Electoral Bias and Policy Choice: Theory and Evidence

Timothy Besley
Ian Preston

THE INSTITUTE FOR FISCAL STUDIES
WP07/06

Forthcoming, *Quarterly Journal of Economics*

Electoral Bias and Policy Choice: Theory and Evidence*

Timothy Besley
London School of Economics, CIAR and IFS

Ian Preston
University College, London and IFS

February 3, 2007

Abstract

This paper develops an approach to studying how bias in favor of one party due to the pattern of electoral districting affects policy choice. We tie a commonly used measure of electoral bias to the theory of party competition and show how this affects party strategy in theory. The usefulness of the approach is illustrated using data on local government in England. The results suggest that reducing electoral bias leads parties to moderate their policies.

*We are grateful to Jim Alt, Robert Barro, Andrew Chesher, Steve Coate, Antonio Merlo, Costas Meghir, Francesco Trebbi and Frank Windmeijer as well as two anonymous referees for useful comments. The paper has also benefitted from discussions with numerous seminar participants.

1 Introduction

One of the central issues in political economy is to understand how electoral incentives shape economic policy outcomes. The dominant model of party competition following Downs (1957) models the consequences of parties competing for votes by making policy promises. The benchmark median voter result suggests that it is voters' preferences rather than political institutions that determine policy outcomes.

The new generation of political economy research gives more weight to how institutions affect policy. This paper focuses on one important institution – the design of the electoral system. It looks specifically at how districting bias in favor of a party affects electoral incentives and policy outcomes. The paper develops a theoretical approach and an empirical application. En route, it discusses how data on the relationship between seats and votes can be incorporated into the study of policy formation.

The basic framework adopted is one in which two parties compete for election by offering policy platforms in a majoritarian electoral system. Voters are either partisan (committed to a particular party) or independent. The latter, which we refer to as swing voters, play a key role in determining election outcomes as parties compete for their voters with their platform choices. Voters are distributed across districts which are heterogeneous, containing different fractions of partisan and swing voters. Such heterogeneity can either be by design, as in the case of partisan gerrymandering, or due to constraints imposed by geography or history.

The distribution of voters across districts affects the way in which votes are translated into seats. A uniform distribution of partisan and swing voters across districts implies that a party will win seats in proportion to its vote share. An electoral system is said to be biased in favor of one party in so far as an equal split of the votes between the parties translates into an unequal division of the seats. As we discuss further in the next section, there is a large empirical literature on the measurement of this kind of bias.

Parties must decide how far to formulate their platforms to please their core (partisan) supporters or appeal to swing voters. The key focus for party strategy is the location of the median district, i.e. the one that they need to win in order to have a majority of seats. If the core support of a party is concentrated, then the median district will be easier to win as a smaller proportion of swing voters will need to be won over in order to win the election. A key insight of the paper is to show that the extent of core support in the median district is equivalent to a statement about bias in the electoral system. This provides a link between the theory of electoral strategy and the empirical study of the relationship between seats and votes.

The theory suggests an empirical approach that we illustrate using English local government data. The estimation strategy is recursive. First we estimate the bias in the electoral system. We then relate this measure of bias to policy outcomes. The results suggest that electoral bias in favor of a party leads to more extreme policy outcomes, i.e. those that favor its core supporters.

The remainder of the paper is organized as follows. The next section relates the paper to a number of literatures in economics and political science. Section four develops the theoretical approach. Section three links this to the seat-votes relationship implied by the model and states our main result. In section five we develop an application. Section six concludes.

2 Related Literature

This paper pulls together ideas from three distinct literatures: models of voting with a mixture of partisan and swing voters, the empirical relationship between seats and votes, and strategic models of party competition. We briefly review how this paper relates to each.

A world comprising partisan and swing voters is a central tenet of the Michigan election studies beginning in the U.S. in the 1950s.[1] Partisans are core supporters who are strongly committed to a given party while swing voters are those whose party attachment is weak. While the nature of partisan attachment in this framework is still much debated (see, for example, Green et al (2004)), the basic idea remains relevant and most survey data collected on voters and

[1] The classic reference – Campbell et al (1960) – has influenced work on voting by political scientists for more than a generation.

voting behavior uses these categories to classify the electorate. This approach is straightforwardly combined with a probabilistic voting model of the kind that is now popular in the political economy literature (see, for example, Persson and Tabellini (2000)).

As we noted above, there is a long-standing interest in political science in modeling the empirical relationship between seats and votes.[2] For a two-party system, this can be represented in general by a seats-votes curve – a real value function mapping the share of the votes received by one party, denoted here by P, into its share of the seats, denoted by S which we write as $S = f(P)$. An electoral system is said to be *biased* when $f\left(\frac{1}{2}\right) \neq \frac{1}{2}$, i.e. a party that wins half of the votes fails to win half of the seats. The slope of the function $f(P)$ represents the *responsiveness* of the electoral system. The case of $f'(P) = 1$ for all $P \in [0, 1]$ represents proportional representation.

The existing political economy literature on electoral systems (for example Persson and Tabellini (2003, 2004)) has tended to adopt a rules driven approach to classifying an electoral system. The seats-votes curve suggests a different way of looking at this, focusing on bias and responsiveness.

Given that data on seats and votes are in plentiful supply, there has been a lot of interest in measuring $f(P)$ empirically and the implied bias and responsiveness. Traditionally this has been done by specifying a functional form for $f(P)$ and estimating its parameters. A popular model is:

$$\log\left(\frac{S}{1-S}\right) = \alpha + \beta \log\left(\frac{P}{1-P}\right). \tag{1}$$

In this case, the parameter α represents bias while β represents responsiveness. Perfect proportional representation is given by $\alpha = 0$ and $\beta = 1$. Majoritarian electoral systems typically yield $\beta > 1$ and to the extent that districting favors one party, then $\alpha \neq 0$.

The empirical model in (1) has a long history. Kendall and Stewart (1950) famously postulated the "cube law" for British elections where $\beta = 3$. The model is developed further in Gelman and King (1994a,b), Quandt (1974) and King (1989). King and Browning (1987) estimate this model on U.S. data. Gelman and King (1994a), among others, have emphasized that bias in the

[2]For an overview, see Taagepera and Shugart (1989).

electoral system is affected by the way in which the boundaries of districts are drawn.

The curve $f(P)$ has not featured to date in theoretical models of partisan politics. In part, this is because, it is has no clear "micro-foundations". This paper shows how the relationship $f(P)$ can be derived in general from an underlying model featuring a mixture of partisan and swing voters who are resident in electoral districts. We also derive necessary and sufficient conditions on primitives that yield the parametric model (1). This gives an exact interpretation of α based on the skewness of a particular distribution relevant to partisan and swing voter support.

The building blocks that we use are similar to those of Coate and Knight (2007), the only other paper of which we are aware that tries to micro-found the seats-votes curve. They too begin with distributions of partisan and swing voters and use their apparatus to look at optimal seats-votes curves for a specified social welfare function. They apply this to U.S. state level data. Their paper links to an earlier literature on optimal partisan gerrymandering in which a party tries to draw district boundaries to maximize its seat share – see, for example, Owen and Grofman (1988).

The final issue concerns the impact of the seats-votes curve on policy. Coate and Knight (2007) adopt a citizen-candidate approach in which each party puts up candidates of a given type who then agree on policy in a legislature. Policy is then affected by the pattern of districting because this has an impact on the composition of the legislature. The mechanism here is somewhat different. We use a model of electoral competition of the kind introduced by Calvert (1985) and Wittman (1983). Parties have policy preferences which reflect the views of their partisan supporters. In this framework, districting affects policy by changing the incentives of parties to appeal to swing voters in order to get elected.

The paper contributes to somewhat broader debates about the value and consequences of political competition.[3] A number of commentators have discussed how voters may benefit from more vigorous competition, but there is no agreed upon framework for studying this issue. A good example is discussion

[3]See Stigler (1972) for an early discussion.

of the pros and cons of one-party systems. In his influential commentary on the one-party south Key (1949) argues that "In the two-party states the anxiety over the next election pushes political leaders into serving the interests of the have-less elements of society," (Key (1949), page 307.) This was institutionalized in part by partisan gerrymandering in favor of the Democratic party. To study this requires a way of thinking about the underlying forces that shape political advantage. To date, the literature has tended to focus on realized political outcomes, such as seat shares, to measure political advantage.[4] But this is conceptually flawed since, in models of strategic party competition, these political outcomes are endogenous. If the parameters of the seats-vote curve $-$ $f(P)-$ can be measured independently of the policy process, then this provides a natural way into modeling the consequences of factors that affect political advantage in favor of a particular party. The framework developed here makes progress on this and suggests a recursive estimation strategy.

The focus of this paper is on bias inherent in districting. This is distinct from the large literature on *incumbent* bias in congressional elections in the United States. That begins from the empirical observation that congressional legislators enjoy a significantly increased chance of being re-elected over time.[5] The policy concern is with whether such bias diminishes the accountability of legislators to the electorate. This makes a lot of sense in an individualistic system like the U.S. where the personal vote matters significantly. In an English context from which we draw our application, parties dominate legislative decision making and this kind of bias is less of an issue. Thus the analysis that we develop is a natural counterpart to the literature on individual legislators.[6]

[4]The U.S. political science literature often uses the Ranney index based on seats and Gubernatorial voting data. Rogers and Rogers (1999) and Besley and Case (2003) show that competitiveness measures based on legislative seats are correlated with policy outcomes in U.S. state level data. This is readily computed using state level data. Holbrook and van Dunk (1993) experiment with a more disaggregated measure using district level data on incumbent's winning margins.

[5]See, for example Erikson (1971).

[6]Some contributions have linked incumbency bias and districting bias. For example, Tufte (1973) argued that incumbents used gerrymandering in the U.S. to create incumbency bias. However, these two kinds of bias are conceptually distinct. Moreover, Ferejohn (1977) argued that behavioral change in the electorate rather than redistricting was most likely responsible

3 Theoretical Preliminaries

We are interested in understanding the incentives for parties to pick policies and how this depends on the relationship between seats and votes. In time-honored tradition, we model two parties competing for office by choosing policy platforms. The model will show how the equilibrium policy platforms depend upon on an empirically relevant model of electoral bias.

3.1 Preferences and Actions

The economy comprises three groups of citizens denoted by $\theta \in \{a, 0, b\}$. These labels denote the loyalties that the citizens have to two parties – labelled a and b with type '0' citizens being independent, i.e. not attached to a party. We will refer to type 0 voters as *swing voters* and the others as *partisan voters*.

The citizens' utility depends on an L-vector of policies $y_1, ..., y_L$ that affect their utility. The set of feasible policies is denoted by a compact set Y.[7] Preferences over these policies are denoted by a bounded utility function $V(\theta, y)$ for a voter of type θ. Partisan voters receive an additional additive component to their utility denoted by Ω_ρ ($\rho \in \{a, b\}$) if their preferred party is in office.[8] Let

$$y^*(\theta) = \arg\max_{y \in Y} V(\theta, y)$$

denote the optimal policy of type θ.

3.2 Politics

Two parties compete for election. We suppose that their memberships comprise only partisan citizens and that they maximize the average welfare of their members. For $z \in [V(0, y^*(0)), V(0, y^*(\rho))]$, consider the following Pareto

for the increase in incumbency bias in the U.S..

[7]The government budget constraint is embodied in Y.

[8]This could correspond to a behavioral model of party loyalty – viewing party attachment as something akin to support for a football team – or else it could represent unmodeled fixed policy preferences on key issues such as state ownership.

efficiency problem for $\rho \in \{a, b\}$:

$$\begin{aligned} \hat{V}(z, \rho) &= \arg\max_{y \in Y} V(\rho, y) \\ &\quad \text{subject to} \\ V(0, y) &\geq z. \end{aligned}$$

This picks the best level of utility for the party subject to delivering a certain utility level for the swing voters. Let $\hat{y}(z, \rho)$ be the policies generated by this program.

We model parties as competing by picking utility levels from the range

$$z \in [\, V(0, y^*(0)),\, V(0, y^*(\rho))\,]$$

– anything else would be ex post Pareto dominated. This begs the question of whether parties can commit to offering something other than their ideal policy. This could be finessed by assuming that candidate selection is a commitment device as in Besley and Coate (1997) or by assuming that parties play a repeated game with the voters as in Alesina (1988).[9]

Let v_ρ be the utility level being offered to the swing voters by party $\rho \in \{a, b\}$. We use $-\rho$ to denote the other party. Let $\underline{V}(z, \rho) = V(\hat{y}(z, -\rho), \rho)$ for $\rho \in \{a, b\}$ be the policy-related utility of a partisan voter when the other party is in office and has offered a utility level of z to the swing voters.

We assume that type a and type b voters remain loyal to their parties. The swing voters weigh up their utility from voting for each party. We model them using a fairly standard probabilistic voting model of the kind used extensively in Persson and Tabellini (2000). Swing voter utility is affected by two "shocks", an idiosyncratic shock ω which has distribution function $F(\omega)$ and an aggregate shock η with distribution function $H(\eta)$. Thus a swing voter with shocks (ω, η) prefers party a over party b if:

$$\omega + \eta + [v_a - v_b] > 0.$$

The final component of this expression is the utility difference $(v_a - v_b)$ between having party a and b in office.

[9]See Besley (2006) for further discussion.

4 Votes and Seats

The model can be used to generate a theoretical seats-votes relationship. Suppose voters are distributed in a continuum of equal sized districts. Each district may contain a different fraction of swing and partisan voters. Specifically, suppose that in any particular district, there are σ voters of type 0 and $(1-\sigma)$ partisan voters. Furthermore $(1+\mu)/2$ of partisan voters support party a in any particular district. Thus $\mu > 0$, denotes an advantage to party a. We now show how the joint distribution of σ and μ across districts affects election outcomes.

First observe that party a wins a seat in which the share of swing voters is σ and it has an advantage in terms of partisan voters of μ if and only if:

$$\sigma\left[1-F\left(-\eta-\left[v_a-v_b\right]\right)\right]+(1-\sigma)\left(\frac{1+\mu}{2}\right)>\frac{1}{2},$$

i.e., it obtains more than half the votes. Rearranging the expression shows that this is so if and only if:[10]

$$\zeta\equiv\mu\left(\frac{1-\sigma}{\sigma}\right)>2F\left(-\eta-\left[v_a-v_b\right]\right)-1.$$

Let $G(\zeta;\Lambda)$ be the distribution function of ζ with parameters $\Lambda=(\bar{\mu},\bar{\sigma},\overline{\sigma\mu})$ where $\bar{\mu}$ is the mean of μ, $\bar{\sigma}$ is the mean of σ and $\overline{\mu\sigma}$ is the mean of $\mu\sigma$.

It is now straightforward to see that party a's seat share is given by:

$$S_a=1-G\left(2F\left(-\eta-\left[v_a-v_b\right]\right)-1;\Lambda\right), \tag{2}$$

which depends on the the distribution of ζ. Party a's overall vote share is equal to:

$$P_a=\frac{1}{2}+\frac{1}{2}\bar{\sigma}\left(2F\left(-\eta-\left[v_a-v_b\right]\right)-1\right)+\frac{1}{2}\left(\bar{\mu}-\overline{\mu\sigma}\right). \tag{3}$$

[10]The new variable ζ defined here can be interpreted as party a's majority from partisan voters expressed relative to the number of swing voters, a natural measure of the party's electoral advantage in a district. The fact, established below, that deriving the properties of the seats-votes relationship requires consideration only of the properties of the univariate distribution of ζ as opposed to those of the bivariate distribution of σ and μ considerably simplifies the analysis.

Substituting (3) into (2) now yields the following theoretical seats-votes relationship:

$$S_a = 1 - G\left(\lambda - \frac{2P_a - 1}{\bar{\sigma}}; \Lambda\right) \tag{4}$$

where $\lambda \equiv \frac{\bar{\mu} - \overline{\mu\sigma}}{\bar{\sigma}}$. This reveals the crucial role played by the shape of the distribution function $G(\cdot;\Lambda)$ and its associated parameters Λ. This seats-votes relationship does not depend on the policy choices of the two parties which have been substituted out since they affect both votes and seats in a similar way.[11] This implies that the relationship in (4) is identified independently of the policy choices made by parties and underpins the recursive empirical strategy outlined below.

Equation (4) can be used to define electoral bias quite generally as the seat share that party a receives when it obtains fifty per cent of the vote. Thus setting $P_a = 1/2$, bias is defined as:

$$S_a - \frac{1}{2} = \frac{1}{2} - G(\lambda;\Lambda).$$

It is clear that the magnitude of the bias depends on how λ deviates from the median of the distribution of ζ, i.e. bias is zero when $G(\lambda;\Lambda) = \frac{1}{2}$.

But how to estimate (4) is not entirely clear. One approach is to adopt the popular parametric log-odds formulation discussed in section 2 above. To rationalize this in our framework requires finding a functional form for $G(\zeta;\Lambda)$ such that (4) becomes:

$$\ln\left(\frac{S_a}{1 - S_a}\right) = \alpha + \beta \ln\left(\frac{P_a}{1 - P_a}\right). \tag{5}$$

Our main theoretical result derives the conditions under which we can move from (4) to (5). The result also shows that appropriate parameters, Λ, can be found to rationalize any such seats-votes curve.[12] We state this as:

Proposition. *For any $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}_+$ there exist $\bar{\sigma} \in [0,1]$, $\bar{\mu} \in [-1,1]$ and*

[11]The critical assumption here is that the policy vector y cannot be targeted to specific districts.

[12]There will typically be many distributions $G(\cdot;\Lambda)$ that can generate a given seat vote curve. All we require for this result is that there there exists one value of Λ.

$\overline{\mu\sigma} \in [-1, 1]$ *such that*

$$\ln\left(\frac{S_a}{1-S_a}\right) = \alpha + \beta \ln\left(\frac{P_a}{1-P_a}\right)$$

for $P_a \in [\frac{\bar{\mu}-\overline{\mu\sigma}-\bar{\sigma}+1}{2}, \frac{\bar{\mu}-\overline{\mu\sigma}+\bar{\sigma}+1}{2}]$ *if and only if*

$$G(\zeta;\Lambda) = \left[1 + \exp(\alpha)\left[\frac{1-\bar{\sigma}\zeta+\bar{\mu}-\overline{\mu\sigma}}{1+\bar{\sigma}\zeta-\bar{\mu}+\overline{\mu\sigma}}\right]^{\beta}\right]^{-1}$$

for $\zeta \in [-1, 1]$.

Moreover the median of the distribution across seats is

$$m \equiv G^{-1}\left(\frac{1}{2};\Lambda\right) = \lambda + \frac{1}{\bar{\sigma}} f(\alpha, \beta)$$

where

$$f(\alpha, \beta) \equiv \left[\frac{1-\exp\{-\alpha/\beta\}}{1+\exp\{-\alpha/\beta\}}\right].$$

provided $m \in [-1, 1]$.

We prove this proposition as the implication of a result for a more general linear seats-votes relationship (of which (5) is a special case) in Appendix I.[13]

The Proposition shows that, for the case where (5) holds, the median can be decomposed into two parts – (i) a factor λ which depends upon the average advantage in partisan votes ($\bar{\mu}$), the average fraction of swing voters ($\bar{\sigma}$) and the covariance between σ and μ and (ii) a factor $f(\alpha, \beta) \in [-1, 1]$ which depends upon the bias and responsiveness coefficients calculated from the seats-votes relationship. This is a useful decomposition since the latter term isolates the skewness of the districting distribution of electoral advantage separately from mean voting preferences. The model identifies districting bias with skewness in the distribution of seats and votes. As α moves further away from zero, the distribution $G(\zeta;\Lambda)$ is more skewed.

[13]Specifically, the Lemma in the Appendix gives necessary and sufficient conditions for any seats-votes curve in the class:

$$\phi(S_a) = \alpha + \beta\psi(P_a).$$

There are three useful and important properties of the function $f(\alpha, \beta)$. First $f(0, \beta) = 0$ in which case the median of the distribution equals the mean and there is no skewness in the distribution. Second, $f(\alpha, \beta)$ is increasing in α and hence is increasing in bias. Third, $f(\alpha, \beta)$ is decreasing in β and thus skewness is decreasing in responsiveness.

Since (4) is derived from a theoretical model, this result will help to provide a link between the measurement of bias and responsiveness in electoral models and policy incentives. We now show that the parameter m –the median of the distribution $G(\cdot; \Lambda)$ – affects the political equilibrium.

5 Political Equilibrium

Party a wins the election if it takes half the seats. This requires that η be sufficiently high that:

$$1 - G\left(2F\left(-\eta - [v_a - v_b]\right) - 1; \Lambda\right) > \frac{1}{2}.$$

or

$$\eta > -[v_a - v_b] - \kappa$$

where $\kappa = F^{-1}\left(\frac{1}{2}(1 + m)\right)$ is an increasing function of the median of the distribution of ζ.[14] The first term, $[v_a - v_b]$, represents the policy advantage of party a if it offers more to the swing voters while the second term, κ, depends upon m. A higher median, m, reflects greater concentration of party a's core support which makes it more likely that party a wins a majority for a given policy advantage.

To see this more clearly, assume that $H(\cdot)$ is uniformly distributed on $\eta \in \left[-\frac{1}{2\xi}, \frac{1}{2\xi}\right]$. Then the probability that party a wins the election for fixed

[14]For example, in the case where $F(\cdot)$ is a uniform distribution on $\left[-\frac{1}{2\phi}, \frac{1}{2\phi}\right]$, then $\kappa = m/\phi$.

(v_a, v_b) is:

$$P_a(\kappa + v_a - v_b) = \begin{cases} 1 \quad \text{if} \;\; \xi[\kappa + v_a - v_b] \geq \frac{1}{2} \\ \frac{1}{2} + \xi[\kappa + v_a - v_b] \\ 0 \;\; \text{if} \;\; \xi[\kappa + v_a - v_b] \leq -\frac{1}{2} \; . \end{cases} \tag{6}$$

A higher value of κ increases the probability that party a wins. Moreover, with a sufficiently high value of κ, party a wins for sure.

Parties compete by picking utility levels for the swing voters. A *political equilibrium* is a pair of promises to the swing voters $(\tilde{v}_a, \tilde{v}_b)$ which form a Nash equilibrium, i.e.

$$\begin{aligned} \tilde{v}_a = \arg\max\nolimits_{v \in [V(0,y^*(0)), V(0,y^*(a))]} \{ \left[\tfrac{1}{2} + \xi[\kappa + [v - v_b]]\right] \left(\Omega_a + \hat{V}(v, a)\right) \\ + \left[\tfrac{1}{2} - \xi[\kappa + [v - v_b]]\right] \underline{V}(v_b, a) \} \end{aligned}$$

and

$$\begin{aligned} \tilde{v}_b = \arg\max\nolimits_{v \in [V(0,y^*(0)), V(0,y^*(b))]} \{ \left[\tfrac{1}{2} - \xi[\kappa + [v_a - v]]\right] \left(\Omega_b + \hat{V}(v, b)\right) \\ + \left[\tfrac{1}{2} + \xi[\kappa + [v_a - v]]\right] \underline{V}(v_a, b) \}. \end{aligned} .$$

At an interior solution, the first order conditions are:

$$\xi \left[\Omega_a + \hat{V}(\tilde{v}_a, a) - \underline{V}(\tilde{v}_b, a)\right] + \left[\frac{1}{2} + \xi[\kappa + [\tilde{v}_a - \tilde{v}_b]]\right] \hat{V}_v(\tilde{v}_a, a) = 0$$

for party a and

$$\xi \left[\Omega_b + \hat{V}(\tilde{v}_b, b) - \underline{V}(\tilde{v}_a, b)\right] + \left[\frac{1}{2} - \xi[\kappa + [\tilde{v}_a - \tilde{v}_b]]\right] \hat{V}_v(\tilde{v}_b, b) = 0$$

for party b.

From these, it is clear that κ, and hence m, affects electoral incentives. This provides a direct link (using the Proposition above) between seats-votes bias and the choice of policy since, as we saw there, the parameter m depends upon α and β.

To get a better feel for how the model works, suppose that y is a scalar $\in [0, 1]$ with party a having preference at $y_a^* = 1$, party b having preference at $y_b^* = 0$ and the swing voters preferring $y_0^* = 1/2$. Preferences are Euclidean, i.e.

$\Omega_r - \|y - y^*_\theta\|$. Solving for the Nash equilibrium, assuming an interior solution[15], yields equilibrium promises:

$$\tilde{v}_a = \frac{1}{2}\left(\Omega_a - \kappa - \frac{1}{2\xi}\right) \text{ and } \tilde{v}_b = \frac{1}{2}\left(\Omega_b + \kappa - \frac{1}{2\xi}\right).$$

This implies that parties offer equilibrium policy platforms equal to:

$$\tilde{y}_a = \frac{1}{2}\left(1 + \frac{1}{2\xi} - \Omega_a + \kappa\right) \text{ and } \tilde{y}_b = \frac{1}{2}\left(1 - \frac{1}{2\xi} + \Omega_b + \kappa\right).$$

In this case $\tilde{y}_a$ and $\tilde{y}_b$ are both increasing in κ, i.e. bias towards a leads *both* parties to move their policy platforms in the direction preferred by party a supporters. Hence in jurisdictions biased towards a, we should see party a moving towards its preferred outcome and in jurisdictions where the bias is towards b, we should see party b moving its policy platform towards outcomes preferred by its core supporters b. This is a key prediction that we test below.[16]

More generally, we can write equilibrium policies predicted by the model as:

$$\begin{aligned} \tilde{y}_{\ell\rho} &= \hat{y}_\ell(\rho, \tilde{v}_\rho) \\ &= Y_\ell(\rho, \kappa) = \hat{Y}_\ell\left(\rho, \lambda + \frac{1}{\bar{\sigma}} f(\alpha, \beta)\right). \end{aligned} \tag{7}$$

for $\rho \in \{a, b\}$ and $\ell = 1, ..., L$. As in the example, electoral bias as represented by α has an effect on policies via its impact on $\tilde{v}_\rho$.

While the Euclidean preference model gives a clean and instructive answer, the theoretical direction of electoral bias on any specific policy is uncertain *a priori*. It depends on the policy positions of committed and swing voters on

[15]This requires that

$$0 \le \frac{1}{2\xi} - \Omega_a + \kappa \le 1 \text{ and } -1 \le \kappa - \frac{1}{2\xi} + \Omega_b \le 0.$$

[16]This example can also be used to think about how policies can be made credible using candidate selection along the lines of Besley and Coate (1997). Suppose that parties pick "leaders" to deliver platforms. They can pick their leader form an array of candidates with ideal points $[Y_a, 1]$ for party a and $[0, Y_b]$ for party b with $Y_a > \frac{1}{2} > Y_b$. This puts an additional constraint on the policy choices which may or may not be binding. For further discussion, see Besley (2006). Coate and Knight (2006) use a model where political selection determines ex post policies. It would be interesting to consider a more general approach (encompassing the model in this paper and their approach) where the selection of candidates on a seat-by-seat basis affects the legislative bargain and/or leadership election after the election.

each issue. Ultimately, therefore, it is an empirical question which policies are affected by electoral bias.

6 An Application

6.1 Background and Data

Our application comes from local governments (called councils) in England for the years 1973-98.[17] As in many countries, local authorities provide of a variety of public services which have an impact on people's daily lives. They deal with public housing, local planning and development applications, leisure and recreation facilities, waste collection, environmental health and revenue collection. Representatives (called councillors) are elected to serve on a geographical basis. The basic geographical unit (district) is a *ward*, generally returning between one and three council members, and usually three. Ward boundaries are determined by a politically independent commission which carries out electoral reviews in each local government area at periodic intervals.[18] In all types of authority,

[17] The system is somewhat complex involving a mixture of single and two tier authorities. There is a single tier of government in London and other metropolitan areas since 1988. Since 1995, there has been a move towards a single tier system throughout England via the creation of shire authorities. However, London and metropolitan areas before 1988 and all shire (rural) areas before 1995 and most of them since responsibilities are split between two levels - a higher level county council and a lower level district council. Where such a split exists the current allocation of functions is roughly as follows. District councils deal with public housing, local planning and development applications, leisure and recreation facilities, waste collection, environmental health and revenue collection. County councils deal with education, strategic planning, transport, highways, fire services, social services, public libraries and waste disposal. Where there is a single tier it typically covers all of these functions (although in London and metropolitan areas transport, fire and waste disposal are handled by joint bodies). In 1990, the break down was 12 inner London boroughs, 20 outer London boroughs, 36 metropolitan districts and 296 shire districts. For an earlier study of seats-votes bias in English local elections and references to previous literature on that subject, see Johnston, Rallings and Thrasher (2002).

[18] Electoral cycles vary depending on type of authority. County councils and London boroughs elect all members at a single election every four years. Metropolitan districts elect by thirds, returning a third of their members on a rotating basis in each of three out of four years.

elections are conducted on a first-past-the-post basis, returning the candidates with most votes, irrespective of whether or not any gains an absolute majority.

The tax raising power of these local authorities is limited with only around 25% of expenditures being funded by local (property) taxes. The remaining budget is financed from (block) grants from central government. There are three main parties competing for office – Conservative, Labour and Liberal Democrats. They communicate platforms to voters mainly through leaflet drops focused on local issues. However, party reputation at the national level may also be a factor in shaping voter loyalties. The Conservative and Labour parties dominate in the national political arena. However, the complexion of local government contests is more heterogeneous. One aspect of this is the relative importance of independent representatives in local politics. We create a catch-all grouping which we call "others" in our data and which comprises independents and other minor parties.[19]

The Conservative party (the party of Margaret Thatcher) is traditionally a right wing party and has a reputation for desiring smaller government and lower taxation at both local and national levels. The Labour party (the party of Tony Blair) is the traditional party of the left while the Liberal Democrats are mostly viewed as mildly left of centre. We would expect these underlying party preferences to show up in the policies chosen under different patterns of political control.

We classify a party as being in control of a local authority if it holds more than 50% of the seats on the council. A small number of local authorities are in the hands of independents in which case we classify them as "other" control. Finally, there are councils that are not controlled outright by anyone – "no overall control". Since many local governments have multiple competing parties, they do not fit the theoretical model particularly well. Hence, the empirical analysis is confined to those local authorities which are essentially "two party" in the sense that there two parties which controlled more than 75%

Shire districts, whether unitary or not, have a choice to opt for either system and changes between the two systems are permitted.

[19]While local candidate reputations clearly play a role in the elections that we are studying, it remains plausible to believe that party attachments are important.

of the seats in every year between 1973 and 1997. This gives a sample of 150 local authorities compared to the universe of 364. To check robustness, we also look at sub-sample of these (108) where the competition is between the Labour and Conservative party. We will use information about seats and votes to construct measures of districting bias in each local authority. The data are available at the district (ward) level. For uncontested wards (approximately 7% of the sample), we impute the vote share with the average for the party over the sample period.[20]

Table 1 reports political outcomes for the 150 local authorities that we study. The table gives the mean of actual political control over our sample period. It also gives the break down of our sample by region and by class of local authority. In spite of restricting the sample to 150 two-party authorities, a broad regional distribution remains as well as a selection of the different authority types. Since we have four party groupings, there are six possible varieties of political competition. The distribution of the sample over these types among the authorities is also given in the Table 1.

The main period for which we have "policy data" is 1980-1998. For each local authority, we have finance data and expenditure. We also get employment in each local authority administration for full time and part time workers. In addition, we have some background socioeconomic data from the Census and other sources. The main economic controls that we use are (log of) household income in each district, the level of unemployment and the (log of) population. Since 1994, we have data available on a range of outcomes from the *Audit Commission*, which was set up in 1992 to monitor the performance of local government. The sample means of the controls and policy variables are given in Tables 1 and 2.

6.2 Empirical Method and Results

The above discussion motivates the following recursive empirical approach. We first use data on seats and votes to estimate parameters of the seats-vote re-

[20]The results are not sensitive to the exact method used to deal with this. Gelman and King (1994b, Appendix A) discusses some alternatives.

lationship α and β. We then use these estimated parameters to see whether electoral bias affects policy outcomes.

6.2.1 Bias

Let j denote a jurisdiction and τ an election date. There are four parameters which determine the crucial jurisdictional median $m_{j\tau}$, namely $\left(\lambda_{j\tau}, \bar{\sigma}_{j\tau}, \alpha_{j\tau}, \beta_{j\tau}\right)$. We assume that $\beta_{j\tau} = \beta$ and $\bar{\sigma}_{j\tau} = \bar{\sigma}$ are common parameters across districts[21], $\lambda_{j\tau} = \lambda_j$ is indexed by j, the local jurisdiction, but constant over time and we decompose $\alpha_{j\tau} = \alpha_j + \epsilon_{j\tau}$ into a fixed authority component α_j and a time-varying authority term ϵ_{jt}. We assume ϵ_{jt} is an innovation unknown to political agents at the time of policy decisions and distributed identically and independently of (λ_j, α_j) so that parties can be assumed to take decisions based on $f(\alpha_j, \beta) = \left[\frac{1-\exp(\alpha_j/\beta)}{1+\exp(\alpha_j/\beta)}\right]$.

We estimate the equation:

$$\ln\left(\frac{S_{aj\tau}}{S_{bj\tau}}\right) = \alpha_j + \beta \ln\left(\frac{P_{aj\tau}}{P_{bj\tau}}\right) + \varepsilon_{kj\tau}. \tag{8}$$

The estimated parameters are then used to construct $f(\alpha_j, \beta)$ which will be used to explain economic policy choices.

The relationship between votes and seats, controlling for local authority effects, is illustrated for the main sample of 150 local authorities and the 108 authorities in the "Labour-Conservative subset" respectively in Figures 1 and 2. The figure is constructed by first regressing both log odds of seats and log odds of votes on a full set of authority dummies, then plotting the relationship between residuals. In each case, we add both the OLS regression and a semi-parametric estimate of the relationship using the same data[22].. The closeness

[21]The proof of the Proposition in Appendix I establishes that it is always possible to find authority specific λ_j and common $\bar{\sigma}$ compatible with the estimated seats-vote relationship. Though it is not necessary to have the same $\bar{\sigma}$ in all jurisdictions, making the assumption here removes an issue of coefficient heterogeneity in the policy regressions.

[22]The semi-parametric estimates fit a non-linear relationship to the residuals using locally weighted regression and should be seen as estimates of the more general form:

$$\ln\left(\frac{S_{aj\tau}}{S_{bj\tau}}\right) = \alpha_j + g\left(\frac{P_{aj\tau}}{P_{bj\tau}}\right) + \varepsilon_{kj\tau}.$$

of the plots suggest a fairly good fit for the linear log-odds relationship between seats and votes.

Results from estimating (8) for elections between 1973 and 1997 are given in Table 3. As we would expect in a majoritarian system, the control of seats varies more than proportionately with changes in votes, i.e. the responsiveness parameter β exceeds one. The point estimate is 1.89 in sample of 150 two party jurisdictions and 2.27 in the smaller sample of 108 Conservative-Labour jurisdictions.

Using the results from (8), we recover the α_j parameter for each of our local authorities.[23] A Wald test strongly rejects the hypothesis that the values of α_j are equal across local authorities in either sample. The next step is to use these estimated values of α_j to construct our measure of districting bias when party r controls in local authority j at time t as:

$$\text{bias}_{rjt} = (2\delta_{rjt} - 1)\, f(\alpha_j, \beta)\,.$$

For the sake of illustration, Figure 3 gives the distribution of (average) bias in favour of the controlling party in each local authority for the sample of 150 two-party authorities. Unsurprisingly, this bias tends to be positive – parties are in control more often when bias is in their favour. This is borne out by looking at the difference in political control in districts with bias and finding that the larger party tends to have bias in its favor.

6.2.2 Policy

Let $y_{r\ell jt}$ be policy outcome ℓ in jurisdiction j in year t when party $r \in \{a, b\}$ is in office where $\ell = 1, ..., L$. Our basic policy equations are of the form:

$$\begin{aligned} y_{r\ell jt} &= \xi_{rj} + \xi_{\ell j} + \xi_{\ell t} + \rho_{\ell r}\,(2\delta_{rjt} - 1)\, f(\alpha_j, \beta) + \delta_{rjt} \\ &\quad + \gamma_\ell x_{\ell jt} + \eta_{r\ell jt} \end{aligned} \tag{9}$$

[23] For five of our local authorities: East Cambridgeshire, Hackney, Islington, Southwark and Surrey Heath, the second placed party received no seats in one or more years. This would not be a problem except for the log formulation of the seats-votes relationship. In these cases, we estimates the bias parameter on the years where seats are not zero for the second placed party and set it to missing in years when the second placed party receives no seats. In practice, this is a very minor issue.

where ξ_{rc} are region dummy variables, $\xi_{\ell c}$ are dummy variables for the type of authority, $\xi_{\ell t}$ are year dummy variables, $\delta_{rjt} = 1$ if party r is in office in district j in year t, and $x_{\ell jt}$ is a vector of exogenous regressors varying at the jurisdiction level that also affect policy.

It is plausible to expect swing voters to have more centrist policy preferences. If this is true, then parties would tend to moderate their policy preferences when facing less electoral bias in their favor. Thus, our empirical specification allows for districting bias to have a different effect on policy depending on which party is in office as would be the case in theory if parties have different policy preferences.

The theory suggests that policy choices depend upon the median of the distribution of ζ in jurisdiction j, i.e.

$$m_j = \lambda_j + \frac{1}{\bar{\sigma}} f(\alpha_j, \beta).$$

This implies that the error term in (9) contains:

$$\eta_{r\ell jt} = \rho_{\ell r} (2\delta_{rjt} - 1) \lambda_j + \varphi_{r\ell jt}.$$

where λ_j is treated as a random effect. We will also allow the distribution of $\varphi_{r\ell jt}$ to be heteroskedastic.[24]

The actual estimation is laid out in detail in Appendix II. It shows how we deal with two further econometric issues. First, we need to allow for the generated regressor bias in the standard errors due to the fact that α_j and β are estimated in (8). Second, since there are four parties rather than two, we estimate these policy equations for different (a, b) pairs in different jurisdictions. Moreover, since we are studying competition between the main two parties in every jurisdiction, we need to introduce a third possibility (no overall control)

[24]Since we assume a fixed β and district-specific α_j Appendix I proves that that it is consistent with theory to have constant σ and then λ_j deterministically related to α_j and β. In that case we would interpret the estimated coefficient on $f(\alpha_j, \beta)$ as coming through λ_j. In that the case the random effect captures authority level specific effects germane to the policy in question. If σ is not constant then λ_j could vary orthogonally to $f(\alpha_j, \beta)$ and this would be absorbed by the random effects. In this case, we have a random coefficient which induces heteroskedasticity in the error term. Our estimation procedure covers both possibilities.

in which neither party has a majority of the seats. This will be the baseline category in the estimations that follow.

We take total local authority expenditures per capita as a core example to illustrate our findings. These are reported in Table 4. We begin in column (1) of Table 4 by showing the relationship between bias and expenditure with only region, class of authority and year dummies as controls, without differentiating between party control and estimating by OLS (with standard errors corrected for the use of generated regressors). The relationship between bias and total spending is positive and significant – incumbents with more bias in their favor set higher expenditures. This relationship holds up when controls are introduced for grants received, income, unemployment and population as is done in column (2). This reduces the coefficient on bias considerably but it remains significant. The coefficients on these controls are sensible – higher grants being correlated with higher expenditures. Column (3) adds in dummies for political control but no interactions between this and bias. Conservative party control reduces expenditures while control by any of the other parties increases it, although not significantly in the case of "other" control. The coefficient on bias falls further in this column, but remains positive and significant, albeit at less than 5%.

Column (4) allows bias to have a different effect depending on which party controls the council. The effect of bias for Labour and Liberal Democratic control is positive and significant. The sign under Conservative party control is negative. Thus bias appears to exaggerate the consequences of party control as measured in column (3). These effects are stronger still when we move to the random effects estimation in column (5) which deals with the full set of issues discussed in the context of equation (9). All effects of bias are larger and more significant. Finally, in column (5), we test the robustness of this finding to looking only at the Conservative-Labour sub-sample. In this instance, bias leads to an increase in expenditure only in Labour controlled councils.[25]

[25]Consistency of our estimates of $\rho_{\ell r}$ relies on consistency in estimation of the α_j and therefore on number of observed election cycles used for estimation becoming large in each jurisdiction. In practice the length of this time dimension is fairly small and we need therefore to recognise the possibility of attenuation bias arising at the second stage as a consequence of inaccuracy in the estimation of the districting biases α_j. Note that, as with classical measurement error, this will make it more difficult for us to establish evidence of an effect.

Together these results suggest that bias in favor of the incumbent does have a significant effect on the expenditure level in a local authority. It appears as if, in line with the theory, less bias leads to parties compromising on their spending preferences – higher spending parties reduce their spending and low spending parties increase it when bias is smaller. The effects are of a reasonable size. A one standard deviation increase in bias increases spending under Labour control as much as a 12% increase in average income and leads to a similar sized reduction in spending under Conservative control.

The remaining tables investigate whether these findings are an artefact of picking total expenditure as an outcome measure. From now on we report only the GLS specification of (9). Table 5 concentrates on taxes.[26] The first column is for the size of the local tax. Council taxes, a form of property taxation, are interesting as they are set locally and are highly visible to residents. The tax varies according property value with each property assigned to a value band. As our policy measure, we use the tax on a particular standardized property value (*Band D*). Column (1) shows that there is a positive and significant effect of bias on the size of this tax under Labour control – a one standard deviation increase in bias leading to a 1% increase in the size of the tax. Column (2) looks at whether bias affects the cost of tax collection. The effect of bias is not significant in this case. Column (3) looks at the percentage of tax collected that is owed to the local authority. There is some evidence here that those jurisdictions with more bias tend to have lower tax collection effort regardless of who controls the local authority.

Having observed in Table 4 that total expenditures are affected by bias, Table 6 looks at their composition. For leisure spending and spending on parks, we find effects that parallel the results in Table 4 with Labour favoring higher spending and Conservatives less. For spending on refuse collection it

Appendix II.4 shows that this bias should be a similar proportion of the true value $\rho_{\ell r}$ across different policies ℓ and suggests a way to estimate the magnitude of the bias. Applying this formula to results in column (5) suggest that the bias towards zero could be about 49% of the true coefficient for Labour, 35% for the Conservatives, 15% for the Liberal Democrats and 21% for others.

[26] These data come from an organization called the Audit commission and are for a shorter time period than the data in Table 4.

appears that bias under Liberal Democratic control increases spending while they spend less on transportation. These results are consistent with parties having different policy preferences.

Table 7 looks at employment by the local authority. In column (1), we look at the (log of) total employment. Again, we find strong effects of bias differentiated by type of party control. Bias under Labour control leads to increases in employment while, under Conservative control, it leads to retrenchment. This accords well with what would expect from first principles. A one standard deviation increase raises full time employment under Labour control by 6% and reduces it by around 3% under Conservative control.

In Table 8, we look at a variety of other policy data from the Audit Commission. Provision of cheap social housing is a typical policy that might appeal to Labour's traditional support but would likely be unattractive to Conservative voters who are more likely to be owner-occupiers. Column (1) shows that more bias under Labour control tends to be associated with greater rent collection in public housing while column (2) shows that there are lower management costs in public housing as bias increases under Labour control. There is weak evidence that bias under Labour control reduces rents on social housing while it increases such rents under Conservative control (column (3)). There is little evidence that bias affects planning and the costs of administering benefits (columns (5) and (6)).

Overall, the results provide convincing evidence that bias matters for policy outcomes. Consistent with the theory, the results suggest that parties try to appeal more to swing voters by moderating their "true" party preference when elections are more competitive (electoral biases are reduced). Particularly persuasive in this respect is the finding that electoral bias increases spending and public employment under left-wing (Labour) control while reducing them under right-wing (Conservative) control.

7 Concluding Comments

A key function of political economy models is to identify how political incentives shape policy choices. Despite a plethora of theoretical models, there are relatively few efforts to build links to empirical estimation. This paper has put forward an approach that links theory and data by exploiting the empirical relationship between seats and votes. The core theoretical idea that is taken to the data is that bias towards one party induced by districting skewed in its favor will make that party more keen to offer policies to suit its core supporters rather than swing voters.

To illustrate the usefulness of the approach, the paper has developed an application to English local government data. We show how the key parameters can be identified and used to explain policy. In line with theory, there is evidence that parties moderate their policy stance when they have less bias in their favor.

The results presented here contribute to debates about the consequences of electoral districting. The pattern of districting is one of the key choices in an electoral system and generates significant policy interest. However, this issue has largely been left alone by the new political economy literature.[27] The findings in this paper suggest that understanding the policy consequences of districting bias does require exploring its implications for the policy strategies formulated by political parties.

[27] Coate and Knight (2007) whose contribution is discussed above is the main exception.

References

[1] Alesina, Alberto, [1988], "Credibility and Policy Convergence in a Two-Party System with Rational Voters," *American Economic Review*, 78(4), 796-806.

[2] Baltagi, Badi H., [1995], *Econometric Analysis of Panel Data*, John Wiley and Sons.

[3] Besley, Timothy, [2005], "Electoral Strategy and Economic Policy" Walras-Bowley Lecture delivered at the World Congress of the Econometric Society, London, August 2005.

[4] Besley, Timothy and Anne Case, [2003], "Political Institutions and Policy Choices: Evidence from the United States," *Journal of Economic Literature*, 41(1), pp. 7-73.

[5] Besley, Timothy and Stephen Coate, [1997], "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 112, 85-114.

[6] Besley, Timothy, Torsten Persson and Daniel Sturm, [2005], "Political Competition and Economic Growth: Theory and Evidence from the U.S. south," NBER Working Paper No.11484.

[7] Calvert, Randall, [1985], "Robustness of the multi-dimensional voting model: Candidate motivations, uncertainty, and convergence," *American Journal of Political Science*, 29, 69-95.

[8] Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, [1960], *The American Voter*, The University of Chicago Press.

[9] Coate, Stephen and Brian Knight, [2006], "Socially Optimal Districting: A Theoretical and Empirical Exploration" forthcoming *Quarterly Journal of Economics*.

[10] Downs, Anthony, [1957], *An Economic Theory of Democracy*, New York: Harper and Bros.

[11] Erikson, Robert S., [1971], "The Advantage of Incumbency in Congressional Elections," *Polity*, 3, 395-405.

[12] Ferejohn, John, [1977], "On the Decline of Competition in Congressional Elections," *The American Political Science Review*, 71(1), 177-76.

[13] Gelman, Andrew and Gary King. [1994a], "A Unified Method of Evaluating Electoral Systems and Redistricting Plans," *American Journal of Political Science*, 38(2), 514-554.

[14] Gelman, Andrew and Gary King, [1994b], "Enhancing Democracy Through Legislative Redistricting," *American Political Science Review*, 88(3), 541-59.

[15] Green, Donald, Bradley Palmquist, and Eric Schickler, [2004], *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*, Yale: Yale University Press,

[16] Holbrook, Thomas M., and Emily van Dunk, [1993], "Electoral Competition in the American States," *American Political Science Review*, 87(4), 955-962.

[17] Johnston, Ron, Colin Rallings and Michael Thrasher, [2002], "Electoral Success, Electoral Bias, and Labour Hegemony: Electoral System Effects in English Metropolitan Boroughs," *Environment and Planning A*, 34(7), 1303-1317.

[18] Kendall, Maurice G. and Alan Stewart, [1950], "The Law of the Cubic Proportion in Election Results," *British Journal of Sociology*, 1, 183-96.

[19] Key, V.O., [1949], *Southern Politics in State and Nation*, reprinted by The University of Tennessee Press, Knoxville: KY.

[20] King, Gary. [1989], "Representation Through Legislative Redistricting: A Stochastic Model," *American Journal of Political Science*, 33(4), 787-824

[21] King, Gary and Robert X. Browning, [1987], "Democratic Representation and Partisan Bias in Congressional Elections", *American Political Science Review*, 81(4), 1251-1273.

[22] Owen, Guillermo and Bernard Grofman, [1988], "Optimal Partisan Gerrymandering," *Political Geography Quarterly*, 7(1), 5-22.

[23] Persson, Torsten and Guido Tabellini, [2000], *Political Economics: Explaining Economic Policy*, Cambridge MA: MIT Press.

[24] Persson, Torsten and Guido Tabellini, [2003], *The Economic Effects of Constitutions,* Cambridge, MIT Press.

[25] Persson, Torsten and Guido Tabellini, [2004], "Constitutional rules and fiscal policy outcomes", *American Economic Review*, 94(1), 25-46.

[26] Quandt, Richard E. [1974], "A Stochastic Model of Elections in Two-Party Systems," *Journal of the American Statistical Association*, 69, No. 346, 315-324.

[27] Rogers, Diane Lim, and John H. Rogers, [2000], "Political competition and state government size: Do tighter elections produce looser budgets?" *Public Choice*, 105 (1), 1-21.

[28] Rogowski, Ronald and Mark Andreas Kayser, [2002], "Majoritarian Electoral Systems and Consumer Power: Price-Level Evidence from the OECD Countries," *American Journal of Political Science*, 46(3), 526-539.

[29] Stigler, George, J. [1972], "Economic Competition and Political Competition", *Public Choice* 13, 91-106.

[30] Taagepera, Rein and Matthew S. Shugart, [1989], *Seats and votes : the effects and determinants of electoral systems,* Newhaven : Yale University Press.

[31] Tufte, Edward R., [1973], "The Relationship Between Seats and Votes in Two-Party Systems," *American Political Science Review*, 67, 540-554.

[32] Wittman, Donald A., [1983], "Candidate Motivation: A Synthesis," *American Political Science Review*, 142-157.

Appendix I Proof of Proposition

The Proposition is proved as an implication of the following lemma.

Lemma. *For any $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}_+$ and for any increasing, surjective ϕ, ψ : $(0,1) \to \mathbb{R}$ there exist $\bar{\sigma} \in [0,1]$, $\bar{\mu} \in [-1,1]$ and $\overline{\mu\sigma} \in [-1,1]$ such that*

$$\phi(S_a) = \alpha + \beta\psi(P_a) \tag{A.1}$$

for $P_a \in [\frac{\bar{\mu}-\overline{\mu\sigma}-\bar{\sigma}+1}{2}, \frac{\bar{\mu}-\overline{\mu\sigma}+\bar{\sigma}+1}{2}]$ if and only if

$$G(\zeta;\Lambda) = 1 - \phi^{-1}\left(\alpha + \beta\psi\left(\frac{1}{2} - \frac{1}{2}\left(\bar{\sigma}\zeta - \bar{\mu} + \overline{\mu\sigma}\right)\right)\right) \tag{A.2}$$

for $\zeta \in [-1,1]$.

Moreover the median of the distribution across seats is

$$m \equiv G^{-1}\left(\frac{1}{2};\Lambda\right) = \lambda + \frac{1}{\bar{\sigma}}\left(1 - \psi^{-1}\left(\frac{1}{\beta}\phi\left(\frac{1}{2}\right) - \frac{\alpha}{\beta}\right)\right) \tag{A.3}$$

provided $m \in [-1,1]$.

Proof. From (3) the widest possible range for P_a as $F(-\eta - [v_a - v_b]])$ ranges from 0 to 1 is

$$\frac{\bar{\mu} - \overline{\mu\sigma} - \bar{\sigma} + 1}{2} \le P_a \le \frac{\bar{\mu} - \overline{\mu\sigma} + \bar{\sigma} + 1}{2}.$$

Suppose that (A.1) holds or equivalently

$$S_a = \phi^{-1}\left(\alpha + \beta\psi(P_a)\right)$$

for $P_a \in [\frac{\bar{\mu}-\overline{\mu\sigma}-\bar{\sigma}+1}{2}, \frac{\bar{\mu}-\overline{\mu\sigma}+\bar{\sigma}+1}{2}]$. This corresponds to (4)

$$S_a = 1 - G\left(\lambda - \frac{2P_a - 1}{\bar{\sigma}};\Lambda\right)$$

and hence (A.2) holds for $\zeta \in [-1,1]$.

If ζ lies outside the range $[-1,1]$ then the number of swing voters is fewer than one or other party's majority among partisan voters and all such seats are won by the dominant party whatever policy platforms v_a and v_b are offered and irrespective of the realization of η. The shape of the seats-votes curve does not therefore restrict the continuation of $G(\zeta;\Lambda)$ to $\zeta \notin [-1,1]$.

We need to show that there exist $\bar{\sigma} \in [0,1]$, $\bar{\mu} \in [-1,1]$ and $\overline{\mu\sigma} \in [-1,1]$ such that these can correspond to the appropriate means if (A.2) holds. To see that this is possible, suppose that $\sigma = \frac{1}{2}$ everywhere so that $\bar{\sigma} = \frac{1}{2}$, $\overline{\mu\sigma} = \frac{1}{2}\bar{\mu}$ and therefore $\zeta = \mu$. Given these values, the mean of ζ must therefore equal $\bar{\mu}$.

Note that we are free to set $G(\zeta;\bar{\mu},\frac{1}{2},\frac{1}{2}\bar{\mu})$ however we want for $\zeta \notin [-1,1]$ provided only that we preserve the properties of a distribution function. Suppose therefore that $G(\zeta;\bar{\mu},\frac{1}{2},\frac{1}{2}\bar{\mu}) = 0$ for $\zeta < -1$ and $G(\zeta;\bar{\mu},\frac{1}{2},\frac{1}{2}\bar{\mu}) = 1$ for $\zeta > 1$.

The difference between the mean of ζ and $\bar{\mu}$ is

$$\Delta(\bar{\mu}) \equiv -G(-1;\bar{\mu},\frac{1}{2},\frac{1}{2}\bar{\mu}) + \left[1 - G(1;\bar{\mu},\frac{1}{2},\frac{1}{2}\bar{\mu})\right] + \int_{G(-1;\bar{\mu},\frac{1}{2},\frac{1}{2}\bar{\mu})}^{G(1;\bar{\mu},\frac{1}{2},\frac{1}{2}\bar{\mu})} G^{-1}(\pi;\bar{\mu},\frac{1}{2},\frac{1}{2}\bar{\mu})\mathrm{d}\pi - \bar{\mu}.$$

Using the properties of $\phi(\cdot)$ and $\psi(\cdot)$,

$$\begin{aligned} G(-1;-1,\frac{1}{2},-\frac{1}{2}) &= 1-\phi^{-1}\left(\alpha+\beta\psi(\frac{1}{2})\right) = 1-\pi_0, \\ G(1;-1,\frac{1}{2},-\frac{1}{2}) &= 1-\phi^{-1}\left(\alpha+\beta\psi(0)\right) = 1, \\ G(-1;1,\frac{1}{2},\frac{1}{2}) &= 1-\phi^{-1}\left(\alpha+\beta\psi(1)\right) = 0, \\ G(1;1,\frac{1}{2},\frac{1}{2}) &= 1-\phi^{-1}\left(\alpha+\beta\psi(\frac{1}{2})\right) = 1-\pi_0 \end{aligned}$$

where $\pi_0 \equiv \phi^{-1}\left(\alpha+\beta\psi(\frac{1}{2})\right)$.

Therefore, by substitution,

$$\begin{aligned} \Delta(-1) &= -[1-\pi_0] + \int_{1-\pi_0}^{1} G^{-1}(\pi;-1,\frac{1}{2},-\frac{1}{2})\mathrm{d}\pi + 1 \\ &\geq (1-\pi_0) - 1 + \pi_0 = 0 \end{aligned}$$

and

$$\begin{aligned} \Delta(1) &= 1-[1-\pi_0] + \int_{0}^{1-\pi_0} G^{-1}(\pi;1,\frac{1}{2},\frac{1}{2})\mathrm{d}\pi - 1 \\ &\leq (1-\pi_0) + \pi_0 - 1 = 0 \end{aligned}$$

using the fact that $-1 \leq G^{-1}(\pi;\bar{\mu},\frac{1}{2},\frac{1}{2}\bar{\mu}) \leq 1$ for all π and all $\bar{\mu}$.

Since $\Delta(\bar{\mu})$ is a continuous function there must therefore be a $\bar{\mu} \in [-1,1]$ such that $\Delta(\bar{\mu}) = 0$ as required.

The formula for the median (A.3) follows simply by inversion.

∎

Letting $\phi(y) = \psi(y) = \ln\left(\frac{y}{1-y}\right)$ in these formulae gives the expressions for $G(\zeta;\Lambda)$ and m in the Proposition.

Appendix II Estimation

Consider estimation of the model for one performance indicator y_{it} in isolation (allowing us to drop the ℓ subscript from the performance equation). Let $s_{it} = \log\left(\frac{S_{ait}}{S_{bit}}\right)$ and $p_{it} = \log\left(\frac{P_{ait}}{P_{bit}}\right)$. Consider firstly the basic model in which incumbent bias effects are the same for all parties. Such a model combines equations

$$\begin{aligned} s_{it} &= \alpha_i + p_{it}\beta + u_{it} \qquad i = 1,\ldots,N \qquad t = 1,\ldots,T_1 \\ y_{it} &= \rho w_{it} f(\alpha_i,\beta) + X_{it}\gamma + \eta_{it} \qquad i = 1,\ldots,N \qquad t = 1,\ldots,T_2 \end{aligned}$$

where T_1 denotes the number of time periods available to estimate the votes-seats relationship, T_2 the number of time periods available to estimate the performance equation and N the number of authorities in the cross section. Here s_{it} denotes the log odds of seat share and v_{it} the log odds of vote share for one party, w_{it} denotes a variable taking values 1, 0 and -1 according to whether that party, no party or its opponent have control of the authority and X_{it} denotes a row vector of all other relevant variables including time and region dummies and dummies for party control.

In matrix form write this as

$$\begin{aligned} s &= D_1\alpha + p\beta + u \\ y &= \rho W f + X\gamma + \eta \end{aligned}$$

in an obvious notation.

II.1 Seats Votes Relationship

Let $Z_1 = (D_1 \quad p)$. Then α and β are estimated at the first stage by

$$\begin{aligned} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} &= (Z_1'Z_1)^{-1}Z_1's \\ &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + (Z_1'Z_1)^{-1}Z_1'u. \end{aligned}$$

We use these estimates to construct estimates of the bias variables

$$\begin{aligned} \hat{f} &= f(\hat{\alpha},\hat{\beta}) \\ &= f(\alpha,\beta) + \nabla(\alpha,\beta)(Z_1'Z_1)^{-1}Z_1'u + o_p(T_1^{-\frac{1}{2}}). \end{aligned}$$

where $\nabla(\alpha,\beta)$ denotes the Jacobian matrix $(\partial f/\partial\alpha' \quad \partial f/\partial\beta)$.

II.2 OLS Performance Regressions

Let $Z_2 = (W\hat{f} \quad X)$. Then second stage OLS estimates are

$$\begin{aligned}
\begin{pmatrix} \hat{\rho}_O \\ \hat{\gamma}_O \end{pmatrix} &= (Z_2'Z_2)^{-1}Z_2'y \\
&= \begin{pmatrix} \rho \\ \gamma \end{pmatrix} + (Z_2'Z_2)^{-1}Z_2'\left[\eta - \rho W\nabla(Z_1'Z_1)^{-1}Z_1'u\right] + o_p(T_1^{-\frac{1}{2}}).
\end{aligned}$$

Hence, as N, $T_1 \to \infty$,

$$\sqrt{N}\left[\begin{pmatrix} \hat{\rho}_O \\ \hat{\gamma}_O \end{pmatrix} - \begin{pmatrix} \rho \\ \gamma \end{pmatrix}\right] \to \mathcal{N}\left(0, A_O^{-1}[B_O + C_O]A_O^{-1}\right)$$

where

$$\begin{aligned}
A_O &= \text{plim}\frac{1}{N}\,Z_2'Z_2 \\
B_O &= \text{plim}\frac{1}{N}\,Z_2'\eta\eta'Z_2 \\
C_O &= \rho^2\text{plim}\frac{1}{N}\,Z_2'W\nabla(Z_1'Z_1)^{-1}Z_1'uu'Z_1(Z_1'Z_1)^{-1}\nabla'W'Z_2
\end{aligned}$$

assuming existence of the appropriate probability limits, taken as $N \to \infty$.

Consistent estimators of standard errors, robust to heteroskedasticity, are based upon corresponding sample moments of the residuals (with appropriate small sample degrees of freedom corrections):

$$\widehat{\text{Avar}}\begin{pmatrix} \hat{\rho}_O \\ \hat{\gamma}_O \end{pmatrix} = \frac{1}{N}\hat{A}_O^{-1}[\hat{B}_O + \hat{C}_O]\hat{A}_O^{-1}$$

where

$$\begin{aligned}
\hat{A}_O &= \frac{1}{N}\,Z_2'Z_2 \\
\hat{B}_O &= \left(\frac{NT_2}{NT_2 - k}\right)\frac{1}{N}\sum_i\sum_t Z_{2it}'Z_{2it}\hat{\eta}_{Oit}^2 \\
\hat{C}_O &= \hat{\rho_O}^2\left(\frac{NT_1}{NT_1 - N - 1}\right)\frac{1}{N}\,Z_2'W\hat{\nabla}(Z_1'Z_1)^{-1}\left[\sum_i\sum_t Z_{1it}'Z_{1it}\hat{u}_{it}^2\right](Z_1'Z_1)^{-1}\hat{\nabla}'W'Z_2.
\end{aligned}$$

Allowing incumbent bias effects to differ by party involves extending ρ to a vector of effects on interactions of $f(\alpha_i, \beta)$ with indicators of party control. Corresponding variance formulae follow by similar reasoning incorporating the appropriate matrix products.

II.3 GLS Performance Regressions

Note however that there is good reason to expect correlation between observations within authorities because of the common influence λ_i. This suggests a random-effects structure. Suppose $E(\eta_i \eta_i') = \Omega_i$ and $\hat{\Omega}_i$ is a consistent estimator of Ω_i. Let $\hat{\Omega}$ denote the block diagonal matrix constructed in the natural way from the estimates $\hat{\Omega}_i$. Then we can calculate a feasible GLS estimator

$$\begin{aligned}\begin{pmatrix} \hat{\rho}_G \\ \hat{\gamma}_G \end{pmatrix} &= (Z_2'\hat{\Omega}^{-1}Z_2)^{-1}Z_2'\hat{\Omega}^{-1}y \\ &= \begin{pmatrix} \rho \\ \gamma \end{pmatrix} + (Z_2'\hat{\Omega}^{-1}Z_2)^{-1}Z_2'\hat{\Omega}^{-1}\left[\eta - \rho W\nabla(Z_1'Z_1)^{-1}Z_1'u\right] \\ &\quad + o_p(T_1^{-\frac{1}{2}}) \end{aligned} \tag{A.4}$$

and, as N, $T_1 \to \infty$,

$$\sqrt{N}\left[\begin{pmatrix} \hat{\rho}_G \\ \hat{\gamma}_G \end{pmatrix} - \begin{pmatrix} \rho \\ \gamma \end{pmatrix}\right] \to \mathcal{N}\left(0, A_G^{-1}[B_G + C_G]A_G^{-1}\right)$$

where

$$\begin{aligned} A_G &= \text{plim}\frac{1}{N}\, Z_2'\Omega^{-1}Z_2 \\ B_G &= \text{plim}\frac{1}{N}\, Z_2'\Omega^{-1}\eta\eta'\Omega^{-1}Z_2 \\ C_G &= \rho^2\text{plim}\frac{1}{N}\, Z_2'\Omega^{-1}W\nabla(Z_1'Z_1)^{-1}Z_1'uu'Z_1(Z_1'Z_1)^{-1}\nabla'W'\Omega^{-1}Z_2 \end{aligned}$$

assuming again that the appropriate probability limits exist.

Consistent estimators of standard errors are based upon corresponding sample moments of the residuals:

$$\widehat{\text{Avar}}\begin{pmatrix} \hat{\rho}_G \\ \hat{\gamma}_G \end{pmatrix} = \frac{1}{N}\hat{A}_G^{-1}[\hat{B}_G + \hat{C}_G]\hat{A}_G^{-1}$$

where

$$\hat{A}_G = \frac{1}{N} Z_2' \hat{\Omega}^{-1} Z_2$$
$$\hat{B}_G = \hat{A}_G \frac{\mathrm{tr}(\hat{\eta}_G \hat{\eta}_G' \hat{\Omega}^{-1})}{NT_2 - k}$$
$$\hat{C}_G = \hat{\rho_G}^2 \left(\frac{NT_1}{NT_1 - N - 1} \right) \frac{1}{N} Z_2' \hat{\Omega}^{-1} W \hat{\nabla} (Z_1' Z_1)^{-1} \left[\sum_i \sum_t Z_{1it}' Z_{1it} \hat{u}_{it}^2 \right] (Z_1' Z_1)^{-1} \hat{\nabla} W' \hat{\Omega}^{-1} Z_2.$$

It remains only to describe the calculation of $\hat{\Omega}$. We set $\hat{\Omega}_i = \hat{\theta}_1 I_{T_2} + \hat{\theta}_2 1_{T_2} 1_{T_2}'$ for all i, estimating $\hat{\theta}_1$ and $\hat{\theta}_2$ from the within- and between-authority sample moments of the pooled OLS residuals $\hat{\eta}_O$ according to the formulae in Baltagi (1995, p.182).

II.4 Small sample bias

Consistent estimation of ρ requires $T_1 \to \infty$ so that estimation error in $\hat{\alpha}$ disappears. For finite T_1, estimation imprecision in $\hat{\alpha}$ creates an effective measurement error issue at the stage of the performance regressions and therefore small sample attenuation bias in estimates of ρ (and γ). Specifically, taking the expectation of the leading term in the expression for (A.4)

$$\begin{aligned}
&E\left[\begin{pmatrix} \hat{\rho}_G \\ \hat{\gamma}_G \end{pmatrix} - \begin{pmatrix} \rho \\ \gamma \end{pmatrix} \right] \\
&= E(Z_2' \hat{\Omega}^{-1} Z_2)^{-1} Z_2' \hat{\Omega}^{-1} \left[\eta - \rho W \nabla (Z_1' Z_1)^{-1} Z_1' u \right] \\
&\simeq \rho (Z_2' \hat{\Omega}^{-1} Z_2)^{-1} \begin{pmatrix} -E \hat{f}' W' \hat{\Omega}^{-1} W \nabla (Z_1' Z_1)^{-1} Z_1' u \\ 0 \end{pmatrix} \\
&= -\rho (Z_2' \hat{\Omega}^{-1} Z_2)^{-1} \begin{pmatrix} E u' Z_1 (Z_1' Z_1)^{-1} \nabla' W' \hat{\Omega}^{-1} W \nabla (Z_1' Z_1)^{-1} Z_1' u \\ 0 \end{pmatrix} \\
&\simeq -\rho (Z_2' \hat{\Omega}^{-1} Z_2)^{-1} \begin{pmatrix} \mathrm{tr} \left(Z_1 (Z_1' Z_1)^{-1} \nabla' W' \hat{\Omega}^{-1} W \nabla (Z_1' Z_1)^{-1} Z_1' \right) E(uu') \\ 0 \end{pmatrix}
\end{aligned} \tag{A.5}$$

The proportional bias in estimation of ρ is independent of the performance indicator in question (except in so far as that affects the sample which can be used for estimation).

Note also that a putative estimate of the bias is available since the variance of u can be estimated using the seats-votes data and the first stage regression results - specifically, $\hat{u}'\hat{u}/NT_1$ consistently estimates $E(u'u)$. This could be used as the basis of small-sample-bias-corrected estimates of ρ and γ. However, while this should reduce small-sample bias it is not clear how it would affect precision of the estimator and it is quite possible that mean square error could increase. We prefer therefore to report the uncorrected but consistent estimates while noting the possible magnitude of bias implied by (A.5).

Table 1: **Sample Characteristics**

Variable	Mean
Characteristics	
ln Average Income	8.7144
	(0.3285)
ln Population	11.6272
	(0.5448)
ln Unemployment	8.1844
	(0.8869)
ln Per Capita Grant	3.8773
	(1.2609)
Political control	
Labour	0.3152
Conservative	0.2818
Liberal Democrat	0.0562
Other	0.0725
Political Competition	
Conservative-Labour	0.7194
Conservative-Lib Dem	0.0866
Labour-Lib Dem	0.0473
Conservative-Other	0.0333
Labour-Other	0.0466
Lib Dem-Other	0.0666
Authority Type	
Inner London	0.0300
Outer London	0.0566
Metropolitan District	0.1020
Shire District	0.8114
Region	
North	0.0807
North West	0.0866
Yorkshire and Humberside	0.0684
East Midlands	0.1117
West Midlands	0.1047
East Anglia	0.0568
South West	0.1009
South East	0.2623
Greater London	0.1280

Standard deviations are in parentheses.

Table 2: **Performance Variables**

Variable	Mean	Standard Deviation
ln Total Expenditure	5.0760	1.0645
Band D Tax	6.4212	0.1651
Tax Collection Costs	17.3802	5.3117
% Tax Collected	0.9423	0.0406
ln Leisure Expenditure	1.9408	0.7480
ln Parks Expenditure	2.0657	0.7982
ln Refuse Expenditure	2.2357	0.2883
ln Transport Expenditure	1.3827	1.6882
ln Total Employment	2.6274	0.8749
ln Full Time Employment	2.3368	0.7728
ln Part Time Employment	1.0997	1.2349
% Rent Collected	0.9852	0.0225
% Housing Management Costs	10.2545	5.1089
Average rent	40.3254	9.7625
Planning decisions within 8 weeks	0.7374	0.1164
Benefit costs	81.2243	32.7658

Table 3: **Seats Votes Regression**

Variable	All Coeff	Lab-Con Coeff
ln odds Vote Share	1.883	2.269
	(0.137)	(0.071)
Sample size	1004	747
R^2	0.868	0.911
Standard error	0.321	0.171
W_{Auth}	985.925	483.793
p value	0.000	0.000

OLS estimates of (8). Standard errors in parentheses are robust to heteroscedasticity. Dependent variable is ln odds Seat Share. All regressions contain a full set of authority dummies and W_{Auth} is a Wald test of absence of authority differences, asymptotically distributed as χ^2_{149}.

Table 4: **Total Expenditure Regressions**

	All					Lab-Con
	OLS				GLS	GLS
	(1)	(2)	(3)	(4)	(5)	(6)
Variable	Coeff	Coeff	Coeff	Coeff	Coeff	Coeff
Pro inc bias	0.407	0.111	0.026	.	.	.
	(0.079)	(0.047)	(0.046)	.	.	.
Pro inc bias * Lab	.	.	.	0.160	0.625	0.844
	.	.	.	(0.064)	(0.159)	(0.238)
Pro inc bias * Con	.	.	.	-0.352	-0.552	-0.504
	.	.	.	(0.111)	(0.163)	(0.256)
Pro inc bias * Lib Dem	.	.	.	0.174	0.409	.
	.	.	.	(0.148)	(0.199)	.
Pro inc bias * Oth	.	.	.	0.313	-0.021	.
	.	.	.	(0.153)	(0.244)	.
Characteristics						
ln Average Income	.	0.058	-0.032	-0.029	0.215	0.371
	.	(0.097)	(0.091)	(0.092)	(0.108)	(0.141)
ln Population	.	-0.315	-0.206	-0.193	-0.099	-0.111
	.	(0.024)	(0.027)	(0.027)	(0.040)	(0.049)
ln Unemployment	.	0.403	0.311	0.301	0.244	0.260
	.	(0.017)	(0.018)	(0.019)	(0.021)	(0.026)
ln Per Capita Grant	.	0.055	0.063	0.060	0.004	-0.004
	.	(0.010)	(0.009)	(0.009)	(0.008)	(0.010)
Political control						
Labour	.	.	0.094	0.087	0.011	0.017
	.	.	(0.016)	(0.016)	(0.018)	(0.021)
Conservative	.	.	-0.111	-0.091	-0.026	-0.048
	.	.	(0.020)	(0.021)	(0.019)	(0.022)
Liberal Democrat	.	.	0.153	0.142	0.065	.
	.	.	(0.027)	(0.027)	(0.028)	.
Other	.	.	-0.003	0.021	-0.026	.
	.	.	(0.026)	(0.030)	(0.039)	.
Sample size	2133.000	2133.000	2133.000	2133.000	2133.000	1535.000
No of years	14.000	14.000	14.000	14.000	14.000	14.000
θ_1	.	.	.	.	0.133	0.140
θ_2	0.079	0.048	0.043	0.043	0.164	0.168
W_{Class}	19855.103	4814.180	5392.707	660.042	207.376	1.466
p value	0.000	0.000	0.000	0.000	0.000	0.690
W_{Region}	140.206	109.713	114.396	6257.870	2168.757	825.494
p value	0.000	0.000	0.000	0.000	0.000	0.000
W_{Year}	1590.053	333.443	249.976	258.298	174.363	104.532
p value	0.000	0.000	0.000	0.000	0.000	0.000

OLS and GLS estimates of (9). Standard errors in parentheses are robust to estimation error in the districting advantage variable. Dependent variable in all regressions is log of total expenditure.

θ_1 denotes the idiosyncratic component and θ_2 the common component to the standard error. W_{Class} is a Wald test of absence of authority class differences, asymptotically distributed as χ^2_3. W_{Region} is a Wald test of absence of regional differences, asymptotically distributed as χ^2_8. W_{Year} is a Wald test of absence of year differences, asymptotically distributed as χ^2 with degrees of freedom one less than the number of years.

Table 5: **Tax Regressions**

Variable	Band D Tax (1) Coeff	Tax Collection Cost (2) Coeff	% Tax Collected (3) Coeff
Pro inc bias * Lab	0.323	4.322	-0.064
	(0.105)	(3.815)	(0.026)
Pro inc bias * Con	-0.190	-4.329	-0.154
	(0.141)	(7.078)	(0.059)
Pro inc bias * Lib Dem	0.087	-2.678	0.000
	(0.118)	(5.140)	(0.031)
Pro inc bias * Oth	-0.489	13.707	0.021
	(0.261)	(10.254)	(0.060)
Characteristics			
ln Average Income	0.020	0.103	-0.011
	(0.082)	(4.236)	(0.027)
ln Population	-0.031	-0.199	-0.018
	(0.034)	(1.421)	(0.009)
ln Unemployment	0.070	-1.359	-0.004
	(0.021)	(1.019)	(0.006)
ln Per Capita Grant	-0.017	1.321	-0.015
	(0.013)	(0.862)	(0.005)
Political control			
Labour	-0.015	0.698	-0.003
	(0.012)	(0.625)	(0.004)
Conservative	-0.035	-0.869	0.021
	(0.015)	(0.916)	(0.007)
Liberal Democrat	0.032	0.272	0.003
	(0.022)	(1.016)	(0.006)
Other	-0.041	1.229	-0.003
	(0.046)	(1.776)	(0.010)
Sample size	720.000	569.000	569.000
No of years	4.000	3.000	3.000
θ_1	0.107	3.522	0.020
θ_2	0.058	2.223	0.015
W_{Class}	16.838	5.835	47.817
p value	0.001	0.120	0.000
W_{Region}	35.313	42.112	69.516
p value	0.000	0.000	0.000
W_{Year}	51.806	3.304	4.392
p value	0.000	0.347	0.222

GLS estimates of (9). Standard errors in parentheses are robust to estimation error in the districting advantage variable.

θ_1 denotes the idiosyncratic component and θ_2 the common component to the standard error. W_{Class} is a Wald test of absence of authority class differences, asymptotically distributed as χ^2_3. W_{Region} is a Wald test of absence of regional differences, asymptotically distributed as χ^2_8. W_{Year} is a Wald test of absence of year differences, asymptotically distributed as χ^2 with degrees of freedom one less than the number of years.

Table 6: **Expenditure Regressions**

Variable	Leisure (1) Coeff	Parks (2) Coeff	Refuse (3) Coeff	Transport (4) Coeff
Pro inc bias * Lab	0.613	1.364	0.208	1.464
	(0.365)	(0.370)	(0.165)	(0.694)
Pro inc bias * Con	-1.368	-0.810	-0.435	0.283
	(0.402)	(0.319)	(0.170)	(0.703)
Pro inc bias * Lib Dem	1.091	-0.040	0.474	-1.376
	(0.491)	(0.344)	(0.214)	(0.945)
Pro inc bias * Oth	0.151	-1.725	-0.567	-1.319
	(0.610)	(0.893)	(0.329)	(1.614)
Characteristics				
ln Average Income	-0.103	0.232	0.016	-1.119
	(0.252)	(0.213)	(0.121)	(0.566)
ln Population	0.140	0.836	0.068	-0.160
	(0.104)	(0.103)	(0.047)	(0.190)
ln Unemployment	0.138	-0.130	-0.079	0.035
	(0.049)	(0.042)	(0.023)	(0.104)
ln Per Capita Grant	-0.026	0.009	0.001	-0.058
	(0.017)	(0.014)	(0.008)	(0.038)
Political control				
Labour	0.136	0.040	0.031	-0.029
	(0.041)	(0.036)	(0.020)	(0.089)
Conservative	0.028	-0.032	-0.033	0.017
	(0.044)	(0.036)	(0.021)	(0.091)
Liberal Democrat	0.038	-0.001	0.099	-0.179
	(0.065)	(0.054)	(0.031)	(0.143)
Other	-0.175	-0.389	-0.142	-0.459
	(0.095)	(0.135)	(0.053)	(0.265)
Sample size	2118.000	2131.000	2132.000	1910.000
No of years	14.000	14.000	14.000	14.000
θ_1	0.400	0.465	0.168	0.595
θ_2	0.364	0.322	0.177	0.783
W_{Class}	24.045	68.670	12.696	3.281
p value	0.000	0.000	0.005	0.350
W_{Region}	25.289	54.578	70.724	275.041
p value	0.001	0.000	0.000	0.000
W_{Year}	59.154	27.438	20.471	75.203
p value	0.000	0.017	0.116	0.000

GLS estimates of (9). Standard errors in parentheses are robust to estimation error in the districting advantage variable. Dependent variable in all regressions is log expenditure.

θ_1 denotes the idiosyncratic component and θ_2 the common component to the standard error. W_{Class} is a Wald test of absence of authority class differences, asymptotically distributed as χ^2_3. W_{Region} is a Wald test of absence of regional differences, asymptotically distributed as χ^2_8. W_{Year} is a Wald test of absence of year differences, asymptotically distributed as χ^2 with degrees of freedom one less than the number of years.

Table 7: **Employment Regressions**

Variable	ln Total Employment (1) Coeff	ln Full time (2) Coeff	ln Part time (3) Coeff
Pro inc bias * Lab	1.342	1.430	0.922
	(0.242)	(0.245)	(0.326)
Pro inc bias * Con	-0.706	-0.671	-0.874
	(0.191)	(0.182)	(0.303)
Pro inc bias * Lib Dem	0.478	0.651	-0.461
	(0.242)	(0.284)	(0.419)
Pro inc bias * Oth	-0.057	0.205	-1.399
	(0.254)	(0.265)	(0.763)
Characteristics			
ln Average Income	0.093	0.030	0.362
	(0.133)	(0.125)	(0.236)
ln Population	0.268	0.310	0.103
	(0.047)	(0.045)	(0.084)
ln Unemployment	0.011	0.006	-0.002
	(0.023)	(0.022)	(0.041)
ln Per Capita Grant	0.005	0.008	0.004
	(0.008)	(0.007)	(0.014)
Political control			
Labour	0.037	0.032	0.083
	(0.022)	(0.022)	(0.036)
Conservative	0.006	-0.007	0.031
	(0.021)	(0.019)	(0.036)
Liberal Democrat	-0.004	-0.027	0.139
	(0.032)	(0.031)	(0.057)
Other	-0.113	-0.094	-0.253
	(0.042)	(0.039)	(0.114)
Sample size	1696.000	1696.000	1696.000
No of years	12.000	12.000	12.000
θ_1	0.170	0.169	0.308
θ_2	0.156	0.150	0.261
W_{Class}	68.755	93.442	2.590
p value	0.000	0.000	0.459
W_{Region}	791.606	563.443	653.319
p value	0.000	0.000	0.000
W_{Year}	28.146	35.488	14.514
p value	0.005	0.000	0.269

GLS estimates of (9). Standard errors in parentheses are robust to estimation error in the districting advantage variable. Dependent variable in all regressions is log employment.
θ_1 denotes the idiosyncratic component and θ_2 the common component to the standard error. W_{Class} is a Wald test of absence of authority class differences, asymptotically distributed as χ^2_3. W_{Region} is a Wald test of absence of regional differences, asymptotically distributed as χ^2_8. W_{Year} is a Wald test of absence of year differences, asymptotically distributed as χ^2 with degrees of freedom one less than the number of years.

Table 8: **Performance Regressions**

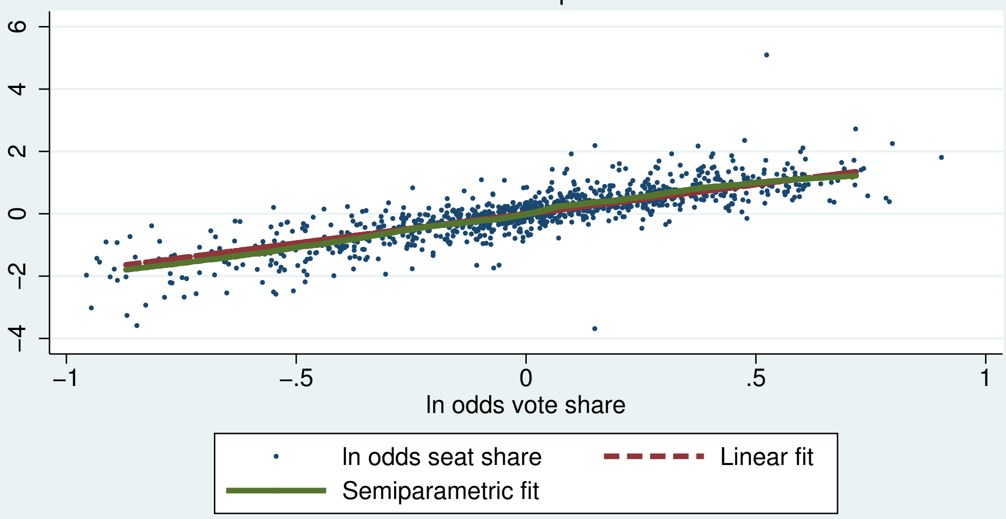
Variable	% Rent Coll (1) Coeff	% Man costs (2) Coeff	Av rent (3) Coeff	Planning (4) Coeff	Benefit costs (5) Coeff
Pro inc bias * Lab	0.052	-9.036	-1.981	-0.153	1.411
	(0.022)	(2.681)	(3.652)	(0.096)	(20.741)
Pro inc bias * Con	0.085	2.781	7.828	0.049	9.752
	(0.059)	(4.595)	(6.836)	(0.191)	(37.352)
Pro inc bias * Lib Dem	0.008	-0.977	-4.436	0.092	26.523
	(0.028)	(2.763)	(4.371)	(0.137)	(28.061)
Pro inc bias * Oth	-0.068	-4.650	-0.258	0.022	22.303
	(0.071)	(8.340)	(13.365)	(0.248)	(55.590)
Characteristics					
ln Average Income	0.007	1.876	4.245	0.193	7.580
	(0.026)	(2.367)	(3.549)	(0.115)	(22.410)
ln Population	0.014	-0.668	0.827	-0.130	10.904
	(0.008)	(0.846)	(1.305)	(0.037)	(7.683)
ln Unemployment	-0.014	0.727	-0.636	0.086	-7.841
	(0.006)	(0.594)	(0.897)	(0.027)	(5.396)
ln Per Capita Grant	0.006	1.142	1.027	-0.023	4.832
	(0.005)	(0.512)	(0.782)	(0.023)	(4.604)
Political control					
Labour	0.010	-0.186	-0.409	0.005	0.370
	(0.004)	(0.335)	(0.488)	(0.018)	(3.241)
Conservative	0.005	0.205	0.539	0.010	-0.407
	(0.006)	(0.516)	(0.761)	(0.027)	(4.718)
Liberal Democrat	-0.001	-0.143	0.724	0.018	0.528
	(0.006)	(0.623)	(0.967)	(0.027)	(5.448)
Other	-0.012	-0.281	-0.638	-0.014	3.243
	(0.012)	(1.387)	(2.196)	(0.043)	(9.363)
Sample size	516.000	517.000	517.000	569.000	567.000
No of years	3.000	3.000	3.000	3.000	3.000
θ_1	0.017	2.281	4.126	0.079	20.266
θ_2	0.013	1.009	1.622	0.068	11.297
W_{Class}	5.580	12.877	2.012	12.820	2.688
p value	0.134	0.005	0.570	0.005	0.442
W_{Region}	9.612	108.087	175.649	22.768	64.867
p value	0.212	0.000	0.000	0.002	0.000
W_{Year}	1.629	6.850	57.180	5.454	4.751
p value	0.653	0.077	0.000	0.141	0.191

GLS estimates of (9). Standard errors in parentheses are robust to estimation error in the districting advantage variable. Dependent variable in all regressions is log expenditure.

θ_1 denotes the idiosyncratic component and θ_2 the common component to the standard error. W_{Class} is a Wald test of absence of authority class differences, asymptotically distributed as χ^2_3. W_{Region} is a Wald test of absence of regional differences, asymptotically distributed as χ^2_8. W_{Year} is a Wald test of absence of year differences, asymptotically distributed as χ^2 with degrees of freedom one less than the number of years.

Figure 1. Seats−votes relationship

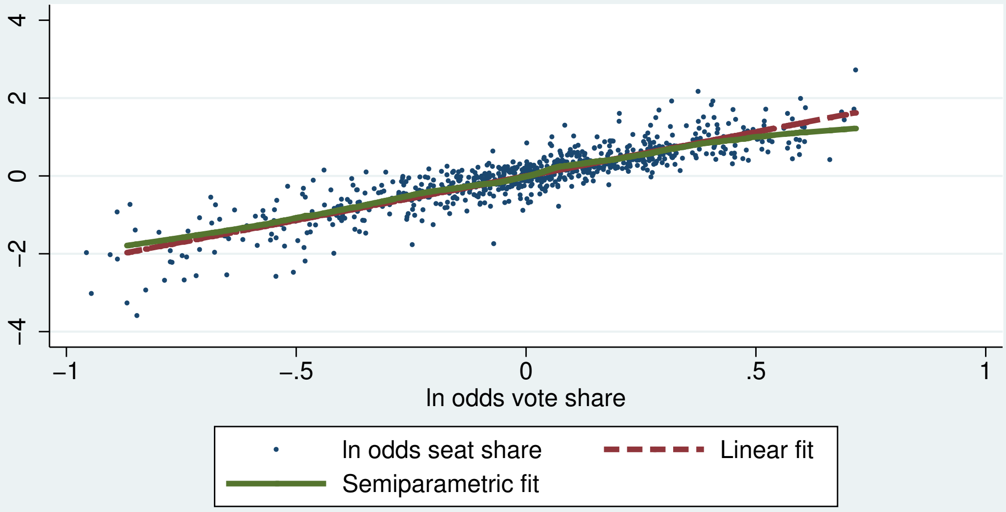
Full sample



Variables on both axes are residuals from regression on a full set of authority dummies.
Linear fit is by OLS regression.
Semiparametric fit is by locally weighted regression.

Figure 2. Seats–votes relationship

Labour–Conservative subsample



Variables on both axes are residuals from regression on a full set of authority dummies.
Linear fit is by OLS regression.
Semiparametric fit is by locally weighted regression.

Figure 3. Distribution of electoral bias