated by OLS, by attributing to each local government the price and income variables of the segment on which it is observed to be located. The error term is correlated with the price and income variables, both if it represents unobserved preference heterogeneity and if it represents random optimisation error. Consequently, an alternative approach must be used. The most satisfying approach consists in an econometric model that has two error terms. One error term represents unobserved heterogeneity, and is interpreted as representing variation in utility functions across local authorities. The second error term represents optimisation or measurement error.

Burtless and Hausman (1978) and Hausman (1985) formulate a model of consumer choice which takes explicit account of non-linearities in the budget set, and discuss some of the econometric issues involved in the estimation of a labour supply function. Moffitt (1986, 1990) presents the two-error ML procedure and compares it to alternative estimation methods,[7] while Moffitt (1984) applies this method to the estimation of a demand function for local public expenditure in the US, where the grant distribution system creates a piecewise-linear non-convex budget constraint. Barnett *et al.* (1991, 1992) study the English local governments' responses to central government grants and restrictive policies. They estimate a demand function for local public expenditure for each of the fiscal years 1983/84 to 1988/89, and carefully model the system of expenditure targets and penalties that was in place until 1986. They assume that local governments maximise a local social welfare function, and allow for heterogeneity amongst local authorities by adopting a two-error Maximum Likelihood technique.

A more detailed analysis of the problems posed by a non-linear and non-convex budget constraint is presented in chapter two, with an application to the English non-metropolitan authorities' expenditures in the 1980s.

1.3 Interjurisdictional spillovers and spatial interactions

When estimating a demand function for local public expenditure, it is usually assumed that the observations on the N local jurisdictions are independent. For instance, a demand function for local public expenditure could be written as:

$$x_i = z_i'\alpha + u_i \qquad (1.6)$$

where z_i is a vector of characteristics of locality i. The error term u in the regression model is assumed to be independently and identically distributed across local governments, and the level of public expenditure in a local jurisdiction is assumed not to be affected by the expenditures in neighbouring jurisdictions.

---

[7] An alternative to the two-error ML procedure consists in finding instruments for the price and virtual income variables - that is variables that are correlated with them but uncorrelated with the error term. However, the weakness of the IV approach is that it does not easily address the problem of kink locations.

The first reason why the above model may not be appropriate for studying local government expenditure is that the error term u might be spatially correlated. Local jurisdictions can be subject to shocks that affect their expenditure decisions, and are spatially autocorrelated - such as regional shocks to income. Alternatively, those common shocks may be the result of central government regional policies or intermediate level of government - e.g., the counties in the UK - fiscal policies. Spatially correlated shocks can be allowed for by appropriately modelling a spatial process in the error term (Anselin, 1988a).

A second, and more serious, problem arises if the taxation and expenditure decisions in a local jurisdiction have spillover effects into neighbouring jurisdictions. First, in the presence of interjurisdictional mobility of the tax base, the tax in a locality affects the size of both the own and the nearby jurisdictions' tax base. A large number of studies in the "tax competition" literature have explored the consequences of mobility of the tax base on local government expenditure and taxation decisions. When the tax base is mobile, local jurisdictions' taxation decisions are interdependent, in that local governments could actively compete against one another for the mobile tax base. The notion that local governments and regions compete by varying the tax rates dates back at least to Oates (1972). Most of the subsequent literature on tax competition has relied on a sort of "perfect competition" model, which is characterised by a perfectly mobile tax base - capital - and a large number of jurisdictions.[8] This implies that each locality faces an infinitely elastic supply of capital, and cannot affect its national rate of return. If both a head tax and a capital income tax can be used to finance local public good provision, welfare-maximising governments will set the capital income tax to zero and finance public good expenditure through the head tax - that is, local taxes become benefit taxes. However, if the head tax is not available, local public goods will, in general, be underprovided. Each jurisdiction does not take into account the flow of capital into other jurisdictions which results from its own taxation.[9]

---

[8] Beck (1983), Wilson (1985, 1986), McLure (1986), Zodrow and Mieszkowski (1986), and Oates and Schwab (1988, 1991) are classic examples.

[9] A similar model of tax competition for freely mobile capital among many identical small jurisdictions is analysed in Edwards and Keen (1994). They consider both the case in which the

On the other hand, some studies have analysed the "small number" case, where a few jurisdictions compete against one another, and the policies of each of them affect the net return to capital. Oates and Schwab (1991) clearly state the distinction between the "perfect competition" and the "small number" case, while concentrating on the former. Wildasin (1988, 1989), Mintz and Tulkens (1986), de Crombrugghe and Tulkens (1990), and Kanbur and Keen (1993) analyse tax competition between a small number of regions. It turns out that optimality requires different conditions as compared with the perfect competition case. The main idea underlying this sort of model is that an increase in a region's tax rate raises the tax base available to the other region - whether the regional tax is an origin-based commodity tax or a capital income tax. As a result, the optimal tax in a region is a function of the taxes in the other region(s), and the tax rates will be correlated. By defining an objective function for the regional governments, a reaction function can be obtained that relates a region's tax to the other region's tax.

Furthermore, there is a second reason why local governments' decisions are interdependent. Public expenditures in one jurisdiction might have external effects on neighbouring jurisdictions. For instance, the benefits of public spending in one locality may spill over into another. The theoretical implications of such spillovers have long been analysed in the literature. In the presence of spillovers generated by a community, Williams (1966) demonstrates that only by chance will the decentralised result be a socially optimal one. The most likely outcome is an underprovision of the public good in the aggregate, even though the income effects of spillovers may more than compensate the substitution effects, leading to a situation in which too much of the public good will be provided in the aggregate. Oates (1972), Boskin (1973), and Gordon (1983) describe systematically the types of externalities that a local

regional government can raise a non-distorting tax to finance the provision of the local public good (setting the capital tax equal to zero and providing an optimal amount of the public good), and the case in which only the distorting tax is available. In the latter case, the evaluation of the effects and consequences of coordination and harmonisation turns out to depend crucially on the objective function of the policymakers. In a slightly different context, Keen (1987) tries to evaluate the welfare effects of commodity tax harmonisation, that is a movement towards a common rate structure. This implies computing some kind of average of pre-existing tax structures. He proves that a harmonisation process, starting from a situation where domestic tax systems are different, results in a welfare gain, that is represents a strict potential Pareto improvement.

government can create for non-residents, analyse the problems that can arise from the decentralised form of decision-making, and explore what the central government can do to lessen the social costs.[10]

However, despite a substantial theoretical literature on interregional expenditure spillovers, there has been little empirical work aimed at estimating the direction and magnitude of those external effects. In fact, only recently have spatial econometric techniques - developed by Cliff and Ord (1981), Anselin (1988a), and Cressie (1991) - been applied to the estimation of spatial interactions between local governments. Case *et al.* (1993) estimate an expenditure equation using a panel data set of the US states' budgets, where spatial correlation in both the dependent variable and in the errors is allowed for. They attribute the presence of spatial effects to the fact that benefits from public expenditures in one state may spill over into neighbouring states. Besley and Case (1995) use data on US states' income taxes from 1960 to 1988 in order to investigate whether neighbouring states' tax changes are correlated with a given state's tax change. A state's tax change is linearly related to state-specific variables and neighbouring states' tax changes. Both two stage least squares and maximum likelihood estimates suggest that neighbours' tax changes have a positive and significant effect on a given state's tax change. Kelejian and Robinson (1993) propose a regional public expenditure model with spatial correlation in the dependent variable, and with two stochastic shocks generated within each region. While one is unique to the region, the other generates spillover effects in other regions. When estimating an equation for police expenditures in a number of US counties - where the spatial correlation component in the error term has been set to zero - they find that police expenditures in a given county are significantly and positively influenced by neighbouring county police expenditures. If counties inflict a negative externality on their neighbours by spending more on police services, the need for police services in a given county tends to increase as such services in neighbouring counties increase.

---

[10] More recently, Wellisch (1994) analyses the optimality properties of decentralised provision of public goods generating spillovers, in the presence of perfect and imperfect household mobility.

As regards the UK, Bivand and Szymanski (1997) analyse the pattern of spatial dependence in local public service provision cost. They present a principal agent model that suggests that yardstick competition regulation of natural monopolies generates spatial dependence if local principals (English local authorities) can pursue idiosyncratic policies when they contract with local agents (DSO, Direct Service Organisation). An externality arises because contracts are often based on comparisons of performance against neighbouring jurisdictions. By comparing the pattern of spatial dependence in the provision of a local public service (garbage collection) before and after the introduction of Compulsory Competitive Tendering (CCT), they show that CCT, by imposing standard contracting rules, substantially reduces the scope for local authorities to pursue idiosyncratic policies, and, as a result, the extent of spatial correlation.[11]

An empirical analysis of interjurisdictional spillovers requires explicit modelling of spatial dependence, and the estimation procedure of the equation of local public expenditure determination must take into account that local jurisdictions' policies are determined simultaneously. Assume, for example, that preferences in locality i can be represented by the local social welfare function (1.7), defined over private good and public good consumption. x_i^n is public expenditure in neighbouring jurisdictions,[12] and u_i is an unobservable shock to local income. All variables are in per capita terms, and local income y_i might be inclusive of grants from central government.

$$\omega_i = \omega(y_i + u_i - x_i, x_i + \phi x_i^n) \qquad (1.7)$$

[11] Besley, Preston and Ridge (1997) allow for spatial dependence in local tax non-payment rates in England. They model poll tax non-payment rates in a short panel of data on the 366 English metropolitan and shire districts. Their analysis allows for neighbourhood influences across authority boundaries, that is the level of non-compliance in a district is a function of non-compliance in geographically contiguous districts.

[12] The way neighbours' spending is defined clearly depends on the underlying economic model. Discussion of the choice of the neighbourhood criterion and of the appropriate "weighting" procedure is in chapter three.

Spatial interactions between neighbouring jurisdictions can occur in two forms. First, expenditures in neighbouring jurisdictions can spill over into jurisdiction i, and the parameter ϕ measures the extent to which this occurs. $\phi>0$ means that expenditures in other jurisdictions are a substitute for expenditures in jurisdiction i, while $\phi<0$ means that they are complementary. Second, the shock to local income can be spatially correlated, as in (1.8):

$$u_i = \lambda u_i^n + \xi_i \tag{1.8}$$

where u_i^n is the shock in neighbouring jurisdictions, ξ_i is a spatially uncorrelated mean-zero error term, and $-1<\lambda<1$ is the spatial autoregressive parameter.

An equation for local public expenditure determination can be obtained by choosing a specific form for the local preferences. For instance, if the local social welfare function is a Cobb-Douglas with parameters β_0 and β_1 for private and public consumption respectively, the optimum level of expenditure in jurisdiction i is:

$$x_i = \gamma y_i + \rho x_i^n + \varepsilon_i \tag{1.9}$$

where: $\gamma \equiv \beta_1/(\beta_0+\beta_1)$, $\rho \equiv -\phi\beta_0/(\beta_0+\beta_1)$, and $\varepsilon_i \equiv \gamma u_i$.

This stylised model shows that public expenditures will be spatially correlated if either ϕ or λ is different from zero. The aim of the analysis in chapter three is to try and determine the extent to which correlation in local public expenditure is the result of genuine behavioural interdependence (ϕ) or rather the result of spatially correlated shocks (λ).

1.4 Non welfare-maximising governments

In sections **1.1** through **1.3**, it has been assumed that governments are benevolent maximisers of their citizens' welfare. Even if officials are only interested in being re-elected for their own well-being - as in Downs (1957) - they must still please voters and try to gain their support.

However, there exists a radically different view within the public choice literature that stresses the role of information in an uncertain environment, and sees governments as monopolists taking advantage of their position in the supply of public goods (Breton, 1974), or as intrinsically untrustworthy revenue-maximisers (Brennan and Buchanan, 1980). According to this view, the relationship between voters and incumbent governments in a representative democracy can be described as a principal-agent one.[13] Officials are the agents and the voters their principals. Once elected, the former are asked to make decisions in the interest of the latter, about, among other things, taxation and public spending. Agency problems arise because incumbents know more about their own level of administrative ability or competency, and about the cost of providing public services, than do voters.

It has been argued that competition - usually in the sense of competition between incumbent and opposition parties - as well as reputation, monitoring, and optimum contract design, should be able to reduce the potential for opportunism by politicians. Consequently, the principal-agent problem in the political sector may be no more severe than in the private sector.[14] However, the important role played by private information in public services provision is well documented in the literature. Recent analyses have focused on optimum selection processes in the presence of imperfectly observable government actions.[15]

It has also been suggested that the existence of competing governments can put a constraint on the power of the government. Decentralised decision-making is seen as a way of introducing competition into politics, and of forcing representatives into behaving efficiently. First, the Tiebout (1956) model implies that citizens can move among competing local jurisdictions. Second, in the presence of asymmetric information between voters and representatives, comparative performance evaluation

---

[13] Recent works following this approach are Besley and Case (1995a, 1995b), Kalt and Zupan (1990), Peltzman (1990, 1992), and Rogoff (1990). Fratianni *et al.* (1993) describe the relationship between the general public and the central bankers as a principal-agent problem.

[14] As argued by Wittman (1989).

[15] Azariadis and Lahiri (1997) is a recent example.

can improve incentives and efficiency in principal-agent relationships. Rewards based on relative performance can be superior to payments based on individualistic performance. Salmon (1987) suggested that one of the main advantages of decentralisation is that of providing local administrators with incentives for behaving efficiently and reducing managerial slack, or X-inefficiency. Voters can use information on local taxes and public services in nearby jurisdictions, when deciding whether or not to re-elect the incumbent government.[16]

When taking into account that voters' and representatives' objectives can be different, the simple model of local public expenditure determination must clearly be amended. If politicians differ in their level of ability or competency, the budget constraint (1.3) should be rewritten as:

$$c_i + (1 - m_i(x_i) + \mu_i)x_i = y_i + L_i(x_i) \qquad (1.10)$$

where μ_i is a parameter reflecting the efficiency of the local government in jurisdiction i in the provision of public services. If the government is still welfare-maximising, the parameter μ_i is not a choice variable for a representative, but an individual characteristic (as in Rogoff, 1990). Voters appraise the incumbent government's competency by looking at their performance in local public good provision. On the other hand, μ_i may as well be interpreted as "governmental profit", represented by the excess of local taxes over service costs (Epple and Zelenitz, 1981). If the government's utility does not coincide with the community's welfare, but rather depends on governmental profit, μ_i becomes a choice variable for the government. Of course, the government cannot claim unbounded profits from the inhabitants of the community, if only because of the household's budget constraint and of the possibility of migration to another jurisdiction if taxes are too high. The second constraint

[16] Besley and Case (1995) present a model of yardstick competition in taxes, in which voters make comparisons between states, and incumbents look to other states' taxing behaviour before changing taxes at home. By using US state data, they find that electoral defeat of an incumbent is positively correlated with a tax increase in the incumbent's own state, and negatively correlated with tax increases in neighbouring states. They also find that tax changes in neighbouring states tend to be positively correlated.

limiting governments' choices is the electoral mechanism. However, as long as electors cannot separately observe all the variables affecting the cost of public services, politicians have the chance of enjoying some governmental profit. In a representative democracy, elections are held at regular intervals, when electors decide whether to re-elect the incumbent government or vote for the opposition candidate. The probability of re-election of the incumbent depends on the electorate's welfare. On the other hand, the higher the governmental profit, the lower the jurisdiction's welfare. The incumbent therefore faces a trade-off between governmental profit and probability of re-election.

Chapter four takes this simple model as the starting point and looks at its implications for the tax-setting decisions of the incumbents and the voting decisions of the electors. Further, the model in chapter four augments the above analysis by looking at the interactions that emerge between local governments in the presence of asymmetric information between voters and incumbents, when local jurisdictions are subject to correlated shocks. Due to spatial correlation in the cost shock, it is optimal for voters to appraise the government's quality by looking at their jurisdiction's relative performance. As a result, both own and neighbours' taxes will have an influence on the votes cast for the incumbent party.

In chapter four, I look at the effects of local taxes on local election results, by using a panel data set of the English local governments. First, the spatial character of the data is exploited in order to find out whether the chances of re-election of an incumbent are affected, besides his own tax, by his neighbours' taxes. Second, I control for the influence of national politics, that is I check whether local election results are affected by the popularity of the central government party.

Chapter Two

Local public expenditure with piecewise-linear non-convex budget constraints

2.1 Introduction

This chapter[1] analyses local authority expenditure decisions when local governments have two sources of funding. In the first place, they can raise revenues by setting a "rate" on a local tax base. The local tax is a property tax, levied on both domestic and non-domestic hereditaments. The second source of revenue is grants from central government. Grants are distributed to local governments according to a formula that is intended to equalise both for differences in availability of local resources and for differences in needs. The presence of a lump-sum grant and of a matching grant - where the matching rate changes with the level of expenditure - makes the budget constraint that local authorities face non-linear. Furthermore, for high levels of the local tax base, the marginal grant rate can be negative and lead to the loss of all entitlement to block grant for high levels of spending ("grant exhaustion"). This implies that for some authorities the budget constraint is a piecewise-linear non-convex one. The two-error Maximum Likelihood procedure developed by Burtless and Hausman (1978) and Hausman (1985) for the study of labour supply with piecewise-linear budget constraints, and by Moffitt (1984) as an application to the AFDC benefits in the US, is applied here in order to consistently estimate a demand function for local public expenditure in England in the fiscal years 1986/87 to 1989/90.

To our knowledge, no previous work in this area tackles the problem of explaining local authority expenditure decisions in the face of the actual piecewise-linear non-

---

[1] The material covered in this chapter draws on Revelli (1998).

convex budget constraints that British local authorities face. Barnett, Levaggi and Smith (1992) study British local governments' expenditure decisions when the grant system produces a piecewise-linear budget constraint. They estimate a demand function for local public expenditure for each of the fiscal years 1983/84 to 1988/89, and carefully model the system of expenditure targets and penalties that was in place until 1986. They assume that local governments maximise a local social welfare function, and allow for heterogeneity amongst local authorities by adopting a two-error Maximum Likelihood technique. Their analysis differs from this in two main respects. First, they consider the English shire counties and metropolitan districts, while we look at the much larger cross-sections of the shire districts. Second, they rule out the grant-exhaustion case and just consider convex budget constraints. They correctly argue that for most of the local authorities they consider, no non-convex kink featured in their budget constraints. However, as far as the non-metropolitan districts are concerned, the number of grant-exhausted authorities steadily increased during the second half of the 1980s (see appendix 2.4, table A2.5), and excluding them from this analysis might create potentially serious selection problems.

The plan of the work is as follows. Section **2.2** describes the operation of the Block Grant system and formalises the budget constraint that local authorities face. Section **2.3** builds up a basic model of local authority expenditure decisions. Section **2.4** examines the econometric procedure used to estimate a demand function for public expenditure when local authorities face non-linear and non-convex budget constraints, and section **2.5** presents the results. Section **2.6** concludes.

2.2 The Block Grant system

The kind of intergovernmental fiscal relations we model are those of the actual rates' system that was in operation in the UK from 1981 - when the reformed Block Grant was introduced - to 1990.[2] Since the system rarely operated in its pure form during that decade, we will refer to the years during which a broadly stable set of rules was in operation. Actually, in the last four years of operation of the local rates (1986/87 to 1989/90), the system came back to its pure form, since expenditure targets and penalties were abolished.[3]

Local expenditures are funded through a local property tax - the rate - and through grants from central government, the Block Grant. According to the Block Grant system, each local authority set a rate (number of pence in the pound) which applied to all properties, with the only difference of a domestic rate relief (a rebate of 18.5 pence per pound) which was determined by the central government. The tax base was represented by the rateable value - a sort of imputed rental value - of the property.

The Block Grant is conceptually the combination of a lump-sum grant and of a matching grant. Two crucial variables must be defined in order to understand how the Block Grant system worked. The first one is the GRE - Grant Related Expenditure assessment. It is the central government's estimate of the level of expenditure required to offer a standard level of public services, given the local jurisdiction's social and economic needs. The second one - the GRP (Grant Related Poundage) - is the tax rate applying to the local tax base. The GRP is a non-linear function of the level of spending. For spending below a threshold - set at 10% above GRE - the GRP increases with expenditure at a constant rate, which expresses how much local taxation must increase to match increases in local spending. The higher the slope, the higher the cost of public services to local residents, and the lower the central government's marginal contribution to it. For spending above the threshold, the GRP schedule becomes steeper, meaning that the marginal cost of additional

---

[2]In 1990, the non-domestic rate was replaced by a national non-domestic rate (NNDR), while the domestic rate was replaced by the Community Charge - which was still set at the local level. Due to its enormous unpopularity, the Community Charge was later replaced by the Council Tax.

[3]A system of targets and penalties was effective during the first half of the decade and it had been introduced in order to curb local expenditures. Given the targets set by the central government, for each percentage point of excess expenditure with respect to the target, an authority had a corresponding abatement in the grant. Penalties became more and more severe during the 1980s, and were at last abolished in 1986, after the introduction of rate-capping. For the present study, rate-capping does not really represent a problem, because a negligible proportion of the non-metropolitan districts were rate-capped in the period we consider (see appendix 2.4, table A2.4).

local expenditure to the local tax-payer is higher. The typical local jurisdiction faced a piecewise-linear budget constraint, with two segments and a kink at a threshold level of spending (*THR*), as depicted in figure 2.1.[4] *X* represents local public expenditure per head and *t* the local rate. Along segment 1, the local authority's post-grant budget constraint is given by:

$$X_j = M_{1j} + B_1 t_j \qquad (2.1)$$

where: M_{1j}=GRE_j-$B_1 t^*$, and t^* is the standard rate poundage for spending at GRE. M_{1j} can be seen as the lump-sum component of the grant along the first segment. B_1 can be interpreted as an implicit standard tax base along segment 1. A local authority locating on the first segment of its budget constraint will have a positive (negative) matching grant if its per capita tax base - B_j - is smaller (larger) than the implicit tax base B_1. For spending above the threshold (segment 2), the local authority's post-grant budget constraint is given by:

$$X_j = M_{2j} + B_2 t_j \qquad (2.2)$$

where: M_{2j}=THR_j-$B_2 t^{**}$, and t^{**} is the standard rate poundage for spending at the threshold. M_{2j} and B_2 have a similar interpretation as the corresponding parameters on the first segment. However, local authorities that receive a negative matching grant can achieve a level of spending where they are "grant-exhausted", that is they do not receive any grant at all. This exhaustion level of spending (*EX*) can be either smaller or larger than the threshold level of spending (*THR*), giving rise to two different budget constraints.

If the exhaustion level of spending is smaller than the threshold, local authorities become grant-exhausted before reaching the threshold, and the budget constraint ends up having only two segments. The second segment is represented by a ray

[4] In fact, a further convex segment can be added to the budget constraint (segment $0M_{1j}$ in figure 2.1). This would occur only in the very unlikely event that expenditure is low enough to be met entirely by grant, and the tax rate is zero. For ease of exposition, our analysis is conducted in terms of budget constraints that do not include this additional segment.

emanating from the origin, whose slope is equal to the local authority's per capita tax base. Since local authorities receive no grant in this case, the level of spending per head must equal local taxes raised per head:

$$X_j = B_j t_j \qquad (2.3)$$

If grant is distributed according to the grant distribution formula we have described, grant per head is given (for spending along segment 1) by:

$$K_j = M_{1j} + (B_1 - B_j)t_j \qquad (2.4)$$

The first term on the right hand side represents the implicit lump-sum grant and the second term on the right hand side the implicit matching grant. It is then easily seen that the funding of local expenditure comes from three sources: lump-sum component of the grant, matching grant, and locally raised funds:

$$X_j = K_j + B_j t_j = M_{1j} + (B_1 - B_j)t_j + B_j t_j \qquad (2.5)$$

But when the local tax base is larger than the tax base implicit in the first segment of the budget constraint, the matching grant is negative and, for spending larger than *EX*, an authority ends up receiving no grant at all (figure 2.2). For the grant-exhaustion level of spending (*EX*) to be smaller than the threshold level of spending - that is, for the budget constraint to have only two segments - the local tax base must be such that:

$$B_j > \frac{M_{1j} + B_1 t^{**}}{t^{**}} = \frac{THR_j}{t^{**}} \qquad (2.6)$$

The two-segment after-grant budget constraint can be expressed as (figure 2.2):

$$X_j = M_{1j} + B_1 t_j \qquad \text{if} \qquad X_j < EX_j = B_j (B_j - B_1)^{-1} M_{1j} \qquad (2.7)$$

$$X_j = B_j t_j \qquad \text{if} \qquad X_j > EX_j = B_j (B_j - B_1)^{-1} M_{1j} \qquad (2.8)$$

On the other hand, if the exhaustion level of spending is larger than the threshold, the budget constraint is made up of three segments. The third segment is represented by a ray emanating from the origin, whose slope is equal to the local authority's per capita tax base. For the grant-exhaustion level of spending (*EX*) to be larger than the threshold level of spending - that is, for the budget constraint to have three segments - the local tax base must be such that:

$$B_2 < B_j < \frac{M_{1j} + B_1 t**}{t**} = \frac{THR_j}{t**} \qquad (2.9)$$

The three-segment post-grant budget constraint can now be expressed as (figure 2.3):

$$X_j = M_{1j} + B_1 t_j \qquad \text{if} \qquad X_j < THR_j \qquad (2.10)$$

$$X_j = M_{2j} + B_2 t_j \qquad \text{if} \qquad THR_j < X_j < EX_j = B_j (B_j - B_2)^{-1} M_{2j} \qquad (2.11)$$

$$X_j = B_j t_j \qquad \text{if} \qquad X_j > EX_j = B_j (B_j - B_2)^{-1} M_{2j} \qquad (2.12)$$

The consequence of grant-exhaustion on the budget constraint is that it turns out not to be convex any more, leaving the possibility open of perverse comparative statics - income and price effects - that is of non-standard responses to central government's policies.

2.3 Local authority expenditure decisions

Assume that preferences in locality *j* can be represented by a continuous, strictly quasi-concave utility function *u* - as in (2.13). *u* is non-decreasing in its arguments, consumption of local public services per head, X, and per capita consumption of private goods, C. Y represents per capita income.[5]

$$u(X_j, C_j) = u(X_j, Y_j - T_j) = u(X_j, Y_j - t_j B_j) \tag{2.13}$$

Local authorities choose the levels of taxation and expenditure in order to maximise local welfare, subject to the non-linear constraint described in section **2.2**. Let us reformulate the budget constraint in terms of public expenditure per head and after-tax income. Along segment i of the budget constraint (i=1,2,3), the trade-off between after-tax income and public expenditure can be expressed as:

$$C_j = Y_j - t_j B_j = (\beta^*_{ij} M_{ij} + Y_j) - \beta^*_{ij} X_j = I_{ij} - \beta^*_{ij} X_j \tag{2.14}$$

where M_{ij} and B_{ij} indicate the intercept and the implicit tax base of the relevant segment of the budget constraint, respectively. Along segment 1, $M_{ij}=M_{1j}$ and $B_{ij}=B_1$; along segment 2, $M_{ij}=M_{2j}$ and $B_{ij}=B_2$; and along the third ("grant-exhausted") segment, $M_{ij}=0$ and $B_{ij}=B_j$.

$\beta_{ij}^*=B_j/ B_{ij}$ can be interpreted as the price of public services to local residents, and the tax base includes both the personal and the business sector.[6] I_{ij} is virtual income corresponding to segment i.

In the public consumption - private consumption space, we can represent a local authority's optimisation problem in figures 2.4 to 2.6. Let us start from figure 2.4. In the absence of grants, the relevant budget constraint is represented by DF'.

[5] Barnett *et al.* (1992) assume that local authorities maximise a Cobb-Douglas local social welfare function, defined over the excess of expenditure per head over a basic minimum, and the gap between actual per capita tax payment and the maximum possible. In the absence of local income data, they proxy the income available for taxation in a local authority by the maximum tax rate levied, multiplied by the local per capita tax base. In this analysis, we use county level disposable income per head as a proxy for district level disposable income per head.

[6] This implies that the burden of the local business tax is shifted onto the domestic sector. Though a complete tax incidence analysis is beyond the scope of this work, it would be naive to assume that local businesses bear the entire burden of the local tax. However, we also estimated a model where the price of public services to local residents is given by the ratio of the domestic tax base to the implicit tax base of the relevant segment of the budget constraint. The coefficient estimates are not significantly different from the ones presented in section **2.5**.

Maximising local welfare would yield an equilibrium such as G. The grant distribution system shifts the budget constraint to DAEF, and a local authority will locate either on the first segment, or on the second segment, or at the kink (E). For the reasons given in section **2.2**, we do not model the choice of the convex kink at A. The price of public services increases after the outward kink at point E. The convex budget constraint is described by:

$$C_j = I_{1j} - \beta^*_{1j} X_j = (\beta^*_{1j} M_{1j} + Y_j) - \beta^*_{1j} X_j \quad \text{if} \quad X_j \leq THR_j \qquad (2.15)$$

$$C_j = I_{2j} - \beta^*_{2j} X_j = (\beta^*_{2j} M_{2j} + Y_j) - \beta^*_{2j} X_j \quad \text{if} \quad X_j > THR_j \qquad (2.16)$$

In figure 2.5, authority *j* is grant-exhausted after the inward kink at F, and the budget constraint is:

$$C_j = I_{1j} - \beta^*_{1j} X_j = (\beta^*_{1j} M_{1j} + Y_j) - \beta^*_{1j} X_j \quad \text{if} \quad X_j \leq THR_j \qquad (2.17)$$

$$C_j = I_{2j} - \beta^*_{2j} X_j = (\beta^*_{2j} M_{2j} + Y_j) - \beta^*_{2j} X_j \quad \text{if} \quad THR_j < X_j < EX_j \qquad (2.18)$$

$$C_j = I_{3j} - \beta^*_{3j} X_j = Y_j - \beta^*_{3j} X_j \quad \text{if} \quad EX_j < X_j \qquad (2.19)$$

Finally, an authority such as the one depicted in figure 2.6 faces a two-segment non-convex budget constraint, since the inward kink occurs along the first segment of the budget constraint:

$$C_j = I_{1j} - \beta^*_{1j} X_j = (\beta^*_{1j} M_{1j} + Y_j) - \beta^*_{1j} X_j \quad \text{if} \quad X_j < EX_j \qquad (2.20)$$

$$C_j = I_{3j} - \beta^*_{3j} X_j = Y_j - \beta^*_{3j} X_j \quad \text{if} \quad EX_j < X_j \qquad (2.21)$$

2.4 Estimation procedure

A demand function for local public expenditure - derived from the optimisation problem described in section **2.3** - cannot be estimated by OLS, by attributing to each local government the price and income variables of the segment on which it is observed to be located. The error term is correlated with the price and income variables, both if it represents unobserved preference heterogeneity and if it represents random optimisation error. If the error term represents unobserved preference heterogeneity, then it affects the slope of the authority's indifference map - the marginal rate of substitution between private consumption and public goods - and it is correlated with the price and income of the segment along which the authority is located. For instance, for a two-segment convex budget constraint, the higher price and virtual income will be observed only if X is high, simply because of the nature of the constraint. Consequently, causality runs from the choice of X to the choice of segment, rather than vice versa. On the other hand, suppose that the random term represents optimisation error. If the error term is sufficiently large, it might move the observed value of X to the segment which is characterised by high price and high virtual income. Hence, the error term is correlated with the parameters of the observed segment. The stochastic structure must then incorporate both sources of error, and the choice of segment must be formally modelled and estimated.

Let us first consider the case of a linear budget constraint. The demand function for local public expenditure in a jurisdiction (j) facing a linear budget constraint with price β_j^* and income I_j can be written as:

$$X_j = f(\beta_j^*, I_j, P_j) + v_j \tag{2.22}$$

where P_j is a vector of characteristics specific of jurisdiction j, and v_j is the error term. The random term v_j can be interpreted as arising from two different sources.

The first one - that we call ε_j - deals with the fact that there may be a divergence between optimal or desired level of public expenditure X_j* and actual level X_j. This random difference can be due either to measurement error in the left-hand side variable, or to unpredicted shocks at the local government level that drive the level of spending away from its optimal level (optimisation error).

The second form of randomness arises from unobserved preference heterogeneity of local governments. Let h_j be the term representing unobserved differences among local authorities, after controlling for a set of local characteristics, P_j.
In order to see how the two random terms enter the demand function, let us specify - following Burtless and Hausman (1978), and Moffitt (1984) - the following constant elasticity demand function:

$$X_j * = \omega_j \left(\beta_j^*\right)^{a_\beta} \left(I_j\right)^{a_I} \qquad (2.22')$$

where X_j* is optimal or desired local public expenditure per head, and ω_j reflects jurisdiction-specific characteristics. Parameters a_β and a_I are assumed constant across local governments. The indirect utility function corresponding to this demand function is:

$$V_j = \frac{\left(I_j\right)^{1-a_I}}{1-a_I} - \omega_j \frac{\left(\beta_j^*\right)^{1+a_\beta}}{1+a_\beta} \qquad (2.23)$$

Since ω_j reflects jurisdiction j's characteristics, and since not all of them are observable, let the preference heterogeneity term (h) enter the demand function in the following way:

$$\omega_j = \exp(a_P' P_j + h_j) \qquad (2.24)$$

Taking the logarithm of equation (2.22'), and using the expression for ω_j:

$$x_j * = a' z_j + h_j = a_P' P_j + a_\beta \ln \beta_j^* + a_I \ln I_j + h_j \qquad (2.25)$$

where:

$x_j{*} = \ln X_j{*}$
$z_j' = [P_j', \ln \beta^*_j, \ln I_j]$
$a' = [a_P', a_\beta, a_I]$

As we said, the actual level of expenditure will diverge from the optimal or desired one because of random optimisation error. As a result, the observed level of public expenditure is given by:

$$x_j = x_j{*} + \varepsilon_j \qquad (2.26)$$

We assume that h and ε are distributed normally and independently with means zero and variances σ_h^2 and σ_ε^2. Clearly, since the two random terms are both additive, their respective variances are not separately identifiable with a linear constraint.
Now consider what happens with kinked constraints. First, take the case of a local government facing a convex budget constraint such as the one depicted in figure 2.4. A local government having preferences represented by indifference curve u_1 will have its optimum at E, where u_1 is tangent to segment 1. Denote the unobserved heterogeneity term corresponding to indifference curve u_1 by $\underline{h_j}$. The optimal level of expenditure (which coincides with the threshold) can be expressed as:

$$thr_j = a' z_{1j} + \underline{h_j} = a_P' P_j + a_\beta \ln \beta^*_{1j} + a_I \ln I_{1j} + \underline{h_j} \qquad (2.27)$$

where: $thr_j = \ln THR_j$

A local government having preferences represented by indifference curve u_2 will have its optimum at E as well, where u_2 is tangent to segment 2. If $\overline{h_j}$ is the value of the heterogeneity error that makes the desired level of spending along segment 2 equal to the threshold, optimal spending can be written as:

$$thr_j = a' z_{2j} + \overline{h_j} = a_P' P_j + a_\beta \ln \beta^*_{2j} + a_I \ln I_{2j} + \overline{h_j} \qquad (2.28)$$

All authorities having an unobserved preference heterogeneity term larger than $\underline{h}_j$ and smaller than $\overline{h_j}$ will locate at the kink:

$$x_j{}^* = a' z_{1j} + h_j \qquad \text{if} \qquad h_j < \underline{h}_j \tag{2.29}$$

$$x_j{}^* = thr_j \qquad \text{if} \qquad \underline{h}_j \le h_j \le \overline{h_j} \tag{2.30}$$

$$x_j{}^* = a' z_{2j} + h_j \qquad \text{if} \qquad \overline{h_j} < h_j \tag{2.31}$$

Intuitively, the heterogeneity term *h* affects the slope of the indifference curve - the marginal rate of substitution - of a local authority. The higher *h*, the steeper the indifference curve and the stronger the preference for public services with respect to private goods.

For a local authority facing the non-convex constraint in figure 2.5, the optimal level of expenditure will be determined according to:

$$x_j{}^* = a' z_{1j} + h_j \qquad \text{if} \qquad h_j < \underline{h}_j \tag{2.32}$$

$$x_j{}^* = thr_j \qquad \text{if} \qquad \underline{h}_j \le h_j \le \overline{h_j} \tag{2.33}$$

$$x_j{}^* = a' z_{2j} + h_j \qquad \text{if} \qquad \overline{h_j} < h_j < \underline{\underline{h}}_j \tag{2.34}$$

$$x_j{}^* = a' z_{3j} + h_j \qquad \text{if} \qquad \underline{\underline{h}}_j < h_j \tag{2.35}$$

where $\underline{\underline{h}}_j$ is the heterogeneity term corresponding to indifference curve *uu*, that is it is the value of h_j which equilibrates (indirect) utility along segments 2 and 3:

$$\frac{(I_{2j})^{1-a_I}}{1-a_I} - \left[\exp(a_P' P_j + \underline{\underline{h}}_j)\right]\frac{(\beta^*_{2j})^{1+a_\beta}}{1+a_\beta} = \frac{(I_{3j})^{1-a_I}}{1-a_I} - \left[\exp(a_P' P_j + \underline{\underline{h}}_j)\right]\frac{(\beta^*_{3j})^{1+a_\beta}}{1+a_\beta}$$

$$\Rightarrow \underline{\underline{h}}_j = \ln\left[\frac{1+a_\beta}{1-a_I}\frac{(I_{3j})^{1-a_I} - (I_{2j})^{1-a_I}}{(\beta^*_{3j})^{1+a_\beta} - (\beta^*_{2j})^{1+a_\beta}}\right] - a_P' P_j \tag{2.36}$$

For the two-segment non-convex constraint in figure 2.6:

$$x_j{}^* = a' z_{1j} + h_j \qquad \text{if} \qquad h_j < \overline{\overline{h_j}} \tag{2.37}$$

$$x_j{}^* = a' z_{3j} + h_j \qquad \text{if} \qquad \overline{\overline{h_j}} < h_j \tag{2.38}$$

where $\overline{\overline{h_j}}$ corresponds to indifference curve $u'u'$, that is it is the value of h_j which equilibrates (indirect) utility along segments 1 and 3:

$$\frac{(I_{1j})^{1-a_I}}{1-a_I} - \left[\exp(a_P' P_j + \overline{\overline{h_j}})\right]\frac{(\beta^*_{1j})^{1+a_\beta}}{1+a_\beta} = \frac{(I_{3j})^{1-a_I}}{1-a_I} - \left[\exp(a_P' P_j + \overline{\overline{h_j}})\right]\frac{(\beta^*_{3j})^{1+a_\beta}}{1+a_\beta}$$

$$\Rightarrow \overline{\overline{h_j}} = \ln\left[\frac{1+a_\beta}{1-a_I}\frac{(I_{3j})^{1-a_I} - (I_{1j})^{1-a_I}}{(\beta^*_{3j})^{1+a_\beta} - (\beta^*_{1j})^{1+a_\beta}}\right] - a_P' P_j \tag{2.39}$$

Since the desired level of public expenditure may diverge from the actual level of public expenditure because of measurement or optimisation error, we must allow for the possibility of this random disturbance to shift a local authority to a segment that is different from the one on which the authority maximises its welfare. Thus, the probability of observing a certain level of public expenditure is given by the sum of the joint probabilities that the heterogeneity term is such as to make a local authority optimise on a certain segment - or at the kink - and that the random term moves the local authority to the observed level of expenditure. Details are reported in appendix 2.1.

The presence of the kinked constraint thus allows us to separately identify the variances of the two error terms. This happens because the unobserved preference heterogeneity term appears in the utility function, while the optimisation error does not. Actually, the latter explains why authorities are not as concentrated (clustered) at convex kinks, and away from non-convex kinks, as the model would otherwise suggest.

However, inspection of the data does suggest that local authorities tend to locate (cluster) around the convex kink at which the price of public services increases. In each of the years 1986 to 1989, the mode of the distribution of the tax rate is located between the tax rate for spending at GRE (t^* in figure 2.1) and the tax rate for spending at the threshold (t^{**} in figure 2.1). The model also predicts that authorities should locate away from non-convex kinks. However, since each authority has a different budget constraint - and, in particular, a different grant-exhaustion point - simple inspection of the distribution of the tax rate could be misleading in this case.

As for functional form, we estimate our model by using the following levels specification:

$$X_j^* = a' Z_{ij} + h_j = a_P' P_j + a_\beta \beta^*_{ij} + a_I I_{ij} + h_j \qquad (2.40)$$

The corresponding indirect utility function is given by:

$$V_j = \exp[\ln(a_\beta + a_I X_j^*) - a_I \beta^*_{ij}] \qquad (2.41)$$

The probability of observing a given level of public expenditure has the same shape as before, with the exception of the following terms:[7]

$$\underline{\underline{h}}_j = \frac{(a_I I_{3j} + a_\beta \beta^*_{3j}) \exp[a_I(\beta^*_{2j} - \beta^*_{3j})] - (a_I I_{2j} + a_\beta \beta^*_{2j})}{1 - \exp[a_I(\beta^*_{2j} - \beta^*_{3j})]} - a_P' P_j - \frac{a_\beta}{a_I} \qquad (2.42)$$

$$\overline{\overline{h}}_j = \frac{(a_I I_{3j} + a_\beta \beta^*_{3j}) \exp[a_I(\beta^*_{1j} - \beta^*_{3j})] - (a_I I_{1j} + a_\beta \beta^*_{1j})}{1 - \exp[a_I(\beta^*_{1j} - \beta^*_{3j})]} - a_P' P_j - \frac{a_\beta}{a_I} \qquad (2.43)$$

---

[7] Actually, for the three-segment budget constraint there exists the possibility that the second segment is skipped altogether by those local authorities which have a very small second segment. In that case, computation of the probability requires a comparison between the direct utility at the convex kink and the indirect utility on the third segment. Details are reported in appendix 2.2.

2.5 Results

The results of estimating the model on the English shire districts are shown in table 2.1, where the two-error Maximum Likelihood estimates are compared to standard OLS estimates. The means and data sources of all variables used in the analysis are in appendix 2.3. The dependent variable is per capita expenditure, defined as current expenditure per head in the financial year. Capital expenditures are excluded.

Explanatory variables are subdivided into three groups. The first one includes income - virtual income - and price, as described in section **2.3**. In particular, virtual income of jurisdiction j on segment i (i=1,2,3) is defined as: $\beta_{ij}^{*}M_{ij}+Y_j$, where β_{ij}^{*} is the price of public services to the local taxpayer, M_{ij} is the lump-sum component of the grant, and Y_j is disposable income per capita.[8] The price of public services is the ratio of the actual tax base to the tax base implicit in the relevant segment of the Grant Related Poundage schedule.

The second group includes political and demographic variables - Labour control dummy, closeness to a metropolitan area dummy, population and urbanisation rate. The Labour control dummy (=1 if the local council is Labour controlled, and =0 otherwise) is included in order to allow for the fact that left-wing governments might tend to spend more than right-wing ones. We also introduce a dummy variable for closeness to a metropolitan area to allow for spillovers from metropolitan areas into surrounding communities.

Finally, the third group includes socio-economic variables, which are intended to be a measure of needs specific to the local community - proportion of elderly people, proportion of people belonging to ethnic minorities, and rate of long term unemployment. While GRE set by central government is intended to be a measure of the specific needs of local jurisdictions, we include a further number of socio-economic variables in order to check whether they still have an explanatory power on the level of expenditure.

---

[8] Income data are not available at the district level. As a proxy, county level disposable income per head is used instead.

Table 2.1 presents the results for the pooled cross-sections, where time-dummies are included.[9] OLS estimates - shown in the last column - are obtained by assigning to each government the parameters of the segment on which it is observed to be located.

However, OLS yields biased estimates in this case, because the error term is correlated with the price and income variables, both if the error represents unobserved heterogeneity of preferences, and if it represents optimisation or measurement error. The price coefficient is underestimated relative to its ML counterpart. Most of the local governments locate on the first and second convex segments of the budget constraint, in the vicinity of the kink. Since the price of public services increases after the kink (due to a reduction in the matching grant rate), a spurious positive correlation between expenditure and price is introduced into the equation.

Compared to the ML estimate, the income coefficient does not show a substantial bias. However, due to measurement error in the income variable, the income coefficient is rather imprecisely estimated.

The first column shows the results of estimating a simple, parsimonious specification in which no independent variables other than income and price are included in the equation. The results show a positive income effect and a negative price effect, both of high significance. The mean income and price elasticity are .25 and -.85 respectively. The second column adds political, demographic and socio-economic variables. The orders of magnitude of the income and price effects are not very much affected. All variables in the political and demographic group have a significant effect on the level of expenditure. In the first place, Labour controlled authorities tend to set higher levels of spending than authorities controlled by other parties.[10] The coefficient on the metropolitan dummy variable suggests that districts surrounding metropolitan areas tend to spend more than districts located in rural areas. This could be due to negative spillovers from metropolitan areas.

---

[9] The estimation results for each individual cross-section are reported in appendix 2.5.

[10] However, it could be argued that the Labour dummy variable in the expenditure equation is endogenous. The higher the preference for local public services in a jurisdiction, the more likely it is that voters will choose a left-wing candidate. The interpretation of this coefficient would therefore require a more detailed model of the local political process.

Alternatively, the closeness to a metropolitan area dummy could be picking up the effect of other unobservable location-specific characteristics.

The urbanisation and population variables also appear to have a positive effect on the level of expenditure per head, even though the magnitude of these effects is very small, while the socio-economic variables do add some explanatory power beyond the measure of GRE embedded in the grant distribution system parameters.

Finally, the estimates for σ_h and σ_ε show significant evidence of both unobserved preference heterogeneity and of random optimisation error.

The fact that the two-error Maximum Likelihood estimates diverge from the ones obtained from the simple OLS estimation can readily be explained by the fact that the former takes into account that the segment (or kink) location is endogenous, while the latter does not. However, MaCurdy *et al.* (1990) have argued that the stronger price and income effects that arise when using ML techniques that are designed to account for piecewise-linear constraints, might simply follow from features of the econometric model that force a number of key parameters to obey certain inequalities. The two-error Maximum Likelihood approach relies on a structural economic model, whose assumptions are imposed on the data. In particular, MaCurdy *et al.* (1990) show in the labour supply context that the requirement of non-negative probabilities for kink locations translates into a non-negativity constraint on the Slutzky term.

However, our empirical results on the price and income coefficients in table 2.1 from the OLS and ML estimates suggest that no binding constraint is being imposed on the data in the estimation of the local public expenditure determination model.

2.6 Conclusions

This chapter has analysed the English local authority expenditure decisions under the Block Grant system, that is the central government grant distribution scheme that was in place from 1981 to 1990. It has been recognised that local authorities faced possibly quite involved budget constraints. If local authorities' preferences can be expressed as a non-decreasing function of private consumption and local expenditure on public services, the actual budget constraint is - according to the level of the local tax base - either a two-segment convex, or a three-segment non-convex, or a two-segment non-convex one. The lump-sum component of the grant and the price of public services - the ratio of the actual tax base to the tax base implicit in the relevant segment of the Grant Related Poundage schedule - change according to the segment a local authority locates onto, and must be explicitly modelled. Since each segment is characterised by different price and virtual income, and since the choice of segment is endogenous - either due to unobserved heterogeneity among local communities, or due to measurement/optimisation error - a demand function for local public expenditure cannot be consistently estimated with Ordinary Least Squares by attributing to each local government the price and income variables of the segment on which it is observed to be located. Instead, a two-error Maximum Likelihood estimation procedure allowed us to estimate the effect of income and price - and a series of other demographic, political, and socio-economic characteristics - on the expenditures of the English non-metropolitan districts in the fiscal years 1986/87 to 1989/90. The results have shown a positive income effect - with an elasticity of .25 - and a negative price effect - with an elasticity of .85 - both of high significance. We have compared these results to standard OLS estimates, with no adjustment for the endogeneity of the price and income variables. The price coefficient is underestimated relative to its ML counterpart, due to spurious positive correlation between expenditure and price. On the other hand, the income coefficient does not show a substantial bias. The two-error ML procedure allows us to identify the respective variances of unobserved heterogeneity and optimisation error. The results show significant evidence of unobserved preference heterogeneity, and some weaker evidence of random optimisation error.

Finally, the econometric procedure that we have used is based on the hypothesis that each local government decides its own level of public expenditure independently of the decisions of the other local jurisdictions. Furthermore, the error term in the local public expenditure determination equation is assumed not to be correlated across local jurisdictions. However, the observations on local governments' expenditures might be spatially correlated. We explore this possibility further in the next chapter, and analyse how the econometric approach must be amended when the observations exhibit spatial autocorrelation.

Figure 2.1

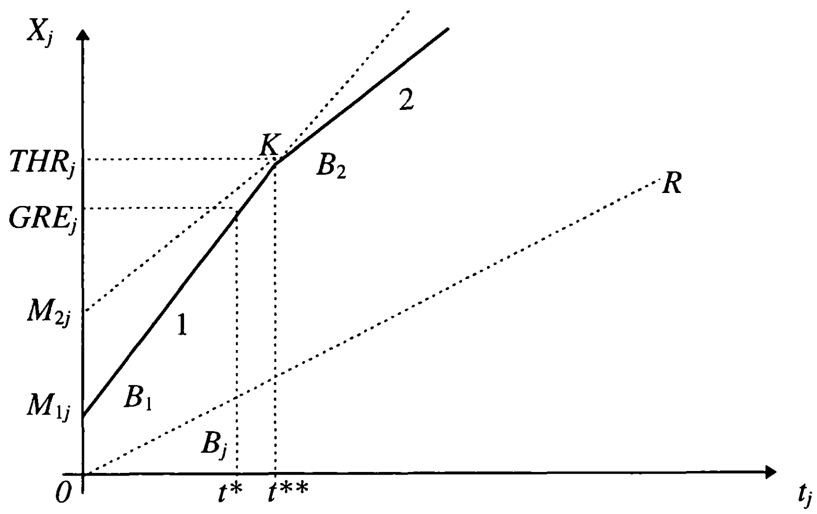


Figure 2.2

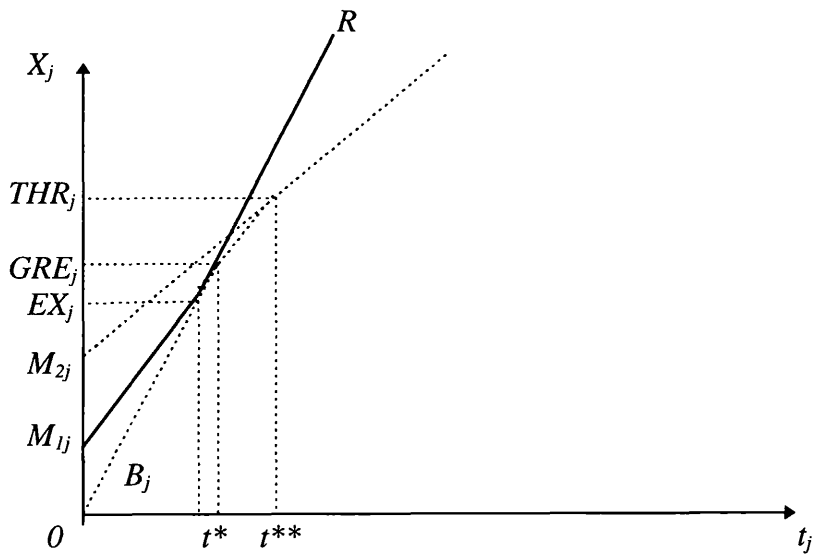


Figure 2.3

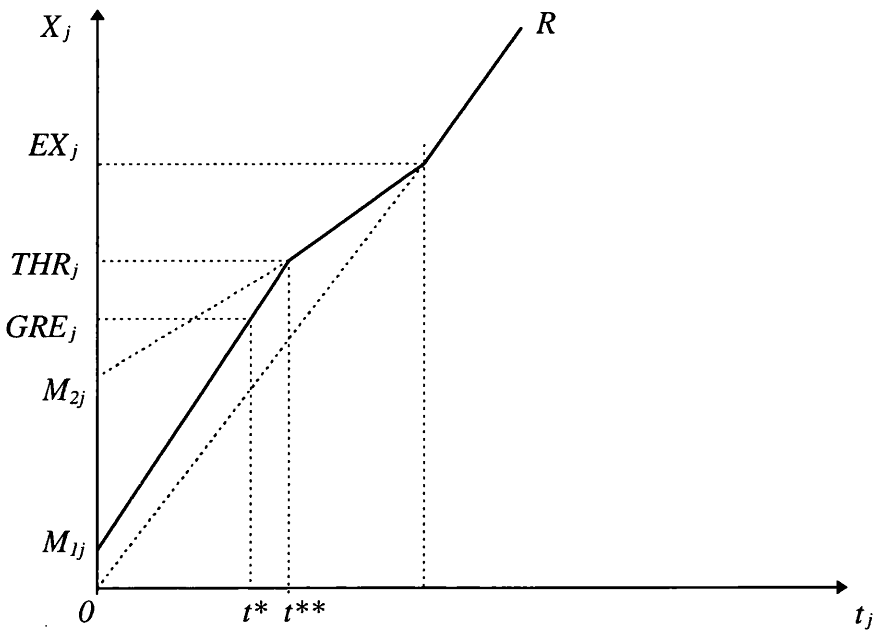


Figure 2.4

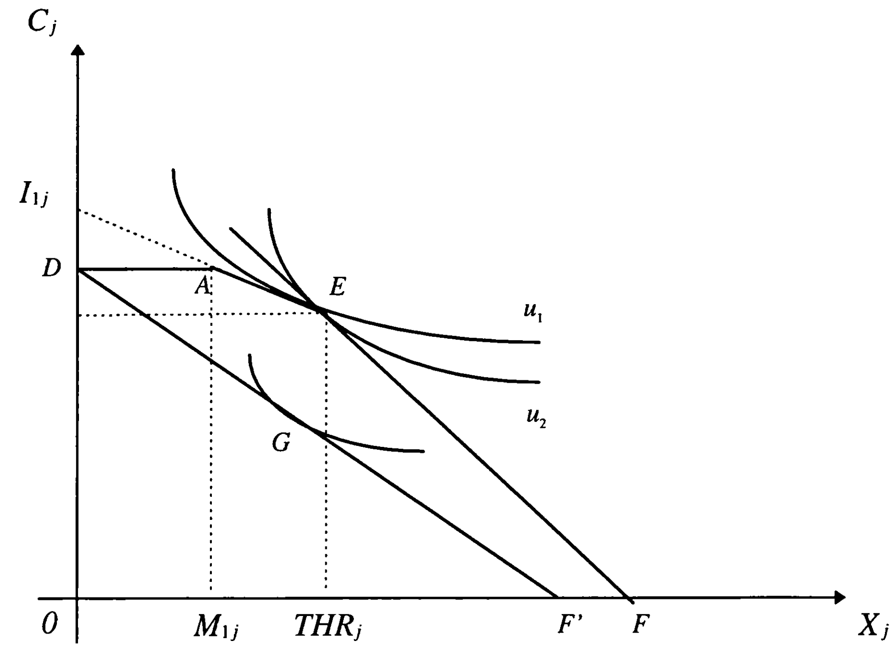


Figure 2.5

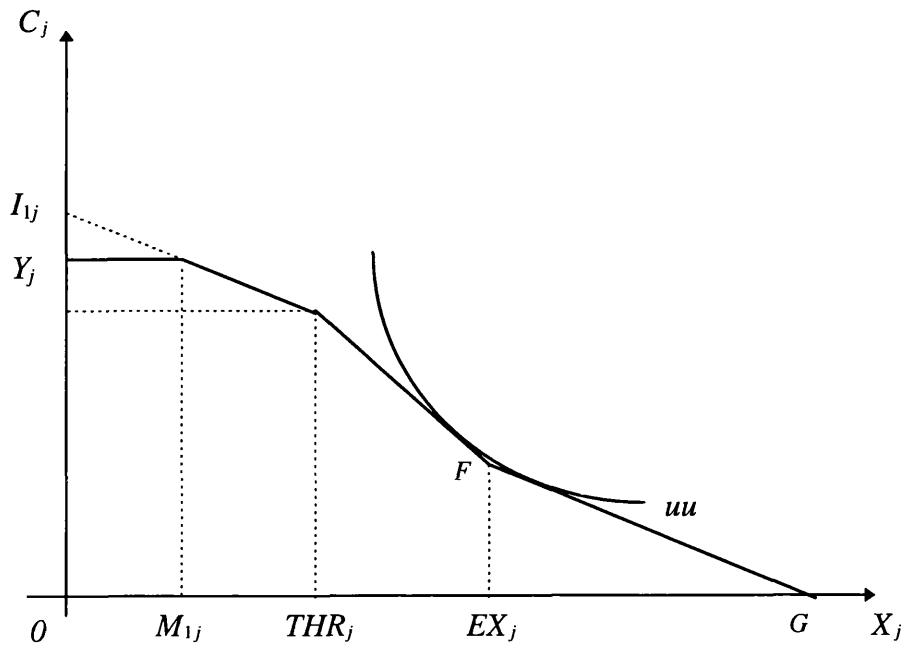


Figure 2.6

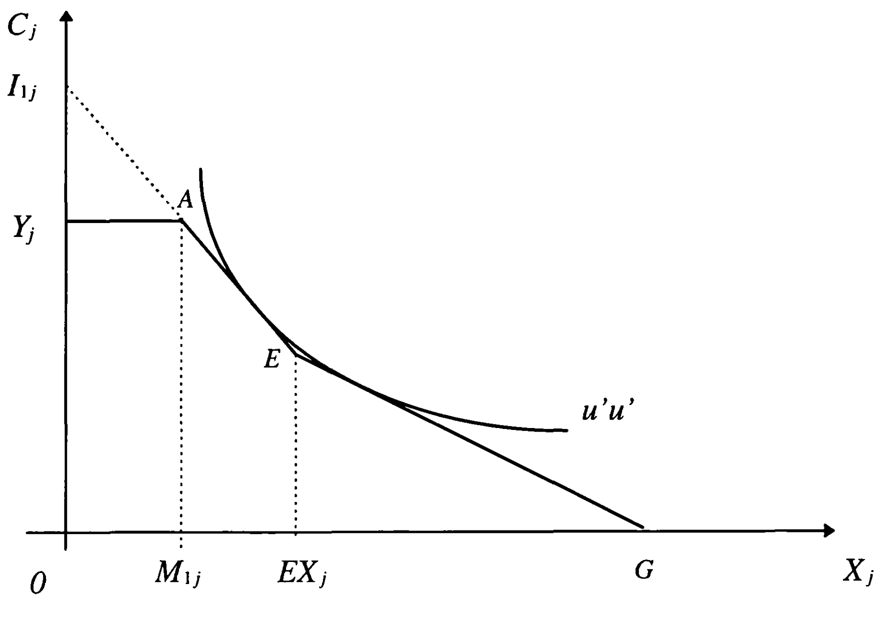


Table 2.1 Shire districts' local expenditure 1986/87 - 1989/90

	two error ML		OLS
income (a_I)	.0021 (2.32)	.0022 (2.77)	.0015 (2.07)
price (a_β)	-66.38 (-24.17)	-43.80 (-22.43)	-23.46 (-10.26)
LabourD		10.08 (8.15)	12.31 (9.86)
population (,000)		.035 (3.14)	.023 (2.11)
urbanisation rate (%)		.125 (5.44)	.094 (5.72)
MetroD		4.68 (4.08)	5.06 (4.82)
age > 65 (%)		1.432 (10.49)	1.232 (11.99)
ethnic minority (%)		.628 (2.75)	.415 (1.66)
long term unemployed (%)		.107 (.714)	.202 (1.57)
σ_h	18.34 (15.80)	10.89 (16.69)	
σ_ε	8.68 (9.57)	6.85 (9.16)	
R^2			.65
observations	1184	1184	1184

Notes:
I) dependent variable = local public expenditure per head;
II) *t* values in parentheses;
III) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;
IV) MetroD=1 if a shire district is a neighbour of a metropolitan authority, and 0 otherwise;
V) time dummies are included.

Appendix 2.1

Let ε_j and h_j be distributed normally and independently with means zero and respective variances σ_ε^2 and σ_h^2. As a result, $v_j=h_j+\varepsilon_j$ is distributed normally with mean zero and variance equal to $\sigma_v^2=\sigma_h^2+\sigma_\varepsilon^2$. ϕ and Φ are the standard normal p.d.f. and conditional c.d.f. the bivariate normal density between h and v is factored into, and $\rho=\sigma_h/\sigma_v$ is the correlation coefficient between h and v.

For a local authority facing a two-segment convex budget constraint as the one depicted in figure 2.4, the probability of a given value of x is:

$$
\begin{aligned}
\Pr\{x_j\} &= \Pr\{h_j < \underline{h}_j, h_j + \varepsilon_j = x_j - a' z_{1j}\} + \Pr\{\underline{h}_j \le h_j \le \overline{h}_j, \varepsilon_j = x_j - thr_j\} \\
&+ \Pr\{\overline{h}_j < h_j, h_j + \varepsilon_j = x_j - a' z_{2j}\} \\
&= \frac{1}{\sigma_v}\phi\left(\frac{x_j - a' z_{1j}}{\sigma_v}\right)\Phi\left[\left(\frac{\underline{h}_j}{\sigma_h} - \rho\frac{x_j - a' z_{1j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right] \\
&+ \frac{1}{\sigma_\varepsilon}\phi\left(\frac{x_j - thr_j}{\sigma_\varepsilon}\right)\left\{\Phi\left[\frac{\overline{h}_j}{\sigma_h}\right] - \Phi\left[\frac{\underline{h}_j}{\sigma_h}\right]\right\} \\
&+ \frac{1}{\sigma_v}\phi\left(\frac{x_j - a' z_{2j}}{\sigma_v}\right)\left\{1 - \Phi\left[\left(\frac{\overline{h}_j}{\sigma_h} - \rho\frac{x_j - a' z_{2j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right]\right\}
\end{aligned}
\tag{A2.1}
$$

If the budget constraint is a three-segment non-convex one (figure 2.5), the probability of a given value of x is:

$$
\begin{aligned}
\Pr\{x_j\} &= \Pr\{h_j < \underline{h}_j, h_j + \varepsilon_j = x_j - a' z_{1j}\} + \Pr\{\underline{h}_j \le h_j \le \overline{h}_j, \varepsilon_j = x_j - thr_j\} \\
&+ \Pr\{\overline{h}_j < h_j < \underline{\underline{h}}_j, h_j + \varepsilon_j = x_j - a' z_{2j}\} + \Pr\{\underline{\underline{h}}_j < h_j, h_j + \varepsilon_j = x_j - a' z_{3j}\} \\
&= \frac{1}{\sigma_v}\phi\left(\frac{x_j - a' z_{1j}}{\sigma_v}\right)\Phi\left[\left(\frac{\underline{h}_j}{\sigma_h} - \rho\frac{x_j - a' z_{1j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right] \\
&+ \frac{1}{\sigma_\varepsilon}\phi\left(\frac{x_j - thr_j}{\sigma_\varepsilon}\right)\left\{\Phi\left[\frac{\overline{h}_j}{\sigma_h}\right] - \Phi\left[\frac{\underline{h}_j}{\sigma_h}\right]\right\} \\
&+ \frac{1}{\sigma_v}\phi\left(\frac{x_j - a' z_{2j}}{\sigma_v}\right) \\
&\times\left\{\Phi\left[\left(\frac{\underline{\underline{h}}_j}{\sigma_h} - \rho\frac{x_j - a' z_{2j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right] - \Phi\left[\left(\frac{\overline{h}_j}{\sigma_h} - \rho\frac{x_j - a' z_{2j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right]\right\} \\
&+ \frac{1}{\sigma_v}\phi\left(\frac{x_j - a' z_{3j}}{\sigma_v}\right)\left\{1 - \Phi\left[\left(\frac{\underline{\underline{h}}_j}{\sigma_h} - \rho\frac{x_j - a' z_{3j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right]\right\}
\end{aligned}
\tag{A2.2}
$$

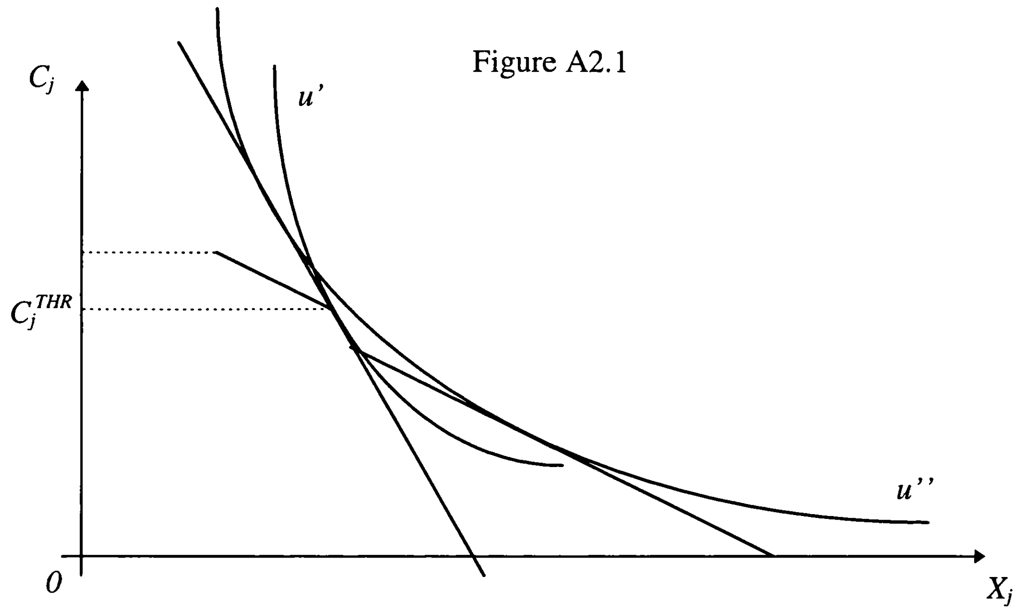
Finally, for a local authority facing a two-segment non-convex budget constraint as the one depicted in figure 2.6, the probability of a given value of x is:

$$
\begin{aligned}
\Pr\{x_j\} &= \Pr\{h_j < \overline{\overline{h}}_j, h_j + \varepsilon_j = x_j - a' z_{1j}\} + \Pr\{\overline{\overline{h}}_j < h_j, h_j + \varepsilon_j = x_j - a' z_{3j}\} \\
&= \frac{1}{\sigma_v}\phi\left(\frac{x_j - a' z_{1j}}{\sigma_v}\right)\Phi\left[\left(\frac{\overline{\overline{h}}_j}{\sigma_h} - \rho\frac{x_j - a' z_{1j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right] \\
&+ \frac{1}{\sigma_v}\phi\left(\frac{x_j - a' z_{3j}}{\sigma_v}\right)\left\{1 - \Phi\left[\left(\frac{\overline{\overline{h}}_j}{\sigma_h} - \rho\frac{x_j - a' z_{3j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right]\right\}
\end{aligned}
\tag{A2.3}
$$

Appendix 2.2

Suppose a local government faces a budget constraint such as the one depicted in figure A2.1. In that case, no point on the second segment can ever be an optimal choice. This happens whenever the heterogeneity term corresponding to indifference curve u'' - that makes segments 2 and 3 indifferent - is smaller than the heterogeneity term corresponding to indifference curve u' - tangent to segment 2 exactly at the convex kink. In this case the relevant comparison is the one between utility at the convex kink and utility on the third segment.

Figure A2.1



The direct utility function corresponding to a linear demand function for public expenditure is given by equation (A2.4), and the indirect utility function on segment 3 is given by equation (A2.5) - Stern (1986).

$$u(X_j*, C_j) = \frac{X_j*+f}{a_I}\exp\left(\frac{a_I(C_j+k)}{X_j*+f}\right) \qquad \text{(A2.4)}$$

where $f = \frac{a_\beta}{a_I}$ and $k = \frac{a_P' P_j + h_j}{a_I} + \frac{a_\beta}{a_I^2}$

$$V_j = \exp\{\ln[a_\beta + a_I(a_P' P_j + a_\beta \beta_{3j}^* + a_I I_{3j} + h_j)] - a_I \beta_{3j}^*\} \qquad \text{(A2.5)}$$

Denoting the heterogeneity term that equates maximum utility on segment three and utility at the kink by $\tilde{h}_j$, the probability of a given level of local public expenditure X_j is now given by:

$$\Pr\{X_j\} = \Pr\{h_j < \underline{h_j}, h_j + \varepsilon_j = X_j - a' Z_{1j}\} + \Pr\{\underline{h_j} \le h_j \le \tilde{h}_j, \varepsilon_j = X_j - THR_j\}$$
$$+ \Pr\{\tilde{h}_j < h_j, h_j + \varepsilon_j = X_j - a' Z_{3j}\}$$
$$= \frac{1}{\sigma_v}\phi\left(\frac{X_j - a' Z_{1j}}{\sigma_v}\right)\Phi\left[\left(\frac{\underline{h_j}}{\sigma_h} - \rho\frac{X_j - a' Z_{1j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right]$$
$$+ \frac{1}{\sigma_\varepsilon}\phi\left(\frac{X_j - THR_j}{\sigma_\varepsilon}\right)\left\{\Phi\left[\frac{\tilde{h}_j}{\sigma_h}\right] - \Phi\left[\frac{\underline{h_j}}{\sigma_h}\right]\right\}$$
$$+ \frac{1}{\sigma_v}\phi\left(\frac{X_j - a' Z_{2j}}{\sigma_v}\right)\left\{1 - \Phi\left[\left(\frac{\tilde{h}_j}{\sigma_h} - \rho\frac{X_j - a' Z_{3j}}{\sigma_v}\right)\frac{1}{\sqrt{1-\rho^2}}\right]\right\}$$

(A2.6)

Appendix 2.3

Table A2.1 Means of the variables used in the analysis

	1986/87	1987/88	1988/89	1989/90
expenditure per head (£)	52.17	56.49	58.92	63.12
GRE per head (£)	53.18	50.43	51.93	58.47
disposable income per head (£)	5,677	6,173	6,831	7,586
grant per head (segment 1, £)	27.47	26.65	26.95	28.93
grant per head (segment 2, £)	35.92	34.49	35.16	38.48
price (segment 1)	.802	.693	.720	.744
price (segment 2)	1.089	.942	.978	1.010
population	98,687	99,348	99,965	100,612
urbanisation rate	57.44			
age>65	21.71			
ethnic minority	2.39			
long term unemployment	23.26			

Table A2.2 Data sources

data	source
expenditure	Chartered Institute of Public Finance and Accountancy (1986-1990) Finance and general statistics
GRE	Association of County Councils (1986-1990) Rate Support Grant - England (London: ACC)
income	Economic Trends (Central Statistical Office)
tax base	Chartered Institute of Public Finance and Accountancy (1986-1990) Finance and general statistics
rates	Chartered Institute of Public Finance and Accountancy (1986-1990) Rate collection statistics
Block Grant parameters	Association of County Councils (1986-1990) Rate Support Grant - England (London: ACC)
population	Chartered Institute of Public Finance and Accountancy (1986-1990) Local government comparative statistics
urbanisation rate	Census of Population (1991)
age>65	Census of Population (1991)
ethnic minority	Census of Population (1991)
unemployment	Census of Population (1991)
political control	University of Plymouth: Local Government Chronicle Elections Centre

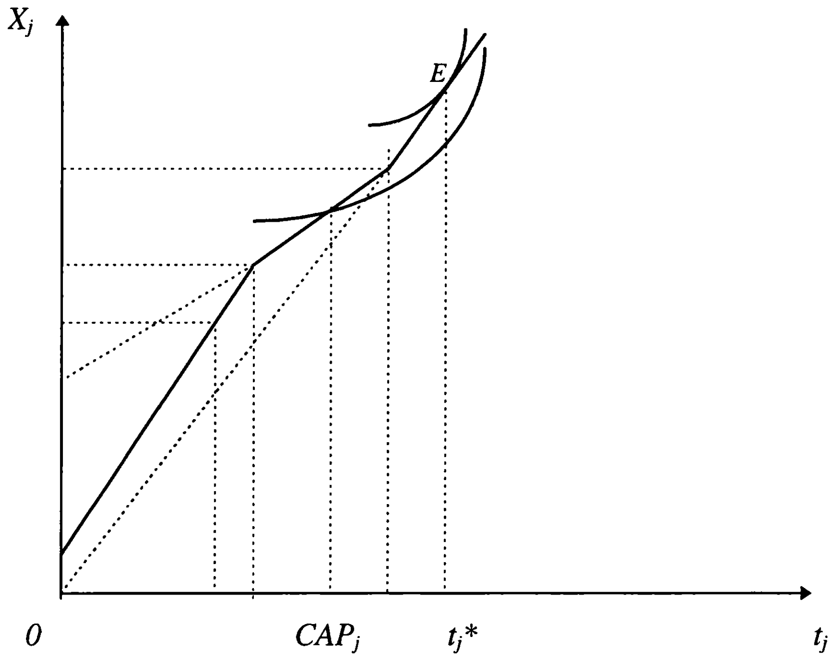
Appendix 2.4

Starting from the financial year 1985/86, the central government assumed for the first time the power of capping - i.e., imposing rate limits on - the local authorities which, in its view, were setting too high rates and, consequently, were spending too much, having regard to general economic conditions. Two justifications were given for expenditure capping. The first was the need for macroeconomic control of the economy. The second was that local ratepayers needed - in the government's view - to be "protected" from the "incorrigible" high-spending authorities.

Capping was retrospective and selective. It took place ex-post, that is after local budgets had been set, and according to criteria that were decided every year, as table A2.3 illustrates. Moreover, the government was given the power of applying different selection principles to different classes of authority, or to those newly selected compared with those already selected for capping in previous years.

Capping has the effect of truncating the budget constraint at the maximum rate, as, for example, CAP_j in figure A2.2. The first consequence of rate-capping is that the observed rate (CAP_j) may be different from the optimal rate t_j* (point E).

Figure A2.2



The second possible consequence of rate-capping is that, in the presence of a non-convex budget constraint, the observed rate can be smaller than the maximum rate, even though the optimal rate exceeds the maximum rate. As is shown in figure A2.3, in the absence of capping the local government would choose t_j* - that is a higher tax than CAP_j. A rate limit of CAP_j in figure A2.3, or even only a threat of capping at CAP_j, would induce to set the rate t_j'.

If on the one hand it is rather unlikely that an authority - after being capped - sets a rate smaller than the maximum one imposed by central government, it may be the case that some local authorities set rates at a lower level than the desired ones in order to avoid being capped. Unfortunately, the latter influence cannot be observed.

As far as the English non-metropolitan districts are concerned, however, rate-capping does not represent a problem for the estimation of an equation of local public expenditure determination. In fact, only a negligible proportion of the shire districts were rate-capped in the financial years 1986/87 to 1989/90, as table A2.4 shows.

Figure A2.3

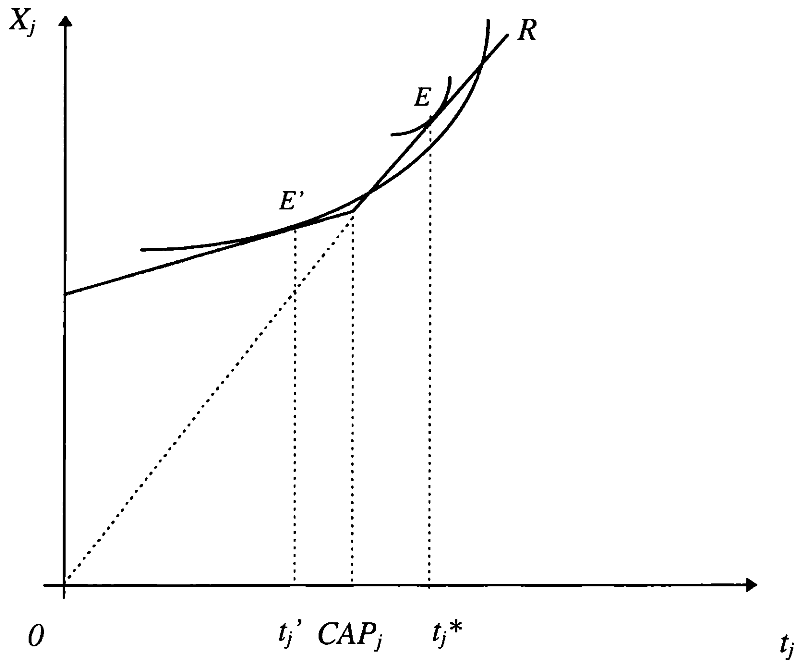


Table A2.3 Criteria for selection of rate limited authorities

	previously selected authorities	newly selected authorities
1985		1984/85 budget was: 1) more than £ 10m, and 2) more than 4% above target, and 3) more than 20% above GRE
1986	1985/86 budget was: 1) more than £ 10.6m, and 2) more than 20% above GRE, and either 3) more than 1% above target, or 4) more than 30% over 81/82 exp.	1985/86 budget was: 1) more than £ 10.6m, and 2) more than 4% above target, and 3) more than 20% above GRE
1987	1986/87 budget was: 1) more than £ 11.1m, and 2) at least 12.5% above GRE, and either 3) at least 4% above 85/86 exp. or 4) at least 20% above 82/83 exp.	1986/87 budget was: 1) more than £ 11.1m, and 2) at least 12.5% above GRE, and 3) at least 4% above 85/86 exp.
1988	1987/88 budget was: 1) more than £ 12.2m, and 2) at least 12.5% above 87/88 GRE	1987/88 budget was: 1) more than £ 12.2m, and 2) at least 12.5% above 87/88 GRE, and 3) at least 6% above 86/87 exp.
1989	1988/89 budget was: 1) at least 12.5% above 88/89 GRE	1988/89 budget was: 1) at least 12.5% above 88/89 GRE, and 2) at least 6% above 87/88 exp.

Table A2.4 Rate-capped shire districts

1986/87	1987/88	1988/89	1989/90
2	4	4	1
Basildon	Basildon	Basildon	Thamesdown
Thamesdown	Thamesdown	Thamesdown	
	Brighton	Middlesborough	
	Middlesborough	Kingston-upon-H	

Table A2.5 Grant-exhausted shire districts

1986/87	1987/88	1988/89	1989/90
14	16	23	26
Chiltern	Chiltern	S.Bedfordshire	S.Bedfordshire
South Bucks	South Bucks	Wokingham	Wokingham
Ellesmere P.	Ellesmere P.	Chiltern	Chiltern
Basildon	Basildon	Milton Keynes	Milton Keynes
Harlow	Harlow	South Bucks	South Bucks
Thurrock	Thurrock	Wycombe	Wycombe
Stevenage	Stevenage	Basildon	Basildon
Welwyn H.	Welwyn H.	Harlow	Harlow
Selby	Daventry	Thurrock	Maldon
Bassetlaw	Selby	Hart	Thurrock
Elmbridge	Bassetlaw	Redditch	Hart
Reigate & B.	Elmbridge	St. Albans	Redditch
N. Warwicksh.	Epsom&Ewell	Stevenage	St. Albans
Crawley	Reigate & B.	Three Rivers	Stevenage
	N. Warwicksh.	Welwyn H.	Three Rivers
	Crawley	Daventry	Welwyn H.
		Selby	Daventry
		Bassetlaw	Selby
		Oxford	Bassetlaw
		Epsom&Ewell	Oxford
		Guildford	Epsom&Ewell
		N. Warwicksh.	Guildford
		Crawley	Waverley
			N. Warwicksh.
			Stratford-on-A
			Crawley

Appendix 2.5

Table A2.6 Shire districts' local expenditure 1986/87

	two error ML		OLS
constant	88.59 (7.53)	-11.40 (-1.09)	-16.06 (-1.50)
income (a_I)	.0012 (.52)	.0027 (1.90)	.0024 (1.60)
price (a_β)	-50.71 (-10.30)	-34.89 (-8.11)	-22.06 (-5.65)
LabourD		9.43 (4.55)	10.92 (5.27)
population (,000)		.033 (1.84)	.025 (1.52)
urbanisation rate (%)		.11 (4.06)	.08 (3.10)
MetroD		2.09 (1.29)	2.49 (1.50)
age > 65 (%)		1.31 (7.09)	1.14 (6.92)
ethnic minority (%)		.72 (1.89)	.55 (1.34)
long term unemployed (%)		.17 (1.18)	.19 (.96)
σ_h	16.36 (8.40)	6.94 (3.58)	
σ_ε	8.08 (4.96)	6.92 (4.65)	
R^2			.70
observations	296	296	296

Notes:
I) dependent variable = local public expenditure per head;
II) *t* values in parentheses;
III) LabourD=1 if the local authority is Labour controlled, 0 otherwise;
IV) MetroD=1 if a shire district is a neighbour of a metropolitan authority, 0 otherwise.

Table A2.7 Shire districts' local expenditure 1987/88

	two error ML		OLS
constant	100.94 (8.10)	17.79 (1.21)	.173 (.02)
income (a_I)	.0016 (.74)	.0009 (.50)	.0002 (.11)
price (a_β)	-69.72 (-15.88)	-54.52 (-13.19)	-25.33 (-5.37)
LabourD		6.59 (2.49)	9.99 (4.12)
population (,000)		.046 (1.94)	.020 (.93)
urbanisation rate (%)		.16 (4.39)	.12 (3.52)
MetroD		8.00 (3.57)	7.27 (3.51)
age > 65 (%)		1.35 (5.78)	1.15 (5.66)
ethnic minority (%)		.39 (.71)	.27 (.56)
long term unemployed (%)		-.14 (-.47)	.07 (.29)
σ_h	18.55 (8.41)	11.58 (7.89)	
σ_ε	7.27 (5.29)	5.82 (5.60)	
R^2			.63
observations	296	296	296

Notes:
I) dependent variable = local public expenditure per head;
II) t values in parentheses;
III) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;
IV) MetroD=1 if a shire district is a neighbour of a metropolitan authority, and 0 otherwise.

Table A2.8 Shire districts' local expenditure 1988/89

	two error ML		OLS
constant	99.13 (10.19)	12.13 (.81)	-7.62 (-.62)
income (a_I)	.0034 (2.23)	.0025 (1.66)	.0016 (1.22)
price (a_β)	-79.28 (-15.48)	-59.73 (-13.96)	-28.42 (-6.16)
LabourD		8.13 (2.95)	11.88 (4.63)
population (,000)		.036 (1.50)	.023 (1.02)
urbanisation rate (%)		.15 (3.81)	.10 (2.96)
MetroD		4.79 (2.06)	5.08 (2.33)
age > 65 (%)		1.38 (5.83)	1.15 (5.45)
ethnic minority (%)		.56 (.93)	.22 (.42)
long term unemployed (%)		-.013 (-.07)	.26 (.96)
σ_h	17.72 (8.07)	11.48 (7.07)	
σ_ε	8.06 (5.80)	6.59 (5.64)	
R^2			.63
observations	296	296	296

Notes:
I) dependent variable = local public expenditure per head;
II) t values in parentheses;
III) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;
IV) MetroD=1 if a shire district is a neighbour of a metropolitan authority, and 0 otherwise.

Table A2.9 Shire districts' local expenditure 1989/90

	two error ML		OLS
constant	97.41 (8.33)	-32.53 (-2.13)	-34.92 (-2.31)
income (a_I)	.0027 (1.55)	.0039 (2.75)	.0026 (1.83)
price (a_β)	-69.39 (-10.13)	-43.42 (-8.50)	-18.58 (-3.43)
LabourD		14.11 (4.71)	17.25 (5.98)
population (,000)		.039 (1.53)	.027 (1.07)
urbanisation rate (%)		.10 (2.65)	.07 (1.93)
MetroD		4.41 (1.92)	4.96 (2.03)
age > 65 (%)		1.85 (7.39)	1.48 (6.25)
ethnic minority (%)		.55 (.81)	.34 (.58)
long term unemployed (%)		.18 (.54)	.31 (1.06)
σ_h	18.19 (6.29)	9.00 (2.74)	
σ_ε	13.22 (5.69)	10.84 (4.86)	
R^2			.64
observations	296	296	296

Notes:
I) dependent variable = local public expenditure per head;
II) *t* values in parentheses;
III) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;
IV) MetroD=1 if a shire district is a neighbour of a metropolitan authority, and 0 otherwise.

Chapter Three

Spatial patterns in local public expenditure

3.1 Introduction

In most empirical studies in the public finance literature, the main determinants of the expenditures of local governments (or states in federal countries) on public services are the traditional economic variables - price, income and grants from other levels of government - as well as local political and demographic characteristics.[1] The error term in the econometric model is assumed to be independently and identically distributed across local governments, and the level of public expenditure in a local jurisdiction is assumed not to be affected by the expenditures in neighbouring jurisdictions.

However, the respective locations of the local authorities might contribute to explaining their expenditure decisions. Since data on local governments are spatial data, it is worthwhile to analyse whether the observed spatial arrangement of local public expenditures is a result of a random assignment, or instead is characterised by some form of spatial dependence.

Finding spatial correlation in local public spending may, by itself, be neither surprising nor very interesting. If the determinants of local public expenditure are spatially correlated - say, income and other variables reflecting local preferences and needs - we will find spatial correlation in local spending. Consequently, we are interested in detecting spatial correlation, after removing those observable effects. We proceed as follows. First, we compute measures of spatial association in the raw level of local public spending (per capita). Second, we test for spatial correlation in the residuals of a regression that removed the effects of observable variables on local

[1] A survey of theoretical and empirical studies of local public expenditure determination is in Wildasin (1986).

spending. Third, we estimate a model of local public expenditure determination, where spatial effects are explicitly allowed for.

Estimates obtained from a regression model that does not account for spatial effects are inefficient in the presence of spatial error dependence, and are biased and inconsistent in the presence of substantive spatial dependence (Anselin, 1988a). The former occurs when the error term in the regression model follows a spatial autoregressive process - that is when the error term in a jurisdiction can be expressed as a spatially weighted average of neighbouring jurisdictions' errors. The latter occurs when the dependent variable itself follows a spatial autoregressive process. The empirical model of local public expenditure should in this case be a mixed regressive-spatial autoregressive model, where the set of explanatory variables also includes a spatial lag of the dependent variable.

Only recently have spatial econometric techniques been applied to the estimation of spatial interactions between local governments. Case, Hines and Rosen (1993) estimate an expenditure equation using a panel data set of the US states' budgets, where spatial correlation in both the dependent variable and in the errors is allowed for. They attribute the presence of spatial effects to the fact that benefits from public expenditures in one state may spill over into neighbouring states. When using a neighbourhood criterion based on geographical proximity - which is consistent with their theoretical model - they find negative spatial correlation in the dependent variable and positive spatial correlation in the errors, which suggests that neighbouring states' expenditures are a substitute for own spending, and states might be subject to common shocks. However, when using a neighbourhood criterion based on similarity of demographic composition of population - irrespective of geographical location - they find a reversed pattern of positive spatial correlation in the dependent variable and negative spatial correlation in the errors.

Besley and Case (1995) use data on US states' income taxes from 1960 to 1988 in order to investigate whether neighbouring states' tax changes are correlated with a given state's tax change. This would arise in a world of asymmetric information between politicians and voters, if the latter consider relative performance evaluation in voting decisions. A state's tax change is linearly related to state-specific variables and neighbouring states' tax changes. Both two stage least squares and maximum likelihood estimates suggest that neighbours' tax changes have a positive and significant effect on a given state's tax change.

Kelejian and Robinson (1993) propose a regional public expenditure model with spatial correlation in the dependent variable, and with two stochastic shocks generated within each region. While one is unique to the region, the other generates spillover effects in other regions. When estimating an equation for police expenditures in a number of US counties - where the spatial correlation component in the error term has been set to zero - they find that police expenditures in a given county are significantly and positively influenced by neighbouring county police expenditures. If counties inflict a negative externality on their neighbours by spending more on police services, the need for police services in a given county tends to increase as such services in neighbouring counties increase.

As regards the UK, Bivand and Szymanski (1997) analyse the pattern of spatial dependence in local public service provision cost. They present a principal agent model that suggests that yardstick competition regulation of natural monopolies generates spatial dependence if local principals (English local authorities) can pursue idiosyncratic policies when they contract with local agents (DSO, Direct Service Organisation). An externality arises because contracts are often based on comparisons of performance against neighbouring jurisdictions. By comparing the pattern of spatial dependence in the provision of a local public service (garbage collection) before and after the introduction of Compulsory Competitive Tendering (CCT), they show that CCT, by imposing standard contracting rules, substantially reduces the scope for local authorities to pursue idiosyncratic policies, and, as a result, the extent of spatial correlation.

Finally, Besley, Preston and Ridge (1997) allow for spatial dependence in local tax non-payment rates in England. They model poll tax non-payment rates in a short panel of data on the 366 English metropolitan and shire districts. Their analysis allows for neighbourhood influences across authority boundaries, that is the level of non-compliance in a district is a function of non-compliance in geographically contiguous districts.

In the following, we test whether local public spending shows a spatial pattern, by looking at the data on the expenditures of the English local governments. Section **3.2** develops the basic framework on which the empirical analysis is based. Section **3.3** formally introduces the notion of spatial correlation and a number of tests for the presence of spatial dependence. Section **3.4** tackles the problem of estimating an equation of local public expenditure determination in the presence of spatial correlation. Section **3.5** describes the data set and the results. Finally, section **3.6** concludes.

3.2 Local public expenditure and spatial correlation

In standard non-spatial empirical models of local public finance, a subcentral government's spending depends on a set of jurisdiction-specific variables - such as income, grants received from other levels of government, and price of public services in terms of private goods - as well as on a series of political, socio-economic and demographic characteristics. In the presence of subsidies from central government, the parameters of the grant distribution formula - such as the lump-sum component and the matching rate of the grant - are bound to be crucial factors in affecting public expenditure decisions. In particular, while a lump-sum grant is equivalent to an increase in the effective income of the locality, a matching grant alters the relative price of public goods with respect to private consumption.

In very general terms, an equation for local public expenditure determination can be written as: $x_i=\alpha' z_i+u_i$, where x_i is public expenditure per head in jurisdiction i, z_i is a vector of own characteristics, α is a vector of parameters, and u_i is an error term, which is assumed to be independently and identically distributed.

The first reason why the above model may not be appropriate for studying local government expenditure is that the error term u might be spatially correlated. Any influence which is omitted from the model and is spatially correlated will lead to spatial dependence in local public expenditure. Furthermore, local jurisdictions may be subject to shocks that affect their expenditure decisions, and are spatially autocorrelated - such as regional shocks to income. Alternatively, those common shocks may be the result of central government regional policies or intermediate level of government - e.g., the counties in the UK - fiscal policies. Spatially correlated shocks can be allowed for by appropriately modelling a spatial process in the error term.

A second, and more serious, problem arises if the expenditure in a local jurisdiction has spillover effects into neighbouring jurisdictions. The presence of spillovers requires explicit modelling of the spatial interdependence, and the estimation procedure of the equation of local public expenditure determination must take into account that local jurisdictions' policies are determined simultaneously.

We first want to explore whether the expenditure on public services in a local jurisdiction is actually correlated with the expenditures in neighbouring jurisdictions. In other words, we want to check whether values at close-by locations are more correlated than values at locations that are far apart. A number of measures of spatial autocorrelation have been suggested.[2] These measures are derived from the null hypothesis that space does not matter, or that the assignment of values to particular locations is not relevant. Under the alternative hypothesis of spatial autocorrelation, either large (small) values of the variable of interest tend to be surrounded by large (small) values - positive autocorrelation - or large (small) values tend to be surrounded by small (large) values - negative autocorrelation. Whereas positive spatial autocorrelation implies a spatial clustering of similar values, negative spatial autocorrelation implies a checkerboard pattern of values. Tests for spatial autocorrelation are based on the magnitude of an indicator that combines the value observed at each location with the values at neighbouring locations. Basically, the tests are measures of the similarity between association in value (correlation) and association in space (contiguity).

The information on the location of the observations is usually summarised in a spatial weights matrix, denoted by W, of dimension equal to the number of observations (NxN). The element corresponding to row r and column c - w_{rc} - is different from zero if observations - spatial sites - r and c are neighbours, and zero

[2] The traditional approach is that of Cliff and Ord (1981). A recent survey is Anselin and Florax (1995).

otherwise. By convention, w_{rc}=0 if $r=c$. The simplest neighbourhood criterion states that two districts are neighbours if they share a border. In this case, w_{rc} is equal to one if they do, and is equal to zero otherwise. However, more sophisticated criteria have been proposed - based, for instance, on inverse distance or inverse distance squared, or computed from the relative length of the shared border.[3] While the original matrix W is symmetric, it is usually standardised such that the elements of a row sum to one.[4] This is obtained by dividing each element in the matrix by its row sum. As a result, the row-standardised weights matrix W^* is not symmetric any more. The most commonly used measure for spatial autocorrelation is the Moran's I statistic (Cliff and Ord, 1981; Anselin, 1988a). The Moran's I statistic for N observations on a variable x is expressed as:

$$I(x) = \frac{N}{S_0} \frac{\sum_{i,j=1}^{N} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \qquad (3.1)$$

where x_i is the level of public expenditure per head observed at location i, $\bar{x}$ is the average expenditure, w_{ij} is the element corresponding to row i and column j in the spatial weights matrix W, and S_0 is the sum of the elements of the weights matrix:[5]

$$S_0 = \sum_i \sum_j w_{ij} \qquad (3.2)$$

---

[3] Furthermore, it has been suggested that geographical proximity may not be the most relevant factor. Local jurisdictions may regard as neighbours other jurisdictions that are similar to them from an economic or demographic point of view, regardless of geographic proximity (as in Case *et al.*, 1993). Since the elements of W cannot be estimated, the choice of the neighbourhood criterion is to some extent arbitrary. More sophisticated methods for defining and constructing neighbourhood matrices can be found in Okabe *et al.* (1992).

[4] Whether or not the neighbourhood matrix should be row-standardised clearly depends on the underlying economic model. The fact that the spatial weights sum to one for each local jurisdiction (row-sum division) means that the total effect of all neighbours is the same, regardless of the number of neighbours.

[5] For a row-standardised spatial weights matrix, the term (N/S_0) in (3.1) obviously disappears.

Inference for the I(x) statistic is carried out by computing a z value, under the hypothesis that the x's are drawn from a normal distribution. Moran's I(x) statistic for the 296 English shire districts' current expenditures per capita in the financial years 1986/87 to 1989/90 - reported in table 3.1 - provides a strong indication of significant positive spatial autocorrelation.

3.3 Tests for spatial dependence in a regression model

3.3.1 We first need to specify a model of local public expenditure determination that allows us to remove the effects of observable variables on local spending, and investigate whether we still find evidence of spatial correlation. If this happens to be the case, we should re-estimate the local public expenditure equation by explicitly allowing for spatial interactions.

In the presence of a non-linear grant distribution system - as the one we have analysed in chapter two - a model of local spending that linearly relates public expenditure per head to financial and demographic variables, is not in principle appropriate. When the budget constraint that local authorities face is non-linear, the demand function for local public spending is non-linear in price and income, and is even non-differentiable at many points. Furthermore, the use of a linear regression equation does not allow for unobserved heterogeneity of preferences.

The grant distribution system that was in place in Britain until 1990 (the Block Grant) creates exactly this sort of problems. As we have described in chapter two, two crucial determinants of local spending depend on the level of public expenditure itself, and are therefore endogenous. These are the price of local public services - net of central government's matching grant - and the lump-sum component of the grant. As regards the former, the central government reduces its marginal contribution to local spending when the level of local expenditure reaches a threshold level, thereby increasing the marginal cost of public services to local residents - a price effect. As for the latter, the lump-sum component of the grant increases for spending above the threshold - an income effect. This creates a piecewise-linear budget constraint (see section **2.2**).

In order to consistently estimate a demand function for public spending under such circumstances, the choice of segment must be modelled and estimated along with the demand function conditional upon choice of segment. The parameters of the conditional demand function for public expenditure can be estimated with a two-error Maximum Likelihood procedure. However, the latter procedure turns out not to be feasible for computational reasons in the presence of spatial autocorrelation.[6] Consequently, we must use a model which is a reasonably good approximation to the non-linear one, and yet allows us to test for spatial correlation and get estimates of the spatial correlation coefficients.

The results in chapter two (section **2.5**) show that the OLS estimates of the marginal demand functions are reasonably close to the two-error Maximum Likelihood estimates. This approach consists in regressing each local government's level of expenditure on the price and income variables of the segment it is observed to be located onto. Furthermore, we know the direction of the bias in the price and income coefficient estimates. As a result, in the following we implement our spatial analysis by using the approximation represented by the linear model.

3.3.2 As the analysis in section **3.2** suggests, there are two possible forms of spatial correlation. The first is an interaction between the dependent variables x, which is due to the fact that local jurisdictions tend to set levels of spending which are correlated with those of their neighbours. Following Anselin (1993), we refer to this kind of interaction as "substantive spatial dependence" or "spatial lag dependence," and express it (in matrix notation) as:

---

[6] The two-error Maximum Likelihood (section **2.4**) allows both for an unobserved heterogeneity term, and for random optimisation error. When we allow for general forms of spatial correlation, observations are not independent any more. As a result, the likelihood function for a vector of N observations (given in appendix 2.1 for the case of independent observations) is now given by the sum of the joint probabilities that the N heterogeneity terms are such as to make local authorities optimise on a certain combination of segments (and kinks), and that the vector of optimisation error terms moves the authorities to the observed vector of spending. Unfortunately, it is not possible to implement this procedure with traditional methods. The number of combinations involved makes the maximisation of the likelihood function extremely cumbersome.

$$X = \rho WX + Z\alpha + \xi \qquad (3.3)$$

where X is a (Nx1) vector of expenditures in N local jurisdictions, and the explanatory variables in the mixed regressive-spatial autoregressive specification (3.3) consist of the usual set of variables in Z - a (NxK) matrix - and of the so-called spatial lag (WX). In (3.3), the dependent variable follows a first order spatially autoregressive process. ρ - where $|\rho|<1$ - is the spatial autoregressive coefficient, and ξ is a vector of spatially uncorrelated error terms. The elements of the neighbourhood matrix W are known and fixed.

The second form of spatial dependence arises when the error term in the regression equation follows a spatial process. This form of spatial pattern is referred to as "spatial error dependence". The hypothesis of first order spatially autoregressive errors implies that the error term in each local jurisdiction can be expressed as a spatially weighted average of its neighbours' errors, plus an i.i.d. disturbance term with mean zero and fixed finite variance (σ^2). In matrix form, we write the spatial error dependence model as:

$$\begin{aligned} X &= Z\alpha + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + \xi \end{aligned} \qquad (3.4)$$

where λ is in this case the parameter measuring spatial dependence in the errors, and $|\lambda|<1$.[7] The corresponding error variance is of the form:

$$E(\varepsilon\varepsilon') = \sigma^2 (I - \lambda W)^{-1}[(I - \lambda W)^{-1}]' \qquad (3.4')$$

---

[7] Kelejian and Robinson (1993) have suggested a form of spatial error components, in which the spatial dependence in the error term is: $\varepsilon=Wv+\eta$, where v and η are uncorrelated error terms. The two error terms can be interpreted as two stochastic shocks generated within each of N regions. One of them (η) is unique to the region, that is it does not lead to spillovers in other regions. The other (v) leads to such spillover effects via a spatial weighting matrix, W.

The two forms of spatial dependence that we have described - models (3.3) and (3.4) - can be seen as special cases of the following general model, which nests the models with spatial correlation in the dependent variable and in the errors:[8]

$$
\begin{aligned}
X &= \rho WX + Z\alpha + \varepsilon \\
\varepsilon &= \lambda W\varepsilon + \xi
\end{aligned}
\qquad (3.5)
$$

The consequences on the model parameter estimates from omitting the two forms of dependence are different. If substantive spatial dependence is present, but the spatial lag term WX is omitted from the model, the estimates for the regression coefficients α will be biased. When spatial error dependence is present, but ignored, the OLS estimates are no longer efficient, but they remain unbiased. The inefficiency arises from the non-diagonal structure of the disturbance matrix. However, since the mean of the error term ε is zero - irrespective of the value of λ - the mean of X is not affected by the spatial error dependence.

As regards the estimation of model (3.3), the correlation of the spatial lag WX with the error term invalidates the optimality of the OLS estimator. The crucial difference with respect to the time series econometrics - where the OLS estimator remains consistent even when a lagged dependent variable is present, as long as the error term is serially uncorrelated - is that the OLS estimator is biased as well as inconsistent, irrespective of the properties of the error term (Anselin, 1988a). This is due to the multidirectional nature of dependence in space, as opposed to the unidirectional nature of dependence in time. Spatial dependence is present in all directions, though it becomes weaker as data locations become more dispersed. Section **3.4** presents the econometric procedure for estimating models (3.3), (3.4), and (3.5).

3.3.3 As regards spatial error dependence, we wish to test the hypothesis that the disturbances are independently distributed, against the alternative hypothesis that

[8] Another way of combining the two forms of spatial dependence is by using a SARMA (spatial autoregressive moving average) specification. This specification combines a spatial lag of the dependent variable with a moving average error: $X=\rho WX+Z\alpha+\theta W\xi+\xi$ (Anselin and Florax, 1995).

they are generated by a first order spatial autoregression. The most popular and widely applied diagnostic for spatial error dependence is an application of the Moran's I statistic to the residuals of an OLS regression of X on Z. For a row-standardised spatial weights matrix, the Moran's I(e) is defined as:

$$I(e) = \frac{e' We}{e' e} \tag{3.6}$$

where e are the OLS regression residuals. For normal error terms, the I(e) statistic is asymptotically normal (Cliff and Ord, 1981). In practice, the statistic is converted to a z value that is then compared to a standard normal distribution. The mean and variance of the I(e) statistic are reported in appendix 3.2.[9]

However, several studies and Monte Carlo simulation experiments[10] have found that the Moran's I for OLS regression residuals has two major drawbacks. First, the test is very sensitive to the presence of other forms of specification error, such as non-normality and heteroscedasticity. Second, Moran's I(e) is not able to discriminate properly between the two forms of spatial autocorrelation: spatial error dependence and spatial lag dependence.

An alternative to the Moran's I(e) is the use of tests based on the Lagrange Multiplier (LM) principle. They are also computed from the OLS regression residuals. In contrast with the Moran's I(e), the two LM tests - LM(err) for spatial error dependence and LM(lag) for spatial lag dependence - should provide an indication of the most likely alternative hypothesis, while being close in power to Moran's I(e) for spatial error dependence and superior for spatial lag dependence.[11]

The LM test for spatial error dependence - which is a scaled Moran's I(e) - is:[12]

---

[9] The I(e) statistic for a moving average alternative hypothesis has the same form. Burridge (1980) proved that the Moran's I(e) test is proportionate to a LM test against either a spatial autoregressive or a spatial moving average model.

[10] Anselin (1988b, 1993), Anselin and Rey (1991), Anselin and Florax (1995).

[11] Anselin and Florax (1995).

[12] This test was originally proposed in Burridge (1980).

$$LM(err) = \frac{\left[\frac{e' We}{(e' e)/N}\right]^2}{tr(W' W + W^2)} \tag{3.7}$$

with *tr* as the matrix trace operator. The LM test for spatial lag dependence is:[13]

$$LM(lag) = \frac{\left[\frac{e' WX}{(e' e)/N}\right]^2}{\frac{(WZa)' M(WZa)}{(e' e)/N} + tr(W' W + W^2)} \tag{3.8}$$

where: $M=I-Z(Z'Z)^{-1}Z'$, and a is the OLS estimate of α. Both tests are asymptotically distributed as χ^2 with one degree of freedom. The Moran' I(e) test statistic and the LM tests only require OLS residuals, and do not necessitate the estimation of the more complex spatial models.[14] Table 3.2 reports the results of the above tests. We run an OLS regression of local public expenditure per head on a matrix of explanatory variables that includes income,[15] grant,[16] price of public services,[17] population, and a series of political and socio-demographic characteristics (a dummy variable for Labour-controlled council, a dummy variable for closeness to a metropolitan area, urbanisation rate, unemployment rate, age and ethnic structure of population).

Of course, one of the reasons why local public expenditure exhibits spatial autocorrelation might be the fact that the determinants of local spending are themselves spatially correlated. Simple application of the Moran principle to the

---

[13] Anselin (1988b).

[14] As a result, we postpone to section **3.5** - i.e., after estimation of the spatial models - the discussion of further tests for spatial dependence.

[15] Income data are not available at the district level. As a proxy, county level disposable income per head is used instead.

[16] Grant is the lump-sum component of the total grant.

[17] The price of public services takes into account the matching grant rate. It is the number of pounds of private income that have to be given up in order to afford one additional pound of expenditure on public services. In the absence of a matching grant, this price would equal one. Given the grant distribution system (the British Block Grant system described in detail in chapter two), the matching grant can be positive or negative, according to the level of the local tax base. If the local tax base is smaller (larger) than a threshold level set by central government, the matching grant is positive (negative), and the price of public services in terms of private income is less (more) than one.

regressors of the above model shows that this is actually the case. The z value of the Moran's I(x) statistic ranges from 9.6 for the Labour incumbency variable, to 14.7 for the rate of urbanisation, and to 16.3 for the ethnic structure of population.

After controlling for those characteristics, however, the Moran's I(e) statistic still points towards some form of positive spatial dependence (table 3.2), but - as we stressed before - it is unable to discriminate properly between lag and error dependence.

On the other hand, the LM tests do not give a clear-cut response either. Since the two LM tests are distributed as χ^2 with one degree of freedom, they both lead us to reject the null hypothesis of absence of spatial dependence.[18]

3.4 Estimation in the presence of spatial correlation

3.4.1 The results in the previous section provide a strong indication of significant positive spatial correlation in the expenditure of the English local authorities. If that is the case, the regression model for local public spending should have a spatial structure. The most general spatial model incorporates spatial correlation in the dependent variable and in the errors:

$$\begin{aligned} X &= \rho \mathbf{W} X + Z\alpha + \varepsilon \\ \varepsilon &= \lambda \mathbf{W} \varepsilon + \xi \end{aligned} \qquad (3.5)$$

Assume that we have T (small) time-series observations on a cross-section of N (large) local districts, and that the NT observations in (3.5) are ordered wave by wave, so that the N observations for period t (t=1,...,T) are grouped together. X, ε, and ξ are (NTx1) vectors. Besides k explanatory variables, matrix Z includes year and district (T+N) indicators. The neighbourhood matrix is in this case: $\mathbf{W}=I_T \otimes W$,

[18] The critical level for p=.99 is 6.63.

where I_T is the (TxT) identity matrix and W is the (NxN) neighbourhood matrix defined in section **3.2.**[19]

As regards estimation of model (3.5), we need to tackle two main problems. The first is obviously the presence of a first-order spatial autoregressive process in the dependent variable and in the errors. The second is that we are dealing with a panel data set and the matrix of explanatory variables Z comprises year and district effects. We treat them as fixed, and estimate them along with the coefficients on the set of explanatory variables. While the time dummies do not create any additional problem, the presence of the district fixed effects requires more careful treatment and is postponed to section **3.4.2.**

The standard way of estimating model (3.5) is by ML methods. Since observations are not independent, the spatially correlated variates must be transformed into uncorrelated ones. This is done by writing (3.5) as:

$$X = (I - \rho \mathbf{W})^{-1} Z\alpha + (I - \rho \mathbf{W})^{-1}\varepsilon \tag{3.9}$$

$$\varepsilon = (I - \lambda \mathbf{W})^{-1}\xi \tag{3.10}$$

The two matrices (I-ρ**W**) and (I-λ**W**) are invertible if $-1<\rho<1$ and $-1<\lambda<1$ respectively (Case, 1991). By using (3.9) and (3.10), model (3.5) can be written as:

$$(I - \lambda \mathbf{W})(I - \rho \mathbf{W})X - (I - \lambda \mathbf{W})Z\alpha = \xi \tag{3.11}$$

If ξ is normally distributed with variance σ^2, the log-likelihood for a sample of NT observations is:

[19] The results of the tests performed in section **3.3** suggest that a row-standardised neighbourhood matrix should be used. A row-standardised weighting matrix based on the criterion that two districts are neighbours if they share a boundary imposes two restrictions. First, all neighbours of a district are given equal weight. Second, all districts are equally influenced by their neighbours taken together. However, since this kind of weighting matrix is widely used in empirical work, our results can be compared with those obtained in previous studies (for instance, Case, 1991, and Case *et al.*, 1993).

$$L = c + \ln(J_1) + \ln(J_2) - \frac{NT}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(H_2X - Z\alpha)' H_1' H_1(H_2X - Z\alpha) \quad (3.12)$$

where: $H_1=I\text{-}\lambda\mathbf{W}$, $H_2=I\text{-}\rho\mathbf{W}$, $J_1J_2=|\det(H_1)|\,|\det(H_2)|$ is the Jacobian of the transformation between ξ and X, and c is a constant.

Each district's right hand side variables (Z) must be uncorrelated with own and neighbours' errors. Consequently, district j's independent variables affect district k's expenditures only through their effect on j's expenditure (through parameter ρ). The correlation which is left after controlling for this effect identifies the second spatial correlation parameter λ.

Special cases of model (3.5) are obtained by setting ρ=0 or λ=0. In the former case, we get the spatial error dependence model of section **3.3**:

$$\begin{aligned} X &= Z\alpha + \varepsilon \\ \varepsilon &= \lambda\mathbf{W}\varepsilon + \xi \end{aligned} \quad (3.4)$$

If ξ is normally distributed with variance σ^2, the log-likelihood for a sample of NT observations is:

$$L = \ln(J_1) - \frac{NT}{2}\ln(2\pi) - \frac{NT}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(X - Z\alpha)' H_1' H_1(X - Z\alpha) \quad (3.13)$$

where: $H_1=I\text{-}\lambda\mathbf{W}$, and $J_1=|\det(H_1)|$ is the Jacobian of the transformation between ξ and X. The log-likelihood in (3.13) can be maximised jointly for all parameters. Alternatively, it can be computationally advantageous to use the following iterative procedure. The α parameters are estimated conditional upon the λ parameter, and λ is estimated conditional upon the α, until convergence is achieved (Anselin, 1988a).

If we set λ=0 in model (3.5), we get the substantive spatial dependence model:

$$X = \rho\mathbf{W}X + Z\alpha + \xi \quad (3.3)$$

For normal error terms, the log-likelihood for model (3.3) is:

$$L = \ln(J_2) - \frac{NT}{2}\ln(2\pi) - \frac{NT}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(H_2X - Z\alpha)'(H_2X - Z\alpha) \qquad (3.14)$$

where: $H_2=I-\rho\mathbf{W}$, and $J_2=|\det(H_2)|$ is the Jacobian of the transformation between ε and X. The log-likelihood needs to be maximised with respect to α, ρ and σ^2. The first order conditions for maximisation of L yield the following estimates for α and σ^2:

$$a = (Z'Z)^{-1}Z'H_2X = a_0 - \rho a_L \qquad (3.15)$$

$$s^2 = \frac{1}{NT}(e_0 - \rho e_L)'(e_0 - \rho e_L) \qquad (3.15')$$

where a_0 and a_L are OLS estimates of a regression of X and $\mathbf{W}X$ on Z, and e_0 and e_L are the corresponding predicted residuals. As a result, the log-likelihood can be concentrated, and needs to be maximised with respect to ρ only:

$$L^c = k - \frac{NT}{2}\ln\left[\frac{1}{NT}(e_0 - \rho e_L)'(e_0 - \rho e_L)\right] + \ln(J_2) \qquad (3.16)$$

where k is a constant. The presence of the Jacobian can considerably complicate the numerical analysis. Hence, the determinant can be expressed as a function of the eigenvalues of the spatial weights matrix (h_i):

$$\ln(J_2) = \ln\prod_{i=1}^{NT}(1-\rho h_i) = \sum_{i=1}^{NT}\ln(1-\rho h_i) \qquad (3.17)$$

3.4.2 As model (3.5) stands ML estimation is not feasible due to the presence of the fixed district and time effects. With k independent variables, this would require estimation of (N+T+k+3) parameters. However, the problem can be solved by finding the transformation that projects out the N district effects (Hsiao, 1986). After finding the estimator of the fixed effects (γ), we can concentrate the likelihood with respect to γ and maximise it using non-linear estimation techniques. Consider first a standard non-spatial panel data model:

$$X = Z\alpha + \xi \qquad (3.18)$$

By partitioning matrix Z, (3.18) can be written as:

$$X = Z_1\beta + Z_2\gamma + \xi \qquad (3.18')$$

where γ are the N district effects, and Z_2 is:

$$Z_2 \equiv i_T \otimes I_N \qquad (3.19)$$

where i_T is a column vector of ones of length T, and I_N is the (NxN) identity matrix. The transformation that projects out the fixed effects is achieved by the matrix:

$$M \equiv \left[I_T - i_T i_T' \tfrac{1}{T}\right] \otimes I_N \qquad (3.20)$$

The OLS estimators of β and γ are:

$$b = (Z_1' M Z_1)^{-1} Z_1' MX \qquad (3.21)$$

$$g = (Z_2' Z_2)^{-1} Z_2' (X - Z_1 b) \qquad (3.22)$$

By using the above expressions, the likelihood function of model (3.18') can be concentrated, and only needs to be maximised with respect to σ^2.

Analogously, estimators of the coefficients in models (3.3), (3.4) and (3.5) allow us to concentrate the likelihood functions (3.12), (3.13) and (3.14) - Besley and Case (1993). In particular, the estimators of β and γ in model (3.5) are given by:

$$\tilde{b} = (Z_1' H_1' M H_1 Z_1)^{-1} Z_1' H_1' M H_1 H_2 X \qquad (3.23)$$

$$\tilde{g} = (Z_2' H_2' H_2 Z_2)^{-1} Z_2' H_2' (H_1 H_2 X - H_1 Z_1 \tilde{b}) \qquad (3.24)$$

(3.23) and (3.24) can be used to concentrate the log-likelihood (3.12). Analogous expressions can be derived for the spatial lag dependence model and the spatial error dependence model.

3.5 Results

We use as dependent variable the current public expenditure per head of the English non-metropolitan districts in the financial years 1986/87 to 1989/90. Capital expenditures are excluded. Explanatory variables are divided into a set which vary over time, and a set which do not (at least in the available data). The former comprise disposable income per head, grant per head, price of public services in terms of private consumption, population and a dummy variable for political control. The latter include data on socio-demographic characteristics.

Since income data are not available at the district level, county level disposable income per head is used instead. The financial variables are as defined in chapter two. In particular, grant per head is the lump-sum component of the grant corresponding to the segment of the budget constraint a local authority locates onto. Price is the ratio of the actual per capita tax base to the tax base implicit in the relevant segment of the budget constraint.

With respect to the empirical analysis in chapter two, however, two main differences should be noted. First, we do not control here for the endogeneity of the price and grant variables. A detailed analysis of how this can be done is in section **2.4**. Second, grant and income enter the expenditure determination equation separately, while in chapter two we computed the virtual income corresponding to each segment and used that variable as a regressor. The reason is that here we want to allow for the possibility that grant and income have different effects on the level of expenditure.[20] The means and data sources of all variables used in the analysis are shown in appendix 3.3. The results are reported in tables 3.3 and 3.4. While in table 3.3 we allow for time effects, but not for district-specific effects, in table 3.4 we allow for both time and district effects, and treat them as fixed - as explained in section **3.4.2**.

In table 3.3, the ML estimates of a conventional linear model of public spending determination with no spatial components (first column) are compared to ML estimates of the spatial models described in section **3.4**, models (3.3), (3.4) and (3.5). The ML estimates of the spatial lag dependence model (3.3) and spatial error dependence model (3.4) are shown in columns 2 and 3 of table 3.3, where ρ and λ represent the spatial lag dependence and the spatial error dependence coefficients respectively. The results show that both models can be used to reject a null hypothesis of absence of spatial interactions. Both models achieve a significant increase in likelihood and a decrease in the regression variance, and both spatial coefficients are positive and highly statistically significant. Estimates of model (3.3) yield a significant spatial component in the dependent variable (ρ=.13), while estimates of model (3.4) yield a large significant spatial component in the errors (λ=.25).

In order to check whether spatial effects really matter in the specification of the local public expenditure determination model, we compute a Likelihood Ratio (LR) test. Twice the difference in log-likelihood of the restricted (no spatial effects) and unrestricted (either the spatial lag dependence model, or the spatial error dependence model) model is distributed as χ^2 with one degree of freedom. In both cases, the test statistic - conditional on the other parameter being zero - exceeds by a wide margin the $\chi^2(1)$ value at p=.99 of 6.63, as shown in table 3.3. We therefore reject the restrictions that equate the autoregressive parameters to zero.

[20] Even though a structural economic model of local public expenditure determination (based upon the constrained maximisation of a local social welfare function on the part of a benevolent government) implies that the effects of local income and grant on the level of spending are the same, several empirical works in the public finance literature have found that the effect of grant turns out to be much larger - the so-called "flypaper effect". As a result, we allow the two variables - income and grant - to affect the level of spending separately.

The most direct way of discriminating between the two models is by means of a model with spatial correlation in the dependent variable and in the errors - as suggested by Case (1991). We therefore estimate model (3.5), that nests the spatial lag dependence and spatial error dependence models, in order to learn more about their relative importance. Estimates of model (3.5) are reported in the fourth column of table 3.3. With respect to the spatial error dependence model, model (3.5) leads to an insignificant increase in likelihood. After spatial autocorrelation in the errors has been accounted for, the autoregressive parameter on the spatially lagged dependent variable turns out not to be significantly different from zero.

However, the presence of district-specific effects in the equation of local public expenditure may bias the β coefficient estimates if the fixed effects are correlated with the explanatory variables. Furthermore, the spatial correlation that emerges from the results in table 3.3 might be due to the fact that the districts' unobservable fixed effects are spatially correlated. As a result, we implement the estimation procedure outlined in **3.4.2** and present estimates of the fixed effects model in table 3.4. In the first place, the fixed effects non-spatial model achieves an increase in log-likelihood over the correspondent model in table 3.3 of over 20%. Second, the parameter estimates exhibit a significant difference with respect to the estimates presented in table 3.3.

In particular, the coefficient estimates on grant and price in the fixed effects model show strong income and price effects, though being substantially smaller than their counterparts in table 3.3.

As for the separate effects of grant and income on the level of local spending, the former turns out to be substantially larger than the latter. This result could be taken as a further example of a "flypaper effect". However, it should be taken into account that income is measured with error at the district level (since only county level income data are available) and the income coefficient is therefore biased downwards. As a result, it is not possible to conclude here whether the divergence between the grant and income coefficients is due to measurement error in income, or rather to a substantially different perception on the part of local authorities of local income versus transferred income.

The coefficient estimate on Labour incumbency in table 3.3 shows an upward bias with respect to the fixed effects model estimate in table 3.4, while population appears to have a significant negative effect on expenditure per capita.

As regards the spatial autoregressive parameters, the results in table 3.4 show that, after controlling for the district-specific fixed effects, none of the spatial models improves significantly upon the non-spatial specification. All tests fail to reject the hypothesis that ρ and λ are equal to zero. As a result, the spatial correlation in local public spending can be attributed to a large extent to spatially autocorrelated unobservable fixed effects.

It should be recognised that the model that we have estimated here has no dynamics, while in reality local public expenditure might adjust in a sluggish way over time. This could be due, for instance, either to the presence of adjustment costs, or to genuinely multiperiod policy making. If this is the case, the econometric model ought to include lagged values of expenditure among the regressors.

Estimation of a dynamic model would certainly contribute to reduce the unexplained part of the regression - thereby reducing the importance of the fixed effects - even though it would almost prohibitively increase the computational burden.

3.6 Conclusions

By using data on the English local governments in the late 1980s, this chapter has explored whether local public expenditure exhibits a spatial pattern. We have computed measures of spatial correlation in the raw data - i.e., the level of local public spending per head - and performed tests based on the OLS regression residuals from the estimation of an equation of local public expenditure determination. First, the Moran's I(X) statistic indicated a significant degree of positive spatial autocorrelation. Second, all tests for spatial correlation - including a Moran's I statistic applied to the OLS residuals, and two LM tests - suggested that the regression model should have a spatial structure.

The results from the estimation of the spatial lag dependence and spatial error dependence models suggest that spatial autocorrelation is an important feature of local governments' expenditure. Both models can be used to reject a null hypothesis of absence of spatial interactions, i.e., both models are superior to a model that arbitrarily constrains the two spatial coefficients to be zero. In order to discriminate between the two models, we have estimated a general model that nests them. When allowing both for a spatially lagged dependent variable and for a spatially autocorrelated error term, it turns out that most of the correlation in public expenditure is captured by correlated shocks.

However, when we control for the presence of district-specific fixed effects, much of the residual spatial correlation disappears. As a result, we can conclude that the significant positive spatial association that we observe in local expenditure could most likely be attributed to spatially autocorrelated unobservable district-specific effects.

Of course, it must be kept in mind that in order to implement the spatial econometric analysis in this chapter, we had to use a linear approximation to the non-linear model presented in chapter two. In spite of the non-linear constraint faced by local governments, we have treated the lump-sum component of the grant and the price of public services as exogenous determinants of the level of expenditure. The model coefficient estimates should therefore be used with some caution.

Nevertheless, we think that it was a price worth paying for two reasons. First, we have been able to show the advantage of the use of panel data over cross-sectional analysis. Once we control for the presence of district-specific effects, the explanatory power of the model increases dramatically and the coefficient estimates change considerably. Second, the above analysis suggests that spatial dependence is indeed an important issue in any empirical study that uses data on local governments' expenditures.

In the next chapter, we explore in much greater detail the process of local policy-making in the presence of spatial correlation. In particular, we suggest that spatial dependence among local jurisdictions can help interpret the relationship between voters and officials at the local authority level.

Table 3.1 Test for spatial correlation in local public expenditure per head

	Moran's statistic
I(*x*)	.202
z	10.85

Notes:
I) the spatial weights matrix for the I(*x*) statistic is row-standardised, and is based on the criterion that two districts are neighbours if they have a common border;
II) the I(*x*) statistic is computed under the null hypothesis of normality;
III) the moments of the I(*x*) statistic are reported in appendix 3.2;
IV) number of observations=1184.

Table 3.2 Tests for spatial dependence based on OLS regression residuals

I(*e*)	LM(*err*)	LM(*lag*)
5.12*	23.83*	17.25*

Notes:
I) the spatial weights matrix is row-standardised, and is based on the criterion that two districts are neighbours if they have a common border;
II) the I(*e*) statistic is computed under the null hypothesis that the errors are normally distributed;
III) the moments of the I(*e*) statistic are reported in appendix 3.2;
IV) I(*e*) is distributed as a standard normal *z*(0,1); LM(*err*) and LM(*lag*) are χ^2 with one degree of freedom;
V) *=significant at the 1% level;
VI) the residuals are obtained from an OLS regression on the pooled cross-sections (number of observations=1184);

Table 3.3 ML estimates of the spatial models

	Non-spatial model ML estimates	Model (3.3) ML estimates	Model (3.4) ML estimates	Model (3.5) ML estimates
λ			.244 (5.58)	.257 (2.71)
ρ		.130 (4.17)		-.013 (-.17)
income	.0022 (3.09)	.0024 (3.15)	.0016 (1.88)	.0015 (1.73)
price	-37.06 (-15.46)	-35.89 (-15.02)	-37.99 (-15.31)	-38.07 (-14.15)
grant	.518 (13.83)	.505 (13.81)	.499 (13.77)	.498 (13.60)
LabourD	15.44 (12.50)	15.27 (12.43)	15.91 (12.94)	15.93 (12.90)
population (,000)	.026 (2.33)	.029 (2.76)	.036 (3.29)	.036 (3.19)
urbanisation rate (%)	.145 (9.13)	.137 (8.68)	.133 (8.38)	.133 (8.34)
MetroD	4.99 (4.66)	4.04 (3.75)	4.38 (3.90)	4.42 (3.53)
age > 65 (%)	.666 (5.88)	.629 (5.65)	.669 (5.75)	.673 (5.56)
ethnic minority (%)	.404 (1.43)	.633 (2.40)	.614 (2.28)	.606 (1.97)
long term unemployed (%)	.292 (2.18)	.299 (2.33)	.283 (2.14)	.281 (2.07)
σ	13.40	13.27	13.14	13.13
LR test - χ^2		17.57 (1)	29.21 (1)	29.32 (2)
observations	1184	1184	1184	1184

Notes:

I) Model (3.3) is the spatial lag dependence model; Model (3.4) is the spatial error dependence model; Model (3.5) is the comprehensive model with spatial dependence in both the dependent variable and the errors;
II) dependent variable = local public expenditure per head;
III) *t* values in parentheses;
IV) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;
V) MetroD=1 if a shire district is a neighbour of a metropolitan authority, and 0 otherwise;
VI) time dummies are included;
VII) individual effects (district dummies) are *not* included.

Table 3.4 ML estimates of the spatial models

	Non-spatial model ML estimates	Model (3.3) ML estimates	Model (3.4) ML estimates	Model (3.5) ML estimates
λ			.035 (.71)	.142 (1.70)
ρ		.008 (.35)		-.121 (-1.41)
income	-.0007 (-.77)	-.0007 (-.59)	-.0007 (-.77)	-.0007 (-.70)
price	-27.63 (-12.61)	-27.63 (-12.59)	-27.76 (-12.54)	-28.01 (-12.78)
grant	.231 (4.80)	.231 (4.79)	.231 (4.85)	.230 (4.81)
LabourD	3.54 (2.13)	3.56 (2.14)	3.70 (2.19)	3.92 (2.33)
population (,000)	-.228 (-5.09)	-.228 (-5.03)	-.226 (-4.99)	-.218 (-4.81)
σ	5.74	5.74	5.73	5.71
LR test - χ^2		.05 (1)	.59 (1)	2.04 (2)
observations	1184	1184	1184	1184

Notes:
I) Model (3.3) is the spatial lag dependence model; Model (3.4) is the spatial error dependence model; Model (3.5) is the comprehensive model with spatial dependence in both the dependent variable and the errors;
II) dependent variable = local public expenditure per head;
III) *t* values in parentheses;
IV) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;
V) time and district effects are included.

Appendix 3.1

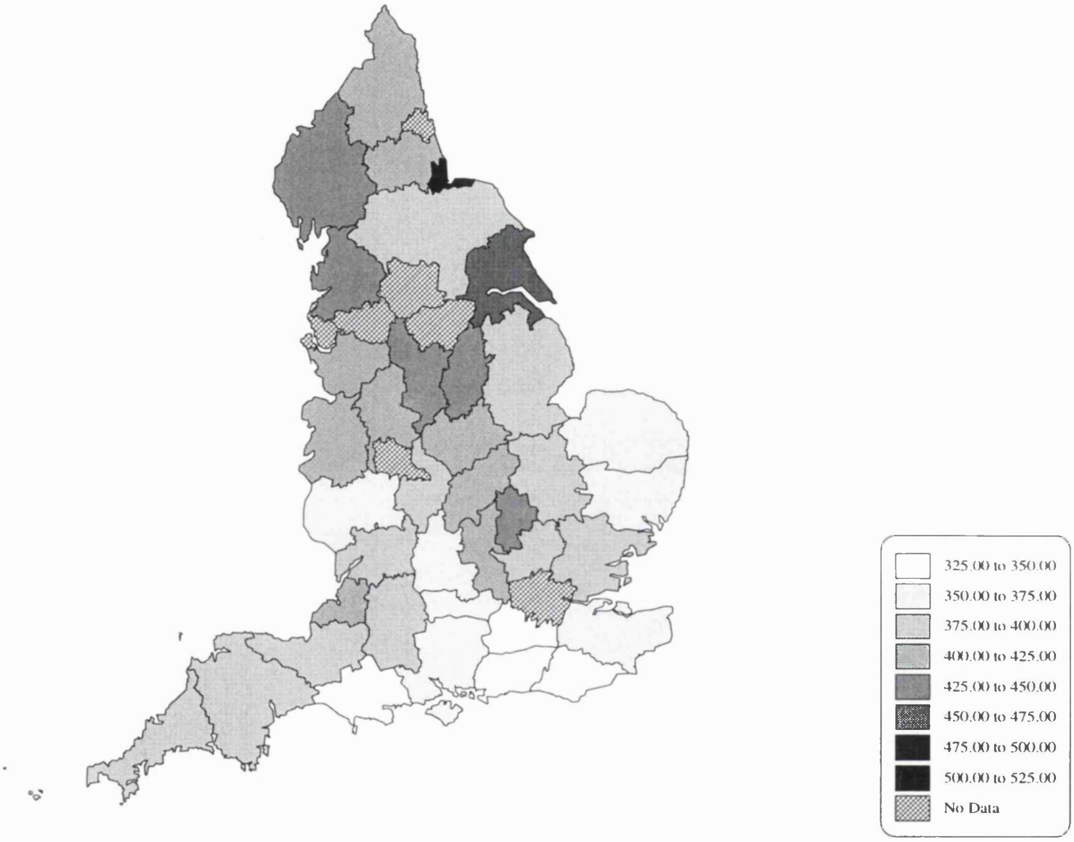
We can get a qualitative and intuitive insight into the extent of spatial dependence by looking at the spatial pattern exhibited by the data. We look at the expenditures of the English non-metropolitan authorities. The two-tier system of local government in England includes 39 shire counties and 296 shire districts. During the 1980s, the local property tax - the rate - was independently set at both levels, according to the share of responsibilities, and corresponding expenditure needs, attributed to each kind of authority.

Consider the last four years of operation of the system of locally varying rates (1986/87 to 1989/90). The spatial pattern exhibited by mean real local expenditure (£ per head) is shown in chart A3.1. The chart shows the 39 English non-metropolitan counties,[21] and provides a first insight into the spatial pattern displayed by local spending.

In the first place, high-spending (low-spending) authorities tend to be surrounded by high-spending (low-spending) authorities, pointing towards positive spatial correlation. In the second place, high-spending authorities seem to be clustered around the metropolitan areas. Third, some regions show a highly homogeneous composition of local authorities - the North and North West as regards high-spending authorities, and the South West and East Anglia for low-spending authorities. Overall, the degree of spatial clustering is impressive.

[21] Data on the metropolitan areas are not available, since the metropolitan counties were abolished in 1985.

Chart A3.1 Mean county expenditure 1986/87 - 1989/90 (£ per head)



Appendix 3.2

The Moran's I statistic for N observations on a variable x is expressed as:

$$I(x) = \frac{N}{S_0} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \tag{A3.1}$$

where: $S_0 = \sum_i \sum_j w_{ij}$.

Moran's I(x) is asymptotically normally distributed (Cliff and Ord, 1981). The mean and variance of I(x), evaluated under the assumption that the x's are the results of N independent drawings from a normal distribution, are:

$$E(I(x)) = -\frac{1}{N-1} \tag{A3.2}$$

$$Var(I(x)) = \frac{N^2 S_1 - N S_2 + 3 S_0^2}{S_0^2 (N^2 - 1)} - \frac{1}{(N-1)^2} \tag{A3.3}$$

where:

$$S_1 = \tfrac{1}{2} \sum_i \sum_j \left(w_{ij} + w_{ji}\right)^2 \tag{A3.4}$$

$$S_2 = \sum_i \left(w_{i\cdot} + w_{\cdot i}\right)^2$$

$$w_{i\cdot} = \sum_j w_{ij}$$

$$w_{\cdot i} = \sum_j w_{ji}$$

A test based on the application of the Moran's statistic to the OLS residuals of a linear regression model (X=$Z\alpha$+ε) is the following:

$$I(e) = \frac{N}{S_0} \frac{e' We}{e' e} \qquad \text{(A3.5)}$$

where e are the OLS residuals. For normal error terms, the distribution of the Moran's I(e) statistic is asymptotically normal (Cliff and Ord, 1981). In order to carry out an operational test (Anselin and Hudak, 1992), we need the expected value and the variance of $I(e)$:

$$E(I(e)) = -\frac{N \cdot tr(A)}{S_0(N-K)} \qquad \text{(A3.6)}$$

where N–K are the degrees of freedom, $A=(Z'Z)^{-1}Z'WZ$, and tr is the matrix trace operator. For K=1 (constant term only), (A3.6) reduces to (A3.2).

$$Var(I(e)) = \frac{N^2}{S_0^2(N-K)(N-K+2)} \left\{ S_1 + 2tr\left(A^2\right) - tr(P) - \frac{2[tr(A)]^2}{N-K} \right\} \qquad \text{(A3.7)}$$

where $P=(Z'Z)^{-1}Z'(W+W')^2Z$. For K=1 (constant term only), (A3.7) reduces to (A3.3).

Appendix 3.3

Table A3.1 Means of the variables used in the analysis

	1986/87	1987/88	1988/89	1989/90
expenditure per head (£)	52.17	56.49	58.92	63.12
disposable income per head (£)	5,677	6,173	6,831	7,586
grant per head (segment 1, £)	27.47	26.65	26.95	28.93
grant per head (segment 2, £)	35.92	34.49	35.16	38.48
price (segment 1)	.802	.693	.720	.744
price (segment 2)	1.089	.942	.978	1.010
population	98,687	99,348	99,965	100,612
urbanisation rate	57.44			
age>65	21.71			
ethnic minority	2.39			
long term unemployment	23.26			

Table A3.2 Data sources

data	source
expenditure	Chartered Institute of Public Finance and Accountancy (1986-1990) Finance and general statistics
income	Economic Trends (CSO)
Block Grant parameters	Association of County Councils (1986-1990) Rate Support Grant - England (London: ACC)
population	Chartered Institute of Public Finance and Accountancy (1986-1990) Local government comparative statistics
urbanisation rate	Census of Population (1991)
age>65	Census of Population (1991)
ethnic minority	Census of Population (1991)
long term unemployment	Census of Population (1991)
political control	University of Plymouth: Local Government Chronicle Elections Centre

Appendix 3.4

Table A3.3 Moran's I(x) statistic - 296 shire districts

	1986/87	1987/88	1988/89	1989/90
I(x)	.177	.202	.169	.189
z	4.85	5.53	4.63	5.18

Notes:
I) the spatial weights matrix for the I(x) statistic is row-standardised, and is based on the criterion that two districts are neighbours if they have a common border;
II) the I(x) statistic is computed under the null hypothesis of normality;
III) the moments of the I(x) statistic are reported in appendix 3.2.

Table A3.4 Tests for spatial dependence - 296 shire districts

	1986/87	1987/88	1988/89	1989/90
I(e)	2.76**	2.67**	2.55**	3.21**
LM(*err*)	4.45*	4.65*	5.52*	7.90**
LM(*lag*)	.64	4.83*	4.30*	7.99**

Notes:
I) the spatial weights matrix is row-standardised, and is based on the criterion that two districts are neighbours if they have a common border;
II) the I(e) statistic is computed under the null hypothesis that the errors are normally distributed;
III) the moments of the I(e) statistic are reported in appendix 3.2;
IV) I(e) is distributed as a standard normal; LM(*err*) and LM(*lag*) are χ^2 with one degree of freedom;
V) *=significant at the 5% level; **=significant at the 1% level.

Table A3.5 Shire districts' expenditures: 1986/87

	Non-spatial model OLS estimates	Model (3.3) ML estimates	Model (3.4) ML estimates	Model (3.5) ML estimates
λ			.235 (2.56)	.378 (2.97)
ρ		.046 (.85)		-.136 (-1.40)
income	.0041 (2.68)	.0042 (2.77)	.0039 (2.33)	.0030 (1.49)
price	-37.29 (-9.54)	-37.03 (-9.62)	-40.43 (-9.57)	-41.80 (-9.89)
grant	.623 (9.50)	.617 (9.64)	.609 (9.33)	.599 (8.99)
LabourD	13.21 (6.59)	13.07 (6.62)	13.21 (6.67)	13.33 (6.78)
population (,000)	.023 (1.38)	.025 (1.53)	.032 (1.90)	.031 (1.90)
MetroD	3.06 (1.85)	2.78 (1.68)	2.45 (1.42)	2.68 (1.50)
urbanisation rate (%)	.122 (4.94)	.120 (4.94)	.116 (4.75)	.113 (4.69)
age > 65 (%)	.559 (3.23)	.546 (3.21)	.567 (3.14)	.593 (3.17)
ethnic minority (%)	.473 (1.17)	.528 (1.31)	.557 (1.39)	.460 (1.11)
long term unemployed (%)	.142 (.70)	.148 (.75)	.169 (.81)	.175 (.85)
σ	10.33	10.13	9.97	9.81
R^2	.71			
observations	296	296	296	296

Notes:

I) Model (3.3) is the spatial lag dependence model; Model (3.4) is the spatial error dependence model; Model (3.5) is the comprehensive model with spatial dependence in both the dependent variable and the errors;

II) dependent variable = local public expenditure per head;

III) t values in parentheses;

IV) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;

V) MetroD=1 if a shire district is a neighbour of a metropolitan authority, and 0 otherwise.

Table A3.6 Shire districts' expenditures: 1987/88

	Non-spatial model OLS estimates	Model (3.3) ML estimates	Model (3.4) ML estimates	Model (3.5) ML estimates
λ			.217 (2.44)	.164 (1.12)
ρ		.141 (2.22)		.046 (.43)
income	.0008 (.43)	.0012 (.70)	.0001 (.03)	.0004 (.22)
price	-34.72 (-7.04)	-33.79 (-7.02)	-35.16 (-6.99)	-34.90 (-6.73)
grant	.530 (6.63)	.523 (6.68)	.517 (6.58)	.520 (6.66)
LabourD	13.70 (5.62)	13.39 (5.64)	13.83 (2.36)	13.74 (5.74)
population (,000)	.023 (1.02)	.026 (1.20)	.028 (1.29)	.028 (1.25)
MetroD	7.47 (3.48)	6.42 (3.00)	6.88 (2.22)	6.74 (3.02)
urbanisation rate (%)	.159 (4.93)	.150 (4.74)	.148 (4.65)	.149 (4.62)
age > 65 (%)	.615 (2.77)	.586 (2.70)	.648 (2.79)	.632 (2.77)
ethnic minority (%)	.239 (.47)	.444 (.89)	.454 (.84)	.464 (.86)
long term unemployed (%)	.174 (.663)	.176 (.67)	.160 (.64)	.165 (.68)
σ	13.46	13.07	13.01	13.03
R^2	.61			
observations	296	296	296	296

Notes:
I) Model (3.3) is the spatial lag dependence model; Model (3.4) is the spatial error dependence model; Model (3.5) is the comprehensive model with spatial dependence in both the dependent variable and the errors;
II) dependent variable = local public expenditure per head;
III) t values in parentheses;
IV) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;
V) MetroD=1 if a shire district is a neighbour of a metropolitan authority, and 0 otherwise.

Table A3.7 Shire districts' expenditures: 1988/89

	Non-spatial model OLS estimates	Model (3.3) ML estimates	Model (3.4) ML estimates	Model (3.5) ML estimates
λ			.235 (2.71)	.257 (1.96)
ρ		.131 (2.11)		-.021 (-.22)
income	.0021 (1.53)	.0023 (1.73)	.0017 (1.11)	.0015 (.92)
price	-40.39 (-8.25)	-39.96 (-8.39)	-41.26 (-8.45)	-41.29 (-8.48)
grant	.496 (6.51)	.484 (6.54)	.487 (6.55)	.487 (6.50)
LabourD	14.89 (5.83)	14.75 (5.94)	15.23 (6.14)	15.26 (6.15)
population (,000)	.024 (1.05)	.028 (1.26)	.031 (1.37)	.031 (1.37)
MetroD	4.65 (2.09)	3.73 (1.70)	4.22 (1.84)	4.30 (1.79)
urbanisation rate (%)	.148 (4.43)	.141 (4.32)	.134 (4.11)	.134 (4.08)
age > 65 (%)	.571 (2.46)	.561 (2.49)	.598 (2.48)	.602 (2.47)
ethnic minority (%)	.286 (.54)	.465 (.92)	.491 (.94)	.487 (.89)
long term unemployed (%)	.322 (1.17)	.328 (1.24)	.303 (1.09)	.299 (1.04)
σ	13.96	13.58	13.46	13.44
R^2	.61			
observations	296	296	296	296

Notes:

I) Model (3.3) is the spatial lag dependence model; Model (3.4) is the spatial error dependence model; Model (3.5) is the comprehensive model with spatial dependence in both the dependent variable and the errors;

II) dependent variable = local public expenditure per head;

III) t values in parentheses;

IV) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;

V) MetroD=1 if a shire district is a neighbour of a metropolitan authority, and 0 otherwise.

Table A3.8 Shire districts' expenditures: 1989/90

	Non-spatial model OLS estimates	Model (3.3) ML estimates	Model (3.4) ML estimates	Model (3.5) ML estimates
λ			.278 (3.27)	.269 (2.67)
ρ		.173 (2.81)		.008 (.20)
income	.0030 (2.00)	.0032 (2.24)	.0020 (1.18)	.0021 (1.16)
price	-36.58 (-6.13)	.34.82 (-6.01)	-35.90 (-6.00)	-35.91 (-6.00)
grant	.465 (6.34)	.456 (6.38)	.435 (6.14)	.437 (6.15)
LabourD	20.58 (7.05)	20.54 (7.29)	21.27 (2.82)	21.26 (7.53)
population (,000)	.039 (1.50)	.046 (1.80)	.053 (.025)	.053 (2.05)
MetroD	4.36 (1.73)	3.00 (1.21)	3.57 (1.36)	3.55 (1.35)
urbanisation rate (%)	.145 (3.95)	.132 (3.65)	.128 (3.55)	.128 (3.56)
age > 65 (%)	.898 (3.48)	.844 (3.37)	.867 (3.17)	.867 (3.17)
ethnic minority (%)	.335 (.55)	.631 (1.02)	.749 (1.23)	.747 (1.17)
long term unemployed (%)	.542 (1.78)	.533 (1.81)	.500 (1.61)	.502 (1.63)
σ	15.82	15.27	15.14	15.14
R^2	.61			
observations	296	296	296	296

Notes:
I) Model (3.3) is the spatial lag dependence model; Model (3.4) is the spatial error dependence model; Model (3.5) is the comprehensive model with spatial dependence in both the dependent variable and the errors;
II) dependent variable = local public expenditure per head;
III) t values in parentheses;
IV) LabourD=1 if the local authority is Labour controlled, and 0 otherwise;
V) MetroD=1 if a shire district is a neighbour of a metropolitan authority, and 0 otherwise.

Chapter Four

Local taxes, spatial effects and election results

4.1 Introduction

The relationship between voters and incumbent governments in a representative democracy has often been studied within a principal-agent framework.[1] Officials can be viewed as agents and the voters as their principals. Once elected, the former are asked to make decisions in the interest of the latter, about, among other things, taxation and public spending. In this context, political agency problems are bound to arise. In particular, incumbents typically know more about their own level of administrative ability or competency, and about the cost of providing public services, than do voters.

It has been argued that competition - usually in the sense of competition between incumbent and opposition parties - as well as reputation, monitoring, and optimum contract design, should be able to reduce the potential for opportunism by politicians. Consequently, the principal-agent problem in the political sector may be no more severe than in the private sector.[2] However, the important role played by private information in public services provision is well documented in the literature.[3] Recent analyses have focused on optimum selection processes in the presence of imperfectly observable government actions.[4]

What we want to examine here is to what extent political agency problems can be solved in a country that is subdivided into a number of local jurisdictions, and where

---

[1] Recent works following this approach are Besley and Case (1995a, 1995b), Kalt and Zupan (1990), Peltzman (1990, 1992), and Rogoff (1990). Fratianni *et al.* (1993) describe the relationship between the general public and the central bankers as a principal-agent problem.

[2] As argued by Wittman (1989).

[3] Dating back at least to Downs (1957).

[4] Azariadis and Lahiri (1997) is a recent example.

decisions are made at the local level. We use a principal-agent model, in which public good provision needs to be managed by a local government.

This work builds on the results of some recent research on "yardstick competition" - analysed, among others, by Shleifer (1985), and first transposed to the analysis of horizontal competition between subcentral authorities by Salmon (1987). Yardstick competition has been originally proposed as a way of improving on "cost of service" regulation of franchised monopolies. Cost of service regulation does not address the problem of cost reduction by the regulated firm, since the firm has no profit incentive to minimise costs, and the regulator does not know the appropriate cost level. Since, under such circumstances, the regulator cannot know whether the firm is run efficiently, what is needed is a benchmark, against which to evaluate the firm's performance. It has been argued that, for a given firm, the regulator should observe the costs of similar firms to infer a firm's attainable cost level, and set the price accordingly. This is what has been called a yardstick competition regulatory scheme.[5]

The yardstick competition literature heavily relies on concepts developed by the theory of incentives. Within firms, it is often the case that the effort of workers (or managers) is not directly and costlessly observable by superiors (or by the owners of the firm). Comparative performance information can in this case improve incentives and efficiency in principal-agent relationships. In the presence of imperfect information, rewards based on relative output can be superior to payments based on individualistic output. Contests or tournaments,[6] and rewards based on performance standards[7] are ways of obtaining information about the environment, by looking at the performances of more than one individual.

In an analogous way, voters in a multijurisdictional world can make comparisons between local jurisdictions to overcome political agency problems. Salmon (1987) suggested that one of the main advantages of decentralisation is that of providing local administrators with incentives for behaving efficiently and reducing managerial slack, or X-inefficiency. In the presence of asymmetric information between voters

---

[5] Bivand and Szymanski (1997) show how local yardstick regulation of natural monopolies can generate externalities and spatial correlation.

[6] Lazear and Rosen (1981) and Nalebuff and Stiglitz (1984).

[7] Holmstrom (1982), Mookerjee (1984), and Meyer and Vickers (1994).

and politicians about, say, the cost of providing public services, voters can use information on the cost of public services in similarly situated jurisdictions, when deciding whether or not to re-elect the incumbent government (as in Besley and Case, 1995b).

In our model, we assume that *n* local jurisdictions are subject to spatially correlated cost shocks. Officials can extract a "surplus" from the electorate - excess of taxes over public service costs - because of the voters' ignorance about the cost shock. On the other hand, officials are heterogeneous. The surplus they extract depends on a continuously distributed stochastic variable that we call "quality." Due to spatial correlation in the cost shock, it is optimal for voters to appraise the government's quality by looking at their jurisdiction's relative performance. As a result, both own and neighbours' taxes will have an influence on the votes cast for the incumbent party. Furthermore, the tax-setting decision in a local jurisdiction will be affected by the tax-setting decisions in neighbouring jurisdictions.

Some recent empirical studies have looked at the influence of local budgets - or state budgets in federal countries - on incumbents' popularity. Peltzman (1990, 1992) analyses voting behaviour in Presidential, Senatorial, and gubernatorial elections in the US from 1950 to 1988, and concludes that voters penalise both federal and state taxation and spending growth. Besley and Case (1995b) look at the effects of different kinds of state taxes on governors' chances of re-election. They find that a governor's electoral defeat is positively correlated with own state tax increases and negatively correlated with tax increases in neighbouring states.

When analysing the determinants of the vote at the local level, however, it should be taken into account that both local and national factors will play a role. Most of the literature on local election results - especially as regards the UK - focuses on the latter and tends to reduce local elections to "national opinion polls," or little more than "referenda" on the current standings of the national parties (Miller, 1988). However, recent studies on cross-sectional British data suggest that local taxes might have an impact on the share of the vote of the incumbent at the local authority level. Gibson (1988), Gibson and Stewart (1992) and Rallings and Thrasher (1997) present evidence on several local elections, in support of the view that "local" voting does exist.

We test for the presence of spatial interactions in local tax-setting and look at the effects of local taxes on local election results, by using a panel data set of the English local governments. First, the use of panel data should represent a substantial improvement upon previous analyses based on cross-sectional data. Second, the spatial character of the data is exploited in order to find out whether the chances of re-election of an incumbent are affected, besides his own taxes, by his neighbours' taxes. Finally, we control for the influence of national politics, that is we check whether local election results are affected by the popularity of the central government party.

The plan of the work is as follows. In section **4.2** we outline a simple political agency model in the context of a multijurisdictional economy. In section **4.3** we look at the implications of yardstick evaluation of performance on local governments' taxation decisions, while in section **4.4** we analyse the influence of own and neighbours' taxes on the votes cast for the incumbent party at the local authority level. Section **4.5** concludes.

4.2 Political agency

4.2.1 Let us consider a country which is subdivided into a fixed number n of local jurisdictions. In each of them, citizens elect representatives that manage the provision of a local public good with no interjurisdictional spillovers.[8] We can think of the government in jurisdiction i (i=1,...,n) as buying the good on a national market at

[8] The existence of spillovers has frequently been assumed in the theoretical local public finance literature, and it has often been offered as a justification for the presence of spatial interactions among local jurisdictions (see Case *et al.*, 1993). Even though such spillovers cannot be ruled out *a priori*, the argument here is that spatial interactions can emerge due to informational externalities, even in the presence of purely local public spending.

price p_t and providing it to the citizens in the local jurisdiction. The public good is supplied at a final per unit cost of c_{it} in jurisdiction i at time t:

$$c_{it} = p_t + \theta_{it} - q_i \tag{4.1}$$

While p_t is known by both taxpayers and representatives, the other two terms are stochastic.[9] The first one, θ_{it}, is a cost shock specific of jurisdiction i in year t, over which the local government has no control. The second one, q_i, is the level of "quality" or "competency" of the incumbent in jurisdiction i. In this context, quality is not a choice variable for an official, but an individual characteristic.[10] q_i is not observed by voters. They only observe c_{it} - the final per unit cost - and p_t, but cannot separately identify θ_{it} and q_i by looking at their jurisdiction alone.

Assume that $E(\theta_{it})=0$, $Var(\theta_{it})=\sigma_\theta^2$, $E(q_i)=0$, $Var(q_i)=\sigma_q^2$, and that θ_{it} and q_i are uncorrelated. Finally, assume that one unit of a homogeneous public good is provided in each community, that is each local government faces an inelastic unit demand for the public good.[11] Voters in jurisdiction i observe the costs of public services in jurisdictions i and j, and the correlation between the cost shocks in the two jurisdictions is:[12]

$$\rho = corr(\theta_{it}, \theta_{jt}) \tag{4.2}$$

At the beginning of period t, voters in jurisdiction i elect a government. Thereafter, nature selects a cost shock and the government selects the cost "surplus" ($-q_i$)

---

[9] In the presence of a grant distribution system - that we ignore here for simplicity - p_t would reflect the central government contribution to local expenditure. A grant reduction represents a common effect on all local jurisdictions in a given year, and is reflected by a higher tax burden on local residents (higher p_t).

[10] We may think of quality as managerial ability, and of high cost of public services as inefficiency due to low competency (as in Rogoff, 1990). In section **4.3**, we extend the model by allowing incumbents to care about re-election, and to maximise a multiperiod objective function that depends on their quality.

[11] The consequences of relaxing this assumption are discussed in the empirical section **4.4**.

[12] Generalisation to any number of neighbouring jurisdictions is straightforward and does not affect the results.

corresponding to his quality. At the end of period t, voters observe c_{it} and c_{jt} and decide on whether to re-elect the incumbent. *Ex ante*, i.e., before observing the realisation of the costs (beginning of period t), the expected value of the incumbent's quality is zero.[13] After observing the actual levels of c_{it} and c_{jt}, voters update their beliefs about the quality of the incumbent. The expectation of the incumbent's quality, conditional upon c_{it} and c_{jt}, is:

$$E(q_i|c_{it},c_{jt}) = \beta(c_{it} - E(c_{it})) + \delta(c_{jt} - E(c_{jt})) = \beta(c_{it} - p_t) + \delta(c_{jt} - p_t) \quad (4.3)$$

$$\beta = \frac{\operatorname{cov}(q_i,c_{it})\operatorname{var}(c_{jt}) - \operatorname{cov}(q_i,c_{jt})\operatorname{cov}(c_{it},c_{jt})}{\operatorname{var}(c_{it})\operatorname{var}(c_{jt}) - \left[\operatorname{cov}(c_{it},c_{jt})\right]^2} = -\frac{\tau}{1-\rho^2(1-\tau)^2} \quad (4.4)$$

$$\delta = \frac{\operatorname{cov}(q_i,c_{jt})\operatorname{var}(c_{it}) - \operatorname{cov}(q_i,c_{it})\operatorname{cov}(c_{it},c_{jt})}{\operatorname{var}(c_{it})\operatorname{var}(c_{jt}) - \left[\operatorname{cov}(c_{it},c_{jt})\right]^2} = \rho\frac{\tau(1-\tau)}{1-\rho^2(1-\tau)^2} \quad (4.5)$$

where $\tau \equiv \sigma_q^2/(\sigma_\theta^2+\sigma_q^2)$ is a measure of the relative variances of θ and q. If $\rho=0$ (no spatial correlation in the cost shock), it follows that $\beta=-\tau<0$, and $\delta=0$.

As a result: $E(q_i|c_{it},c_{jt})=\tau(p_t-c_{it})$. In this case, c_{it} and c_{jt} are not correlated. The cost level in j is not informative for voters in i as regards the quality of the incumbent. Only information on c_{it} is used to infer the level of the cost shock, and hence the incumbent's quality.

On the other hand, if $\rho>0$ the weights on own and neighbours' costs depend on the relative variances of θ and q. A high cost of public services in jurisdiction i always reduces the conditional expectation of the incumbent's quality ($\beta<0$). On the other hand, since in this case the cost shocks in the two jurisdictions are positively correlated, a high cost of public services in jurisdiction j increases it ($\delta>0$). If costs of

[13] Strictly speaking, this is true only of a newly elected government, and not of a government that was in office in the previous term. However, to keep things simple, we do not introduce any dynamics at this stage, and assume that voters forget about the previous election outcome.

public services are high everywhere, voters may be convinced that this is due to a high cost shock, and not to inefficient use of resources on the part of the incumbent. If local jurisdictions are subject to correlated shocks ($\rho \neq 0$), using information on own and neighbours' costs improves on the way voters in i appraise the quality of their incumbent, by decreasing the variance of q_i.[14] The higher the correlation between the cost shocks in jurisdictions i and j, the lower the variance of q_i conditional on c_{it} and c_{jt}.

4.2.2 We assume that voters minimise the expected cost of public services.[15] A representative voter in jurisdiction i will desire to re-elect the incumbent if the expected cost of public services under the incumbent in period t+1 is less than that with another candidate.[16] The expected cost of public services if the incumbent (I) is re-elected is:

$$E_t[c^I_{i(t+1)} | c_{it}, c_{jt}] = p_t - \beta(c_{it} - p_t) - \delta(c_{jt} - p_t) \qquad (4.6)$$

The expected cost of public services if the opposition candidate is elected is:

$$E_t[c_{i(t+1)}] = p_t \qquad (4.7)$$

---

[14] By defining $\sigma^2=\sigma_q^2+\sigma_\theta^2$, it is easy to show that: $Var(q_i|c_{it},c_{jt})=\tau\sigma^2[1-\tau(1-\rho^2(1-\tau)^2)^{-1}]<Var(q_i|c_{it})=\tau(1-\tau)\sigma^2<Var(q_i)=\tau\sigma^2$.

[15] It could be argued that citizens can move among local jurisdictions in response to tax differentials. According to the Tiebout hypothesis, "the consumer-voter moves to that community whose local government best satisfies his set of preferences" (Tiebout, 1956: 418). However, Epple and Zelenitz (1981) have argued that if the government in each jurisdiction maximises governmental profit (excess of tax revenue over expenditure) and if there is no political mechanism by which residents can affect local tax rate and public spending, the right of citizens to choose among many communities cannot completely eliminate governmental monopoly power. On the other hand, in this paper we focus on the government's efficiency in public good provision, and argue that it is optimal for voters to unseat an inefficient government.

[16] Voters do not know anything about the "opposition" candidate, apart from the parameters of the distribution he comes from, i.e., the distribution of quality q.

If all voters in a jurisdiction are identical, and in the absence of any random influence on the election outcome, the incumbent's probability of re-election is:[17]

$$P_{it}^{E}=\begin{cases}1 & \text{if} \quad \beta(c_{it}-p_t)+\delta(c_{jt}-p_t)\geq 0\\ 0 & \quad\quad \beta(c_{it}-p_t)+\delta(c_{jt}-p_t)<0\end{cases} \tag{4.8}$$

(4.8) can be rewritten as:

$$P_{it}^{E}=\begin{cases}1 & \text{if} \quad c_{it}\leq\gamma c_{jt}+(1-\gamma)p_t\\ 0 & \quad\quad c_{it}>\gamma c_{jt}+(1-\gamma)p_t\end{cases} \tag{4.9}$$

where: $\gamma\equiv-\delta/\beta=\rho(1-\tau)$. This voting rule implies that voters compare their own cost of public services to a threshold:

$$\tilde{c}_{it}=\gamma c_{jt}+(1-\gamma)p_t \tag{4.10}$$

The threshold is a weighted average of jurisdiction *j*'s cost of public services and of the *a priori* expected level of the cost. If $\rho=0$, then clearly $\gamma=0$, and the first term of the threshold disappears because the other jurisdiction's cost is not informative for voters in jurisdiction *i*. Voters re-elect the incumbent if the cost in *i* is no larger than p_t. If $\rho\neq0$, the weights depend on the correlation among the cost shocks, and on the relative variances of q and θ. In general, the higher the correlation (in absolute value) between c_{it} and c_{jt}, the larger the weight placed on the latter - relative to the weight placed on the a priori mean of c_{it} - in deciding whether to vote for or against the incumbent. If $\rho>0$ (which clearly implies $\gamma>0$), a higher than average cost of public services in jurisdiction *j* makes voters in jurisdiction *i* raise the threshold above p_t.

4.2.3 In section **4.2.2**, we have assumed that the voting rule is deterministic. According to (4.9), the incumbent is re-elected with probability one if the cost of

---

[17] We introduce voter heterogeneity and a stochastic voting rule in section **4.2.3** below.

public services is lower than the threshold, while he is not re-elected with probability one if the cost is higher than the threshold. The latter - equation (4.10) - is a weighted average of the expected cost, and of the cost in neighbouring jurisdictions. However, the voting decision is bound to depend on other factors than the cost of public services - such as general economic conditions and local political party affiliation.[18] Furthermore, voters in a jurisdiction might care differentially about the cost of public services, be heterogeneously informed, and have different attitudes towards the incumbent. We allow for voter heterogeneity by introducing a zero mean additive random term, η_{it}, which is independently and identically distributed across the n_i voters of jurisdiction i. The probability of re-election can then be rewritten as:

$$P_{it}^{E} = \Pr\{E_t[c_{i(t+1)}^{I} | c_{it}, c_{jt}] - E_t[c_{i(t+1)}] + \eta_{it} < 0\} = \Pr\{\eta_{it} < \beta(c_{it} - p_t) + \delta(c_{jt} - p_t)\} \tag{4.11}$$

If the random term η_{it} comes from a logistic distribution, the probability of re-election is:[19]

$$P_{it}^{E} = \frac{\exp[\beta(c_{it} - p_t) + \delta(c_{jt} - p_t)]}{1 + \exp[\beta(c_{it} - p_t) + \delta(c_{jt} - p_t)]} \tag{4.12}$$

where P^E equals the expected share of the vote of the incumbent (S). By rearranging terms and taking logarithms, (4.12) can be rewritten as:

$$s_{it} = h_t + \beta c_{it} + \delta c_{jt} \tag{4.13}$$

[18] If this is the case, the number of factors affecting the voting decision grows and will probably tend to reduce the role of efficiency in public goods provision. In our empirical application, we will take those factors into account.

[19] Since the normal distribution and the logistic distribution are very similar, we choose the latter because of its mathematical convenience.

where $h_t \equiv -(\beta+\delta)p_t$, and $s_{it} \equiv log[S_{it}/(1-S_{it})]$. The impact of neighbours' taxes on the incumbent's expected share of the vote is positive if δ is positive - that is in the presence of positive spatial correlation in the cost shock ($\rho>0$) - while the impact of the own tax is always negative ($\beta<0$). In the presence of a positively spatially correlated cost shock, we should find that the incumbent's share of the vote is negatively correlated with his own tax, and positively correlated with the taxes in neighbouring jurisdictions.

4.3 Opportunistic behaviour and yardstick competition

4.3.1 According to the political agency model in section **4.2**, a high local tax is a signal of low quality of the incumbent, and tends to reduce his chances of re-election.[20] If politicians realise that tax increases are harmful to their popularity, and if they care about re-election, they will try and keep taxes low in order to be re-elected.[21] They will weight the costs of reducing local taxes today against the benefits of increasing their re-election chances. Incumbents may therefore maximise an objective function that incorporates a trade-off between the "utility" officials get from the excess of tax revenue over service costs in the current period (governmental profit, denoted by μ_{it}), and the probability of being in office in future periods (P^E):

$$\Psi_t^i = \max_{\mu_{it}} \left\{ v(\mu_{it}; q_i) + P^E(c_{it}, c_{jt}) \Psi_{t+1}^i \right\} \qquad (4.14)$$

[20] If local governments are unable to borrow and in the absence of grants from central government - or if p_t is net of central government grants - c_{it} must be equal to locally raised tax revenues.

[21] It could be argued that though voters observe the local tax burden, they do not know the level of grant the local government is given. Consequently, a local tax increase due to a reduction in the grant might wrongly be attributed to local inefficient behaviour. However, if grant is not distributed selectively, all jurisdictions in a given year will be affected in the same way by a central government restrictive policy, making comparative performance evaluation superior to individual performance evaluation.

The governmental profit μ_{it} is the excess of tax revenue (c_{it}) over provision cost (p_t+θ_{it}). The function v represents the utility incumbents get from μ_{it}. It depends on the incumbent's quality, in that it has a maximum at μ_{it}=$-q_i$ (the profit that incumbents would set if there were no re-election incentives). For the reasons and under the circumstances outlined above, the probability of re-election P^E depends negatively on own tax (c_{it}), and positively on neighbours' tax (c_{jt}). In (4.14), we have normalised the payoff from not being re-elected to zero.

As compared with an incumbent that does not run for re-election, an incumbent who wishes to be re-elected tends - other things equal - to set a lower tax. The former simply maximises $v(\mu_{it};q_i)$ and selects a cost surplus equal to $-q_i$, while the latter faces the trade-off between cost surplus and re-election probability. The first order condition for a maximum shows that the value of μ_{it} that maximises (4.14) is smaller than $-q_i$:

$$\frac{\partial v(\mu_{it};q_i)}{\partial \mu_{it}}+\frac{\partial P^E(c_{it},c_{jt})}{\partial \mu_{it}}\Psi^i_{t+1}=0 \Rightarrow \mu_{it}*<-q_i \qquad (4.15)$$

We would therefore expect to find that the taxes observed in local jurisdictions where the incumbents run for re-election are lower than the taxes observed in jurisdictions where the incumbents do not (or cannot) run for re-election. Furthermore, the optimal surplus also depends on neighbours' surplus through the probability of re-election P^E. In the presence of positive spatial correlation in the cost shock, high taxes in nearby jurisdictions make voters in jurisdiction i raise the threshold, that is the level of the tax above which they vote against the incumbent. If taxes are high in nearby jurisdictions, even a high tax in jurisdiction i may be politically acceptable. An incumbent can raise the tax if its neighbours are doing the same, because this will be interpreted by electors as the effect of a correlated cost shock. On the other hand, if its neighbours set low taxes, an incumbent will have to do the same in order to be re-elected. Failure to do so will be interpreted by the electorate as a signal of low quality, and will result in the incumbent's electoral defeat. The marginal effect of neighbours' surplus on jurisdiction i's surplus can be obtained by totally differentiating (4.15):

$$\left[\frac{\partial^2 v}{\partial \mu_{it}^2}+\frac{\partial^2 P^E}{\partial \mu_{it}^2}\right]d\mu_{it}+\left[\frac{\partial^2 P^E}{\partial \mu_{jt}\partial \mu_{it}}\right]d\mu_{jt}=0 \Rightarrow \frac{d\mu_{it}}{d\mu_{jt}}>0 \qquad (4.16)$$

The first term in brackets is negative due to concavity of the objective function (second order condition characterising the optimal choice of μ_{it}). The second term in brackets is positive, as long as the incumbent government in jurisdiction i aims at being re-elected.[22] Consequently, we would expect to find that the tax in jurisdiction i tends to be positively correlated with the taxes in neighbouring jurisdictions.

The political agency yardstick competition model has therefore two main implications. First, the incumbent's share of the vote should be negatively correlated with his own tax, and positively correlated with the taxes in nearby jurisdictions. Second, we should expect to find positive spatial correlation in local taxes.

Since spatial correlation in taxes could also arise because of fiscal spillovers between jurisdictions, the estimation of a popularity equation augmented with spatial effects is required in order to discriminate between the yardstick competition model and a fiscal spillover model.

4.3.2 We want to test whether the re-election incentive has an impact on local government taxation decisions, and whether local authorities look at their neighbours when setting their own taxes. As regards the re-election incentive, ideally we would like to identify the difference in local tax-setting behaviour between an incumbent that wishes to be re-elected, and an incumbent who does not (or cannot) run for re-election, that is whether "electoral accountability" affects choices about local taxes.[23]

---

[22] However, a low quality incumbent might find it optimal to get the maximum governmental profit in the current period, and be out of office in the next period. The advantage of yardstick competition would in this case consist in finding out the lower quality incumbents - because it is too costly for them to reduce the tax to their neighbours' level. However, we expect the re-election incentive to dominate, and to affect incumbents' behaviour in the way described by equation (4.16).

[23] As in Besley and Case (1995a).

Unfortunately, no such information is present in the available data,[24] and no term limitations exist in the electoral system of the English local governments. However, we can identify the difference in local tax-setting behaviour in election vs. non-election years. This is feasible, because elections do not occur every year, but rather at regular intervals, and the data set includes local authorities that vote at different points in time.[25] Actually, the English local governments have different electoral systems.[26] London boroughs have all-out elections where all council seats come up for re-election every four years. Metropolitan districts have annual elections for 1/3 of the council, with a fourth non-electoral year (which used to be reserved for metropolitan county elections before 1986). Finally, non-metropolitan districts are free to select from all-out elections every four years and annual elections by thirds.[27]

Chart 4.1 reports the path of the average local property tax (the rate) for the non-metropolitan districts having all out elections every four years (1979, 1983, and 1987 in our sample) and for those voting by thirds.[28]

Even though we are not controlling here for local characteristics, changes in the grant distribution system or for any other shock to local tax-setting, the behaviour of the two classes of authorities appears to be slightly different.

As regards the districts voting by thirds, the rate tends to increase slightly more markedly in the non-election years: 1981, 1985, and 1989. As regards the districts having all-out elections, the rate appears not to increase much - or even to decrease - as the elections approach. In order to check more rigorously whether the electoral

---

[24] We do not have any information on candidates and their willingness to run for re-election, but only on political parties.

[25] If electors attach greater importance to events that occur close to the election years, incumbents will endeavour to hold down tax increases at or close to elections, in order to boost their probability of re-election. This is justified as long as electors have a "decaying memory" of past events (as in Nordhaus, 1975).

[26] The English system of local government underwent several changes in the last two decades. After 1986, metropolitan areas (Greater London, Greater Manchester, Merseyside, South Yorkshire, Tyne and Wear, West Midlands, and West Yorkshire) have a single tier system (32 London boroughs and 36 metropolitan councils), while non-metropolitan areas have a two-tier system of 39 shire counties and 296 shire districts.

[27] Some authorities actually switched from one system to the other during the 1970s and the 1980s.

[28] The rate is expressed as a percentage of the imputed rental value of the property (rateable value).

system and the timing of elections affect the budgetary behaviour of local governments, we turn to the estimation of a tax-setting model.

Due to the radical changes brought about by the Local Government Act 1985 in metropolitan areas (abolition of the top tier and transfer of their functions to the second tier or to joint boards), we use data on the English non-metropolitan districts' taxes (1979 to 1990) in order to find statistical evidence of the effects of re-election incentives on local taxation decisions, and of the presence of spatial interactions between neighbouring districts. The means and data sources of the variables used in the analysis are in appendix 4.1. We estimate the following tax-setting equation:

$$c_{it} = \lambda \sum_{j=1}^{N} w_{ij} c_{jt} + z_{1it}'\phi_1 + z_{2i}'\phi_2 + f_i + h_t + \varepsilon_{it} \qquad (4.17)$$

$$\varepsilon_{it} = \tau \sum_{j=1}^{N} w_{ij} \varepsilon_{jt} + u_{it} \qquad (4.18)$$

In equation (4.17), jurisdiction i's tax rate at time t (c_{it}) depends on two vectors of characteristics of district i (z_{1it} and z_{2i}). z_{1it} comprises a set of district-specific time-varying financial, political and electoral variables, while z_{2i} contains a set of socio-economic and demographic characteristics of district i which are constant over time - at least in the available data. Equation (4.17) also has a spatially weighted average of neighbouring districts' taxes as an explanatory variable. λ is the coefficient of the first-order spatial autoregressive process in the dependent variable, and w_{ij} is the (i,j) element of the neighbourhood matrix W.[29]

We allow for spatially autocorrelated errors in equation (4.18). The spatial dependence takes the form of a first-order spatial autoregressive process in the error

[29] Unlike the model in section **4.2**, it is usually the case that each of n local jurisdictions has more than one neighbour, and that the n jurisdictions face spatially correlated cost shocks. One way of using this information is by means of a weights matrix, denoted by $W(nxn)=\{w_{ij}\}$, where n is the number of local jurisdictions. The spatial weights are such that $w_{ij}=0$ if district j is not a neighbour of district i, and $\sum w_{ij}=1$. All neighbours of a district are given equal weight (1/number of neighbours) and all districts are equally influenced by their neighbours taken together.

term, where τ is the spatial autoregressive coefficient and u is an uncorrelated error term.

Finally, the specification (4.17) contains a district-specific effect f_i - an unobserved characteristic of the local jurisdiction that influences its fiscal decisions and does not change over time - and a time effect h_t - intended to control for variables that might have a common effect on the local districts in a given year.

Equation (4.17) cannot be estimated by OLS, whether or not the error term follows a spatial autoregressive process. This is due to the fact that neighbours' and own taxes are determined simultaneously, and any influence is bound to be reciprocal. Furthermore, if the error term actually follows a spatial process - as in (4.18) - then neighbours' taxes are correlated with the error term. One way of dealing with those endogeneity problems is by instrumenting the neighbours' taxes. Neighbours' own covariates (z_{1jt} and z_{2j}) can be used as instruments in order to identify the spatial autoregressive parameter λ. The estimation results are presented in tables 4.1a and 4.1b.

Table 4.1a presents the results of estimating a model with no spatial components ($\lambda=\tau=0$). The first column shows the results of OLS estimation of the tax-setting equation in levels, while the results in the second column are obtained after taking first differences. This allows us to identify the coefficients on the time-varying characteristics (z_{1it}), but not the ones on the characteristics of the local districts which are (approximately) constant over time (z_{2i}).

Table 4.1b presents estimates of the model (4.17)-(4.18), where spatial dependence is allowed for. The first and third columns present OLS estimates of the equation in levels and in first differences respectively. In the second and fourth columns, neighbours' own covariates (z_{1jt} and z_{2j} in the levels equation, and Δz_{1jt} in the differenced equation) are used as instruments for the spatially lagged dependent variable. In the IV estimation, all remaining variables are assumed to be strictly exogenous, although none of the overidentifying restrictions that follow from this assumption are exploited.

The results in tables 4.1a and 4.1b suggest, in the first place, that incumbents tend to behave in an "opportunistic" way, i.e., they tend to reduce the local tax in election years. The coefficient on the election year dummy in the equation in first differences definitely points towards a significant reduction in the local tax in the election years. On the other hand, the electoral system appears to have a smaller effect on the tax-setting decisions of districts voting by thirds. Furthermore, the effect of elections on local taxes is robust to the introduction of neighbourhood effects.

As for "partisan" differences between incumbents belonging to different political parties, the political control dummy suggests that Labour controlled councils tend to set higher local taxes than councils controlled by other parties. However, when the presence of the district-specific fixed effect is controlled for in the first-differenced equation, the effect of Labour incumbency on local taxation is considerably reduced.

Equation (4.17) also includes two measures of availability of local resources: domestic tax base per head and non-domestic (industrial and commercial) tax base per head. In the presence of a grant distribution system that equalises both for differences in the local tax base and for differences in needs (the British block grant system described in chapter two), the cost (tax-price) of public services to local residents is higher, the higher the local per capita tax base. The domestic tax base per head has the expected negative sign, though the coefficient estimate is imprecise. On the other hand, the non-domestic tax base does not appear to have a negative effect on the rate. This lends some support to the thesis that the domestic sector could have had a "free-ride" on the business sector in the funding of local public services.[30] The small variability of the tax base over time, however, makes the estimates of these coefficients rather inaccurate in the first-differenced equation.

The most striking result in table 4.1b is the estimate of the coefficient on the neighbours' rate. It suggests that taxes in neighbouring districts are positively correlated. The coefficient is still statistically significant - even though less precisely estimated - when neighbours' taxes are instrumented with neighbours' covariates. Since neighbours' taxes are instrumented, the correlation that emerges in the IV estimation in table 4.1b (columns 2 and 4) cannot be attributed to correlated shocks.[31]

---

[30] This was one of the arguments that motivated the reforms of local government finance in the 1990s.

[31] At first sight, this result may appear to be at variance with the findings in chapter three, where we concluded that the most likely source of spatial autocorrelation in local public expenditure is the

If neighbours were simply subject to correlated shocks, the OLS estimation of λ in equation (4.17) would pick up the correlation between ε_{it} and ε_{jt}. The bias arises because the regressor (c_{jt}) is correlated with the unobservables (ε_{it}). However, z_{1jt} and z_{2j} (neighbours' explanatory variables) can be used as instruments for c_{jt}, since they are correlated with c_{jt}, but are uncorrelated with ε_{it}. The correlation between regressor and error term is eliminated, and the IV estimate of λ is picking up only the correlation in the component of neighbours' taxes that is attributable to neighbours' observable variables.

According to the political agency model in section **4.3**, we should expect to find positive spatial correlation in local taxes, since incumbents set their own local taxes after looking at the taxes set by their neighbours. Consequently, the results of the estimation are at least compatible with the insights that the previous stylised theoretical analysis provided. First, incumbents care about re-election, and tend to slow down tax increases in election years. Second, when deciding their own taxes, they appear to take into account the taxes set by their neighbours.

Of course, yardstick competition is not the only possible explanation of positive spatial correlation in local taxes. The results are compatible with a tax competition model, where a mobile tax base moves among local jurisdictions in response to tax differentials. In order to be able to discriminate between alternative models, we estimate a popularity equation where spatial effects are explicitly allowed for.

If spatial interactions emerge simply due to a fiscal externality, the popularity of the government in a local jurisdiction should not be affected by the taxes set in neighbouring jurisdictions. However, if voters use information on nearby jurisdictions in order to solve the asymmetric information problem, then we should expect neighbours' taxes to be a determinant in a local government popularity function.

---

presence of omitted influences - and in particular district-specific effects - which have a spatial pattern. However, the local tax rate is the most visible variable of the local government policy, or even the only variable that voters care about in their voting decision - as we have assumed here. It is therefore conceivable that the local tax rate is the policy instrument that incumbents can more easily manoeuvre for strategic purposes, and that reacts more promptly to changes in nearby jurisdictions.

4.4 Rates and local elections

By using data on the English local governments' elections, we now want to check whether own and neighbours' taxes have an effect on the incumbent's share of the vote, as the political agency model in section **4.2** suggests.[32]

English local government elections are normally held on the first Thursday in May, that is after the budget for the coming financial year has been made. For electoral purposes, the districts are divided into wards, each represented by a varying number of councillors (usually three). About two thirds of the shire district councils are elected *en bloc* every four years. In the other districts, one third of the councillors are elected in each of the three years between the county council elections.[33] The only factor in common is that all councillors sit for a four-year period.

We use data on the local councils which were controlled by one political party, and for which we have three or more continuous observations during the period 1979 to 1990.[34] We estimate separate popularity equations for the districts having by thirds elections every year, and for those having all out elections every four years.

As regards the former, we use an unbalanced panel data set, both in the sense that we have more observations on some districts than on others, and because observations correspond to different points in time. After excluding districts with no overall political control and those for which less than three observations are available, we are left with 87 local authorities, for a total of 602 observations. We consider a dynamic popularity equation of the form:

$$s_{it}^{k} = \phi s_{it-1}^{k} + \alpha_0' x_{it} + \alpha_1' x_{it-1} + f_i^k + h_t^k + \xi_{it} \qquad (4.19)$$

---

[32] Electoral data were obtained from the University of Plymouth, Local Government Chronicle Elections Centre. Data about local taxes are published by the Chartered Institute of Public Finance and Accountancy (see appendix 4.1).

[33] District authorities have the opportunity to choose their method of election. This means that in some areas neighbouring councils operate with different electoral cycles.

[34] By continuous observations we mean observations on consecutive elections.

where $s_{it}^k \equiv log[S_{it}^k/(1-S_{it}^k)]$, and S_{it}^k is the share of the vote of the incumbent party. The subscript i refers to the district, t refers to the election year, and $t-1$ is the year when the previous election was held. The dependent variable is now also indexed by the political party (k) the local incumbent belongs to. In equation (4.19), we allow for an influence from older data by including party k's share of the vote and values of the explanatory variables in the last election ($t-1$). The vector x contains own and neighbours' taxes.[35] For each local government, we compute a spatially weighted average of neighbours' taxes.[36] Finally, equation (4.19) contains district-party specific and time-party specific effects. The district-party specific effect - f_i^k - is included in order to allow for local political party affiliation, which is assumed to be roughly constant over time. It controls for the fact that in each jurisdiction a number of voters may stick with their preferred party, even if there are large increases in local taxes. The time-party specific effect - h_t^k - allows for an influence on local elections from national politics. If local taxes are only rarely a decisive factor in determining local election results, we should expect local election results to be driven by national issues and by the popularity of the Prime Minister's party. h_t^k is intended to pick up the national politics effect in year t on local councils controlled by party k.

Under the assumption of lack of serial correlation in the error term ξ, values of the share of the vote lagged two electoral periods or more are valid instruments for the lagged dependent variable in the first-differenced equation. The assumption of white noise errors is essential for the consistency of an estimator that uses lags as instruments. Consequently, we report two tests of lack of serial correlation: a test on the second-order residual serial correlation coefficient, and a Sargan test of over-

---

[35] Actually, we observe the local tax set by each local government, but unfortunately we do not have information on the levels of expenditure on various public services. Furthermore, no reliable measure of the quality of public services is available. If the quality of public services has a positive effect on the incumbent's popularity, and if it is positively correlated with the local tax (as higher taxes are needed in order to supply public services of a better quality) omitting it from the popularity equation creates a positive correlation between the local tax and the error term. We therefore tend to underestimate the impact of the local tax on the incumbent's share of the vote. However, we control for district level fixed effects, and the quality of most public services in a district is likely not to change much over time.

[36] All geographical neighbours are included, irrespective of their electoral system.

identifying restrictions.[37] As regards own and neighbours' taxes, they cannot be assumed strictly exogenous in equation (4.19). The shock to the incumbent's popularity (ξ) may be correlated with the local tax if incumbents try to counterbalance those shocks by varying it. For instance, an incumbent seeking re-election in a time of low popularity can try and gain consensus by cutting the local tax. On the other hand, if an incumbent is confident of re-election (either because a certain target of popularity is surpassed, or because he expects not to be opposed by any serious contender), he can pursue his surplus goal, raise the local tax, and expect to remain in office. Consequently, values of own and neighbours' taxes lagged two electoral periods or more are used as instruments in the equation in first differences.[38]

The estimation results are presented in table 4.2. Column 1 presents OLS estimates of a model that uses own and neighbours' taxes (and a number of district-specific socioeconomic characteristics that are - approximately - constant over time) as explanatory variables, but does not account for the influence of national politics. The effect of own taxes on the incumbents' share of the vote is negative and significant. However, neighbours' taxes do not seem to have any impact on an incumbent's popularity. On the other hand, a high unemployment rate and a high proportion of the population in priority housing need - that are intended to be proxies for welfare in the local jurisdiction - tend to be associated with a low share of the vote of the incumbent party. Column 2 presents GMM estimates of the same model. We take first differences and use lags of the share of the vote lagged two electoral periods or more as instruments for the lagged differenced dependent variable, and lags of own tax (c) and neighbours' tax (c^n) dated (t–2) and earlier as instruments for c, $c(t-1)$, c^n, and $c^n(t-1)$. The own tax shows a large negative effect of high significance, while neighbours' taxes only have a small positive effect on the incumbent's popularity, that

---

[37] Arellano and Bond (1991).

[38] Due to the by thirds electoral system, there might also be cases when the election cannot result in a change in political control. This occurs when a party has secured a vast majority of the seats in the previous by thirds elections. Consequently, we would expect the electoral threat to be weaker and the local tax to be higher under such circumstances. The tax at time t would therefore be positively correlated with the shock to the incumbent party's share of the vote in the elections held at times t–1, t–2, and t–3. As long as the error term ξ is not serially correlated, however, an IV approach that uses levels of the tax dated t–2 and earlier as instruments, yields consistent coefficient estimates in the first-differenced equation.

does not offset the negative effect of own taxes. While an increase in the own tax by ten percentage points is estimated to reduce the incumbent's share of the vote by almost three percentage points, an increase in the neighbours' tax by the same amount is expected to increase it by slightly less than one percentage point.

Columns 3 and 4 show the results of estimating the same model with national politics influences. In order to allow for the effect of political parties' popularity at the national level, we include time dummies interacted with political control dummies. Allowing for an influence from national politics reduces considerably the effect of own and neighbours' taxes on the election results. When we control for national politics influences, all tests for the validity of the instruments used (lack of serial correlation) pass comfortably. The tests on residual serial correlation confirm the presence of first-order serial correlation - which is to be expected, given that the model has been transformed to first differences - but not of second-order serial correlation - which would reject the validity of the instruments used.

Overall, these results suggest that even though local taxes may have a negative effect on local incumbents' electoral prospects, accounting for national politics dramatically reduces this effect. Furthermore, the taxes set by neighbouring governments do not seem to affect the electoral outcomes.

However, we want to check whether the national politics effects that we have estimated are compatible with the popularity of the Conservative central government in the 1980s, as expressed by the national opinion polls. If local elections are mainly driven by national issues, we would expect the Conservative controlled local governments to have good electoral performances when the popularity of the Prime Minister's party is high. We first obtain the estimated national politics effects in terms of the share of the vote from table 4.2, and we plot them together with the national opinion polls data (from table 4.3) in chart 4.2.

Chart 4.2 shows that the estimated national politics effects on Conservative controlled local governments reflect very closely the popularity of the Prime Minister's party. In particular, the two local elections when Tory controlled councils performed best on average (May 1982 and May 1987) were associated with the popularity peaks of the Conservative central government, while the popularity deficit of the Conservative central government in 1984 to 1986 appears to have affected negatively, on average, the Conservative incumbents.

As regards the districts having all out elections, three such elections occurred in the period for which data are available (1979, 1983, and 1987). After excluding those districts which had no overall political control, we are left with 122 local authorities, for a total of 366 observations. Table 4.4 presents the results of the estimation of the popularity equation (4.19). It turns out that own taxes have a negative effect, and neighbours' taxes have a positive effect on the incumbent's share of the vote. Although this result is robust to the introduction of national politics, the effect of neighbours' taxes does not entirely offset the negative effect of own taxes, and both coefficients are rather imprecisely estimated.

Furthermore, own and neighbours' taxes are strongly spatially autocorrelated. A Moran's I(x) test on the tax rate variable rejects the null hypothesis of absence of spatial correlation with a level of confidence well in excess of 99%. This simply confirms the results obtained in section **4.3**.[39]

The national politics effects again confirm that Conservative-controlled local governments may have benefited from the popularity of the Conservative central government in the 1983 elections.

The above results are broadly consistent with the model in section **4.2**, in terms of the effects of own and neighbours' policies on the share of the vote of the incumbent party at the local authority level. However, they cast serious doubts on the relative importance of local issues - and local taxes in particular - in determining local election results. When we control for the influence of national politics, the effect of local taxes is considerably weakened. As a result, the political agency model in section **4.2** - which attributes systematic influences on voting behaviour exclusively to local factors - is likely not to be an accurate representation of actual voting behaviour.

In spite of the influence of national politics on local elections, though, there is still a large variation in party performance among local authorities. It is therefore perfectly

---

[39] The imprecision in the estimate of the effect of neighbours' taxes might also be due to the low variance of the spatially weighted average of neighbours' taxes vis-à-vis the variance of the own tax rate, combined with the high correlation between the two variables.

reasonable that local tax-setting decisions are affected by electoral considerations. Actually, local incumbents could manoeuvre local taxes in order to offset popularity spillovers from national politics. The spatial interactions that emerge from the estimation of the tax-setting equation in section **4.3** are therefore not inconsistent with the above results.

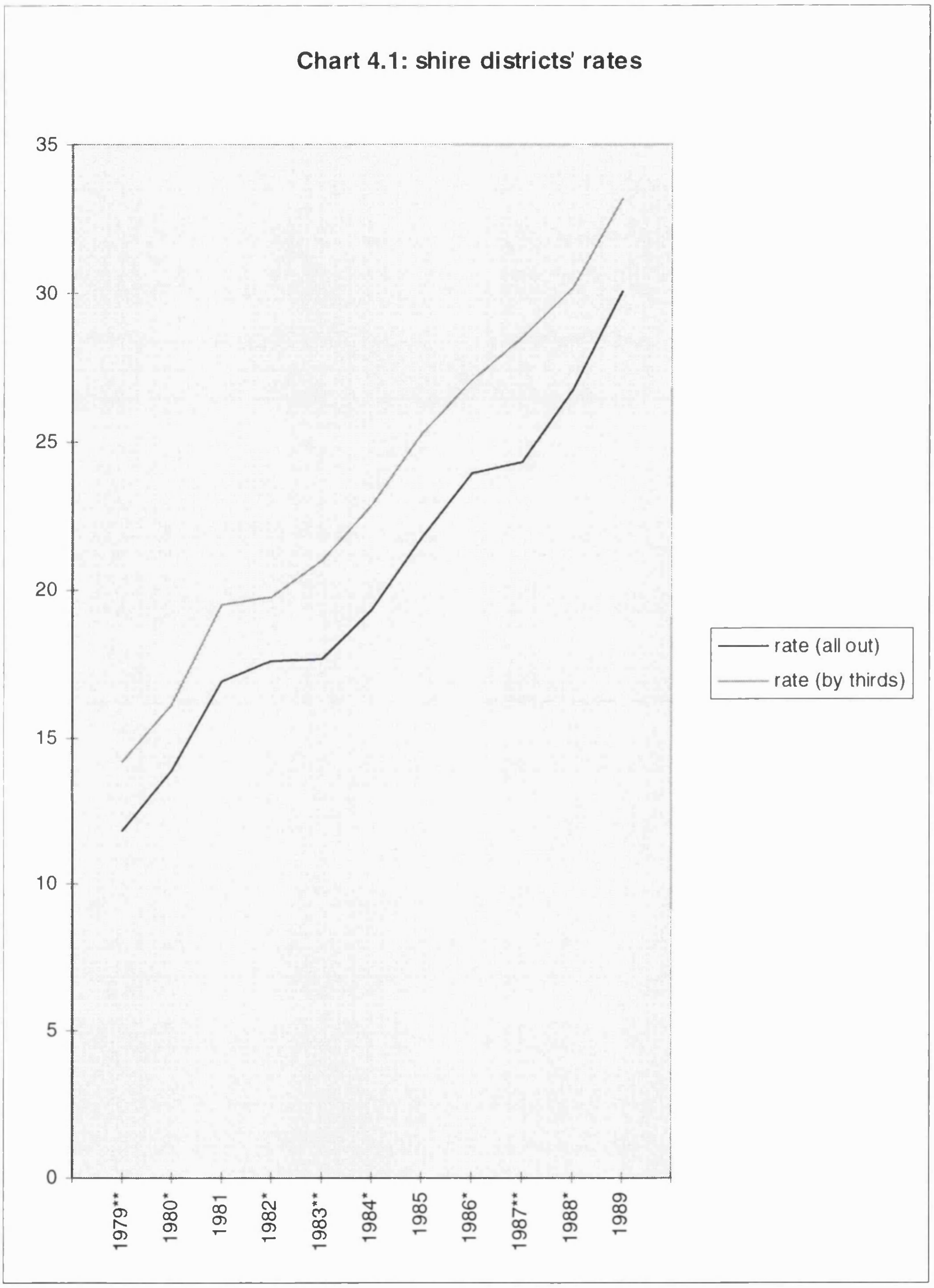
Of course, yardstick competition might not be the only force driving the observed spatial pattern in local taxation. If the tax base can move to nearby jurisdictions in response to tax differentials, local governments could engage in tax competition and give rise to spatial patterns in local tax rates. The mobility of the tax base imposes a further constraint on governments' policies, and a model which incorporates this additional trade-off might be able to discriminate between alternative explanations of spatial correlation in local tax-setting behaviour.

4.5 Conclusions

This work has examined whether political agency problems arising from asymmetric information between voters and representatives can be overcome in a country that is subdivided into a number of local jurisdictions, and where decisions are made at the local level. We have used a principal agent model in which public good provision needs to be managed by a local government, and the local jurisdictions are subject to spatially correlated cost shocks. Voters are assumed to appraise the incumbent government's quality by looking at their jurisdiction's relative performance. In the presence of positive spatial correlation in the cost shock, the incumbent's share of the vote in a jurisdiction would be negatively correlated with the own tax, and positively correlated with the taxes in neighbouring jurisdictions. The model predicts that incumbents seeking re-election look at their neighbours when setting their own taxes, in order to avoid electoral defeat.

We have tested the model on the English local governments' taxes and elections in the 1980s. It turns out that local taxes are positively spatially autocorrelated, and the

IV procedure used to estimate the tax-setting equation guarantees that this correlation cannot be attributed only to spatially correlated shocks. On the other hand, the estimation results of the popularity equation strongly suggest that both local issues and national politics play a role in determining local government election results. As regards the non-metropolitan districts having elections every year, the local tax has a significant negative impact on the incumbent's re-election chances, independently of neighbouring authorities' taxes. However, when national politics spillovers are accounted for, the effect of local taxes decreases dramatically and the local election outcomes seem to be driven to a large extent by national politics issues. As regards the districts having elections every four years, own taxes have a negative effect, and neighbours' taxes have a positive effect on the incumbent's share of the vote. Although this result is robust to the introduction of national politics effects, the effect of neighbours' taxes is imprecisely estimated and does not offset the negative effect of own taxes.



Notes:
1) the rate is expressed as the number of pence per pound of rateable value;
2) all out = shire districts having all out elections every four years (174 districts);
3) by thirds = shire districts having by thirds elections (44 districts);
4) the remaining 78 districts changed electoral system during the period considered here;
5) * = by thirds election year;
6) ** = both all out and by thirds election year.

Table 4.1a Tax-setting equation

	OLS^a	OLS^b
Labour control dummy	7.292 (1.016)	2.304 (.430)
Election year dummy	-.556 (.532)	-.500 (.107)
By-thirds election year dummy	.274 (.501)	.320 (.251)
Domestic tax base (dom. rateable value per head)	-.102 (.037)	-.025 (.028)
Non-domestic tax base (non-dom. rateable value per head)	.019 (.018)	.012 (.040)
urbanisation rate	1.593 (1.334)	
unemployment rate	-20.267 (10.087)	
% under 30	-68.226 (73.350)	
% over 65	-18.253 (71.245)	
% lone parents	93.645 (64.024)	
% ethnic minority	-46.679 (49.893)	
% housing benefit cases	93.778 (51.532)	
% not self-contained accommodation	-151.413 (47.561)	
observations	3552	3256
districts	296	296
R^2	.62	.34

Notes
1) dependent variable is the local tax;
2) standard errors robust to cross-section and time-series heteroskedasticity are in parentheses;
3) OLS^a: equation in levels; OLS^b: equation in first differences;
4) time dummies are included;
5) the time period is 1979 to 1990.

Table 4.1b Tax-setting equation

	OLS^a	IV^a	OLS^b	IV^b
λ	.303 (.073)	.240 (.110)	.329 (.044)	.397 (.247)
Labour control dummy	6.833 (1.070)	6.928 (1.093)	2.226 (.429)	2.209 (.420)
Election year dummy	-.285 (.518)	-.341 (.532)	-.482 (.110)	-.478 (.112)
By-thirds election year dummy	.136 (.490)	.165 (.489)	.352 (.254)	.359 (.256)
Domestic tax base (dom. rateable value per head)	-.082 (.039)	-.086 (.040)	-.028 (.028)	-.028 (.028)
Non-domestic tax base (non-dom. rateable value per head)	.031 (.019)	.029 (.019)	.015 (.039)	.015 (.040)
urbanisation rate	.814 (1.373)	.976 (1.396)		
unemployment rate	-20.196 (9.962)	-20.211 (9.931)		
% under 30	-27.872 (75.200)	-36.263 (77.451)		
% over 65	7.875 (73.479)	2.442 (74.010)		
% lone parents	96.453 (62.004)	95.869 (62.328)		
% ethnic minority	-36.115 (51.503)	-38.312 (51.481)		
% housing benefit cases	91.954 (52.181)	92.333 (52.001)		
% not self-contained accommodation	-97.927 (50.533)	-109.048 (53.291)		
observations	3552	3552	3256	3256
districts	296	296	296	296
Sargan test		19.09 (df=23) p=.70		.68 (df=4) p=.95

Notes
1) dependent variable is the local tax;
2) standard errors robust to cross-section and time-series heteroskedasticity are in parentheses;
3) instruments used are the following. IV^a: z_{1jt}, z_{2j}; IV^b: Δz_{1jt}.
4) df (degrees of freedom) for the Sargan test is the number of overidentifying restrictions;
5) all variables other than the neighbours' taxes are assumed to be strictly exogenous, although none of the overidentifying restrictions that follow from this assumption are exploited;
6) OLS^a, IV^a: equation in levels; OLS^b, IV^b: equation in first differences;
7) time dummies are included;
8) the time period is 1979 to 1990.

Table 4.2 Popularity equation (by thirds elections)

	1 OLS	2 GMM	3 OLS	4 GMM
share $s(t\text{-}1)$	.449 (.044)	.170 (.096)	.519 (.050)	.279 (.046)
local tax c	-.736 (.418)	-1.068 (.490)	-.219 (.436)	-.132 (.463)
local tax $c(t\text{-}1)$	.748 (.412)	.786 (.660)	.358 (.438)	.522 (.633)
neighbours' tax c^n	-.035 (.267)	.341 (.284)	-.050 (.259)	.269 (.281)
neighbours' tax $c^n(t\text{-}1)$	.094 (.323)	-.062 (.321)	.124 (.312)	.078 (.321)
unemployment rate	-.801 (.644)		-.638 (.578)	
% priority housing need	-.478 (.167)		-.420 (.149)	
% housing benefits	1.187 (1.432)		.699 (1.256)	
D80 (Conservative party)			-.133 (.097)	
D82 (Conservative party)			.392 (.125)	.495 (.157)
D83 (Conservative party)			.117 (.091)	-.194 (.072)
D84 (Conservative party)			-.059 (.090)	-.180 (.066)
D86 (Conservative party)			-.070 (.107)	-.052 (.101)
D87 (Conservative party)			.592 (.113)	.571 (.061)
D88 (Conservative party)			-.068 (.102)	-.492 (.076)
Wald test (year/party dummies)			154.9 (p=.00)	107.8 (p=.00)
test for 1st order serial correlation		-2.53 (p=.01)		-2.62 (p=.00)
test for 2nd order serial correlation		-1.56 (p=.12)		-1.32 (p=.19)
Sargan test		46.4 (p=.01)		28.6 (p=.33)
R^2	.37		.50	
observations	602	515	602	515

Notes

1) dependent variable is the logarithm of the logit of the incumbent's share of the vote;

2) standard errors robust to cross-section and time-series heteroskedasticity are in parentheses;

3) columns 1 and 2 include time dummies (not shown); columns 3 and 4 include time dummies interacted with Conservative party incumbency dummy. Those indicators are intended to pick up the influence of national politics on local election results.

4) GMM takes first differences and uses lags of the share dated $(t\text{-}2)$ and earlier as instruments for the lagged dependent variable, and lags of c and c^n dated $(t\text{-}2)$ and earlier as instruments for c, $c(t\text{-}1)$, c^n, and $c^n(t\text{-}1)$;

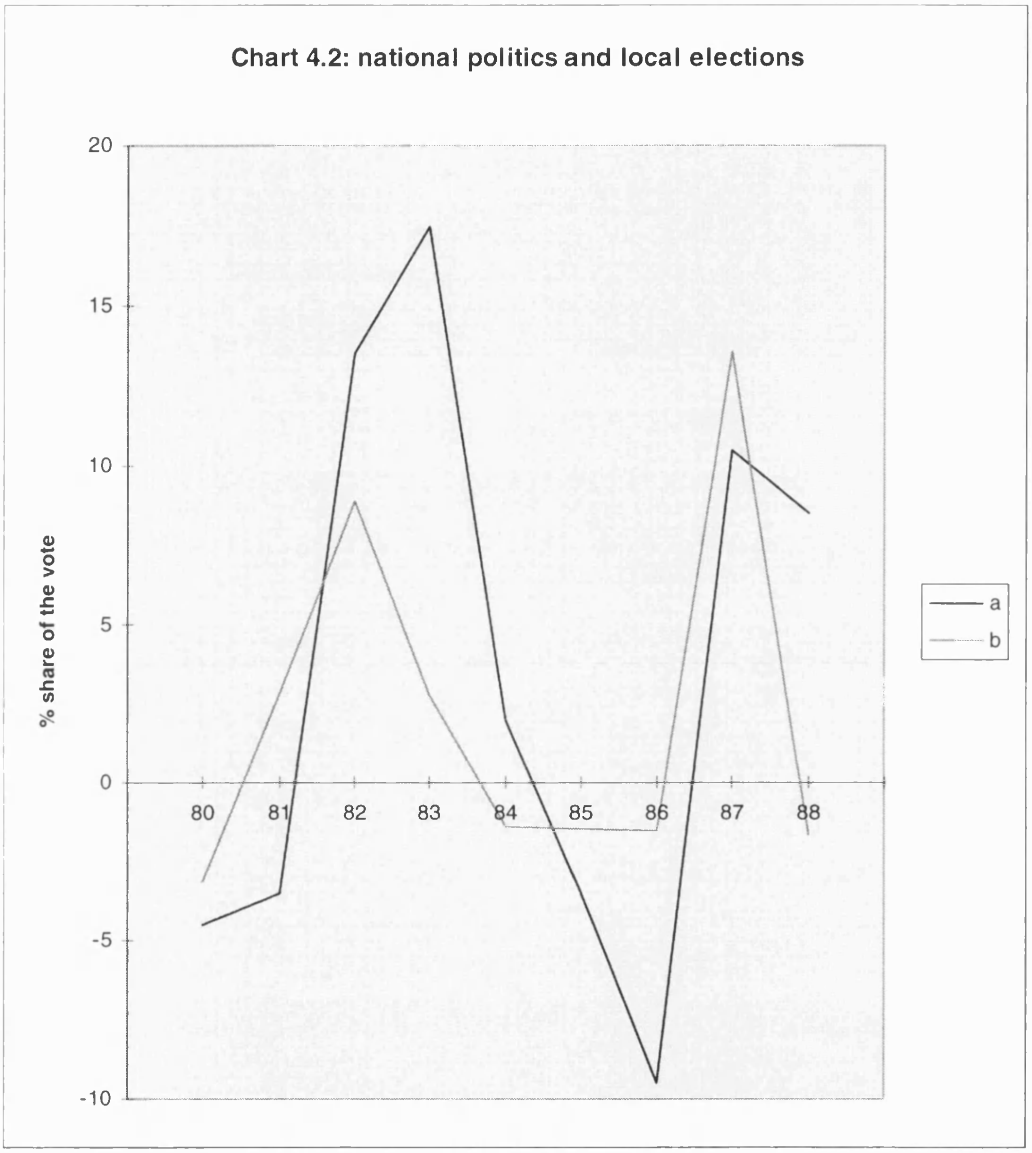
5) we report the one step rather than the two step GMM estimates because of the greater reliability of standard errors in small samples - see Arellano and Bond (1991);

6) one cross-section is lost in taking first differences in columns 2 and 4.

Table 4.3 Gallup poll: size of government party lead over main opposition party

Parliament (*Conservative*)	Government % lead
May 79	-1.5
Sept 79	-4.5
Jan 80	-9
May 80	-4.5
Sept 80	-9.5
Jan 81	-13.5
May 81	-3.5
Sept 81	-4.5
Jan 82	-2
May 82	+13.5
Sept 82	+13.5
Jan 83	+12.5
May 83	+17.5
Sept 83	+21
Jan 84	+3.5
May 84	+2
Sept 84	+1
Jan 85	+6
May 85	-3.5
Sept 85	-1.5
Jan 86	-4.5
May 86	-9.5
Sept 86	-5
Jan 87	-4.5
May 87	+10.5
Sept 87	+11
Jan 88	+8
May 88	+8.5
Sept 88	+6
Jan 89	+7.5
May 89	+2.5
Sept 89	-5

Source: Punnett (1994: 481)



Notes:
a: size of Conservative central government party lead over main opposition party (from table 4.3);
b: estimated effect of national politics on Conservative controlled local governments (in terms of percentage share of the vote of the incumbent; computed from table 4.2, col. 3, after transformation of the dependent variable).

Table 4.4 Popularity equation (all out elections)

	1 OLS	2 IV	3 OLS	4 IV
share $s(t\text{-}1)$	.677 (.041)	.614 (.111)	.670 (.040)	.502 (.117)
local tax c	-.505 (.639)	-1.295 (1.019)	-.599 (.636)	-1.276 (.936)
local tax $c(t\text{-}1)$	.129 (.761)	.770 (1.215)	.539 (.714)	1.372 (1.053)
neighbours' tax c^n	.396 (.872)	.680 (.971)	.592 (.872)	1.383 (.917)
neighbours' tax $c^n(t\text{-}1)$	-.189 (.887)	-.054 (.897)	-.367 (.902)	-.649 (.844)
unemployment rate	-.244 (.665)		-.181 (.669)	
% priority housing need	-.053 (.211)		-.093 (.218)	
% housing benefits	1.693 (1.850)		.308 (1.947)	
D79 (Conservative party)			-.341 (.112)	
D83 (Conservative party)			.121 (.131)	.463 (.119)
D87 (Conservative party)			.044 (.127)	-.008 (.093)
Wald test (year/party dummies)			21.7 (p=.00)	21.8 (p=.00)
instruments used		own and neighbours' demographics, share of the vote lagged twice		own and neighbours' demographics, share of the vote lagged twice
R^2	.58		.62	
observations	366	244	366	244

Notes

1) dependent variable is the logarithm of the logit of the incumbent's share of the vote;

2) standard errors robust to cross-section and time-series heteroskedasticity are in parentheses;

3) columns 1 and 2 include time dummies (not shown); columns 3 and 4 include time dummies interacted with Conservative party incumbency dummy. Those indicators are intended to pick up the influence of national politics on local election results;

4) one cross-section is lost in taking first differences in columns 2 and 4.

Appendix 4.1

Table A4.1 Means of the variables and data sources

Tax-setting equation

variable	sample size	mean	standard deviation	source
local tax	3552	.21	.10	*CIPFA(I)*
neighbours' tax	3552	.20	.07	*CIPFA(I)*
Labour control	3552	.18	.38	*LGCEC*
Dom. tax base (£ per head)	3552	73.60	16.51	*CIPFA(II)*
Non dom. tax base (£ per head)	3552	60.50	22.48	*CIPFA(II)*
urbanisation rate	3552	.58	.34	*CP*
unemployment rate	3552	.23	.05	*CP*
% under 30	3552	.25	.02	*CP*
% over 65	3552	.22	.04	*CP*
% lone parents	3552	.05	.01	*CP*
% ethnic minority	3552	.03	.02	*CP*
% housing benefit cases	3552	.06	.02	*CP*
% not self-contained accommodation	3552	.03	.01	*CP*

Popularity equation (by-thirds elections)

variable	sample size	mean	standard deviation	source
share of the vote	602	.46	.08	*LGCEC*
share of the vote (Conservative)	347	.46	.06	*LGCEC*
local tax	602	.22	.11	*CIPFA(I)*
neighbours' tax	602	.19	.07	*CIPFA(I)*
unemployment rate	602	.23	.04	*CP*
% priority housing need	602	.002	.001	*CP*
% housing benefits	602	.06	.02	*CP*

Popularity equation (all-out elections)

variable	sample size	mean	standard deviation	source
share of the vote	366	.49	.11	*LGCEC*
share of the vote (Conservative)	225	.47	.06	*LGCEC*
local tax	366	.18	.09	*CIPFA(I)*
neighbours' tax	366	.19	.07	*CIPFA(I)*
unemployment rate	366	.23	.05	*CP*
% priority housing need	366	.002	.001	*CP*
% housing benefits	366	.05	.02	*CP*

Data sources

CIPFA(I): Chartered Institute of Public Finance and Accountancy, Finance and general statistics;

CIPFA(II): Chartered Institute of Public Finance and Accountancy, Local government comparative statistics;

LGCEC: Local Government Chronicle Elections Centre, University of Plymouth;

CP: Census of Population (1991).

Conclusions

This work has developed empirical models for the analysis of data on local government taxation and expenditure.

In particular, chapter two has analysed local authority expenditure decisions under the Block Grant system, that is the central government grant distribution scheme that was in place in Britain from 1981 to 1990. It has been recognised that local authorities faced possibly quite involved budget constraints. If local authorities' preferences can be expressed as a non-decreasing function of local private income and local expenditure on public services, the actual budget constraint is - according to the level of the local tax base - either a two-segment convex, or a three-segment non-convex, or a two-segment non-convex one. The lump-sum component of the grant and the price of public services - the ratio of the actual tax base to the tax base implicit in the relevant segment of the Grant Related Poundage schedule - change according to the segment a local authority locates onto, and must be explicitly modelled. Since each segment is characterised by different price and virtual income, and since the choice of segment is endogenous - either due to unobserved heterogeneity among local communities, or due to measurement/optimisation error - a demand function for local public expenditure cannot be consistently estimated with Ordinary Least Squares by attributing to each local government the price and income variables of the segment on which it is observed to be located. Instead, a two-error Maximum Likelihood estimation procedure allowed us to estimate the effect of income and price - and a series of other demographic, political, and socio-economic characteristics - on the expenditures of the English non-metropolitan districts in the fiscal years 1986/87 to 1989/90. The results have shown a positive income effect and a negative price effect, both of high significance. We have compared these results to standard OLS estimates, with no adjustment for the endogeneity of the price and income variables. The price coefficient is underestimated relative to its ML counterpart, due to spurious positive correlation between expenditure and price. On the other hand, the income coefficient does not show a substantial bias. The two-error ML procedure allows us to identify the respective variances of unobserved heterogeneity and optimisation error. The results show significant evidence of unobserved preference heterogeneity, and some weaker evidence of random optimisation error.

Chapter three has explored whether local public spending exhibits a spatial pattern. We have computed measures of spatial correlation in the raw data - i.e., the level of local public expenditure per head - and performed tests based on the OLS regression residuals from the estimation of an equation of local public expenditure determination. All tests for spatial correlation suggested that the regression model should have a spatial structure. The results from the estimation of the spatial lag dependence and spatial error dependence models suggest that spatial autocorrelation is an important feature of local governments' expenditure. Both models can be used to reject a null hypothesis of absence of spatial interactions, i.e., both models are superior to a model that arbitrarily constrains the two spatial coefficients to be zero. In order to discriminate between the two models, we have estimated a general model that nests them. When allowing both for a spatially lagged dependent variable and for a spatially autocorrelated error term, it turns out that most of the correlation in public expenditure is captured by correlated shocks. However, when we control for the presence of district-specific fixed effects, much of the residual spatial correlation disappears. As a result, we can conclude that the significant positive spatial association that we observe in local expenditure could most likely be attributed to spatially autocorrelated unobservable district-specific effects.

Finally, chapter four has examined whether political agency problems arising from asymmetric information between voters and representatives can be overcome in a country that is subdivided into a number of local jurisdictions, and where decisions are made at the local level. We have used a principal agent model in which public good provision needs to be managed by a local government, and the local jurisdictions are subject to spatially correlated cost shocks. Voters are assumed to appraise the incumbent government's quality by looking at their jurisdiction's relative performance. In the presence of positive spatial correlation in the cost shock, the incumbent's share of the vote in a local jurisdiction would be negatively correlated with the own local tax, and positively correlated with the local taxes in neighbouring jurisdictions. The model predicts that incumbents seeking re-election look at their neighbours when setting their own taxes, in order to avoid electoral defeat. We have tested the model on the English local governments' taxes and elections in the 1980s. It turns out that local taxes are positively spatially autocorrelated, and the IV procedure used to estimate the tax-setting equation guarantees that this correlation cannot be attributed to spatially correlated shocks. On the other hand, the estimation results of the popularity equation strongly suggest that both local issues and national politics play a role in determining local government election results.

Overall, the above results have two main implications. First, a spatial modelling approach is appropriate for the analysis of local government tax-setting decisions. Though this work has mainly been concerned with the yardstick competition that arises in a political agency framework, clearly other motives for spatial interactions exist. In the presence of mobility of the tax base - mainly mobility of businesses in response to tax differentials - local governments might engage in tax competition, and give rise to spatial patterns in local taxation.

Second, when studying the determinants of local election results, both local and national factors should be taken into account. Treating local elections either as referenda on the standing of the national parties, or as contests exclusively driven by local issues, is likely to lead to unsatisfactory interpretations of local voting behaviour.

References

Anselin, L. (1988a) *Spatial econometrics: methods and models* (Dordrecht: Klower Academic Publishers).

Anselin, L. (1988b) "Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity," *Geographical Analysis*, **20**, 1-17.

Anselin, L. (1993) "Discrete space autoregressive models," in: M. Goodchild ed., *Environmental modeling with GIS* (New York: Oxford University Press), 454-469.

Anselin, L., and R. Florax (1995) "Small sample properties of tests for spatial dependence in regression models: some further results," in: L. Anselin and R. Florax eds., *New directions in spatial econometrics* (Heidelberg: Springler-Verlog), 21-74.

Anselin, L., and S. Hudak (1992) "Spatial econometrics in practice. A review of software options," *Regional Science and Urban Economics*, **22**, 509-536.

Anselin, L., and S. Rey (1991) "Properties of tests for spatial dependence in linear regression models," *Geographical Analysis*, **23**, 112-131.

Arellano, M., and S. Bond (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, **58**, 277-297.

Azariadis, C., and A. Lahiri (1997) "Do rich countries choose better governments?" University of California, LA, mimeo.

Banerjee, S., and S. Rashid (1996) "Tournaments and sharecropping," *Economica*, **63**, 343-352.

Barnett, R., M. Barrow, and P. Smith (1991) "Representation without taxation: an empirical assessment of the accountability argument underlying the reform of local government finance in England," *Fiscal Studies*, **3**, 30-46.

Barnett, R., R. Levaggi, and P. Smith (1991) "Does the flypaper model stick? A test of the relative performance of the flypaper and conventional models of local government budgetary behaviour," *Public Choice*, **69**, 1-18.

Barnett, R., R. Levaggi, and P. Smith (1992) "Local authority expenditure decisions: a maximum likelihood analysis of budget setting in the face of piecewise linear budget constraints," *Oxford Economic Papers*, **44**, 113-134.

Barr, J., and O. Davis (1966) "An elementary political and economic theory of the expenditures of local governments," *Southern Economic Journal*, **33**, 149-165.

Beck, J. (1983) "Tax competition, uniform assessment, and the benefit principle," *Journal of Urban Economics*, **13**, 127-146.

Bergstrom, T., and R. Goodman (1973) "Private demands for public goods," *American Economic Review*, **63**, 280-296.

Bergstrom, T., D. Rubinfeld, and P. Shapiro (1982) "Micro-based estimates of demand functions for local school expenditures," *Econometrica*, **50**, 1183-1205.

Besley, T., and A. Case (1995a) "Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits," *Quarterly Journal of Economics*, **110**, 769-798.

Besley, T., and A. Case (1995b) "Incumbent behavior: vote seeking, tax setting and yardstick competition," *American Economic Review*, **85**, 25-45.

Besley, T., I. Preston, and M. Ridge (1997) "Fiscal anarchy in the UK: modelling poll tax noncompliance," *Journal of Public Economics*, **64**, 137-152.

Bivand, R. (1984) "Regression modeling with spatial dependence: an application of some class selection and estimation methods," *Geographical Analysis*, **16**, 25-37.

Bivand, R., and S. Szymanski (1997) "Spatial dependence through local yardstick competition: theory and testing," *Economics Letters*, **55**, 257-265.

Borcherding, T., and R. Deacon (1972) "The demand for the services of non-federal governments," *American Economic Review*, **62**, 891-901.

Blundell, R. (1988) "Econometric issues in public sector economics," in: P. Hare ed., *Public sector economics: a reader* (Oxford: Basil Blackwell), 219-247.

Boskin, M. (1973) "Local government tax and product competition and the optimal provision of public goods," *Journal of Political Economy*, **81**, 203-210.

Brennan, G., and J. Buchanan (1980) *The power to tax: analytical foundations of a fiscal constitution* (Cambridge: Cambridge University Press).

Breton, A. (1974) *The economic theory of representative government* (Chicago: Aldine).

Burridge, P. (1980) "On the Cliff-Ord test for spatial correlation," *Journal of the Royal Statistical Society B*, **42**, 107-108.

Burridge, P. (1981) "Testing for a common factor in a spatial autoregression model," *Environment and Planning A*, **13**, 795-800.

Burtless, G., and J. Hausman (1978) "The effect of taxation on labor supply: evaluating the Gary negative income tax experiment," *Journal of Political Economy*, **86**, 1103-1130.

Case, A. (1991) "Spatial patterns in household demand," *Econometrica*, **59**, 953-965.

Case, A., J. Hines, and H. Rosen (1993) "Budget spillovers and fiscal policy interdependence," *Journal of Public Economics*, **52**, 285-307.

Cliff, A., and J. Ord (1981) *Spatial processes. Models and applications* (London: Pion).

Craig, S., and R. Inman (1982) "Federal aid and public education: an empirical look at the new fiscal federalism," *Review of Economics and Statistics*, **64**, 541-552.

Cressie, N. (1991) *Statistics for spatial data* (New York: Wiley).

de Crombrugghe, A., and H. Tulkens (1990) "On Pareto improving commodity tax changes under fiscal competition," *Journal of Public Economics*, **41**, 335-350.

Denny, K., J. Hall, and S. Smith (1995) *Options for business rate reform* (London: The Institute for Fiscal Studies).

Department of the Environment (1986) *Paying for Local Government* (Cmnd. 9714, London: HMSO).

Downs, A. (1957) *An economic theory of democracy* (New York: Harper).

Duncan, A., and P. Smith (1995) "Modelling local government budgetary choices under expenditure limitation," *Fiscal Studies*, **16**, 95-110.

Edwards, J., and M. Keen (1994) "Tax competition and Leviathan," IFS Working Paper No. W94/7.

Epple, D., and A. Zelenitz (1981) "The implications of competition among jurisdictions: does Tiebout need politics?" *Journal of Political Economy*, **89**, 1197-1217.

Foster, C., R. Jackman, and M. Perlman (1980) *Local government finance in a unitary state* (London: George Allen & Unwin).

Fratianni, M., J. Von Hagen, and C. Waller (1993) "Central banking as a political principal-agent problem," CEPR Discussion Paper No. 752.

Gibson, J. (1988) "Rate increases and local elections: a different approach and a different conclusion," *Policy and Politics*, **16**, 197-208.

Gibson, J., and J. Stewart (1992) "Poll tax, rates and local elections," *Political Studies*, **40**, 516-531.

Gordon, R. (1983) "An optimal taxation approach to fiscal federalism," *Quarterly Journal of Economics*, **98**, 567-586.

Gramlich, E. (1977) "Intergovernmental grants: a review of the empirical literature," in: W. Oates ed., *The political economy of fiscal federalism* (Lexington: Lexington Books), 219-240.

Gramlich, E., and D. Rubinfeld (1982) "Micro estimates of public spending demand functions and tests of the Tiebout and median voter hypothesis," *Journal of Political Economy*, **90**, 536-560.

Hausman, J. (1985) "The econometrics of non-linear budget sets," *Econometrica*, **53**, 1255-1282.

Holmstrom, B. (1982) "Moral hazard in teams," *Bell Journal of Economics*, **13**, 324-340.

Hsiao, C. (1986) *Analysis of panel data* (Cambridge: Cambridge University Press).

Kalt, J., and M. Zupan (1990) "The apparent ideological behaviour of legislators: testing for principal-agent slack in political institutions," *Journal of Law and Economics*, **33**, 103-131.

Kanbur, R., and M. Keen (1993) "Jeux sans frontieres: tax competition and tax coordination when countries differ in size," *American Economic Review*, **83**, 877-892.

Keen, M. (1987) "Welfare effects of commodity tax harmonization," *Journal of Public Economics*, **33**, 107-114.

Kelejian, H., and D. Robinson (1993) "A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model," *Papers in Regional Science*, **72**, 297-312.

Lazear, E., and S. Rosen (1981) "Rank-order tournaments as optimum labor contracts," *Journal of Political Economy*, **89**, 841-864.

MaCurdy, T., D. Green, and H. Paarsch (1990) "Assessing empirical approaches for analyzing taxes and labour supply," *Journal of Human Resources,* **3**, 415-490.

McLure, C. (1986) "Tax competition: is what's good for the private goose also good for the public gander?" *National tax Journal*, **39**, 341-348.

Meyer, M., and J. Vickers (1994) "Performance comparisons and dynamic incentives," Economics Discussion Paper No. 96, Nuffield College Oxford.

Miller, W. (1988) *Irrelevant elections?* (Oxford: Oxford University Press).

Mintz, J., and H. Tulkens (1986) "Commodity tax competition between member states of a federation: equilibrium and efficiency," *Journal of Public Economics*, **29**, 133-172.

Moffitt, R. (1984) "The effects of grants-in-aid on state and local expenditures," *Journal of Public Economics*, **23**, 279-305.

Moffitt, R. (1986) "The econometrics of piecewise-linear budget constraints," *Journal of Business and Economic Statistics*, **4**, 317-328.

Moffitt, R. (1990) "The econometrics of kinked budget constraints," *Journal of Economic Perspectives*, **4**, 119-139.

Mookherjee, D. (1984) "Optimal incentive schemes with many agents," *Review of Economic Studies*, **51**, 433-446.

Mueller, D. (1989) *Public Choice II* (Cambridge: Cambridge University Press).

Nalebuff, B., and J. Stiglitz (1983) "Prizes and incentives: towards a general theory of compensation and competition," *Bell Journal of Economics*, **14**, 21-43.

Nordhaus, W. (1975) "The political business cycle," *Review of Economic Studies*, **42**, 169-190.

Oates, W. (1972) *Fiscal federalism* (New York: Harcourt Brace Jovanovich).

Oates, W., and R. Schwab (1988) "Economic competition among jurisdictions: efficiency enhancing or distortion inducing?" *Journal of Public Economics*, **35**, 333-354.

Oates, W., and R. Schwab (1991) "The allocative and distributive implications of local fiscal competition," in: D. Kenyon, and J. Kincaid eds., *Competition among states and local governments* (Washington: the Urban Institute Press), 127-145.

Okabe, A., B. Boots, and K. Sugihara (1992) *Spatial tessellations: concepts and applications of Voronoi diagrams* (New York: Wiley).

Peltzman, S. (1990) "How efficient is the voting market?" *Journal of Law and Economics*, **33**, 27-63.

Peltzman, S. (1992) "Voters as fiscal conservatives," *Quarterly Journal of Economics*, **107**, 327-361.

Pommerehne, W. (1978) "Institutional approaches to public expenditure: empirical evidence from Swiss municipalities," *Journal of Public Economics*, **9**, 255-280.

Preston, I., and M. Ridge (1995) "Demand for local public spending: evidence from the British Social Attitudes Survey," *Economic Journal*, **105**, 644-660.

Punnett, R. (1994) *British government and politics* (Aldershot: Dartmouth).

Rallings, C., and M. Thrasher (1997) *Local elections in Britain* (London: Routledge).

Revelli, F. (1998) "Local public expenditure with piecewise-linear non-convex budget constraints," *Environment and Planning C* (forthcoming).

Rogoff, K. (1990) "Equilibrium political budget cycles," *American Economic Review*, **80**, 21-36.

Romer, T., and H. Rosenthal (1979) "Bureaucrats versus voters: on the political economy of resource allocation by direct democracy," *Quarterly Journal of Economics*, **93**, 563-587.

Rubinfeld, D. (1987) "The economics of the local public sector," in: A. Auerbach and M. Feldstein eds., *Handbook of Public Economics* (Amsterdam: North-Holland), 571-645.

Salmon, P. (1987) "Decentralization as an incentive scheme," *Oxford Review of Economic Policy*, **3**, 24-43.

Shleifer, A. (1985) "A theory of yardstick competition," *Rand Journal of Economics*, **16**, 319-327.

Smith, P. (1993) "Budgeting in the United Kingdom," in: A. Ott ed., *Public sector budgets. A comparative study* (Aldershot: Edward Elgar), 208-226.

Stern, N. (1986) "On the specification of labour supply functions," in: R. Blundell and I. Walker eds., *Unemployment, search, and labour supply* (Cambridge: Cambridge University Press), 143-189.

Tiebout, C. (1956) "A pure theory of local expenditures," *Journal of Political Economy*, **64**, 416-424.

Wellisch, D. (1994) "Interregional spillovers in the presence of perfect and imperfect household mobility," *Journal of Public Economics*, **55**, 167-184.

Wildasin, D. (1986) *Urban public finance* (New York: Harwood Academic Publishers).

Wildasin, D. (1988) "Nash equilibria in models of fiscal competition," *Journal of Public Economics*, **35**, 229-240.

Wildasin, D. (1989) "Interjurisdictional capital mobility: fiscal externality and a corrective subsidy," *Journal of Urban Economics*, **25**, 193-212.

Williams, A. (1966) "The optimal provision of public goods in a system of local government," *Journal of Political Economy*, **74**, 18-33.

Wilson, J. (1985) "Optimal property taxation in the presence of interregional capital mobility," *Journal of Urban Economics*, **18**, 73-89.

Wilson, J. (1986) "A theory of interregional tax competition," *Journal of Urban Economics*, **19**, 296-315.

Wittman, D. (1989) "Why democracies produce efficient results," *Journal of Political Economy*, **97**, 1395-1424.

Zodrow, G., and P. Mieszkowski (1986) "Pigou, Tiebout, property taxation and the underprovision of local public goods," *Journal of Urban Economics*, **19**, 356-370.