

The Structure of Utility in Spatial Models of Voting

Royce Carroll
Rice

Jeffrey B. Lewis
UCLA

James Lo
University of Mannheim

Keith T. Poole
University of Georgia

Howard Rosenthal
NYU*

March 13, 2013

Abstract

Empirical models of spatial voting allow legislators' locations in a policy or ideological space to be inferred from their roll call votes. These are typically random utility models where the features of the utility functions other than the ideal points are assumed rather than estimated. In this paper, we first consider a model in which legislators' utility functions are allowed to be a mixture of the two most commonly assumed utility functions: the quadratic function and the Gaussian function assumed by NOMINATE. Across many roll call data sets, we find that legislators' utility functions are estimated to be very nearly Gaussian. We then relax the usual assumption that each legislator is equally sensitive to policy change and find that extreme legislators are generally more sensitive to policy change than their more centrally located counterparts. This result suggests that extremists are more often ideologically rigid while moderates are more likely to consider influences that arise outside liberal-conservative conflict.

*This paper was supported by NSF Grant SES-0611974. James Lo also acknowledges support for this project by SFB 884 on the "Political Economy of Reforms" at the University of Mannheim (project C4), funded by the German Research Foundation (DFG).

1 Introduction

Over the past twenty-five years, the study of Congress has increasingly involved the analysis of roll call voting data. Empirical models of spatial voting, often referred to as ideal point estimators, allow legislator locations in an abstract policy or ideological space to be inferred from their roll call votes. These models have provided new insights about the US Congress in particular and legislative behavior more generally (see, for example, Poole and Rosenthal, 1985, 1991, 1997; McCarty et al., 1997, 2006; Poole, 2000, 2005; Heckman and Snyder, 1997; Myagkov and Kiewiet, 1996; Jenkins, 1999; Rosenthal and Voeten, 2004; Clinton et al., 2004; Desposato, 2006; Hix et al., 2007; Schonhardt-Bailey, 2003; Shor et al., 2010). Recently ideal point models have also been applied to voting in non-legislative voting bodies such as the United Nations (Voeten, 2000) and the EU Council of Ministers (Hagemann, 2007), elections (for example, Herron and Lewis, 2007; Jessee, 2009), business settings (Ichniowski et al., 1997), and courts (for example, Martin and Quinn, 2002). There are now a number of alternative models, estimators, and software that researchers can use to recover a latent issue or ideological space from voting data. These approaches are often tailored to particular problems, such as voting in small chambers (Londregan, 2000; Peress, 2009), measuring dynamics (Martin and Quinn, 2002), or application to very large data sets (Lewis, 2001). While these models have many features in common, they also differ in some basic assumptions about exactly how the spatial locations of alternatives are translated into choices. These assumptions are not simply of technical significance, but imply substantively different notions of actor behavior which in turn could have consequences for how legislative institutions function.

In this paper, we explore if these assumptions can be relaxed. That is, rather than assuming a particular utility function for our actors, can we estimate important features of the actors' utility function? It is known that there are limits to how far we can relax our assumptions and still identify legislators' ideal points from their roll call votes (Kalandrakis,

2010). Moreover, it is not obvious that the data to which these estimators are typically applied is sufficiently rich to pin down those features of utility functions that are identified (in the econometric sense). We find that, in fact, some important features of voter utility that have been fixed by assumption in nearly all of the previous literature can be estimated. However, we also find that while these features have important implications for how actors translate the recovered issue space into choices over particular pairs of alternatives, they have relatively little effect on the recovered ideal points.

Nearly all ideal point estimators employ the random utility framework of McFadden (1973). Accordingly, an actor’s choice between two alternatives (*yea* and *nay*) is governed by a systematic spatial voting component and an independent additive random utility shock applied to each of the two alternatives. Generally, simple Euclidean spatial preferences are assumed (Enelow and Hinich, 1984; Hinich and Munger, 1994, 1997). That is, actors are assumed to be more likely to choose the alternative that is located closer to their ideal policy than the alternative that is located farther from their ideal policy. The general form of the random utility function is:

$$U_j(O_j; X) = F(\|O_j - X\|) + \epsilon_j \tag{1}$$

where X is the actor’s most preferred policy outcome in some d -dimensional policy space, O_j for $j \in \{y, n\}$ is the location of the policy outcome associated with the (y)ea or (n)ay alternatives in the same policy space, F is a given monotonically decreasing function, $\|\cdot\|$ denotes Euclidean distance, and ϵ_j is an alternative-specific utility shock. Note that actors do not always choose the closer of the two alternatives because the presence of the non-spatial shocks can reverse the preference for one alternative for another that is implied by the spatial component of the utility function.

As described in greater detail in section 2.2, the (*ex ante*) probability that the spatial preferences will be reversed by the idiosyncratic shocks is a function of three factors: the variability of difference in the utility shocks ($\epsilon_y - \epsilon_n$), the distance between the actor’s ideal

point (X) and each of the alternative (O_j), and the utility function (F). Although often overlooked, the choice of F has important implications for choice behavior. For example, if F is a concave function (e.g., $F : x \mapsto -x^2$), then, holding O_y and O_n fixed, the likelihood of an actor voting for the farther away alternative goes to zero as X moves away from both O_y and O_n . On the other hand, if F has convex tails (e.g., $F : x \mapsto \exp(-\frac{1}{2}x^2)$), then the probability of choosing the further away alternative goes to 1/2 as X is moved away from both O_y and O_n . As we discuss in greater detail in section 2, this difference has implications for how actors respond to different alternatives.

Typically, the central objective in fitting empirical models of spatial voting is to estimate the ideal point (X) of each actor.¹ In this paper, we focus on the estimation of features of F . In what follows, we assume that the underlying issue space is unidimensional. As we will develop further in Section 2, because we are simultaneously estimating features of the alternatives, the ideal points, and features of F , identification in both the strict econometric sense and in the sense that the estimation leans heavily on strong distributional assumptions presents a serious challenge. Indeed, a central object of this research is to probe the limit of what it is possible to learn about F even conditional on the usual (strong) assumptions about the distributions of the idiosyncratic shocks.

We begin this exploration in the context of the one-dimensional model not only because it is the most commonly employed in the literature and because it has been shown to capture a great deal of the empirical structure of roll call voting matrices in many contexts, but also because we have the best chance of learning about F under the assumption that there is a dominant single spatial dimension. How literally one wishes to take our one dimensional plus random shocks approximation to what is undoubtedly a more complex world is certainly debatable. We follow nearly every random utility model employed since McFadden (1974) in assuming that the systematic (this case spatial) utility and the stochastic utility shocks are additively separable and that the shocks are independent of the systematic utility and have

¹In general, the location of each (O_y, O_n) pair cannot be identified without additional assumptions. However, ideal point estimators provide estimates of vote-specific parameters that are functions of (O_y, O_n).

a simple parametric form (usually normal or logistic). All previous models of ideal point estimation make these same assumptions and the locations of the ideal point in an Euclidean space cannot be identified without assuming the existence of and placing structure on the distribution of utility shocks (see Poole 2005).

When we write about the features of F that we estimate as elements of a legislator's spatial utility function, there is obviously an "as if" quality to the description. If we allowed for dependence between the systematic features of the utility functions and the idiosyncratic shocks as might arise due to vote buying, the projection of a higher dimensional space onto our single dimension, and so forth, we could arrive at different conclusions about the shape of F or about how that shape varied across legislators even though the way in which legislators' ideal points are translated into choices (the likelihood of the data) could be little changed. Indeed, we would very quickly lose identification on the ideal points (as would other ideal point estimators) if we remove too much of the assumed structure. However, our characterization of F does allow us to parameterize how the probability of choosing one alternative over another varies as those alternatives or the legislator's ideal point are varied. In this sense, the shape of F in combination with assumptions made about the utility shocks can be thought of as simply providing a way to approximate the frequency with which legislators with different positions in the one-dimensional space will choose various alternatives. From this perspective, what we are asking is not what is the shape of F (which depends strongly on how the idiosyncratic shocks are motivated and defined), but: How do spatial locations map (probabilistically) into choices? If we summarize legislators by their ideal points in a single dimension, what function best maps those positions into observed choices? And, does that choice function vary in important ways across legislators? While we could present our inquiry entirely in terms of probabilistic choice functions approximated by our utility models, for convenience, we will describe our problem in terms of estimating the shape of the spatial utility function (F) as if legislators have such functions and are truly following our simple random utility model.

First, we consider the overall shape of the actor’s spatial utility function, $F(\|O_j - X\|)$. We consider a model that is parameterized such that the quadratic utility function used in estimators such as Jackman’s IDEAL (2004), Martin and Quinn (2002), and many others, is nested in the Gaussian utility model used in Poole and Rosenthal’s NOMINATE (1985). This model allows us to investigate whether voting data can be used to discriminate between these leading assumptions about the shape of spatial utility functions. Perhaps surprisingly, we find that in voting data from bodies as small as the U.S. Senate it is possible to discriminate between these two utility functions. Based on this model, presented in section 2, we estimate that legislators’ utility functions are very nearly Gaussian throughout almost the entire history of the U.S. Congress.

Second, we consider whether F might vary across legislators. In particular, we estimate a model in which legislator utility has the same basic Gaussian shape (as described below), but legislators vary in the overall intensity of their spatial preferences. In the econometric sense, more intense preferences are associated with utility functions with lower variance (i.e. “skinnier” Gaussian functions). In terms of legislator behavior, actors having more intense spatial preferences are relatively more likely to select the closer alternative than are actors with less intense spatial preferences. Consistent with theoretical expectations that policy extremists are more sensitive to policy outcomes than moderates, we find evidence that extremist legislators have higher-intensity utility functions than their moderate counterparts.

Our methodology for this project employs a Markov Chain Monte Carlo (MCMC)-based version of Poole and Rosenthal’s NOMINATE model that allows for easy calculation of auxiliary quantities of interest and measures of estimation uncertainty for all estimated quantities. As noted above, we modify the model to estimate additional parameters that allow for variation in the shape and distribution of the utility function.

We begin our paper by developing a model that nests Quadratic and Gaussian utility. We provide a substantive motivation of how the proposed change in the utility function can affect choice and then proceed to discuss our methodology and results. Next, we discuss a

separate model that allows the intensity of each legislator’s utility function to vary. After presenting a stylized example motivating the importance of the problem, we test the theory that legislators with extreme preferences (ideal points) will also have more intense preferences on four recent U.S. Senates and the U.S. Supreme Court. We conclude with a motivating example and discussion of potential directions for further research.

2 Gaussian vs. Quadratic Utility

2.1 Introduction

What does it mean for utility functions to be Gaussian or quadratic? Figure 1 plots a pair of corresponding spatial utility functions as assumed by NOMINATE’s Gaussian utility model and IDEAL’s quadratic utility model.² The functions correspond to each other in the sense that they both imply Euclidean preferences with the same ideal point. The curves have been further harmonized to yield similar utility levels for outcome locations in the neighborhood of the ideal point. First, note that the two functions are very similar in the region where both functions are concave. In fact, the quadratic utility function is the first-order exponential approximation of the Gaussian utility function, a relationship that we exploit in our test of utility functions.

The key differences between the two distributions are seen in the tails of the plotted curves. In the tails, the marginal loss in utility is decreasing under Gaussian utility, while it is increasing at an increasing rate under quadratic utility. Thus, under quadratic utility, legislators are increasingly more disposed to support the closer alternative the farther away both the bill and status quo are from their ideal points. In contrast, the convex nature of the tails in the Gaussian utility function implies that as the bill and status quo are moved sufficiently far from the legislator’s ideal points, the utility differences between the bill and status quo are decreasing.

²The utility function employed by NOMINATE is referred to as Gaussian or normal because it has the same shape as the normal or Gaussian probability density function.

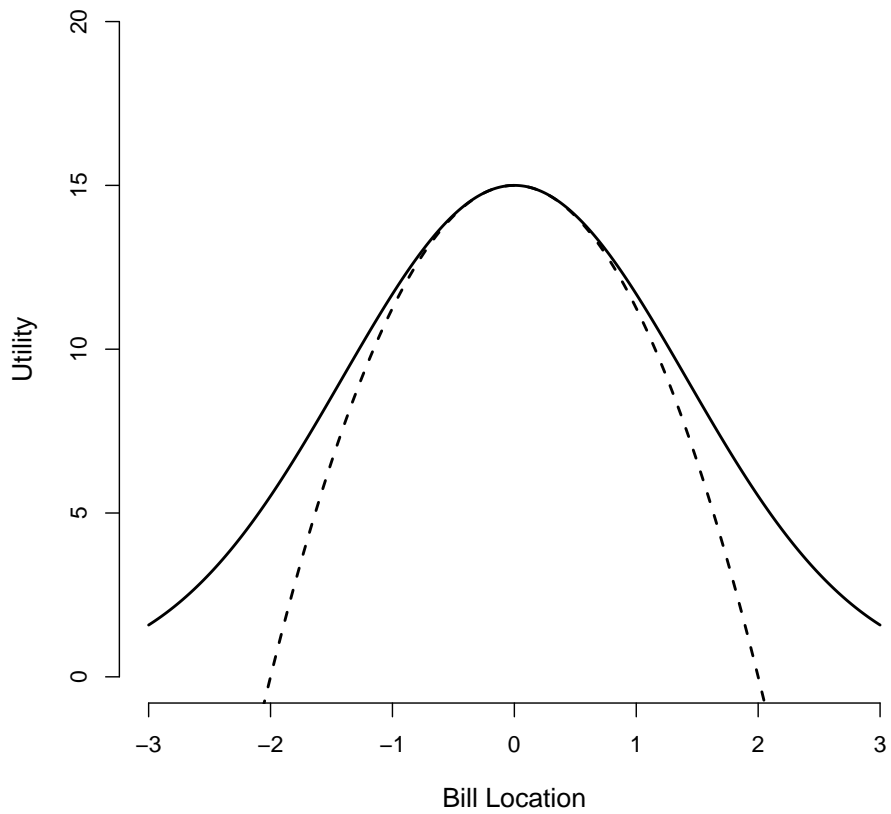


Figure 1: NOMINATE and IDEAL utility functions. *Lines show the deterministic utility functions assumed by NOMINATE (solid line) and IDEAL (dotted line) for a legislator with an ideal point of 0.*

Stated more informally, consider the example of a legislator who is voting on a bill authorizing the construction of a number of F-22 fighters. The legislator has an ideal point of building no new fighters, while the status quo and an amendment propose the construction of 1,000 and 998 fighters respectively — numbers that are both far from the legislator’s bliss point. Gaussian utility implies the legislator will be almost indifferent between the bill and the alternative, while quadratic utility implies that the legislator will perceive an enormous difference — more than, say, if the amendment were moving from a status quo of 60 fighters to 58.

As noted in Section 1, if one does not want to take our one-dimensional random utility model literally, the probabilistic choice function that the combination of F and assumed independent stochastic choice shocks can be taken as primitive. In terms of the F-22 example, a legislator with Gaussian preferences wishing to build no new fighters, could have a greater probability of choosing to build 1,000 over 998 fighters than she does of choosing to build 60 over 58, whereas for a similar legislator with quadratic preferences, the probability of choosing 60 over 58 must exceed the probability of choosing 1,000 over 998. As shown in Figure 1, in the limit as two alternatives are moved away from the legislator’s ideal point, the probability of choosing the closer alternative grows monotonically to one under quadratic preferences, but not under Gaussian preferences. Under Gaussian preferences, in the limit as the alternatives move away from a legislator’s ideal point, her probability of choosing the closer alternative goes to one-half. Whether the non-monotonicity shown in Figure 1 is manifest over the range of alternatives available to the legislators or not, it is the voting behavior of legislators located far from the alternatives that we expect to be most affected by the choice of F .

In developing NOMINATE, Poole and Rosenthal chose the Gaussian deterministic utility function for substantive reasons. Their belief was “that political actors (were) relatively insensitive to small changes in distance from their ideal points; at somewhat greater distances, utility should change sharply; finally, at very great distances, changes in distance should

have little effect on utility” (Poole and Rosenthal, 1983). Later work in psychology provided independent support for the normal utility model (Poole, 2005). In particular, the psychological experiments of Shepard, Nosofsky, and Ennis (Shepard, 1986; Nosofsky, 1986; Ennis, 1988) that examined how people judged similarity between stimuli such as light and sound intensity, revealed that these judgments appeared to use an exponential response function (Shepard, 1987). More specifically, let A represent a distance measure between two different stimuli, where $A = 0$ if the two stimuli are identical. Shepard found that given two competing stimuli with distance A , individuals tend to report the distance e^{-kA} instead, where k is a scaling constant. When perceptual error is added to the Shepard model, the expected value of this response function becomes Gaussian — that is, e^{-kA^2} .

The Shepard–Nosofsky–Ennis model thus implies normally-distributed utility functions in spatial models of voting, because the concept of preference can be reduced to the psychological notion of comparing similarities. In spatial models of voting, legislators with ideal point X_i use that standard to judge other legislative stimuli, O_j . The distances between the ideal point and the stimuli are then perceived as $e^{-kA^2} = e^{-k(X_i - O_j)^2}$, where A is the distance measure between the ideal point and stimulus. Spatial models of voting can therefore be regarded as being consistent with the Shepard–Nosofsky–Ennis similarity model.

In this experimental work that lends empirical support to the exponential response function, the characteristics of the stimuli are observable and fixed by the experimenter. In our setting, the characteristics of the alternatives (the stimuli) are unknown latent quantities to be estimated. Because both the nature of the response function (as parameterized in our model by F) and values assigned by the stimuli must be inferred, there is considerably more freedom in our setting to fit any response function to the observed data by adjusting the values of alternatives being voted upon (as well as the locations of a legislators’ ideal points). So much so, that one might suppose that there would be very little ability to empirically discriminate between a model which fixed F to be Gaussian and one which fixed F to be quadratic. However, as we show in Section 3, it is possible to discriminate between these two

alternative models in typical empirical settings.

2.2 Estimation

In this section, we begin with an overview of the quadratic and Gaussian choice utility models in one dimension. We then describe our model, which estimates an additional parameter α that allows convex combinations of the two utility models. Let p denote the number of legislators ($i = 1, \dots, p$), and q denote the number of roll call votes ($j = 1, \dots, q$), and $l = \{y, n\}$ represent the two possible choices on each vote, yea and nay. Let legislator i 's ideal point be represented by X_i and let O_{jy} and O_{jn} represent the yea and nay locations of roll call j . Then in the quadratic utility model, the utility that legislator i derives from voting yea on roll call j is:

$$U_{ijy}^{Quad} = -(X_i - O_{jy})^2 + \epsilon_{ijy}$$

While the utility of voting nay is:

$$U_{ijn}^{Quad} = -(X_i - O_{jn})^2 + \epsilon_{ijn}$$

In the Gaussian utility model, given the signal to noise parameter β and weight w (which is fixed to $w=0.5$), the corresponding yea and nay utilities for legislator i on vote j are:

$$U_{ijy}^{Norm} = \beta \exp\left\{-\frac{1}{2}w^2(X_i - O_{jy})^2\right\} + \epsilon_{ijy}$$

$$U_{ijn}^{Norm} = \beta \exp\left\{-\frac{1}{2}w^2(X_i - O_{jn})^2\right\} + \epsilon_{ijn}$$

Note that we can take the first order exponential expansion of the utility from a yea vote under Gaussian utility as:

$$U_{ijy}^{Norm} = \beta \sum_{i=0}^{\infty} \frac{(-\frac{1}{2}w^2(X_i - O_{jy})^2)^i}{i!} + \epsilon_{ijy}$$

Hence, the quadratic utility function is a first order approximation of the Gaussian utility function.

Under both models, the difference between the two errors is assumed to have a standard normal distribution, that is, $\epsilon_{ijn} - \epsilon_{ijy} \sim N(0, 1)$. This leads to the standard probit formulation of the probability that legislator i votes yea on the j th roll call as:

$$Pr_{ijy} = Pr(U_{ijy} > U_{ijn}) = Pr(\epsilon_{ijn} - \epsilon_{ijy} < u_{ijy} - u_{ijn}) = \Phi[u_{ijy} - u_{ijn}]$$

where u_{ijy} and u_{ijn} are the deterministic components of U_{ijy} and U_{ijn} respectively. Correspondingly, the probability that legislator i votes Nay on the j th roll call is $Pr_{ijn} = \Phi[u_{ijn} - u_{ijy}] = 1 - Pr_{ijy}$.

The mixture model we estimate is similar to the two models presented here, with the exception of an additional α parameter to be estimated that is permitted to vary from 0 to 1. We take the exponential expansion of the Gaussian utility function, separate the first order approximation from the component, and estimate the following:

$$U_{ijy}^{Mix} = \beta \sum_{i=0}^1 \frac{(-\frac{1}{2}w^2(X_i - O_{jy})^2)^i}{i!} + \alpha\beta \sum_{i=2}^{\infty} \frac{(-\frac{1}{2}w^2(X_i - O_{jy})^2)^i}{i!} + \epsilon_{ijy}$$

$$U_{ijn}^{Mix} = \beta \sum_{i=0}^1 \frac{(-\frac{1}{2}w^2(X_i - O_{jn})^2)^i}{i!} + \alpha\beta \sum_{i=2}^{\infty} \frac{(-\frac{1}{2}w^2(X_i - O_{jn})^2)^i}{i!} + \epsilon_{ijn}$$

Note the close relationship between the utilities from the mixture model presented here compared to the quadratic and Gaussian utilities presented earlier. When $\alpha = 1$, the utility function of the mixture model is identical to the Gaussian model. When $\alpha = 0$, the utility function of the mixture model is identical to the quadratic model with the exception of a constant, which disappears in the vote choice probability function when the Yea and Nay

utilities are subtracted from one another. Hence, estimation of α allows for a convex combination of quadratic and Gaussian utilities to be used in scaling the data, and we set the prior accordingly as $p(\alpha) \sim \text{Uniform}(0, 1)$. α close to zero or one can be interpreted as evidence supporting the quadratic or Gaussian utility functions respectively. Intermediate values of α imply a hybrid choice function in which properties of both models are combined. Note that our prior rules out values of α outside of the unit interval. We do this because we are seeking to discriminate between the two leading models and thus treat those models as polar. We also do this because if we allow α outside of the unit interval, we are not imposing the more fundamental restriction that utility is decreasing in the distance between the legislator and the alternative.³

Given the $p \times q$ matrix of observed votes V , Bayesian inference for the legislators' ideal points, bill parameters, and auxiliary parameters proceeds by simulating the posterior density given by

$$p(\alpha, \beta, X, O|V) \propto p(V|\alpha, \beta, X, O)p(\alpha, \beta, X, O).$$

We emphasize that the parameter w is not estimated but is fixed at $1/2$ as is the case in W-NOMINATE and DW-NOMINATE. We choose a diffuse prior density, $p(\alpha, \beta, X, O) = p(\alpha)p(\beta)p(x)p(O)$, that is typical of Bayesian ideal point models. In particular, we assume that the ideal point and roll call parameters (X_i for all i , and O_{kj} for $k \in \{y, n\}$ and all j) are distributed normally and independently with mean 0. The corresponding variances are parameter specific (i.e., σ_k^2 for $k \in \{\alpha, \beta, X, O\}$) and are given uninformative χ^{-2} priors (zero prior degrees of freedom). The prior distribution of α is taken to be uniform on the $[0, 1]$ interval and the prior distribution of β is given an improper uniform prior.⁴

³We have also estimated models which only Gaussian or quadratic utility is possible. That is models in which the prior over α is Bernoulli with its parameter, π , indicating the probability that the data are generated by a Gaussian or quadratic model. The conclusions drawn from those estimation are analogous to those presented below.

⁴Computer code for drawing simulated samples from the posterior distributions of all the models described is available by request.

The likelihood is given by:

$$p(V|\alpha, \beta, X, O) \propto \prod_{i=1}^p \prod_{j=1}^q \prod_{l=1}^2 P_{ijl}^{C_{ijl}}$$

where $C_{ijl} = 1$ if choice l is the actual choice of legislator i on roll call j and is zero otherwise. Because the priors are diffuse and no restrictions are placed on any of the ideal points or the locations of any of the alternatives, we are following Shapiro (1986) in estimating an “over-parameterized” model—one in which the underlying scale is not fixed in advance. We use Gibbs sampling to draw from the posterior and then post-process the results to establish an (arbitrary) scale of our estimated issue spaces⁵. Note that because the value of α is not a function of the choice of scale, the posterior draws of α , the key quantity of interest, are not affected by the post processing of the ideal points and the bill and *status quo* locations.

2.3 Results

In this section, we present three sets of empirical results. We begin with Monte Carlo tests that validate our estimator’s ability to distinguish between Gaussian and quadratic forms of utility. These results suggest that the estimator is able to distinguish between different utility functions as expected. We apply our estimator to the 109th Senate, and find strong evidence in support of a Gaussian utility function. We then proceed to apply the estimator to all U.S. Congresses. In the vast majority of cases, we continue to find strong evidence of a Gaussian utility function consistent with the results of the 109th Senate.

Our primary means of validating the estimator is through the use of Monte Carlo simulation. We generated two separate roll call matrices with 100 legislators and 500 roll calls using Gaussian and quadratic utility functions.⁶ Recall that α values of 0 are consistent with quadratic utility while α values of 1 are consistent with Gaussian utility. We then applied the mixture estimator to the two data sets over 60,000 iterations, discarding the first 10,000

⁵See Jackman (2008a) and Hoff et al. 2002 for discussions of this general approach to establishing the scale of a latent space in similar settings

⁶For these simulations we set $O_y \sim U(0.3, 1)$, $O_n \sim U(-1, -0.3)$, and $X \sim U(-1, 1)$.

as a burn-in and thinning every 10th iteration. Posterior distributions from both data sets are shown in Figure 2. When choice data generated under the assumption of Gaussian utility is scaled, α had a posterior mean of 0.985 and a posterior standard deviation of 0.014. When the choice data were scaled, α had a posterior mean of 0.054 and a posterior standard deviation of 0.043. These results are fully consistent with the theoretical expectations of the two models, providing evidence that our estimator is able to distinguish data from two similar but distinct utility functions.⁷

We then applied the estimator to the 109th Senate (2005-07), with some additional changes designed to bias our results toward the quadratic utility model as favorably as possible. Our concern here is ensuring that the results estimated by our Markov Chain have reached convergence. We began our estimation by constraining $\alpha = 0$ throughout the estimation and generating posterior samples of both legislator coordinates and bill parameters following convergence. We then take these posterior means as the start values of our estimation, and start α at 0. The idea here is to start the estimation at the parameters that are most favorable to obtaining low values of α ; if the α parameter subsequently moves to higher levels, then we can be reasonably sure that the global maximum truly does converge at higher levels of α .

Our results for the 109th Senate are summarized in the posterior density and trace plots of α in Figure 3. The posterior plot shown for the 109th Senate is very similar to the plot shown for the Gaussian utility Monte Carlo, with a posterior mean of 0.996 and a posterior standard deviation of 0.004 for α . The traceplot suggests rapid convergence to high values of α , despite the quadratic utility-biased starting values used to begin the simulation. These high values persist throughout almost all of the samples of α drawn. It should be emphasized that while

⁷All of the analysis presented in this paper assumes a one-dimensional policy space. A possible concern is that a one-dimensional Gaussian utility function may provide a better approximation to a higher-dimensional utility function (either quadratic or Gaussian) than does a one-dimensional quadratic utility function. To test for this possibility, we conduct a second variant of this Monte Carlo test in which we generate two-dimensional voting data using quadratic utility. In this model α had a posterior mean of 0.038 and a posterior standard deviation of 0.096. Thus, the superior fit of Gaussian utility that we show later in Figure 4 for higher dimensional legislatures such as the 90th House and Senate does not appear to be an artifact of approximating a higher-dimensional policy space by a one-dimensional space.

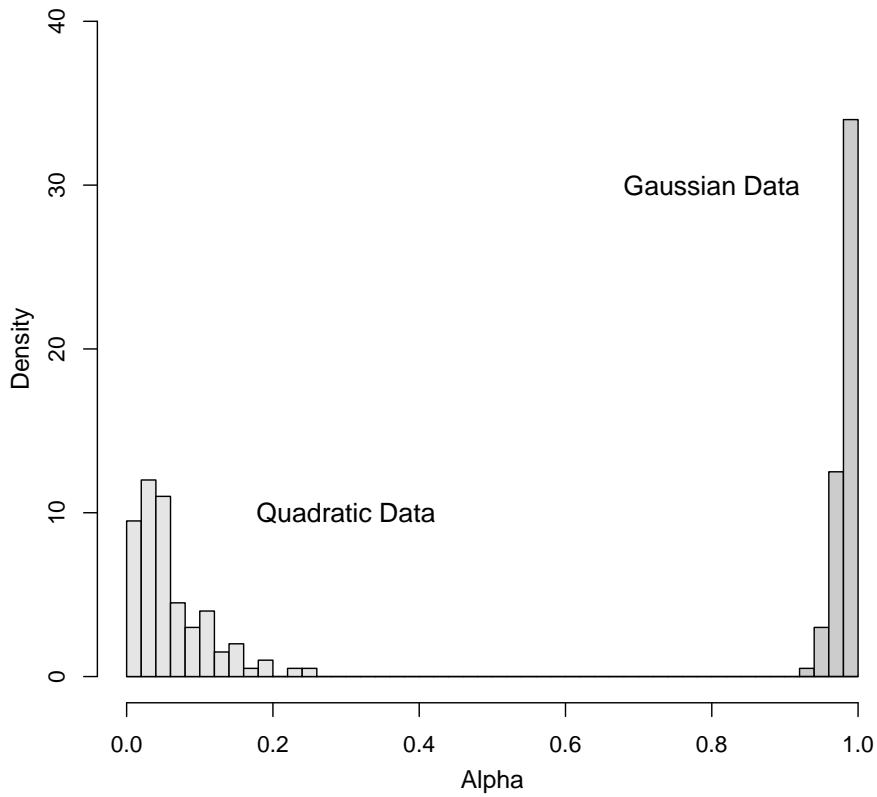


Figure 2: Posterior Distributions of the Alpha Parameter. *Distributions are estimated Alpha parameters from two Monte Carlo data sets of 100 legislators and 500 roll calls. The distribution to the right is estimated from data generated from Gaussian utility, while the distribution to the left is estimated from data generated using quadratic utility.*

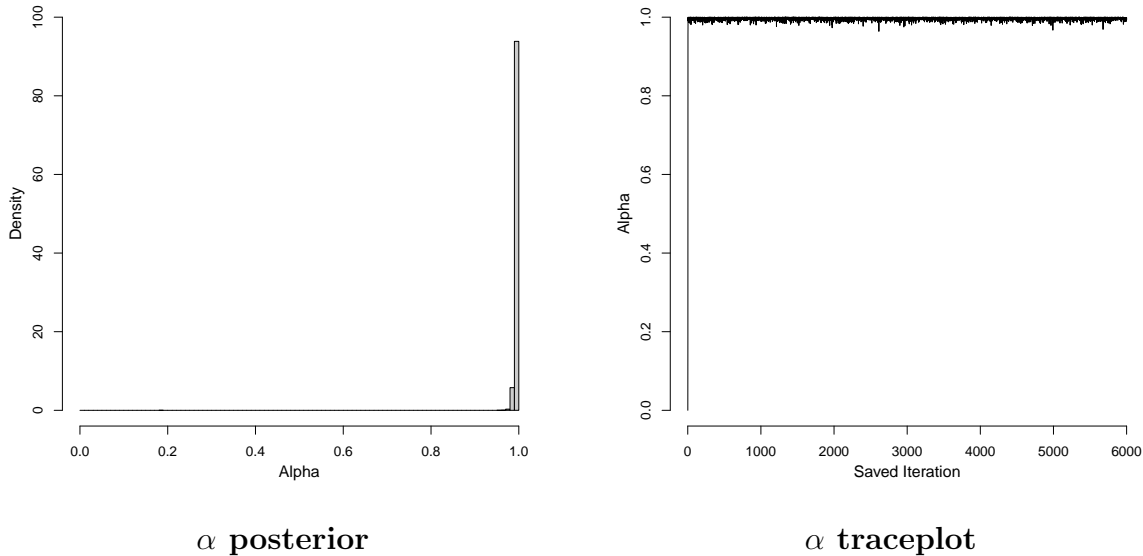


Figure 3: Posterior and trace plot of α : 109th Senate. *The posterior plot is consistent with Gaussian utility posterior from Monte Carlo simulation. Despite using start values biased towards low values of alpha, the traceplot suggests steady state convergence at much higher levels. $\alpha = 1$ is consistent with Gaussian utility while $\alpha = 0$ is consistent with quadratic utility.*

the results from our mixture model suggest that choice probabilities are maximized using a Gaussian utility model, they do not imply that the ideal points recovered via quadratic utility are “wrong”. The ideal points recovered from the quadratic model and mixture model do correlate highly, at 0.950, with those from the mixture model. The correlation for the Gaussian model with the mixture model is slightly higher, at 0.958, suggesting that the pattern of errors is more consistent with Gaussian than with quadratic utility.

These results naturally lead to the question of how α has varied across legislatures over time and whether the results of the 109th Senate are exceptional or common across legislatures. We attempt to answer this question by applying the mixture estimator to every U.S. Senate and House roll call matrix and obtaining estimates of their α parameters.⁸ Figure 4 summarizes these results, plotting the posterior means and empirical 95 per cent credible

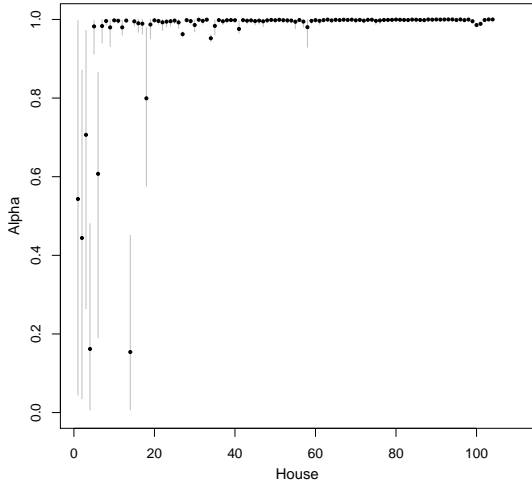
⁸For these estimates, 6,000 iterations were used with a burn in of 1,000 iterations. Starting values for ideal points were drawn $U(-1,1)$, while bill locations were drawn $U(-0.7,0.7)$. For all estimates, α was initialized at a start value of 0.5.

intervals of α for the U.S. House and Senate. In general, these results suggest that high values of α , consistent with Gaussian utility, are pervasive throughout most U.S. Congresses. The major exception appears to be the early Congresses — we hypothesize that this is mainly due to the limited amount of information available for those Congresses due to the lower numbers of legislators and bills. The hypothesis draws support from the observation that normal utility is strongly supported, starting with the 20th House, for the House of Representatives which is typically at least four times the size of the Senate. We plot the posterior densities of two early Congresses — the 5th Senate (1797-1799) and the 6th House (1799-1801) — in Figure 5. These densities can be distributed approximately normal, as in the case of the 6th House, but they can also be skewed, as in the case of the 5th Senate. In the House, which has many more legislators than the Senate, the Gaussian model is always strongly supported after the 20th Congress (1827-1829). A similar pattern for the Senate kicks in only at the 80th Congress (1947-1949). The precision of our estimates also appears to be attributable to the quantity of data available in each estimate. Table 1 shows the results of a simple OLS regression where the dependent variable is the length of the 95% credible intervals for each Congress and the independent variables are the number of legislators, the number of roll calls, and their interaction. In both cases, precision of the estimates increases with both the number of legislators and the number of votes.

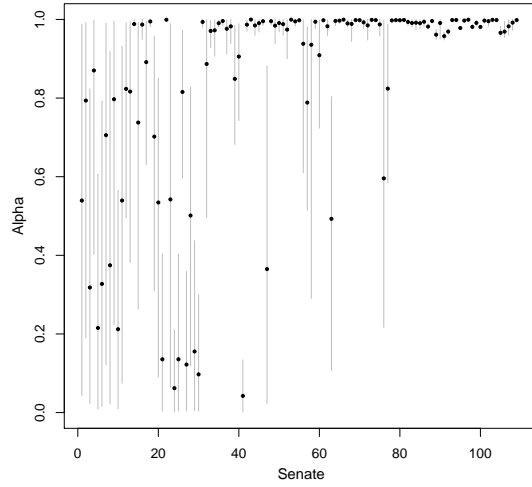
While these results suggest that it can be difficult to distinguish between quadratic and Gaussian utilities in smaller legislatures, judicial settings such as courts may provide a substantively important venue to further evaluate assumptions about utility functions. We applied the mixture estimator to U.S. Supreme Court decisions from 1953-2008, a roll call matrix that includes 31 different justices and 4,333 decisions.⁹ Our results for the Supreme Court are consistent with Gaussian utility, with a posterior mean of 0.998 for α and a posterior standard deviation of 0.002.¹⁰

⁹There are 7,285 total decisions, but all unanimous decisions are discarded before estimates are taken because they do not contribute any useful metric information.

¹⁰In an earlier draft of this paper, we conducted this estimation with only U.S. Supreme Court Data from 1994-97. With only nine justices and 213 votes, the mixture estimator did a poor job of distinguishing

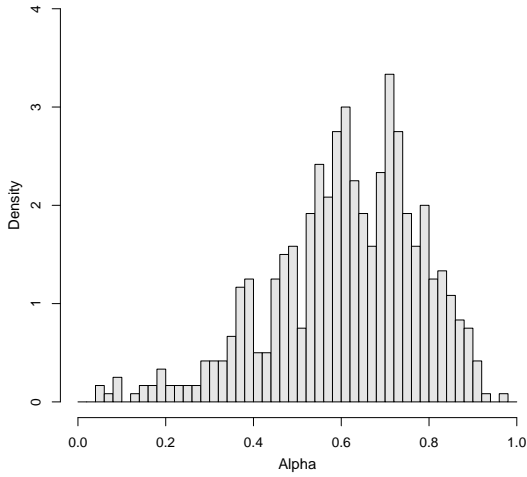


U.S. House

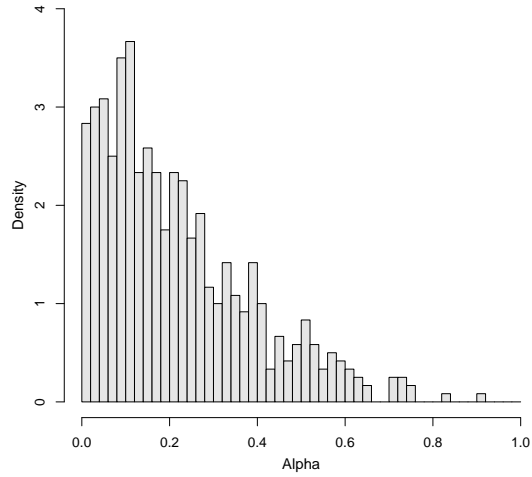


U.S. Senate

Figure 4: Estimates of α over time: U.S. House and Senate. *Points represent the posterior mean of the α draws for each Congress. The lines show the range of the empirical 95 per cent credible intervals of α . $\alpha = 1$ is consistent with Gaussian utility while $\alpha = 0$ is consistent with quadratic utility.*



6th House



5th Senate

Figure 5: Posterior distribution α : 6th House and 5th Senate. *The 6th House had 112 legislators voting on 95 bills, leading to a mean α of 0.607 and a standard deviation of 0.170. The 5th Senate had 43 legislators voting on 194 bills, leading to a mean α of 0.215 and a standard deviation of 0.166.*

	House	Senate
Number of Legislators	-0.135 (0.015)	-1.017 (0.134)
Number of Votes	-0.102 (0.015)	-0.125 (0.034)
Votes x Legislators	0.00024 (0.000038)	0.00118 (0.00038)
Intercept	56.8 (5.51)	112 (10.6)
R^2	0.48	0.53
N	109	109

Table 1: Predicting 95% credible interval length for α : *Standard Errors in Parenthesis. Confidence intervals correspond to those plotted earlier in Figure 4, multiplied by 100.*

Finally, we attempted to determine whether our estimates of a Gaussian α is unique to the U.S. legislature. Given the psychological foundations of the normal utility function in the Shepard–Ennis–Nosofsky stimulus response model, we hypothesize normal utility is likely prevalent in choice data in general. To test this hypothesis, we conducted the same estimation for four different sources of choice data: the U.S. Supreme Court, the French Fourth Republic, the European Parliament, and the California Legislature. Our results, summarized in Table 2, are consistent with our hypothesis that choice data outside the U.S. legislative context also appears to be fit well with normal utility functions.

3 Extremists and Intensity

3.1 What does Intensity Mean?

The mixture model presented in the previous section is predominantly a story of choices at the extremes — that is, how likely are legislators to select the closer alternative when both the bill and the status quo are situated far from their ideal point? Separate but related to this is the question of whether extremists have utility functions that are distinctly different

between quadratic and Gaussian utilities. This new result suggests that there is a lower bound to the size of the roll call matrix needed before reasonable estimates of α can be estimated

Source	Posterior mean of α	σ_α
U.S. Supreme Court, 1953-2008	0.998	0.002
European Parliament, 1979-84	0.986	0.001
European Parliament, 1994-99	1.000	0.000
France First Legislature, 1946-51	1.000	0.000
France Second Legislature, 1951-56	1.000	0.000
France Third Legislature, 1956-58	0.999	0.001
California State Assembly, 1993-94	0.998	0.002
California State Assembly, 1997-98	0.998	0.002
California State Assembly, 2001-02	0.999	0.001
California State Assembly, 2005-06	0.999	0.001
California State Senate, 1993-94	1.000	0.000
California State Senate, 1997-98	1.000	0.000
California State Senate, 2001-02	1.000	0.000
California State Senate, 2005-06	1.000	0.000

Table 2: Estimates of α outside the U.S. Congress. *The results here suggest values of α that are consistent with Gaussian utility in a wide variety of settings. Results shown to three significant digits.*

from those of moderates. In the context of the Gaussian utility model, this suggests that the weight parameter w is not constant as assumed by NOMINATE, but instead varies across legislators. More intense preferences are associated with utility functions with lower variance and larger weight parameters. In particular, we are interested in the possibility that the deterministic component of the utility functions of extremists exhibit greater intensity than that of moderates.

The hypothesis that the choice behavior of extremists may differ systematically from that of moderates is perhaps most forcefully developed in social psychology in the social judgment theory of Sherif and Hovland (1961) and Keisler et al. (1969). Under this theory, behavior is related to “involvements” — that is, more politically involved people may have different utility functions. Moreover, operationally, Sherif et al. (1965) define involvement as membership in a group with a position on an issue. They postulate that individuals will partition the dimension into the three latitudes of acceptance, rejection, and non-commitment. The theory furthermore claims that involvement increases the latitude of rejection. To translate to our model, we hypothesize that extremists will have more sharply peaked utility functions than

moderates — that is, a higher individual weight parameter w_i . A series of empirical studies (Sherif, 1952; Hovland et al., 1957; Sherif et al., 1965) all developed the finding that “those with extreme positions use broader categories for rejection than for acceptance and that their category for rejection is wider than the rejection category of more moderate subjects.”

To observe the substantive impact of such a difference, consider a situation where Justices Kennedy and Scalia are deciding between voting with (O_y) or dissenting (O_n) from a majority opinion, as depicted in Figure 6. The locations of the majority and dissenting opinions are set at 0.4 and 0.6 respectively. The deterministic components of Kennedy and Scalia’s Gaussian utility functions are depicted in dotted lines, centered on their respective ideal points of 0.2 and 0.8. In the case of the utilities depicted with dotted lines, the standard deviation of the Gaussian utility function for both Kennedy and Scalia is set at 0.3. We also show the utility function of a counterfactual “high-intensity” Scalia on a solid line, whose ideal point is still located at 0.8, but has a utility function with noticeably smaller variance.

In the absence of stochastic utility, each justice joins the opinion closest to his ideal point — Kennedy is predicted to join in the majority, while both the low and high intensity Scalias are predicted to join the dissenting opinion. But note that in the presence of random shocks to utility, the justices have different propensities to cast votes that do not conform to this expectation. Both Kennedy and the low-intensity Scalia have the same probability of voting for the alternative that is further from their ideal point, as their deterministic utility differences between the majority opinion and dissent are identical. However, the high-intensity Scalia obtains much more utility from choosing to dissent than the low-intensity Scalia. This in turn suggests that the high-intensity Scalia is much less likely to vote for the alternative that is further from their ideal point than either Kennedy or the low-intensity Scalia.

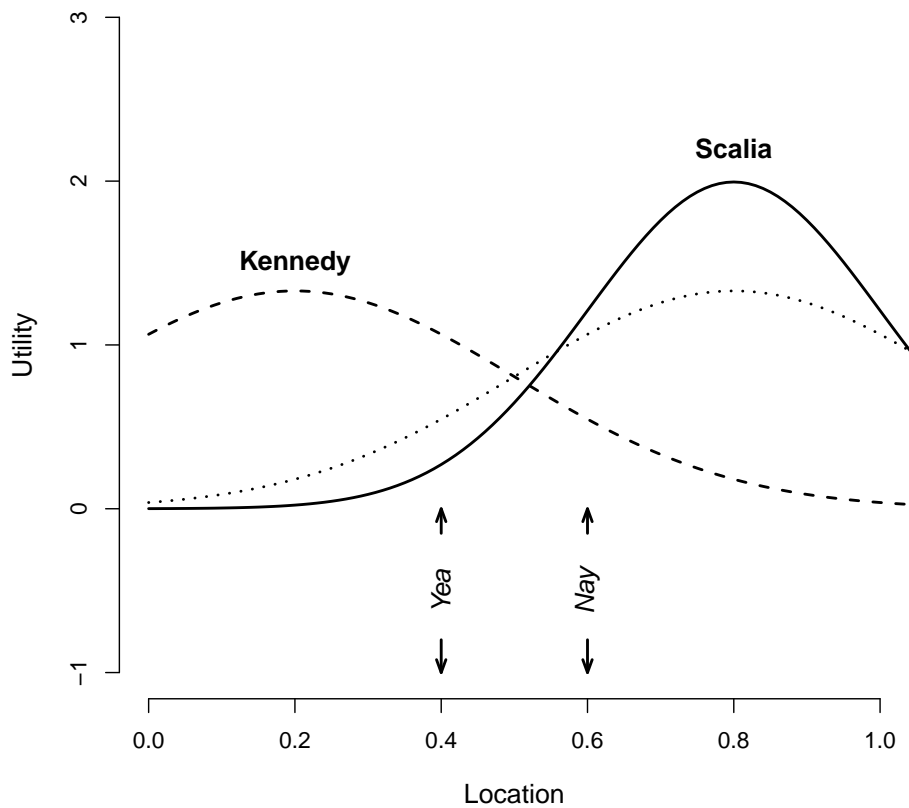


Figure 6: Utility of Functions of Kennedy and Two Scalias. *Kennedy's ideal point is set at 0.2, while both Scalias' ideal points are set at 0.8. While both the low and high-variance Scalias prefer to dissent, the low-variance Scalia experiences a much higher utility difference between voting with the majority and dissenting, and is thus much less likely to err.*

3.2 Estimation and Results

The intensity model we estimate to determine if extremists have more sharply-peaked utility functions than moderates is similar to the Gaussian utility choice model presented earlier. Recall from the Gaussian utility model, given the signal to noise parameter β and weight w , the corresponding yea and nay utilities for legislator i on vote j were:

$$U_{ijy}^{Norm} = \beta \exp\left\{-\frac{1}{2}w^2(X_i - O_{jy})^2\right\} + \epsilon_{ijy}$$

$$U_{ijn}^{Norm} = \beta \exp\left\{-\frac{1}{2}w^2(X_i - O_{jn})^2\right\} + \epsilon_{ijn}$$

Rather than using a global weight parameter w , the intensity model allows each legislator's weight parameter w_i to be estimated separately, resulting in the new utility functions:

$$U_{ijy}^{Intensity} = \beta \exp\left\{-\frac{1}{2}w_i^2(X_i - O_{jy})^2\right\} + \epsilon_{ijy}$$

$$U_{ijn}^{Intensity} = \beta \exp\left\{-\frac{1}{2}w_i^2(X_i - O_{jn})^2\right\} + \epsilon_{ijn}$$

This again leads to the standard probit formulation of the probability that legislator i votes Yea on the j th roll call as:

$$Pr_{ijy} = Pr(U_{ijy} > U_{ijn}) = Pr(\epsilon_{ijn} - \epsilon_{ijy} < u_{ijy} - u_{ijn}) = \Phi[u_{ijy}^{Intensity} - u_{ijn}^{Intensity}]$$

where u_{ijy} and u_{ijn} are the deterministic components of U_{ijy} and U_{ijn} respectively. Correspondingly, the probability that legislator i votes Nay on the j th roll call is $Pr_{ijn} = \Phi[u_{ijn} - u_{ijy}] = 1 - Pr_{ijy}$.

The choice probability function nicely illustrates the identification issues that we high-

lighted in the introduction. Note that we could also allow variation in the importance of the idiosyncratic shocks across members rather than in variation in the importance of the spatial utility (what we are referring to as intensity). In either case, we would make choice more a function of spatial location for some members and less for others. Indeed, in the quadratic utility models increasing w and decreasing the variance of the shocks are isomorphic. In the Gaussian model increasing the w and decreasing the variance of the idiosyncratic shocks are not isomorphic, but they are similar. Thus, we can only say that spatial utility is more important because we are assuming that idiosyncratic shocks are similarly important across members. Of course, all utility calculations are relative and so this is not a surprising situation. However, if we focus on the probabilistic choice function our inability to separate variation in behavior across legislators that is due to spatial or idiosyncratic considerations is of little consequence.

We use the same diffuse priors on the legislator and vote parameters described in Section 2. Given the $p \times q$ matrix of observed votes V , Bayesian inference for the legislators' ideal points, bill parameters, and auxiliary parameters proceeds by simulating the posterior density given by:

$$p(\alpha, \beta, w, X, O|V) \propto p(V|\alpha, \beta, w, X, O)p(\alpha, \beta, w, X, O)$$

where the likelihood and priors are the same as those shown in the previous model.

Our model is closely related to a “robust” ideal point estimator proposed by Bafumi et al. (2005), and later extended by Lauderdale (2010). Both models examine the extent to which individual legislators vote on the basis of the dominant spatial preference dimensions. However, our work differs in two significant ways. First, our model builds on work from the preceding section and is based on the Gaussian utility models used in a large number of empirical applications, whereas the earlier papers exclusively assume quadratic utility. Secondly, our substantive focus examines the relationship between intensity and extremism, an issue left unexplored in the other papers.

Similar to our earlier model, our primary means of validating the estimator is through

the use of Monte Carlo simulation. In this instance, we generate a roll call matrix with 100 legislators and 1000 roll calls using the Gaussian utility function.¹¹ However, in contrast to our earlier Monte Carlo, we vary the weights for each legislator, drawing true weights from the distribution $w \sim U(0.5, 1.5)$. Our estimated ideal points correlate with the true ideal points at $r = 0.999$, and our estimated weights correlate with the true weights at $r = 0.917$ — providing evidence that our estimator is able to recover the parameters of interest successfully.

We note that weights in our Monte Carlo simulations are drawn independently from ideal points. Figure 7 shows a plot of recovered ideal points against the recovered weights. In both the true data and in our estimates, a lowess smoother detects no systematic relationship between the extremism of a legislator and the intensity of their preferences. This is important for our model in demonstrating that extremism and intensity are different concepts that can be separately identified in estimation. More informally, it confirms the fact that the model allows for four types of legislators to be estimated: intense and moderate legislators, intense and extreme legislators, noisy and moderate legislators, and noisy and extreme legislators. Thus, our estimator is able to detect a wide range of possible relationships, both linear and non-linear, between ideological positioning and intensity.

We applied the intensity estimator separately to the 106th to 109th Senates (1999-2007), with the results of our estimation shown in Figure 11. On the X-axis, we plot the Z-transformed ideal points of the senators, calculated as the posterior means of X_i . The posterior mean weight parameter w_i is plotted on the Y-axis, and an 80 per cent confidence band for each legislator’s weight is displayed. A lowess smoother is then applied to the points on each graph in an effort to detect nonlinear patterns in the distribution of weights.

Building on the social judgment theory of Sherif and Hovland (1961), we hypothesized that extremists would have more sharply peaked utility functions than moderates — that is, a higher individual weight parameter w_i . In all four cases we examined, the lowess-smoothed weights generally appear consistent with this hypothesis, though there appears to be some

¹¹For these simulations we set $O_y \sim U(-1.1, 1.1)$, $O_n \sim U(-1.1, 1.1)$, and $X \sim U(-1, 1)$.

Monte Carlo Analysis

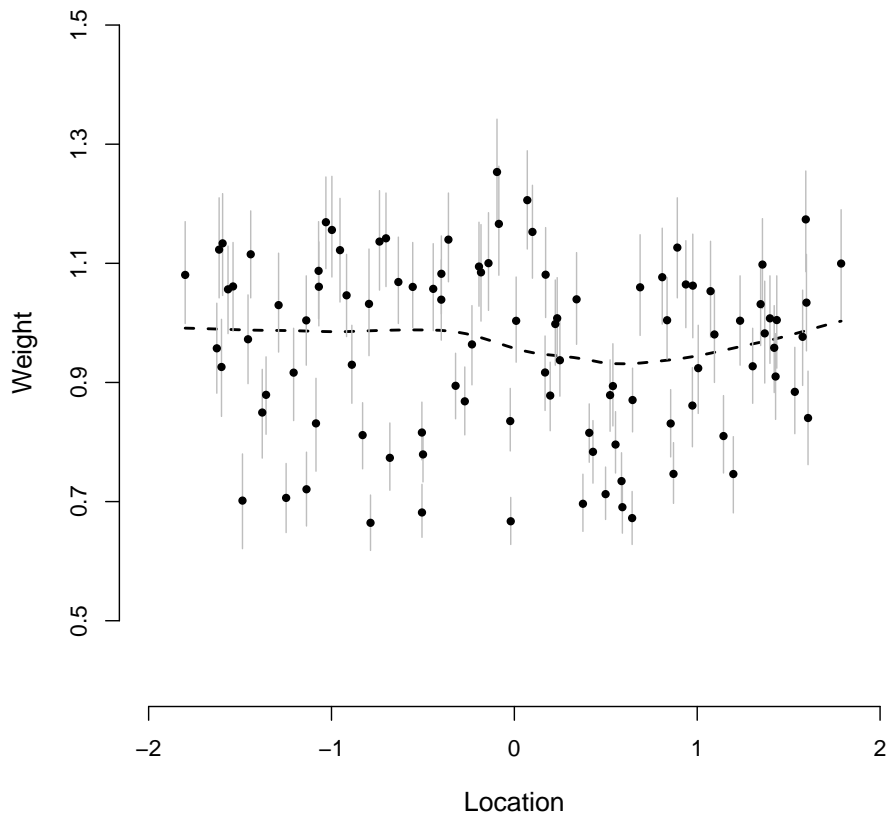


Figure 7: Ideal points vs. Estimated Weights: Monte Carlo Analysis. *Lines reflect the 80 per cent confidence bands on each simulated legislator's estimated weight. Recovered ideal points correlate with true ideal points at $r = 0.999$, while recovered weights correlate with true weights at $r = 0.917$. The lowess smoothed-line shows no relationship between weights and extremism, which are simulated to be independent.*

	Feingold's $w = 0.4$	Harkin's $w = 0.92$
Feingold's $x = -3.23$	-159.5	-248.5
Harkin's $x = -1.31$	-168.8	-245.8

Table 3: Log-Likelihood of Feingold's 109th Senate votes evaluated at combinations of his and Tom Harkin's estimated ideal point (x) and weight parameter (w). *Feingold's estimated ideal point of $x = -3.23$ and weight of $w = 0.4$ are notable outliers, but how much more likely were those outlying values to produce Feingold's votes than the values of x and w estimated for another liberal senator, Tom Harkin, that are not outliers? Note that the difference in log-likelihood is largely the result of the difference in estimated weight (w).*

deviation from this trend at the conservative end of the 109th Senate.¹²

In three of the four cases shown above, Russ Feingold (D-WI) appears as a notable outlier, both in terms of the extremity of his estimated ideal point and his weight parameter. These estimates result from Feingold occasionally deviating from other Democrats and voting with Republicans despite being an otherwise liberal-voting senator. To understand the impact of his occasional "party maverick" behavior on the estimates, we compare the choice probabilities for Feingold in the 109th Senate to that of the second most liberal senator, Tom Harkin (D-IA). In drawing this comparison, we seek to understand why Feingold's estimates deviate so substantially from those of Harkin, despite being similar in their ideological rank positions.

Note that both Feingold's weight and ideal point estimate differ considerably from that of Harkin. How do differences in each of these parameters between Feingold and Harkin affect the fit of the model? We examine this question by calculating how well Feingold's voting record fits the model if his weight and ideal point parameters were set to those of Harkin. Table 3 shows the log-likelihood for Feingold's votes for combinations of weights and ideal points that we estimated for Feingold and Harkin. The upper left cell shows Feingold's log-likelihood under his estimated ideal point and weight of $x = -3.23$ and $w = 0.4$. The other cells show new likelihood values for combinations of Feingold's and Harkin's parameter

¹²Although they are not shown here, we repeated this test for the 100-105th US Senate (1987-1999) as well. Results in those tests are also consistent with the hypothesis that extremists exhibit more sharply-peaked utility functions.

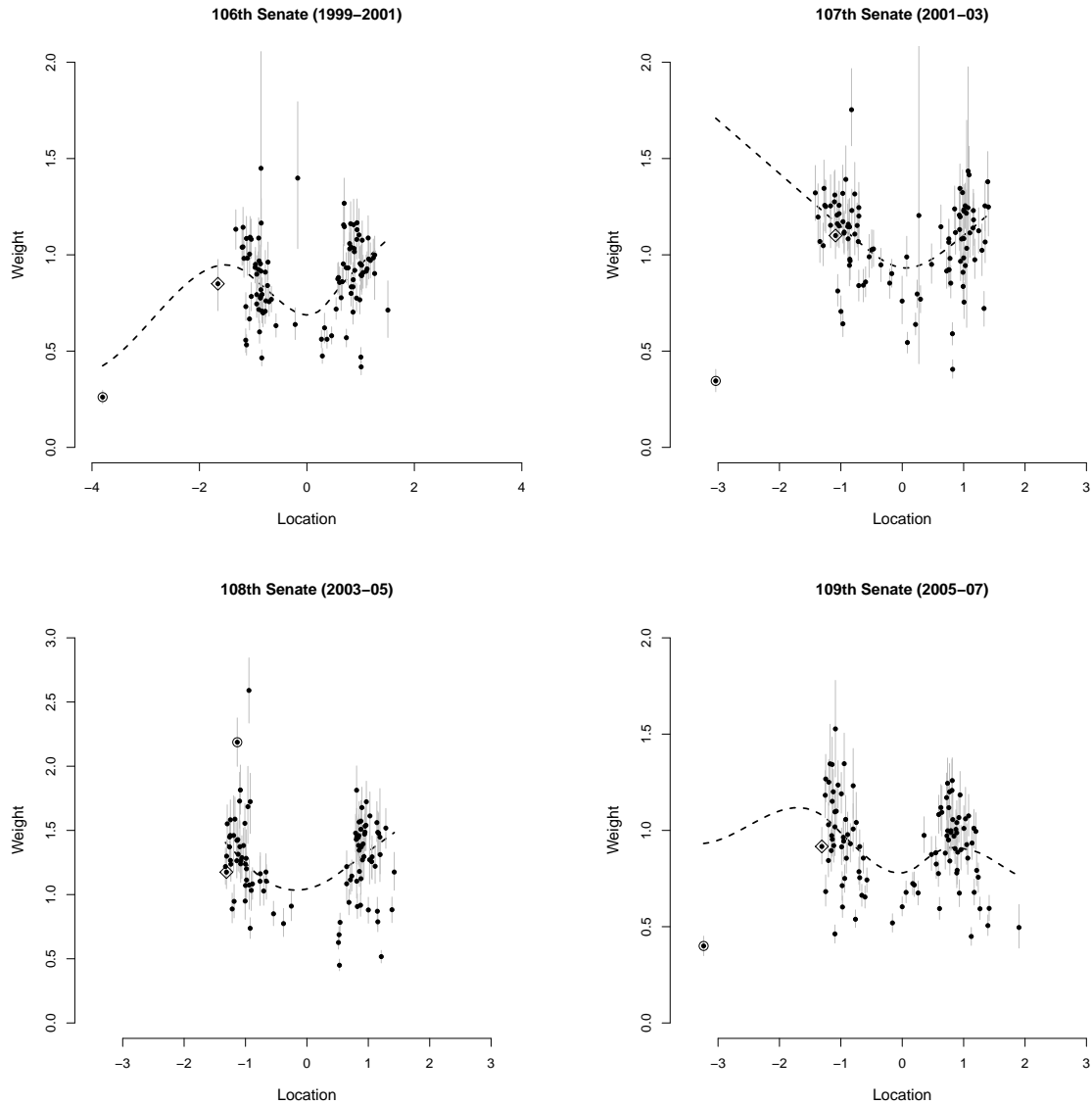


Figure 8: Ideal points vs. Estimated Weights: 106th to 109th Senate. *Lines reflect the 80 per cent confidence bands on each Senator's estimated weight. The dashed line is a lowess smoother of the points and provides evidence that extremists have larger weight parameters than their moderate counterparts. The circled estimate corresponds to Russ Feingold (D-WI) while the estimate enclosed in the diamond corresponds to Tom Harkin (D-IA). In three of the four plots shown here, Feingold is a notable outlier.*

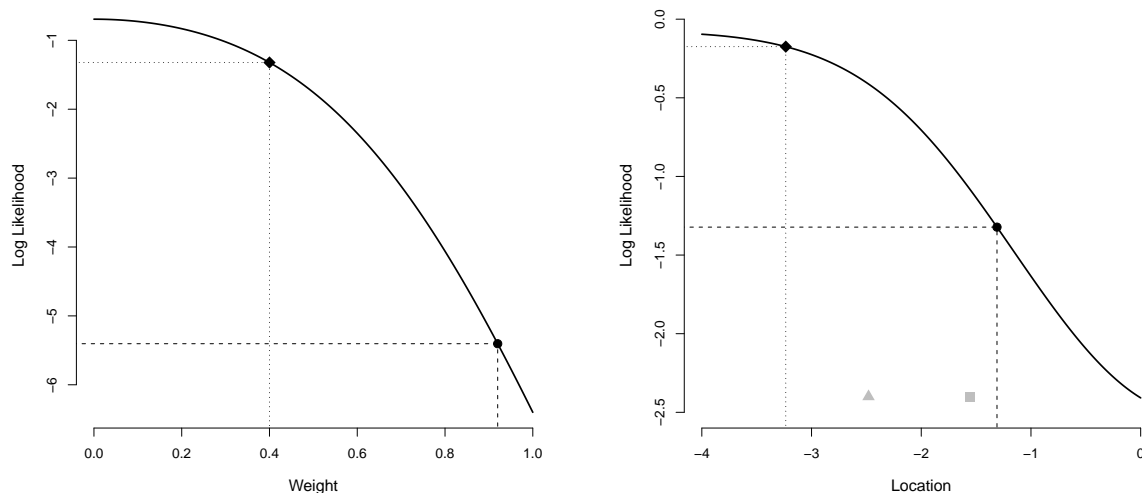


Figure 9: Log-Likelihood of Nay vote by Feingold (D–WI), H.R. 6111 concurrence with House. *The left panel shows the log-likelihood of a Nay vote for different values of w with $x = -3.23$, while the right panel shows the log-likelihood of a Nay vote for different values of x with $w = 0.92$. The diamond indicates the log-likelihood for Feingold, while the point indicated by the circle indicates what Feingold’s log-likelihood would have been if his weight/location was identical to that of Tom Harkin (D–IA), the second most liberal senator in the 109th Senate. The triangle and square on the right panel show the estimated yea and nay locations, respectively.*

estimates. Moving Feingold’s ideal point to Harkin’s position while maintaining the same weight (i.e. the bottom left cell) results in a decrease in the log-likelihood of approximately 9 (about 8 thousand times less likely). However, shifting Feingold’s weight to Harkin’s weight while holding constant Feingold’s ideal point (i.e. the top right cell) decreases the log likelihood by 89 (about 10^{38} times less likely). This shows that the change in weight is the dominant factor explaining the discrepancy in fit.

What sort of voting behavior can account for the large differences in fit for Feingold’s voting record when using Feingold’s weight and ideal point estimates, compared to those of Harkin? Stated differently, what sorts of votes drive the difference between the top left and bottom right corner that we observe in Table 3? Earlier, we characterized Feingold as an occasional ideological maverick who sometimes votes with Republicans and against those in his party. As an illustrative example, consider a Senate motion to concur with the

House Amendment to the Senate Amendment to H.R. 6111, a bill that amended the Internal Revenue Code of 1986. The motion passed with a vote of 79 to 9, with all Democrats other than Feingold voting for the bill; Feingold joined eight Republican senators in voting against it. For this bill, we again separately consider how shifts in weight and ideal point affect the probability of the vote.

The left panel of Figure 9 shows the log-likelihood of a Nay vote at Feingold’s weight parameter value versus the value estimated for Harkin, holding Feingold’s ideal point fixed at its estimated value. The diamond indicates Feingold’s log-likelihood for this vote, while the point indicated by the circle indicates what Feingold’s log-likelihood would be for this vote if his weight was shifted to that of Harkin. We observe that a shift from Harkin’s weight of 0.92 to Feingold’s estimated weight of 0.40 increases the log-likelihood by 4.08. Stated differently, this implies that the lower weight makes the observed maverick vote 59 times more likely to occur.¹³

The right panel of Figure 9, we conduct a plots the same log-likelihood of a Nay vote as Feingold’s ideal point is changed from its estimated value to the value estimated for Harkin holding Feingold’s weight parameter fixed at its estimated value. The diamond and circle indicate Feingold’s and Harkin’s estimated ideal point, respectively. As Feingold’s ideal point shifts from Harkin’s ideal point of -1.31 to Feingold’s estimated ideal point of -3.23, the log-likelihood associated with the vote increases by 1.15, implying that the Nay vote is 3.15 times more likely to occur. In short, reductions in weight and ideal point both substantially improve the fit for Feingold’s maverick votes, but the effect of varying w is much greater.¹⁴

Despite its smaller size, we repeat this analysis for the 1953-2008 U.S. Supreme Court, as shown in Figure 10. Our results for the court suggest a pattern similar to the Senate, with

¹³It should be noted that this is not the most extreme instance of this phenomenon — on one such vote, the change in log-likelihood increased the log-likelihood by 11.77, making the maverick vote 128,888 times more likely to happen.

¹⁴We also examine this behavior for two other presumed Congressional mavericks. John McCain (R-AZ) served in all four legislatures shown here, and his estimated spatial location is largely unaffected by the inclusion of a weight parameter. In contrast, William Proxmire (D-WI) exhibits similar characteristics to Feingold in the 100th Senate — an extreme position (-2.297) coupled with an extremely low weight ($w=0.214$). We thank a reviewer for suggesting this possibility.

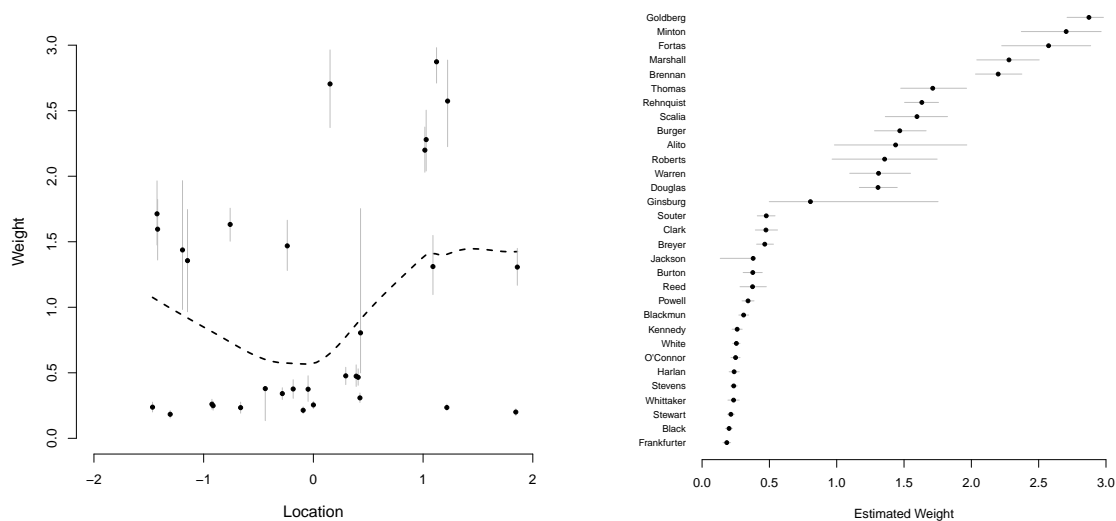


Figure 10: Ideal points vs. Estimated Weights: U.S. Supreme Court 1953-2008. *The left panel shows estimated weights for justices by ideal point, with lines reflecting 80 per cent confidence bands on each estimated weight. The lowess-smoothed line is consistent with earlier results. Weights for each justice are individually identified and sorted in the dotchart on the right panel, also with 80 per cent confidence intervals attached.*

the lowess-smoothed weights suggesting that extremists have more sharply peaked utility functions than moderates. Additionally, we provide a dotchart of the weights that individually identifies each justice separately. The dotchart reveals no obvious pattern in the weight estimates, though a significant number of justices across the ideological spectrum have low estimated weights.

4 Application: Agenda Setting with probabilistic spatial voting over one dimension

Because most theoretical analysis of legislative bargaining and outcomes from within the spatial voting framework assumes simple deterministic Euclidean preferences, it is not clear that our expectations about legislative outcomes or behavior depends in any way on the form of the legislators' utility functions. Indeed, because the ideal point estimates generated under various assumptions about the form of legislators' utility functions are generally similar,

if legislative outcomes and bargaining only depend upon legislators' ideal points then the broader theoretical importance of the preceding discussion would be rather limited.

To demonstrate how the shape of legislators' spatial utility functions (or corresponding probabilistic choice functions) might affect predictions of about legislative behavior and outcomes, we consider a probabilistic voting version of a commonly employed modeling framework for analyzing legislative outcomes (in particular, Cox and McCubbins 1993, Cox and McCubbins 2005, and Krehbiel 1998). In this framework, legislators have Euclidean preferences over a single spatial dimension. We also assume a one-dimensional policy space. However, we assume that in addition to spatial considerations, legislators have non-spatial utility components of the sort described in Section 2.

Ideal point estimates produced by NOMINATE have been used to estimate important quantities such as “gridlock intervals” or “blockout zones” that arise from this modeling framework (for recent examples see Krehbiel et al. 2005, Wiseman and Wright 2008, Stiglitz and Weingast 2010, Richman 2011 or Jenkins and Monroe 2012). However, those estimates are based on the assumption that legislators' preferences are fully determined by their spatial location. By extending the framework to include the non-spatial utility components that are a hallmark of empirical models of spatial voting, we are able to characterize likely legislative outcomes in a way that is fully consistent with the empirical models of legislator choice (like NOMINATE).

Because of non-spatial considerations, our legislators do not always vote for the closer alternative. We assume that the proposer is aware not only of the location of the *status quo* policy and the locations of each legislator's ideal point, but also knows the non-spatial utility values that legislators associate with the *status quo* and with the proposal that she might offer. While the proposer can observe the non-spatial utility values, she cannot affect them. She can only choose the spatial location of the proposal. That is, in the context of any particular legislating opportunity, we assume that the non-spatial utility components are fixed and do not vary as a function of spatial location of the legislation.

In our model, voting is probabilistic *ex ante*. From the standpoint of an outside observer (or an actor prior the revelation of the idiosyncratic utility components), the choice of one alternative over another is not fully determined by the legislator's ideal point and the location of the two alternatives under consideration. As in the empirical models of voting that we considered in the preceding sections, *ex ante* the probability with which one alternative is selected over another is determined by the shape of the spatial utility function and the distribution of the utility shocks.

The model proceeds in the usual way with the proposer selecting the bill (if any) that most improves upon the *status quo* subject to the constraint that it is preferred to the *status quo* by a majority of the legislators and will not be vetoed by the executive. For simplicity, we omit institutions such as the filibuster, a bicameral legislator, or a veto override. This simple setting is sufficient to demonstrate how the distribution of legislative outcomes depends upon how spatial and non-spatial considerations are combined or, more specifically, the form of the random utility functions (described in Section 2). As we will demonstrate shortly, the distribution of legislative proposals and outcomes generated by legislators with Gaussian preferences and legislators with quadratic preferences differ.

Because our chief objective in this article is to probe the limits of how much can be learned about the shape of actors' utility functions, we leave for future work to consider all of the theoretical implications that differently shaped utility functions might imply for the distribution of legislative outcomes or other quantities of interest. Here we will provide a simple motivating example of one such difference.

Consider an $N = 101$ member legislature in which members ideal points are spaced evenly between -1 and 1 implying a floor median of $x_m = 0$. Suppose a leader with proposal power is located at $x_l = -0.5$, and is faced with a legislative *status quo* at $s = 0$ and a president with veto power at $x_p = 0.5$. In the usual world of Euclidean preferences without non-spatial utility considerations, there is no legislative proposal by the party leader that improves upon the *status quo* (from her perspective). However, if legislators also have non-spatial utility

components that can be treated as stochastic from the perspective of the analyst, then it will be possible for the proposer to improve upon the status quo with positive probability and the features of the utility functions that we estimate in this paper affect the distribution of optimal proposals.

To study how the distribution of optimal proposals depends upon the assumed utility function, we conduct a Monte Carlo experiment in which we randomly draw a set of stochastic utility shocks for each alternative and for each of the 101 legislators and the president.¹⁵ Then given these shocks, we consider a series of proposals, beginning at the proposer's ideal point and moving towards the status quo. For each proposal considered, we ask whether a majority of legislators and the president would support the proposal (note that the non-spatial utility components are fixed in each iteration of the experiment). The first proposal that is supported by a majority of the legislators is the optimal bill for the leader to offer. If we find no proposal between the proposer's ideal point and the status quo that is preferred by a majority of the legislators and the president to the status quo, then it is not possible for the proposer to improve upon the status quo in that instance. In that case, the addition of the non-spatial utility components does not alter the prediction of the standard theoretical framework which omits such non-spatial considerations.

We run the experiment separately for the Gaussian and quadratic utility functions 10,000 times. Under both Gaussian and quadratic utility and given the assumed locations of the proposer, the legislators, the president, and status quo, we find that the proposer can offer a bill that passes and which improves upon the status quo (from her perspective) approximately 33 percent of the time. Thus, once additional non-spatial utility components are allowed, the sharp prediction of the standard framework disappears and profitable proposals for this status quo in the black-out interval become possible about one-third of the time. We provide insight as to why the proposer is able to move policy so frequently even under very unfavorable circumstances and also why the probability of a successful proposal is similar under Gaussian

¹⁵In this experiment, utility function parameters are set such that $\beta = 5$ and $w = 1$.

and quadratic utility in Appendix A. In the appendix, we characterize the conditions under which, and frequency with which, a successful proposal is possible in a three-person legislature that is otherwise similar on the legislature considered in our experiment.

While the proposer can move an advantageous proposal in one-third of instances under both quadratic and Gaussian utility, the distribution of the optimal proposals is not the same under both utility functions. The left panel of Figure 11 shows a quantile–quantile (QQ) plot of the winning proposals made under quadratic and Gaussian utility. The dashed 45-degree line illustrates what the QQ-plot would look like if the distribution of winning proposals under Gaussian and quadratic utility were identical. Instead, all the points on the QQ plot fall below the 45 degree line demonstrating that the proposer generally does better (i.e. achieve winning proposals closer to their own ideal point) under Gaussian utility.

How can we explain superior legislative outcomes for the proposer under Gaussian utility? We explore this question on the right panel of Figure 11, which plots the fraction of trials, for each utility function and each ideal point, in which a yea vote is observed across all of the winning proposals. The dark line shows the vote probabilities under normal utility, while the dashed line shows the probabilities under quadratic utility. For both lines, a lowess smoother is used to reduce noise from the simulation. Locations of the proposer, status quo, legislative floor median, and president are shown below. The key trend to note is that under quadratic utility, the probability of a yea vote is strictly decreasing as legislator’s ideal points increase. In contrast, the flatter tails under normal utility create regions where the probability of voting yea is somewhat flat and even begins to increase as legislator’s ideal points move very far away from the proposer. The ability of the normal utility function to more frequently induce spatial “mistakes” in the far tails of the ideological distribution makes it somewhat easier for the proposer to achieve outcomes closer to her own ideal point.

This simple experiment does not describe all of the ways in which the shape of legislators’ utility functions might affect legislative bargaining. However, it does demonstrate how a better understanding of how spatial preferences are translated into voting decisions could

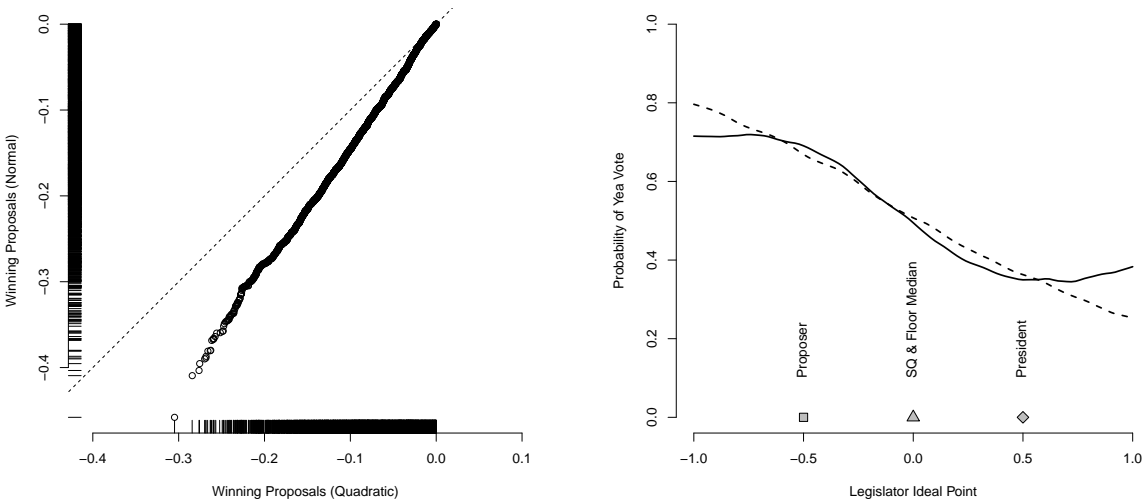


Figure 11: Results of Monte Carlo simulations from our legislative bargaining model with stochastic utility. *The left panel shows a Q-Q plot of the distribution of optimal proposals under quadratic utility against the distribution of optimal proposals under Gaussian utility. Notice that the distribution of optimal proposals under Gaussian utility is more favorable to the proposer. The right panel shows the frequency with which legislators at each point in the ideological spectrum support these proposals. The solid line represents the probability of support under Gaussian utility. The dashed lines show the probability of support under quadratic utility.*

affect theoretical expectations about the legislative process and policy outcomes.

5 Conclusion

Empirical models of spatial voting are typically random utility models of Euclidean spatial voting, where voters assign utility to each of two alternatives associated with each roll call. However, the functions used to assign these utilities are usually assumed rather than estimated. In this paper we attempt to examine the effects of variations in utility functions in the estimation of ideal point estimation methods.

We began by considering the assumed utility functions of two leading implementations of ideal point estimation, Poole and Rosenthal’s (1985) NOMINATE and Jackman’s IDEAL (2008b). We noted that despite many similarities between the Gaussian and quadratic utility functions, the two functions imply different behavior by legislators when choices are located

far from the legislator's ideal point. Exploiting the fact that the quadratic utility function is the first-order exponential approximation of the Gaussian utility function, we then introduce a test designed to determine which utility function best fits a particular roll call data set. Our application of the estimator to the U.S. Congress suggests that Gaussian utility functions generally tend to fit the data better than quadratic utility functions. This trend appears to hold true in a wide variety of contexts outside the U.S. Congress, including the U.S. Supreme Court.

We then examined the possibility that extremists and moderates may have different utility functions. Building on the work of Sherif and Hovland (1961) and Keisler et al. (1969), we hypothesized that extremists would have more sharply-peaked utility functions than moderates. Substantively, this hypothesis implies that extremists are relatively more likely to select the closer alternative than moderates. We then introduce a variation of our original model that allows separate weight parameters to be estimated for each individual legislator. In applying this intensity estimator to four recent U.S. Senates, we found evidence that was supportive of our hypothesis. This trend appears to hold when the same estimator is applied to data from the U.S. Supreme Court.

A Agenda setting in a three-person legislature with observed stochastic utility shocks

Consider a three-person legislature in which the legislators have the preferences assumed in Section 2. Assume there is a single spatial dimension. Suppose the legislators' ideal points, the proposer's ideal point, and the *status quo* policy are as shown in Figure 12. Despite the fact that the *status quo* is coincident with the ideal point of the median legislator, we show that the proposer, P , will be able to offer a winning bill, B^* , that moves policy towards P over half of the time because of the non-spatial stochastic utility components.

It is straightforward to show that the probability that a profitable proposal exists must be at least one-half regardless of the shape of legislators' utility functions.

To begin, define legislator j 's difference in utility between a bill, B , and the *status quo*, 0, as

$$\Delta U_j(B) + \epsilon_j$$

where $\Delta U_j(B) = U(B, X_j) - U(0, X_j)$ is the difference in spatial utility (either Gaussian or quadratic) and ϵ_j is the difference in the normally-distributed non-spatial utility shocks as defined in Section 2. Note that $\Pr(\epsilon_j > 0) = 1/2$. Without loss of generality, we assume that Legislator j will vote for the bill whenever $\Delta U_j(B) + \epsilon_j \geq 0$. Now consider a bill $B^* < 0$ (and, thus, preferred to the *status quo* by the proposer) that is sufficiently close to the *status quo* that $|\Delta U_j(B^*)| \leq |\epsilon_j|$ for all j . Such a B^* is sure to exist because $\lim_{B^* \rightarrow 0} \Delta U_j(B^*) = 0$ for all j . Given B^* , every legislator j 's vote is determined by the sign of ϵ_j . If $\epsilon_j \geq 0$ then legislator j will support B^* and she will support the *status quo* otherwise.

Table 4 shows all of the possible combinations of signs of ϵ_j for $j = a, b, c$. In patterns 1, 2, 3, and 4, the number of positive ϵ 's is greater than one and, thus, B^* is majority preferred. Because the ϵ 's are independently distributed, each of these patterns happens with probability one-eighth. Consequently, it is immediately clear that a feasible

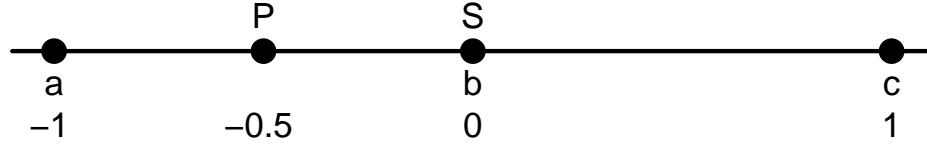


Figure 12: Locations of legislators’ ideal points, the proposer’s ideal point and the status quo. Notice that because the status quo coincides with the median legislator’s ideal point, the proposer cannot improve upon the status quo in the absence of non-spatial utility components.

bill, B^* , is possible in at least one-half of all instances.¹⁶

Now consider patterns 5, 6, 7, and 8 and whether it might be possible to find a B^* to form winning coalitions in any of those instances. First note that no such B^* is possible for patterns 5 or 6. With the status quo located at 0 and $\epsilon_j < 0$ for $j \in \{b, c\}$, there is no way to select a B^* that will offset the negative ϵ ’s of legislators b or c and, therefore, no majority coalition can be formed to move policy towards the proposer. This is because for legislators b and c , $\Delta U(B) \leq 0$ for any $B < 0$ and, thus, there is no B that the proposer prefers to the status quo that can offset their negative ϵ ’s.

Patterns 7 and 8 present a different situation. In these cases it will sometimes be possible to find a B^* that offsets legislator a ’s negative ϵ_a without offsetting legislators b ’s or c ’s positive ϵ_b or ϵ_c . Figure 13 shows the conditions under which the proposer can find a B^* that will offset legislator a ’s negative ϵ_a while not (completely) offsetting the positive ϵ_b in pattern 7—shown in the panels (b) and (d)—and the positive ϵ_c in pattern 8—shown in panels (a)

¹⁶While we have constructed B^* to be close to the status quo, our object is only to establish the existence of a majority-preferred bill that is also preferred by the proposer. As shown in the experiment in the Section 5, the proposer is often able to secure the passage of a bill that is more favorable to her than the B^* that we use to establish the existence of a feasible proposal.

Pattern number	Difference in utility shocks for (a, b, c)	Feasible B^*	Probability of occurring	Probability of feasible B^*
1	(+,+,+)	Yes	1/8	1
2	(+,+,-)	Yes	1/8	1
3	(+,-,+)	Yes	1/8	1
4	(-,+,+)	Yes	1/8	1
5	(-,-,-)	No	1/8	0
6	(+,-,-)	No	1/8	0
7	(-,-,+)	Maybe	1/8	$p_7 > 0$
8	(-,+,-)	Maybe	1/8	$p_8 > 1/2$

Table 4: Possible arrangements of utility shocks and whether the status quo can be moved under those conditions. *In most cases, the potential to move the status quo depends only on whether the idiosyncratic (non-spatial) utility shocks favor the B^* over the status quo. In four of the eight possible patterns of non-spatial utility advantage (or disadvantage) a feasible B^* exists regardless of the exact distribution of the idiosyncratic non-spatial utility shocks or the shape of the utility functions. In two patterns no B^* is feasible. In two other cases, a feasible B^* will exist with a probability (p_7 or p_8) that depends upon the shape of the utility functions. Notice that a feasible B^* will exist in over one-half of all cases.*

and (c). In these cases, the shape of the utility functions matter. Panels (a) and (b) consider the case of Gaussian utility while panels (c) and (d) consider the case of quadratic utility. In each panel, B^* is selected so that $\Delta U_a(B^*) + \epsilon_a = 0$ (note the point $(B^*, -\epsilon_a)$ shown in each panel). For legislator $j \in \{b, c\}$'s positive ϵ_j not to be offset by $\Delta U_j(B^*)$ requires $\epsilon_j \geq \underline{\epsilon}_j$ (note the point $(B^*, -\underline{\epsilon}_j)$ in each panel). Thus, if the point $(B^*, -\epsilon_j)$ falls below $\Delta U_c(B^*)$ then B^* will receive two votes and will pass.

We consider bounds on the probability of a successful bill (B^*) when pattern 7 or 8 obtains, p_7 and p_8 respectively. The dotted line in each of the panels of Figure 13 shows the reflection of ΔU_a about the x -axis. In three of the four panels, this dotted falls below $\Delta U_j(B)$ for all $B \in \{a, b\}$. In those cases, we can write

$$\Pr(\epsilon_j > \underline{\epsilon}_j) = \Pr(\epsilon_j > |\epsilon_a|) + \Pr(|\epsilon_a| > \epsilon_j > \underline{\epsilon}_j).$$

Because $\Pr(|\epsilon_a| > \epsilon_j > \underline{\epsilon}_j) > 0$ in those cases, we know that

$$\Pr(\epsilon_j > \underline{\epsilon}_j) > \Pr(\epsilon_j > |\epsilon_a|).$$

Because the ϵ s are symmetrically distributed about 0 and independently distributed across legislators, the probability that ϵ_j will be larger in absolute value than ϵ_a conditional on $\epsilon_j > 0$ and $\epsilon_a < 0$ is one-half. Thus,

$$\Pr(\epsilon_j > \underline{\epsilon}_j) > 1/2.$$

Therefore, $p_8 > 1/2$ in both the quadratic and the Gaussian utility cases (panels (b) and (d)). It also follows that $p_7 > 1/2$ in the Gaussian case (panel (a)). Because the dotted line falls above $\Delta U_c(B^*)$ in panel (c), $0 < p_7 < 1/2$ in the quadratic case.

Given these minimum values for p_7 and p_8 , we can now bound from below the probability that the proposer will be able to find a profitable B^* to offer at $1/2 + (1/8)(1/2) + (1/8)(1/2) = 5/8$ in the Gaussian case and $1/2 + (1/8)(1/2) = 9/16$ in the quadratic case. Thus, regardless the shape of the utility function, the proposer in this example is able to pull policy away from a status quo located at the ideal point of the median voter over 56 percent of the time.

Adding a presidential veto reduces this probability by a factor 1/2 when the legislator shocks are in patterns 1, 2, 3, and 4. This is because in those cases the president will have a positive ϵ (and thus support B^*) only 1/2 of the time. The addition of the presidential veto reduces the probability that a feasible B^* exists when legislator shock patterns 7 and 8 obtain by a factor that depends upon the location of the president and the shape the president's utility function. However, with a presidential veto p_7 and p_8 remain above zero, and the overall probability that the proposer can offer a winning bill is always greater than 25 percent regardless of the locations of the president or whether her utility function is Gaussian or quadratic.

References

- Bafumi, J., A. Gelman, D. Park, and N. Kaplan (2005). Practical issues in implementing and understanding bayesian ideal point estimation. *Political Analysis* 13(2), 171–187.
- Clinton, J., S. D. Jackman, and D. Rivers (2004). The statistical analysis of roll call data: A unified approach. *American Political Science Review* 98, 355–370.
- Cox, G. and M. McCubbins (1993). *Legislative leviathan: Party government in the House*, Volume 23. University of California Press.
- Cox, G. and M. McCubbins (2005). *Setting the agenda: Responsible party government in the US House of Representatives*. Cambridge University Press.
- Desposato, S. (2006). Parties for Rent? Ambition, Ideology, and Party Switching in Brazil’s Chamber of Deputies. *American Journal of Political Science* 50(1), 62–80.
- Enelow, J. M. and M. Hinich (1984). *The Spatial Theory of Voting*. New York: Cambridge University Press.
- Ennis, D. M. (1988). Confusable and Discriminable Stimuli: Comment on Nosofsky and Shepard. *Journal of Experimental Psychology* 117, 408–411.
- Hagemann, S. (2007). Applying Ideal Point Estimation Methods to the Council of Ministers . *Political Analysis* 8, 279–296.
- Heckman, J. J. and J. M. Snyder (1997). Linear Probability Models of the Demand for Attributes With an Empirical Application to Estimating the Preferences of Legislators. *Rand Journal of Economics* 28, 142–189.
- Herron, M. C. and J. B. Lewis (2007). Did Ralph Nader Spoil Al Gore’s Presidential Bid? A Ballot-Level Study of Green and Reform Party Voters in the 2000 Presidential Election. *Quarterly Journal of Political Science* 2, 205–226.

- Hinich, M. J. and M. J. Munger (1994). *Ideology and the Theory of Political Choice*. Ann Arbor: University of Michigan Press.
- Hinich, M. J. and M. J. Munger (1997). *Analytical Politics*. Ann Arbor: Cambridge University Press.
- Hix, S., A. Noury, and G. Roland (2007). *Democratic Politics in the European Parliament*. Cambridge Univ Pr.
- Hoff, P., A. Raftery, and M. Handcock (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association* 97(460), 1090–1098.
- Hovland, C., O. Harvey, and M. Sherif (1957). Assimilation and Contrast Effects in Reactions to Communications and Attitude Change. *Journal of Abnormal and Social Psychology* 55, 244–252.
- Ichniowski, C., K. Shaw, and G. Prennushi (1997). The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines. *American Economic Review* 87(3), 291–313.
- Jackman, S. (2008a). pscl: classes and methods for r developed in the political science computational laboratory.
- Jackman, S. (2008b). pscl: Political Science Computational Laboratory.
- Jenkins, J. (1999). Examining the Bonding Effects of Party: A Comparative Analysis of Roll-Call Voting in the U.S. and Confederate Houses. *American Journal of Political Science* 43(4), 1144–1165.
- Jenkins, J. and N. Monroe (2012). Buying negative agenda control in the us house. *American Journal of Political Science*.
- Jessee, S. (2009). Spatial voting in the 2004 presidential election. *American Political Science Review* 103(1), 59–81.

- Kalandrakis, T. (2010). Rationalizable Voting. *Theoretical Economics* 5(1), 93–125.
- Keisler, C., B. Collins, and N. Miller (1969). *Attitude Change: A Critical Analysis of Theoretical Approaches*. New York: Wiley Press.
- Krehbiel, K. (1998). *Pivotal politics: A theory of US lawmaking*. University of Chicago Press.
- Krehbiel, K., A. Meirowitz, and J. Woon (2005). Testing theories of lawmaking. *Social Choice and Strategic Decisions*, 249–268.
- Lauderdale, B. (2010). Unpredictable voters in ideal point estimation. *Political Analysis* 18(2), 151–171.
- Lewis, J. B. (2001). Estimating Voter Preference Distributions from Individual-Level Voting Data. *Political Analysis* 9, 275–297.
- Londregan, J. B. (2000). Estimating Legislators’ Preferred Points. *Political Analysis* 8, 35–56.
- Martin, A. D. and K. Quinn (2002). Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999. *Political Analysis* 10, 134–153.
- McCarty, N. M., K. T. Poole, and H. Rosenthal (1997). *Income Redistribution and the Realignment of American Politics*. Washington DC: SEI Press.
- McCarty, N. M., K. T. Poole, and H. Rosenthal (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge: MIT Press.
- McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers of Economics*. New York: Academic Press.
- Myagkov, M. and D. Kiewiet (1996). Czar Rule in the Russian Congress of People’s Deputies? *Legislative Studies Quarterly* 21(1), 5–40.

- Nosofsky, R. M. (1986). Attention, Similarity, and the Context Theory of Classification. *Journal of Experimental Psychology* 115, 39–57.
- Peress, M. (2009). Small Chamber Ideal Point Estimation. *Political Analysis* 17, 276–290.
- Poole, K. and H. Rosenthal (1983). A Spatial Model For Legislative Roll Call Analysis. GSIA Working Paper 5-83-84, Carnegie–Mellon University.
- Poole, K. T. (2000). Non-parametric Unfolding of Binary Choice Data. *Political Analysis* 8, 211–237.
- Poole, K. T. (2005). *Spatial Models of Parliamentary Voting* . Cambridge: Cambridge University Press.
- Poole, K. T. and H. Rosenthal (1985). A Spatial Model for Legislative Roll Call Analysis. *American Journal of Political Science* 29, 357–84.
- Poole, K. T. and H. Rosenthal (1991). Patterns of Congressional Voting. *American Journal of Political Science* 35, 228–278.
- Poole, K. T. and H. Rosenthal (1997). *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.
- Richman, J. (2011). Parties, pivots, and policy: The status quo test. *American Political Science Review* 105(01), 151–165.
- Rosenthal, H. and E. Voeten (2004). Analyzing Roll Calls with Perfect Spatial Voting: France 1946–1958. *American Journal of Political Science* 48(3), 620–632.
- Schonhardt-Bailey, C. (2003). Ideology, Party and Interests in the British Parliament of 1841-47. *British Journal of Political Science* 33(4), 581–605.
- Shapiro, A. (1986). Asymptotic theory of overparameterized structural models. *Journal of the American Statistical Association* 81(393), 142–149.

- Shepard, R. N. (1986). Discrimination and Generalization in Identification and Classification: Comment on Nosofsky. *Journal of Experimental Psychology* 115, 58–61.
- Shepard, R. N. (1987). Toward a Universal Law of Generalization for Psychological Science. *Science* 242, 1317–1323.
- Sherif, C., M. Sherif, and R. Nebergall (1965). *Attitudes and Attitude Change: The Social Judgement-Involvement Approach*. Philadelphia: W.B. Saunders.
- Sherif, C. H. M. (1952). Judgmental Phenomena and Scales of Attitude Measurement: Item displacement in Thurstone scales. *Journal of Abnormal and Social Psychology* 47, 822–832.
- Sherif, M. and C. Hovland (1961). *Social Judgement: Assimilation and Contrast Effects in Communication and Attitude Change*. New Haven: Yale University Press.
- Shor, B., C. Berry, and N. McCarty (2010). A Bridge to Somewhere: Mapping State and Congressional Ideology on a Cross-institutional Common Space. *Legislative Studies Quarterly* 35(3), 417–448.
- Stiglitz, E. and B. Weingast (2010). Agenda control in congress: Evidence from cutpoint estimates and ideal point uncertainty. *Legislative Studies Quarterly* 35(2), 157–185.
- Voeten, E. (2000). Clashes in the Assembly . *International Organization* 54(2), 185–215.
- Wiseman, A. and J. Wright (2008). The legislative median and partisan policy. *Journal of Theoretical Politics* 20(1), 5–29.

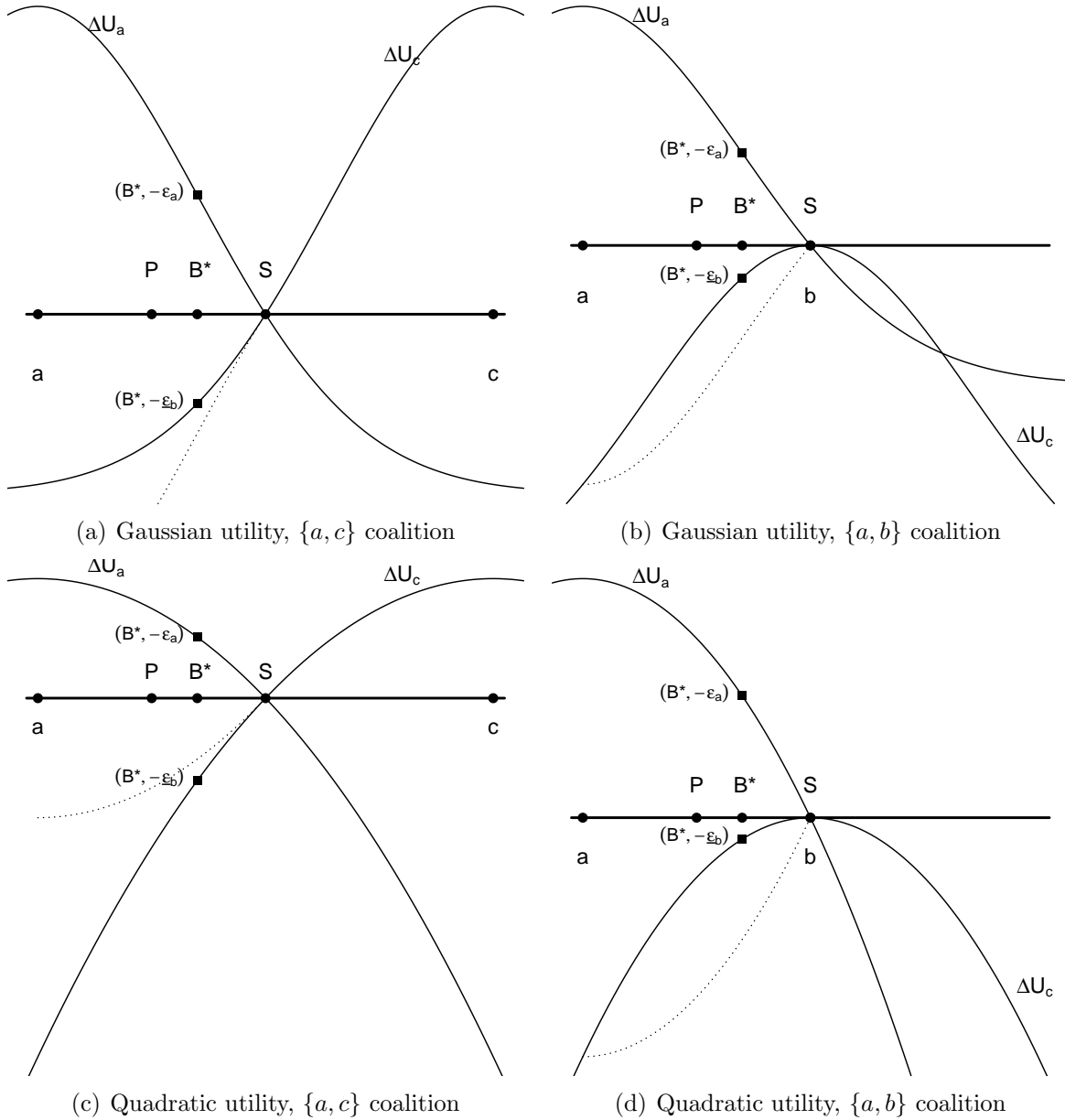


Figure 13: Differences in spatial utility between any bill, B , and a status quo at 0 for given pairs of legislators. Each panel considers conditions under which the proposer can choose a B^* that a and one of b or c will vote for when legislator a 's net non-spatial utility favors the status quo ($\epsilon_a < 0$) and either legislator b or c 's net non-spatial utility favors the bill ($\epsilon_b > 0$ or $\epsilon_c > 0$). If the proposer can offer a B^* such that legislator a 's non-spatial utility difference is exactly offset and legislator b 's or c 's non-spatial utility difference is not completely offset by their spatial preference for the status quo then a winning coalition of legislator a and either b or c is possible. This happens whenever $\epsilon_j \geq \underline{\epsilon}_j$ for $j \in \{b, c\}$. See text for a detailed description.