Electoral Bias and Policy Choice: Theory and Evidence

Timothy Besley
Ian Preston

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, 2006

Abstract

This paper develops a new approach to studying how electoral bias in favor of one party due to the pattern of districting affects policy choice. We tie a commonly used measure of bias to the theory of party competition and show how this affects policy choice. 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.

1 Introduction

One of the central issues in political economy is to understand how electoral incentives shape economic policy outcomes. The classic view of party competition initiated by Downs (1957) models electoral competition between parties that compete for votes by making policy promises. The median voter result that emerges suggests that it is voter preferences rather than political institutions that determine policy outcomes.

The new generation of political economy research gives more weight to institutional differences in determining policy outcomes. A subset of this research focuses on the policy consequences of electoral systems.[1] An influential line of empirical work in this area is that by Persson and Tabellini (2003, 2004) which looks at how electoral rules shape policy outcomes. They find that cross-country differences in electoral rules are correlated with a variety of policy differences. They argue that this is consistent with how we would expect such rules to shape political incentives.

This paper develops a new approach that links theory and data to study how electoral incentives work. The main focus is on how the pattern of districting affects policy choices. The possibility of partisan gerrymandering is a major concern in the architecture of democratic systems. A famous example is the Supreme Court decision Baker vs. Carr in 1962 which outlawed electoral malapportionment which favored the Democratic party in the U.S. south. Despite its importance, there is no accepted framework for studying its consequences in the context of models of electoral strategy. The literature that measures such bias has preceded largely in parallel from models that try to understand how policy decisions get made. The main theoretical result in this paper ties the measurement of districting bias to a theoretical model of policy choice and hence makes precise how electoral bias affects policy choice. We apply these ideas to English local government data and show that malapportionment in favor of a particular party does have policy consequences, leading to more extreme policy outcomes.

The remainder of the paper is organized as follows. The next section relates

---

[1] See, for example, Myerson (1993).

the paper to a number of literatures in economics and political science. Section four develops the theoretical approach. Section three develops the seat-votes relationship implied by the model and states our key identification result. In section five we develop an application. Section six concludes.

2   Related Literature

This paper ties together ideas in three distinct areas: models of voting with a mixture of partisan and swing voters, the empirical relationship between votes and seats, and strategic models of party competition.

We model voting in a world comprising partisan and swing voters. This old idea was central to the so-called Michigan election studies beginning in the U.S. in the 1950s. The classic reference Campbell et al (1960) has influenced work on voting by political scientists for more than a generation. The main idea is that the electorate comprises a group of partisan voters who vote for a particular party. The remainder are independent or swing voters. While the nature of partisan attachment is still much debated (see Green et al (2004)) the basic framework remains relevant and most survey collected data on voting behavior uses these categories and hence classifies the electorate in these terms. Districts within a polity can vary in terms of their composition – containing different fractions of independents and partisans. This approach sits well with a probabilistic voting model of the kind that is now popular in the political economy literature (see, for example, Persson and Tabellini (2000)).

The influential work by Persson and Tabellini (2003, 2004) works by classifying electoral systems according to their *rules*, specifically whether they have majoritarian or proportional representation systems. An alternative approach is to look at the empirical relationship between seats and votes.[2] A variety of models have been developed to look at this empirically. One popular model relates $s = \ln$(seat share for one party / seats for others) and $p = \ln$(vote share for one party / vote share for others) as follows:

$$s = \alpha + \beta p$$

[2] Taagepera and Shugart (1989) provides an excellent overview of the political science literature on electoral systems.

and derives estimates for α and β.[3] The parameters have nice interpretations. The parameter α denotes districting *bias* since it indicates whether one party has an advantage over another when votes are divided equally. The parameter β denotes *swing* (or responsiveness) and measures the degree to which the system deviates from proportional representation ($\beta = 1$).[4]

Gelman and King (1994), among others, have emphasized that bias in the electoral system is affected by the way in which the boundaries of districts are drawn. This could be because of the way in which independents and partisans for each party are allocated to each district. The theory developed below shows precisely how the distribution of partisan voters across districts determines α.[5]

But how do electoral systems measured in this way affect electoral strategy? If parties maintained fixed positions, electoral bias would affect policy only in so far as it changes the probability that one party wins. However, if policy making is strategic, then we might expect changes in strategies to attract support. To study this in theory, we use a model of electoral competition similar to those introduced by Calvert (1985) and Wittman (1983). Parties are assumed to represent their partisans and to behave strategically to attract swing voters. Our model shows how an increase in bias, as measured by α above, will lead a dominant party to make less effort to court swing voters.

The analysis in this paper contributes to broader debates about the pros and cons of vigorous political competition.[6] 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.) Malapportionment in favor of one party constitutes one factor that limits electoral competitiveness and hence could be one factor inhibiting competition. Indeed, the Supreme court judgement Baker *vs.* Carr that we referred to above was premised on this

---

[3] King and Browning (1987), for example study this relationship for the United States.

[4] Kendal and Stewart (1950) famously postulated the "cube law" for British elections where $\beta = 3$. The model is developed further in Gelman and King (1994), Quandt (1974) and King (1989).

[5] As as we aware, this paper and Coate and Knight (2005a,b) are the first to give an underlying foundation for the the seats-votes curve in such terms.

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

being the case.

But it is difficult to establish empirically whether political competition matters. Strategic models of political competition suggest that realized political outcomes such as seat shares for one party are likely to be endogenous.[7] One possibility is to find an instrument for political competition using an feature of the political landscape. For example, Besley, Persson and Sturm (2005) use the civil rights acts which struck down voting restrictions in the U.S. south as instruments. An alternative would be to show that some parameters of the electoral system are identified independently of policy choices and hence can be used to measure the state of competition. This is the strategy that we use here.

Here, we focus on bias in inherent in districting. This related to, but 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.[8] 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 that dominate legislative decision making and hence is less of a factor. Thus the analysis here is a natural counterpart to the literature on individual legislators.[9]

---

[7] Notwithstanding, there is a long tradition of looking at the policy consequences of electoral competitiveness in such terms. 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.

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

[9] 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 for the increase in incumbency bias in the U.S..

3 Theoretical Preliminaries

We are interested in understanding the incentives for parties to pick policies and how this depends on the seats-votes relationship. There are two parties competing for office by choosing policies. Policies affect the extent to which they can attract support from independent (or swing) voters.

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 not being attached to a party. We assume that there are σ voters of type 0 and $(1 - \sigma)$ voters of the other two types. 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 Y. Citizens also have an additional additive component to their utility which depends on which party is in office. The latter is also the primary basis of party loyalty.[10] Let $\Omega_\rho (> 0)$ denote the utility that a partisan voter has from having his own preferred party in office, i.e.,

$$V(\theta, y) + \delta(\theta, \rho) \Omega_\rho \text{ for } \rho \in \{a, b\} \text{ and } \theta \in \{a, 0, b\}$$

where $\delta(0, \rho) = 0$ always and $\delta(\theta, \rho) = 1$ if $\rho = \theta$ and zero otherwise. 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

---

[10]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.

members. For $z \in [V(0, y^*(0)), V(0, y^*(\rho))]$, consider the following Pareto efficiency problem for $\rho \in \{a, b\}$:

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

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 Besley, Persson and Sturm (2005) or by assuming that parties play a repeated game with the voters as in Alesina (1988).

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:

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

Voters are distributed in a continuum of districts. We assume that the fraction of swing voters is the same in each. However, each may contain a different

fraction of partisan voters. Specifically, suppose that $(1 + \mu)/2$ of partisan voters support party a in any particular district where μ is distributed on $[\underline{\mu}, \bar{\mu}]$ where $\underline{\mu} \geq -1$ and $\bar{\mu} \leq 1$, with mean λ. Let $G(\mu; \lambda)$ be the distribution function of μ across districts. The shape of this distribution will be important in affecting electoral incentives. It reflects the pattern of each party's core support across districts and the extent to which it is concentrated in particular seats.

4 Votes and Seats

The model can be used to generate a theoretical seats-votes relationship. First observe that party a wins a seat in which it has an advantage in terms of partisan voters of μ if and only if:

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

i.e., it obtains more than half the votes. Using this, it is straightforward to see that it wins a share of the seats given by:

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

which depends on the distribution of partisan voters across districts. Party a's vote share is equal to:

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

This depends only on the average partisan voters support λ and not its distribution. Substituting (2) into (1) now yields the following theoretical seats-votes relationship:

$$S_a = 1 - G\left(\left(\lambda - \frac{2P_a - 1}{(1 - \sigma)} \right); \lambda \right). \tag{3}$$

This reveals the crucial role played by the shape of the distribution function $G(\cdot; \lambda)$ and the underlying "structural" parameters of the electoral system – λ, σ and the distribution of μ. This 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. This implies that (3) is identified independently

8

of the policy choices made by parties. This underpins the empirical strategy below, where we estimate parameters of (3) which are then used to explain policy choices.

But how to estimate (3) is not entirely clear. We are interested in particular in whether we can justify the classic log-odds relationship between seats and votes in this theoretical framework, i.e. when (3) is the form:

$$\ln\left(\frac{S_a}{1-S_a}\right) = \alpha + \beta \ln\left(\frac{P_a}{1-P_a}\right) \tag{4}$$

where S_a is the share of the seats accruing to party a and P_a is the share of the votes for that party. As we discussed above this an intuitive and convenient approach to measuring features of electoral systems. 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 $\alpha \neq 0$. (The latter represents the fact that one party is favored in seats even if $P_a = \frac{1}{2}$.) Even if (4) can be derived we also need to be sure that we can find parameters (σ, λ) that rationalize this seats-votes relationship.

We now show that we can indeed make the move from (3) to (4) provided that the function $G(\cdot; \lambda)$ takes a particular form. We also show that a (σ, λ) pair can be found to rationalize any such curve. The result is provided in:

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

$$G(\mu, \lambda) = \frac{\exp(\alpha)\left[\frac{\frac{1}{1-\sigma}-\mu+\lambda}{\frac{1}{1-\sigma}+\mu-\lambda}\right]^{-\beta}}{1 + \exp(\alpha)\left[\frac{\frac{1}{1-\sigma}-\mu+\lambda}{\frac{1}{1-\sigma}+\mu-\lambda}\right]^{-\beta}}$$

for $\mu \in [\lambda - \frac{1}{1-\sigma}, \lambda + \frac{1}{1-\sigma}] \cap [-1, 1]$ then

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

for $P_a \in [\frac{1}{2}(1-\sigma)(1+\lambda), \sigma + \frac{1}{2}(1-\sigma)(1+\lambda)]$. Moreover the median of the distribution across seats is

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

9

where

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

provided m_G lies in the interior of the range for μ.

We prove this proposition as the implication of a result for a more general linear seats-votes relationship (of which (4) is a special case) in Appendix I. The Proposition shows that, for the case where (4) holds, the median can be decomposed into two parts – (i) the mean of μ (equal to λ) (ii) a factor $f(\alpha, \beta) \in [-1, 1]$ which depends upon the bias and swing coefficients calculated from the seats-votes relationship. This is a useful decomposition since the latter term isolates the skewness of the distribution of the partisan voters separately from the mean. The model identifies districting bias with skewness in the distribution of seats and votes. As α moves further away from zero, the distribution $G(\cdot, \lambda)$ is more skewed in a well-defined sense.

Since (3) is derived from a theoretical model, this result will help to provide a link between the measurement of bias and swing in electoral models and policy incentives. We now show that the parameter m_G affects political competition and hence political equilibrium.

5   Political Equilibrium

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

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

or

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

where $\kappa = F^{-1}\left(\frac{1}{2} \left(1 + \frac{1-\sigma}{\sigma} m_G \right) \right)$ is an increasing function of the median of the distribution of μ.[11] 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 is a

---

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

measure of the state of political competition. If the median of the distribution of partisan voters is higher – reflecting greater concentration of party a's core support, then this is good for party a and makes it more likely that this party will win. To see this more clearly, assume that $H\left(\cdot\right)$ 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 (v_a, v_b) is:

$$
P_a\left(\kappa + v_a - v_b\right) = \begin{cases} 1 & \text{if } \xi\left[\kappa + v_a - v_b\right] \geq \frac{1}{2} \\ \frac{1}{2} + \xi\left[\kappa + v_a - v_b\right] & \\ 0 & \text{if } \xi\left[\kappa + v_a - v_b\right] \leq -\frac{1}{2} \ . \end{cases} \tag{5}
$$

With a sufficiently large advantage, party a wins for sure. Otherwise, there will be an interior solution. From now onwards, we focus on this case. A higher value of κ increases (reduces) the advantage of party a (b).

Parties compete by picking utility levels for the swing voter. Policy choices form a Nash equilibrium, i.e.

$$
\tilde{v}_a = \arg\max_{v \in [V(0, y^*(0)), V(0, y^*(a))]} \{\left[\frac{1}{2} + \xi\left[\kappa + [v_a - v_b]\right]\right]\left(\Omega_a + \hat{V}(v, a)\right) \\ + \left[\frac{1}{2} - \xi\left[\kappa + [v_a - v_b]\right]\right]\underline{V}(v_b, a)\}
$$

and

$$
\tilde{v}_b = \arg\max_{v \in [V(0, y^*(0)), V(0, y^*(b))]} \{\left[\frac{1}{2} - \xi\left[\kappa + [v_a - v_b]\right]\right]\left(\Omega_b + \hat{V}(v, b)\right) \\ + \left[\frac{1}{2} + \xi\left[\kappa + [v_a - v_b]\right]\right]\underline{V}(v_a, b)\} \ .
$$

We make:

Assumption1 $\quad: \hat{V}_v(v, \rho) + \underline{V}_v(v, \rho) < 0$

$$\text{for all } v \in [V(0, y^*(0)), \max\{V(0, y^*(a)), V(0, y^*(b))\}] \ .$$

This requires that there be sufficient conflict of interest between the parties and the swing voters. As shown in Besley (2005), this assumption guarantees that the game played between the two parties is supermodular and hence that a pure strategy Nash equilibrium exists. The Nash equilibrium is said to be extremal if either $\tilde{v}_\rho = V(0, y^*(0))$ or $\tilde{v}_\rho = V(0, y^*(\rho))$ for some $\rho \in \{a, b\}$ and is interior otherwise.

This model motivates the importance of electoral bias to policy making. To see this, consider an interior solution for both parties. The first order conditions

11

in this case are:

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

for party a and

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

for party b.

From this, it is clear that κ and hence m_G 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 parameters α and β affect m_G.

A *political equilibrium* is a pair of promises to the swing voters $(\tilde{v}_a, \tilde{v}_b)$ which form a Nash equilibrium. It is interior when it satisfies this pair of equations above. Appendix III gives a simple worked example where the function $\hat{V}(\tilde{v}_\rho, \rho)$ $(\rho \in \{a, b\})$ is linear. This illustrates the equilibrium. It also shows that $\tilde{v}_a$ is increasing in κ and $\tilde{v}_b$ is decreasing in κ.

Given an equilibrium pair of promises to the swing voters, the parties pick policies to achieve these utility levels. Hence:

$$
\begin{aligned}
y_{\ell\rho} &= \hat{y}_\ell(\rho, \tilde{v}_\rho) \\
&= Y_\ell(\rho, \kappa) = \hat{Y}_\ell\left(\rho, \lambda + \frac{1}{1-\sigma} f(\alpha, \beta)\right).
\end{aligned}
\tag{6}
$$

for $\rho \in \{a, b\}$ and $\ell = 1, ..., L$. This shows that electoral bias as represented by α has an effect on policies indirectly, i.e. via its impact on $\tilde{v}_\rho$.

The theoretical direction of electoral bias on any specific policies is a priori ambiguous. First, it depends on how it impacts on the promise to swing voters that the party makes and second, on the party's preferences over policies. There may be some issues on which a party does not mind giving ground to the swing voters while on others the conflict of interest between parties and swing voters is strong. This would be true for policies which affects the rents that parties earn from holding office. Hence, while we do expect bias to matter – especially on policies that matter to swing voters, it is an empirical question where this has bite.

12

6   An Application

This section develops an application of the theoretical framework using data from local government in the U.K..

6.1   Background and Data

Our application comes from local government in England for the period 1973-98.[12]   These are charged with responsibility for provision of a variety of local public services.   In all parts of England, they deal with public housing, local planning and development applications, leisure and recreation facilities, waste collection, environmental health and revenue collection.   Councillors are elected to serve on District Councils on a geographical basis.   The basic geographical unit 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.[13]   In all types of authority, elections are on a first-past-the-post basis, returning the candidates with most votes, irrespective of whether or

---

[12] 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.

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

not any gain an absolute majority.

The taxing authority of these local governments is limited with only around 25% of finance being raised locally. The remainder is covered by grants from central government. There are three main parties – Conservative, Labour and Liberal Democrats. The remainder of seats are placed under a catch-all category "others" which mainly comprises independents. The Conservative party (the party of Margaret Thatcher) is a right wing party and has a reputation for desiring smaller government and lower taxation. 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 preferences to show up in the policy chosen under different patterns of political control.

We classify a party as being in control 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 analysis is confined to those which are "two party" authorities in the sense that two parties controlled more than 75% of the seats in every year between 1973 and 1997. This gives a sample of 150 local authorities compared to the universe of 364. We will use data to construct various measures of districting bias in each local authority.[14]

Table 1 reports key political outcomes for this period for the 150 local authorities that we use. The table gives the mean of actual political control over our sample. 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 given in the table.

Our focal period for which we have "policy data" is 1980-1998. For each local authority, we have finance data and expenditure. We also get employment

---

[14] 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).

in each local authority administration for full time and part time workers. We also have some background socioeconomic data constructed 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, set up in 1992 to monitor local government. The sample means of the key controls and policy outcome variables used in the study are given in Tables 1 and 2.

6.2 Empirical Method and Results

The above discussion motivates the following two-stage empirical approach. We first use data on seats and votes to estimate parameters of the seats-vote relationship α and β. We then use the 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 to the distribution of partisan voters across districts, $(\lambda_{j\tau}, \sigma_{j\tau}, \alpha_{j\tau}, \beta_{j\tau})$. We assume that $\beta_{j\tau} = \beta$ and $\sigma_{j\tau} = \sigma$ are common parameters across districts[15], $\lambda_{j\tau} = \lambda_j$ 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 independently of $(\lambda_j, \sigma, \alpha_j)$ 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 then estimate the equation:

$$\ln\left(\frac{S_{aj\tau}}{S_{bj\tau}}\right) = \alpha_j + \beta \ln\left(\frac{P_{aj\tau}}{P_{bj\tau}}\right) + \varepsilon_{kj\tau}. \tag{7}$$

[15] The proof of the Proposition in Appendix I establishes that it is always possible to find authority specific λ_j and common σ compatible with the estimated seats-vote relationship. Though it is not necessary to have common σ making the assumption removes an issue of coefficient heterogeneity in the policy regressions.

The estimated parameters are then used to construct $f(\alpha_j, \beta)$. This is used in the next section to explain economic policy choices.

Our data come from England where there are four main party groupings – Conservatives, Labour, Liberal Democrats and Independents. As we discussed above, we confine our estimation to 150 local authorities that exhibit "essential" two party competition over our data period. This gives 150 bias parameters α_j and a swing parameter β. In the smaller sub-set of Conservative-Labour competition we estimate only 108 bias parameters α_j. The relationship between votes and seats is illustrated for these two cases in Figures 1 and 2. In each case, we also plot the regression line. These plots suggest a fairly good fit for the log-odds relationship between seats and votes.

Results from estimating (7) 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 1.72 in the larger sample and 2.21 in the smaller one.

Using the results from (7), we recover the α_j parameter for each of our local authorities.[16] A Wald test strongly rejects the hypothesis that the values of α_j are equal across local authorities in either sample. We 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:

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

For sake of illustration, Figure 3 gives the distribution of (average) bias 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

---

[16]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.

larger party tends to have a large advantage over its second placed rival.

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:

$$
\begin{aligned}
y_{r\ell jt} &= \xi_{rj} + \xi_{\ell j} + \xi_{\ell t} + \rho_{\ell r} \left(2\delta_{rjt} - 1\right) f\left(\alpha_j, \beta\right) + \delta_{rjt} \\
&\quad + \gamma_\ell x_{\ell jt} + \eta_{r\ell jt}
\end{aligned}
\tag{8}
$$

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.

On the whole swing voters might be expected to have more centrist 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 will be true if parties have different policy preferences.[17]

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

$$
\kappa_j = \lambda_j + \frac{1}{1-\sigma} f\left(\alpha_j, \beta\right).
$$

This implies that the error term in (8) contains:

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

where λ_j is treated as a random effect. We will also allow the distribution of $\varphi_{r\ell jt}$ to be heteroskedastic.[18]

---

[17]This is clear from the example solved in Appendix III.

[18]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.

The actual estimation is laid out in detail in Appendix II. It shows how we deal with two further 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 (7). Second, since there are four parties rather than two, we estimate these policy equations for different (a, b) pairs in different jurisdictions.[19] Moreover, since we are studying competition between the main two parties in every jurisdiction, we need to introduce a third possibility (no overall control) 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.[20] 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

---

[19] To check consistency, we will look at the smaller set of jurisdictions which have only Labour-Conservative competition.

[20] The omitted category here is no overall control.

the random effects estimation in column (5) which deals with the full set of issues discussed in the context of (8). 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 expenditure increases only in Labour controlled councils.[21]

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 increase 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 results in (8). Table 5 concentrates on taxes. These data come from the audit commission and hence are a smaller sample of years than Table 4. The first column is for the size of the local tax. Council taxes are interesting as they are set locally and highly visible. The tax varies by property value. We use the tax on a standardized property (*Band D*). Column (1) shows that there is a positive and significant effect of bias on the size of the 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, as measured on a consistent basis by the Audit commission. The

---

[21]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. 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.

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 authority.

Having observed in Table 4 that total expenditures are affected by bias, Table 6 looks at the composition of expenditures. 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 refuse spending it appears that bias under Liberal Democratic control increases spending while they spend less on Transport. As expected, these results indicate that parties do have differing policy priorities.

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. 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 be unattractive to Conservative voters. 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 while it increases rent 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 suggest convincing evidence that bias matters for policy outcomes. Consistent with the theory, the results also suggest that parties try to appeal more to swing voters by tempering their "true" party preference when elections are more competitive (districting 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

Political economy models try to identify how political incentives affect policy choices. Despite a plethora of theoretical models, there are few efforts to build links to empirical estimation. This paper suggests an approach to doing so and identifies the consequences of districting bias for policy choice. In doing so it provides a link between three branches of previously distinct research – empirical models of voter behavior, the relationship between seats and votes and strategic models of politics. The core idea in the theory is that bias in favor of one party will make that party less keen to court swing voters and will therefore slant policies more towards its core supporters.

To illustrate the usefulness of the approach, we have 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 face a stricter electoral incentive. Finding other ways of linking theory and data in political economy models provides a fruitful and challenging agenda for further research.

The results presented here also contribute to debates about the consequences of electoral districting. Beyond the broad choice of electoral rules, the pattern of districting is one of the key choices in an electoral system. It is no surprise, therefore, that it generates significant discussion. This issue has largely been left alone by the new political economy literature. However, Coate and Knight (2005a,b) is taking up the challenging problem of optimal districting building from micro-foundations. The findings in this paper suggest that understanding the policy consequences fully requires considering the consequences of districting for strategies employed by political parties.

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] Coate, Stephen and Brian Knight, [2005a], "Socially Optimal Districting" unpublished typescript.

[9] Coate, Stephen and Brian Knight, [2005b], "Socially Optimal Districting: An Empirical Investigation" unpublished typescript.

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

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

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

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

[14] Gelman, Andrew and Gary King, [1994], "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] Myerson, Roger B., [1993], "Incentives to Cultivate Favored Minorities under Alternative Electoral Systems," *American Political Science Review*, 87(4): 856-69.

[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 $\sigma \in [0, 1]$ and $\lambda \in [-1, 1]$ such that if*

$$G(\mu, \lambda) = 1 - \phi^{-1}\left(\alpha + \beta\psi\left(\frac{1}{2} + \frac{1}{2}(1 - \sigma)(\lambda - \mu)\right)\right) \tag{A.1}$$

for $\mu \in [\lambda - \frac{1}{1-\sigma}, \lambda + \frac{1}{1-\sigma}] \cap [-1, 1]$ then

$$\phi(S_a) = \alpha + \beta\psi(P_a) \tag{A.2}$$

for $P_a \in [\frac{1}{2}(1 - \sigma)(1 + \lambda), \sigma + \frac{1}{2}(1 - \sigma)(1 + \lambda)]$.

Proof. From (2) the widest possible range for P_a as $F(-\eta - [v_a - v_b]])$ ranges from 0 to 1 is

$$\frac{1}{2}(1 - \sigma)(1 + \lambda) \leq P_a \leq \sigma + \frac{1}{2}(1 - \sigma)(1 + \lambda).$$

Suppose that (A.2) holds or equivalently

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

for $P_a \in [\frac{1}{2}(1 - \sigma)(1 + \lambda), \sigma + \frac{1}{2}(1 - \sigma)(1 + \lambda)]$. This corresponds to (3)

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

and hence (A.1) holds for $\mu \in (-\frac{\sigma}{1-\sigma}, \frac{\sigma}{1-\sigma})$. Since, by its definition, $\mu \in [-1, 1]$ then (A.2) can hold across the required range of values for P_a only if $\sigma \leq \frac{1}{2}$.

We need to show that there exists a $\sigma \in (0, \frac{1}{2}]$ and a $\lambda \in [-1, 1]$ such that λ can be the mean value of μ. Suppose that $G(\cdot)$ takes the form (A.1) not only on $(-\frac{\sigma}{1-\sigma}, \frac{\sigma}{1-\sigma})$ but across the widest possible support $[\mu_0, \mu_1] \equiv [\lambda - \frac{1}{1-\sigma}, \lambda + \frac{1}{1-\sigma}] \cap [-1, 1]$. The difference between the mean of μ and λ is

$$\Delta(\lambda) \equiv \mu_0 G(\mu_0, \lambda) + \mu_1[1 - G(\mu_1, \lambda)] + \int_{G(\mu_0, \lambda)}^{G(\mu_1, \lambda)} G^{-1}(\pi, \lambda)d\pi - \lambda.$$

Let $\sigma = \frac{1}{2}$. Then $[\mu_0, \mu_1] = [-1, 1]$ if $\lambda \in [-1, 1]$. Hence

$$\Delta(\lambda) \equiv 1 - G(-1, \lambda) - G(1, \lambda) + \int_{G(-1, \lambda)}^{G(1, \lambda)} G^{-1}(\pi, \lambda)d\pi - \lambda.$$

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

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

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

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

and

$$
\begin{aligned}
\Delta(1) &= 1 - [1 - \pi_0] + \int_0^{1-\pi_0} G^{-1}(\pi, \lambda)\mathrm{d}\pi - 1 \\
&= \int_0^{1-\pi_0} G^{-1}(\pi, \lambda)\mathrm{d}\pi + \pi_0 - 1 \\
&\leq (1 - \pi_0) + \pi_0 - 1 = 0
\end{aligned}
$$

using the fact that $0 \leq G^{-1}(\pi, \lambda) \leq 1$ for all π.

Since $\Delta(\lambda)$ is a continuous function there must therefore be a $\lambda \in [-1, 1]$ such that $\Delta(\lambda) = 0$ as required. $\qquad\square$

Letting $\phi(y) = \psi(y) = \ln\left(\frac{y}{1-y}\right)$ in these formulae gives the expression for $G(\mu, \lambda)$ in the Proposition and the expression for m_G follows by inversion.

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

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

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

26

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 $N, T_1 \to \infty$,

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

27

where

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

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):

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

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

$$
\begin{pmatrix} \hat{\rho}_G \\ \hat{\gamma}_G \end{pmatrix} = (Z_2'\hat{\Omega}^{-1}Z_2)^{-1}Z_2'\hat{\Omega}^{-1}y
$$

$$
= \begin{pmatrix} \rho \\ \gamma \end{pmatrix} + (Z_2'\hat{\Omega}^{-1}Z_2)^{-1}Z_2'\hat{\Omega}^{-1}[\eta - \rho W\nabla(Z_1'Z_1)^{-1}Z_1'u]
$$

$$
+ o_p(T_1^{-\frac{1}{2}}) \tag{A.3}
$$

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

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

assuming again that the appropriate probability limits exist.

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

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

where

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

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.3)

$$
\begin{aligned}
E & \left[\left(\begin{array}{c} \hat{\rho}_G \\ \hat{\gamma}_G \end{array} \right) - \left(\begin{array}{c} \rho \\ \gamma \end{array} \right) \right] \\
&= E(Z_2'\hat{\Omega}^{-1}Z_2)^{-1}Z_2'\hat{\Omega}^{-1}\left[\eta - \rho W \nabla (Z_1'Z_1)^{-1}Z_1'u \right] \\
&\simeq \rho(Z_2'\hat{\Omega}^{-1}Z_2)^{-1}\left(\begin{array}{c} -E\hat{f}'W'\hat{\Omega}^{-1}W\nabla(Z_1'Z_1)^{-1}Z_1'u \\ 0 \end{array} \right) \\
&= -\rho(Z_2'\hat{\Omega}^{-1}Z_2)^{-1}\left(\begin{array}{c} Eu'Z_1(Z_1'Z_1)^{-1}\nabla'W'\hat{\Omega}^{-1}W\nabla(Z_1'Z_1)^{-1}Z_1'u \\ 0 \end{array} \right) \\
&\simeq -\rho(Z_2'\hat{\Omega}^{-1}Z_2)^{-1}\left(\begin{array}{c} \mathrm{tr}\left(Z_1(Z_1'Z_1)^{-1}\nabla'W'\hat{\Omega}^{-1}W\nabla(Z_1'Z_1)^{-1}Z_1' \right)E(uu') \\ 0 \end{array} \right)
\end{aligned}
$$
$$\tag{A.4}$$

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.4).

Appendix III Political Equilibrium: A Linear Example

Let:

$$\hat{V}(v_a, a) = \Gamma_a - \gamma_{aa} v_a$$
$$and$$
$$\underline{V}(v_b, b) = \Gamma_a - \gamma_{ab} v_b$$

and

$$\hat{V}(v_b, b) = \Gamma_b - \gamma_{bb} v_b$$
$$and$$
$$\underline{V}(v_a, a) = \Gamma_b - \gamma_{ba} v_a.$$

We assume that $\gamma_{aa} > 0$ and $\gamma_{bb} > 0$. The counterpart of Assumption two is $\gamma_{aa} > -\gamma_{ab}$ and $\gamma_{bb} > -\gamma_{ba}$. Then, the first order conditions can be solved (at an interior solution) as:

$$v_a = \Delta_a - \frac{\kappa}{2} + \rho_a v_b$$

and

$$v_b = \Delta_b + \frac{\kappa}{2} + \rho_b v_a$$

where $\rho_a = \frac{1}{2}\left(1 + \frac{\gamma_{ab}}{\gamma_{aa}}\right)$, $\rho_b = \frac{1}{2}\left(1 + \frac{\gamma_{ba}}{\gamma_{bb}}\right)$ and $\Delta_\rho = \frac{\Omega_\rho}{\gamma_{\rho\rho}} - \frac{1}{2\xi}$. Thus, the reaction functions are linear with slope coefficient ρ_a and ρ_b respectively. Then:

$$\tilde{v}_a = \frac{\Delta_a + \rho_a \Delta_b}{(1 - \rho_a \rho_b)} - \kappa \frac{1}{2}\left(\frac{1 - \rho_a}{1 - \rho_a \rho_b}\right)$$

and

$$\tilde{v}_b = \frac{\Delta_b + \rho_b \Delta_a}{(1 - \rho_a \rho_b)} + \kappa \frac{1}{2}\left(\frac{1 - \rho_b}{1 - \rho_a \rho_b}\right).$$

where $1 - \rho_a \rho_b > 0$.

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 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**

	All	Lab-Con
Variable	Coeff	Coeff
ln odds Vote Share	1.721	2.209
	(0.075)	(0.071)
Sample size	1004	747
R^2	0.859	0.909
Standard error	0.344	0.175
W_{Auth}	896.498	570.023
p value	0.000	0.000

Dependent variable is ln odds Seat Share
Standard errors in parentheses are robust to heteroscedasticity.
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	0.360	0.121	0.081			
	(0.060)	(0.041)	(0.045)			
Pro inc bias * Lab				0.158	0.565	0.528
				(0.054)	(0.132)	(0.201)
Pro inc bias * Con				-0.157	-0.389	0.034
				(0.105)	(0.141)	(0.257)
Pro inc bias * Lib Dem				0.386	0.616	
				(0.159)	(0.192)	
Pro inc bias * Oth				0.072	-0.226	
				(0.094)	(0.208)	
Political control						
Labour			0.089	0.084	0.012	0.035
			(0.016)	(0.016)	(0.018)	(0.020)
Conservative			-0.116	-0.101	-0.034	-0.039
			(0.021)	(0.023)	(0.019)	(0.022)
Liberal Democrat			0.148	0.131	0.059	
			(0.027)	(0.029)	(0.031)	
Other			0.007	0.011	-0.044	
			(0.027)	(0.029)	(0.049)	
Characteristics						
ln Average Income		0.048	-0.038	-0.003	0.243	0.374
		(0.097)	(0.091)	(0.092)	(0.108)	(0.141)
ln Population		-0.314	-0.204	-0.195	-0.091	-0.103
		(0.024)	(0.027)	(0.027)	(0.040)	(0.048)
ln Unemployment		0.401	0.310	0.306	0.248	0.266
		(0.017)	(0.018)	(0.018)	(0.021)	(0.026)
ln Per Capita Grant		0.056	0.063	0.061	0.005	-0.002
		(0.010)	(0.009)	(0.009)	(0.008)	(0.009)
Sample size	2133	2133	2133	2133	2133	1535
No of years	14	14	14	14	14	14
θ_1					0.134	0.141
θ_2	0.079	0.048	0.043	0.043	0.164	0.168
W_{Class}	19892.359	4834.774	5406.033	5254.296	1453.682	1024.631
p value	0.000	0.000	0.000	0.000	0.000	0.000
W_{Region}	136.507	108.652	114.649	119.543	32.123	22.045
p value	0.000	0.000	0.000	0.000	0.000	0.002
W_{Year}	1574.072	335.368	250.922	237.400	188.986	112.280
p value	0.000	0.000	0.000	0.000	0.000	0.000

Dependent variable in all regressions is log of total expenditure

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_3^2

W_{Region} is a Wald test of absence of regional differences, asymptotically distributed as χ_8^2

W_{Year} is a Wald test of absence of year differences, asymptotically distributed as χ^2 with degrees of freedom one less than the no 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	0.204	3.134	-0.046
	(0.090)	(3.447)	(0.023)
Pro inc bias * Con	0.089	-2.307	-0.162
	(0.140)	(6.226)	(0.055)
Pro inc bias * Lib Dem	0.126	-2.919	-0.008
	(0.108)	(4.763)	(0.029)
Pro inc bias * Oth	-0.264	13.117	-0.010
	(0.184)	(7.039)	(0.040)
Political control			
Labour	-0.012	0.731	-0.004
	(0.012)	(0.620)	(0.004)
Conservative	-0.048	-0.992	0.022
	(0.016)	(0.890)	(0.007)
Liberal Democrat	0.037	0.149	0.003
	(0.021)	(0.991)	(0.006)
Other	-0.026	2.066	-0.008
	(0.046)	(1.853)	(0.011)
Characteristics			
ln Average Income	0.024	0.082	-0.009
	(0.083)	(4.236)	(0.027)
ln Population	-0.027	-0.158	-0.020
	(0.033)	(1.408)	(0.009)
ln Unemployment	0.075	-1.359	-0.003
	(0.021)	(1.019)	(0.006)
ln Per Capita Grant	-0.018	1.428	-0.017
	(0.013)	(0.859)	(0.005)
Sample size	720	569	569
No of years	4	3	3
θ_1	0.105	3.510	0.020
θ_2	0.059	2.224	0.015
W_{Class}	26.320	34.514	65.166
p value	0.000	0.000	0.000
W_{Region}	38.791	8.292	6.999
p value	0.000	0.308	0.429
W_{Year}	261.570	12.051	28.217
p value	0.000	0.007	0.000

Standard errors in parentheses are robust to estimation error in the districting advantage variable.

All estimates by GLS

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_3^2

W_{Region} is a Wald test of absence of regional differences, asymptotically distributed as χ_8^2

W_{Year} is a Wald test of absence of year differences, asymptotically distributed as χ^2 with degrees of freedom one less than the no of years

Table 6: **Expenditure Regressions**

Variable	Leisure (1) Coeff	Parks (2) Coeff	Refuse (3) Coeff	Transport (4) Coeff
Pro inc bias * Lab	0.892	1.716	-0.032	0.775
	(0.322)	(0.331)	(0.145)	(0.619)
Pro inc bias * Con	-1.406	-0.641	-0.232	0.976
	(0.373)	(0.307)	(0.157)	(0.690)
Pro inc bias * Lib Dem	1.257	-0.057	0.649	-1.807
	(0.427)	(0.367)	(0.217)	(0.913)
Pro inc bias * Oth	0.056	-1.752	-0.245	-0.966
	(0.521)	(0.783)	(0.246)	(1.222)
Political control				
Labour	0.127	0.042	0.035	0.004
	(0.041)	(0.036)	(0.019)	(0.087)
Conservative	0.027	-0.043	-0.039	-0.004
	(0.045)	(0.037)	(0.021)	(0.092)
Liberal Democrat	0.018	0.007	0.091	-0.126
	(0.068)	(0.056)	(0.034)	(0.147)
Other	-0.152	-0.521	-0.110	-0.516
	(0.117)	(0.186)	(0.057)	(0.284)
Characteristics				
ln Average Income	-0.072	0.285	0.031	-1.170
	(0.252)	(0.215)	(0.121)	(0.569)
ln Population	0.157	0.841	0.069	-0.176
	(0.104)	(0.104)	(0.047)	(0.191)
ln Unemployment	0.137	-0.129	-0.071	0.049
	(0.049)	(0.042)	(0.023)	(0.104)
ln Per Capita Grant	-0.026	0.007	0.001	-0.055
	(0.017)	(0.014)	(0.008)	(0.038)
Sample size	2118	2131	2132	1910
No of years	14	14	14	14
θ_1	0.398	0.463	0.168	0.598
θ_2	0.363	0.322	0.177	0.781
W_{Class}	13.619	8.307	59.713	175.619
p value	0.003	0.040	0.000	0.000
W_{Region}	23.668	43.454	9.530	35.776
p value	0.001	0.000	0.217	0.000
W_{Year}	77.978	30.718	46.889	86.686
p value	0.000	0.006	0.000	0.000

Dependent variable in all regressions is log expenditure
Standard errors in parentheses are robust to estimation error in the districting advantage variable.
All estimates by GLS
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_3^2
W_{Region} is a Wald test of absence of regional differences, asymptotically distributed as χ_8^2
W_{Year} is a Wald test of absence of year differences, asymptotically distributed as χ^2 with degrees of freedom one less than the no 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	1.269	1.340	1.026
	(0.186)	(0.184)	(0.283)
Pro inc bias * Con	-0.557	-0.546	-0.733
	(0.164)	(0.156)	(0.285)
Pro inc bias * Lib Dem	0.379	0.514	-0.447
	(0.208)	(0.203)	(0.384)
Pro inc bias * Oth	-0.184	-0.108	-0.798
	(0.232)	(0.215)	(0.482)
Political control			
Labour	0.048	0.044	0.086
	(0.021)	(0.021)	(0.035)
Conservative	0.003	-0.010	0.029
	(0.021)	(0.020)	(0.036)
Liberal Democrat	0.001	-0.021	0.142
	(0.032)	(0.031)	(0.058)
Other	-0.134	-0.128	-0.247
	(0.054)	(0.049)	(0.114)
Characteristics			
ln Average Income	0.095	0.031	0.381
	(0.133)	(0.125)	(0.237)
ln Population	0.275	0.317	0.104
	(0.047)	(0.045)	(0.084)
ln Unemployment	0.018	0.012	0.005
	(0.023)	(0.022)	(0.041)
ln Per Capita Grant	0.007	0.010	0.004
	(0.008)	(0.007)	(0.014)
Sample size	1696	1696	1696
No of years	12	12	12
θ_1	0.169	0.167	0.305
θ_2	0.156	0.149	0.262
W_{Class}	564.097	369.132	530.787
p value	0.000	0.000	0.000
W_{Region}	25.097	41.809	2.813
p value	0.001	0.000	0.902
W_{Year}	77.454	127.919	17.852
p value	0.000	0.000	0.120

Dependent variable in all regressions is log employment

Standard errors in parentheses are robust to estimation error in the districting advantage variable.

All estimates by GLS

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 no of years

Table 8: **Performance Regressions**

Variable	% Rent Coll (1) Coeff	% Man costs (2) Coeff	Av rent (3) Coeff	Planning (4) Coeff	Benefi t costs (5) Coeff
Pro inc bias * Lab	0.047	-6.664	-5.544	-0.148	1.799
	(0.019)	(2.193)	(3.346)	(0.087)	(18.712)
Pro inc bias * Con	0.062	3.701	9.736	0.151	-12.820
	(0.047)	(3.760)	(5.804)	(0.170)	(33.028)
Pro inc bias * Lib Dem	0.014	-1.393	-5.308	0.051	19.155
	(0.027)	(2.614)	(4.148)	(0.127)	(25.069)
Pro inc bias * Oth	-0.018	1.056	3.528	0.040	18.171
	(0.039)	(4.840)	(7.959)	(0.163)	(37.650)
Political control					
Labour	0.010	-0.217	-0.310	0.004	0.213
	(0.004)	(0.330)	(0.485)	(0.018)	(3.221)
Conservative	0.006	0.073	0.564	0.000	1.224
	(0.006)	(0.487)	(0.721)	(0.026)	(4.598)
Liberal Democrat	-0.001	-0.156	0.829	0.013	-1.420
	(0.006)	(0.596)	(0.940)	(0.026)	(5.220)
Other	-0.005	0.534	-0.066	-0.012	4.149
	(0.011)	(1.227)	(1.972)	(0.045)	(9.667)
Characteristics					
ln Average Income	0.004	1.851	3.763	0.186	8.483
	(0.026)	(2.376)	(3.544)	(0.115)	(22.387)
ln Population	0.015	-0.777	0.837	-0.134	10.581
	(0.008)	(0.839)	(1.299)	(0.037)	(7.561)
ln Unemployment	-0.015	0.715	-0.694	0.086	-7.681
	(0.006)	(0.595)	(0.896)	(0.028)	(5.381)
ln Per Capita Grant	0.006	1.063	1.193	-0.022	4.907
	(0.005)	(0.511)	(0.780)	(0.023)	(4.559)
Sample size	516	517	517	569	567
No of years	3	3	3	3	3
θ_1	0.017	2.278	4.111	0.080	19.957
θ_2	0.013	1.022	1.613	0.068	11.456
W_{Class}	2.877	86.741	127.023	11.505	58.584
p value	0.411	0.000	0.000	0.009	0.000
W_{Region}	8.109	13.353	89.527	12.128	5.739
p value	0.323	0.064	0.000	0.096	0.571
W_{Year}	12.336	34.990	206.185	1.100	1.508
p value	0.006	0.000	0.000	0.777	0.681

Standard errors in parentheses are robust to estimation error in the districting advantage variable.

All estimates by GLS

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_3^2

W_{Region} is a Wald test of absence of regional differences, asymptotically distributed as χ_8^2

W_{Year} is a Wald test of absence of year differences, asymptotically distributed as χ^2 with degrees of freedom one less than the no of years
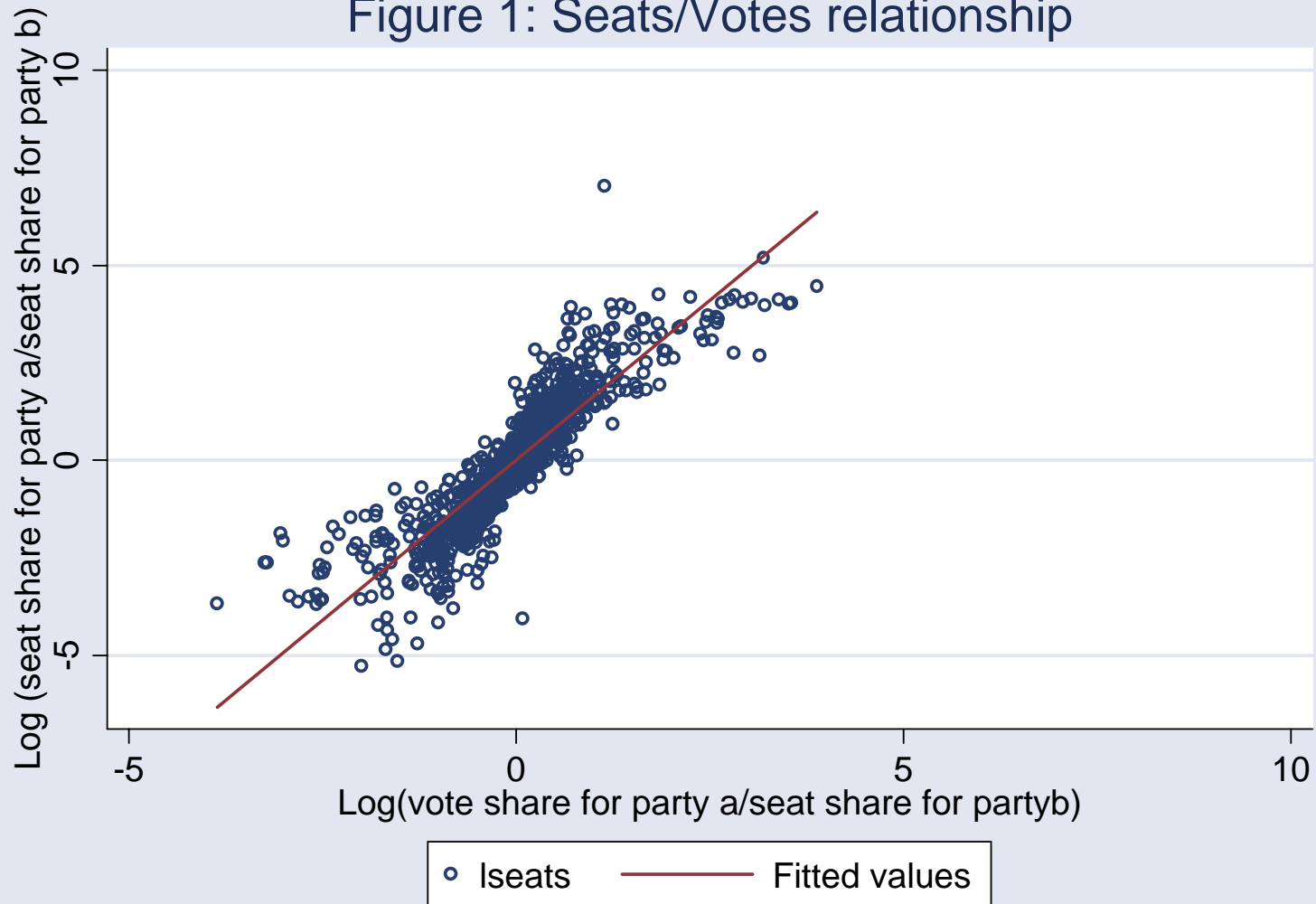
Figure 1: Seats/Votes relationship

Figure 2: Seats/Votes relationship for Con-Lab Sample

Figure 3: Distribution of Electoral Bias