University of California
Los Angeles

Essays on Preferences in Political Science

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Political Science

by

James Lo

2010

The dissertation of James Lo is approved.

Keith T. Poole

Kathleen Bawn

Timothy Groseclose

Jeffrey B. Lewis, Committee Chair

University of California, Los Angeles

2010

To my family …

for all of their support over the years.

TABLE OF CONTENTS

List of Figures

List of Tables

Acknowledgments

Many people contributed to my experience at UCLA and the writing of this dissertation. I would like to begin by thanking the members of my committee for their assistance over the years. Tim Groseclose taught me a considerable amount about ideal point models over the years, and this is reflected by the fact that two papers in this dissertation deal with ideal point estimation. Kathy Bawn played a pivotal role in guiding my progress through the program, especially in her capacity as my primary advisor during the early years of my doctorate. Keith Poole not only helped considerably in advising my dissertation, but also invented the field of ideal point analysis. My experience at UCLA was considerably enhanced through my work as his research assistant under NSF grant SES-0611974. Finally, Jeffrey Lewis has been everything that a dissertation advisor can be, and much more. He was deeply involved in every aspect of my academic development during my time at UCLA, and taught me how good academic work can and should be done. I can only hope that one day, I have the opportunity to contribute as much to my students as Jeff has contributed to me.

Chapter 3 of my dissertation, "The Structure of Utility in Spatial Models of Voting", is work that is coauthored with Royce Carroll, Jeffrey Lewis, Keith Poole, and Howard Rosenthal. I am deeply honored to work with such wonderful collaborators and have learned much from their contributions. This work was completed with the support of NSF grant SES-0611974. Funding for my degree was also provided through a Social Sciences and Humanities Research Council

(SSHRC) doctoral fellowship.

Numerous classmates at UCLA also contributed greatly to my wonderful experience here, and there are far too many contributors for me to name. In particular, Greg Anderson, Shahin Berenji, Lisa Blaydes, Mac Bunyanunda, Elizabeth Carlson, Richard Goldberg, Seth Hill, Diana Ichpekova, Kristen Kao, Dov Levin, Gid Lukens, Jeff Paris, Anoop Sarbahi, and Matt Spence helped to make my stay a particularly pleasant one. In addition, three faculty members who were not on my dissertation committee provided extremely valuable guidance during my time at UCLA — Marc Trachtenberg, Lynn Vavreck, and John Zaller.

Finally, I would like to thank my family for their support over the years, especially my brother Julian and my parents. My mother and father never had the chance to attend college, but they were committed to ensuring that their children had those opportunities that they never enjoyed themselves. In particular, they made the decision to move to a land that was foreign to them so that their children might be able to pursue opportunities that would otherwise be denied to them. I am awed by the sacrifices that they have made for us, and my completion of this degree is a testament to their support.

VITA

1978	Born, Hong Kong.
1997–2002	B.Eng. (Computer Engineering), McMaster University.
2002–2004	M.A. (Political Science), Brock University.
2004–2010	PhD. (Political Science), UCLA.

PUBLICATIONS

Poole, Keith, Jeffrey Lewis, James Lo, Royce Carroll. Forthcoming. "Scaling Roll Call Votes with W-Nominate in R." *Journal of Statistical Software.*

Carroll, Royce, Jeffrey Lewis, James Lo, Keith Poole, and Howard Rosenthal. (2009) "Comparing NOMINATE and IDEAL: Points of difference and Monte Carlo tests." *Legislative Studies Quarterly.* 34(4).

Carroll, Royce, Jeffrey Lewis, James Lo, Keith Poole, and Howard Rosenthal. (2009) "Measuring Bias and Uncertainty in DW-NOMINATE Ideal Point Estimats via the Parametric Bootstrap." *Political Analysis.* 17(3)

Lo, James. (2003) "Nuclear Deterrence in South Asia: Theory and Practice" *International Journal.* 58(3).

Abstract of the Dissertation

Essays on Preferences in Political Science

by

James Lo

Doctor of Philosophy in Political Science

University of California, Los Angeles, 2010

Professor Jeffrey B. Lewis, Chair

This dissertation consists of three distinct chapters. First, I estimate the rate at which public health insurance "crowds out" private health insurance as it expands to higher income levels. I find that estimates of crowd out are highly sensitive to the income level of the targeted expansion. Next, I examine what we might learn about the shape of utility functions that are typically assumed in ideal point models, and find evidence that normal utility is prevalent throughout much of the choice data used in political science. Finally, I examine the sensitivity of ideal point models to the assumption of independence. I find that traditional ideal point models that make this assumption are generally robust when compared to estimates that do not make this assumption.

CHAPTER 1

Introduction

This dissertation consists of three separate papers on preferences in political science. My first paper deals with how individuals react when presented with a choice between public and private health insurance. My central finding is that the income level of the individual presented with such a choice heavily conditions their preference for choice of insurance, and that this impact occurs in unusual ways. My remaining two papers deal with ideal point models, which attempt to estimate to preference of individuals in a multi-dimensional space using vote choice data. These models have been heavily used in the political science literature in the study of Congress, but many features of these models have typically been fixed by assumption. My papers assess the impact of two such assumptions — the shape of the utility function and the assumption of independence across votes.

My first paper examines the impact of the State Children's Health Insurance Plan (SCHIP). SCHIP is Medicaid expansion that has increasingly been extended to higher income populations. However, as public health insurance coverage expands, some individuals with private health insurance may choose to substitute and "crowd out" private coverage for public insurance instead, thus decreasing the marginal effect of public health expenditures on the uninsured population. Little is known about the extent of crowd-out and its impact on total coverage as public health insurance expansions increase to higher income levels. To explore this problem, I examine Illinois' All Kids program, an SCHIP expansion that covers all children in Illinois regardless of income. I then use a procedure developed by Abadie et al. (2010) to construct a synthetic version of Illinois that estimates the levels of insurance coverage that would have been observed in Illinois in the absence of this policy intervention.

Comparing coverage patterns between the observed and synthetic Illinois, I find that for children between 200-300% of the Federal Poverty Line, the All Kids program produced an increase in health insurance coverage of 2% with 60% crowd-out. For children between 300-400% of FPL, All Kids produced no increase in overall coverage and no crowd-out. Finally, I find that for children between 400-500% FPL, SCHIP produced a 3-4% increase in coverage with crowd-in.

My second paper, to be coauthored later with Royce Carroll, Jeffrey Lewis, Keith Poole, and Howard Rosenthal, examines the shape of utility functions in ideal point estimation techniques. Empirical models of spatial voting allow legislators' locations in an abstract policy or ideological space to be inferred from their roll call votes. These are typically random utility models of Euclidean spatial voting, where voters assign utility to each of two alternatives associated with each roll call. The specific functional forms of the utility functions are generally assumed rather than estimated. In this paper, we attempt to infer important features of these utility functions. We first consider a model in which legislators' utility functions are assumed to be a mixture of the two most commonly assumed utility functions (the Gaussian function assumed by NOMINATE and the quadratic function assumed by IDEAL and many other estimators). Applying this estimator to large number of roll call data sets, we find that in nearly every case 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 is substantively important to the formation and interpretation of law, because it suggests that extremists are ideologically rigid whereas moderates are more likely to consider influences that arise outside liberal–conservative conflict. Finally, we considered a third model extension examining the possibility that legislators have asymmetric utility functions. Our results tentatively suggest that, conditional on party, as legislators become more conservative their sensitivity to policy alternatives on the right increases.

My final paper examines the potential impact of two sources of complexity in ideal point models. Specifically, I am interested in understanding how deliberate decisions not to vote on legislation might affect ideal point estimates, and how the dependence across votes that is part of the amendment process might be incorporated into ideal point models. I begin by first introducing a multinomial ideal point estimator, which allows more than two choices on each roll call to be accounted for. I then use this estimator to consider the impact of 'Present' votes in the Illinois State Senate, a procedure Barack Obama was attacked for during the 2008 Presidential campaign. I find that by incorporating Present votes into ideal point models, the precision of the estimates improves by a considerable margin. Next, I consider an ideal point model that accounts for the dependencies across votes introduced by the amendment procedure, and apply it to the 109^{th} House. I find some evidence that the traditional assumption of independence across votes may overstate the precision of ideal point estimates.

CHAPTER 2

How Income Affects SCHIP — An Analysis Using Synthetic Controls

2.1 Introduction

Following the defeat of President Clinton's health care plan in 1993, Democrats sought to pass smaller publicly-funded health initiatives capable of generating bipartisan support. In particular, Congress attempted to pass legislation providing free or subsidized coverage to particular groups. The State Children's Health Insurance Plan (SCHIP), emerged as one such proposal that ultimately passed Congress in 1997. Prior to SCHIP, the primary public health insurance plan in the U.S. for low income children was the 1992 Medicaid expansion, which provided all children living in families below the Federal Poverty Line (FPL) with health insurance through a joint cost sharing program between states and the Federal Government. Under SCHIP, states could receive additional funds to expand public health insurance coverage to children living in families between 100-200% of the FPL, and in some cases obtain waivers to extend coverage beyond that level.

Partial coverage through Medicaid expansions however is not without its problems. As early as 1996 (Currie and Gruber, 1996a), economists have noted that eligibility for coverage does not result in full coverage due to uneven program adoption. Furthermore, as public health insurance coverage expands, some individuals with private health insurance may choose to substitute that coverage for public insurance instead. Public expansions can therefore "crowd out" private coverage (Cutler and Gruber, 1996), offsetting many of the coverage gains expected in the expansion populations. Stated differently, covering 100 individuals with public insurance does not increase the number of individuals covered with health insurance by 100, because some number of those individuals would have purchased private insurance instead, while others will not take advantage of their eligibility for public insurance.

These issues present three questions with important policy implications. First, how much has SCHIP increased health coverage among children? Secondly, how much crowd-out has occurred? Finally, as public coverage expands to higher income groups, how does the income level of the expansion population affect our answers to these questions? While answers to the first two questions have been well researched for the lowest income populations, the impact of income has largely been absent from the literature. In the few instances where researchers have examined income (notably Dubay and Kenney (1996) and Card and Shore-Sheppard (2004), significant differences in coverage and crowd-out across different income levels are found. Furthermore, expansion of public health insurance to higher income levels remains a possibility, as Congress recently increased income eligibility limits to 300% of FPL in its 2009 SCHIP reauthorization bill.

In examining the impact of further expansions of SCHIP, I focus my analysis on Illinois' SCHIP program. Illinois represents an ideal test case for two reasons. First, Illinois is unique among the U.S. states in extending SCHIP coverage to children of all income levels while other states continue to restrict SCHIP enrollment based on family income. Its SCHIP program was initially typical of many other public insurance programs, providing health insurance to children living under 185% FPL. This program expanded dramatically in 2006 under Governor Rod Blagojevich's All Kids program, which expanded SCHIP to children of all income levels. Illinois therefore allows us to examine the effect of public expansions at income levels not implemented anywhere else in the U.S. Secondly, Illinois' demographic characteristics and health insurance coverage profile closely approximates the average values found elsewhere in the U.S, a point discussed later in the paper when interpreting Table 1. Thus, patterns found in Illinois might plausibly be reproduced more broadly in the U.S. if SCHIP is expanded.

Drawing causal inference from the Illinois experience is complicated by the lack of an obvious counterfactual control region that allows us to observe Illinois' coverage levels in the absence of the All Kids intervention. In this paper, I overcome this problem by using "synthetic" controls, an idea first demonstrated in Abadie and Gardeazabal's 2003 analysis of Basque political terrorism (Abadie and Gardeazabal, 2003), and extended in a recent working paper by Abadie et al. (2010). The key insight of this approach is that a synthetic control region can be constructed as a convex combination of multiple 'donor' units unaffected by Illinois' health insurance expansion. In this example, the donor units are the other U.S. states that have not extended SCHIP coverage to higher income levels. The synthetic unit is constructed such that its relevant demographic characteristics and health insurance profile closely resembles that of Illinois prior to the introduction of the All Kids program in 2006. The synthetic control unit thus allows us to simulate what Illinois' health insurance profile would look like in the absence of All Kids after 2006, and causal inference can be drawn by comparing observations from the real Illinois with those of the synthetic unit in the post-treatment period.

The remainder of this paper is divided into five sections. I begin with a discussion of theoretical reasons why we might expect to observe a relationship between crowd-out and income. While the existing literature overwhelmingly argues that expansions of public health insurance to higher income populations inevitably leads to higher crowd-out, I argue that there are theoretical reasons why this expectation may not hold true. Next, I summarize the extensive literature on crowd-out, which largely ignores the impact of the targeted income level on coverage and crowd-out. In particular, I focus on three notable research designs that are prominent in the literature and explain why they may be unsuitable when applied to SCHIP expansions at higher income levels. I then discuss the synthetic control methodology used in this paper and how this procedure overcomes the problems raised in the earlier literature. This methodology is then applied to examine how All Kids affects insurance coverage and turnout in Illinois between 200-500% FPL. I conclude with some thoughts on the policy and political implications of the my findings in this paper.

2.2 Theory

Reflecting the general consensus among health economists, an oft-cited report by the Kaiser Family Foundation (2007) argues that as eligibility levels for public programs are expanded to cover higher incomes, the possibility of substitution from private to public insurance increases. The report argues that because private insurance is often not available to lower income populations, expanding eligibility to higher income levels is likely to increase participation in public programs among income groups where private coverage is already available. While no evidence or citation is provided, the report speculates that crowd-out may be as high as 50% at the 300% FPL.

While the causal mechanism described in the Kaiser report is plausible, the report ignores three important countervailing forces that are likely to mitigate or even reverse crowd-out as SCHIP expands to higher income levels. First, there may be social stigma attached to SCHIP as a low income entitlement program. This stigma imposes a social cost to insurance substitution that is likely to be increasingly costly as an individual's income level increases. Beyond the social stigma attached with welfare, SCHIP's status as a low income entitlement is likely to produce misinformation about the program's eligibility rules. For example, Haley and Kenney (1996) find that while large numbers of people have heard of the Medicaid and SCHIP programs, a significantly smaller number of people were aware that children could participate in these programs without receiving welfare. This misunderstanding is increasingly likely in higher income populations because individuals with higher income are less likely to have access to a welfare caseworker that can correct these false impressions.

Secondly, at higher income levels the quality of private insurance offered through the employer is likely higher than that offered in lower income jobs. In making the decision to substitute private for public insurance, individuals must consider not only the quality of the public insurance option, but also the quality of the private insurance they are substituting away from.[1] Furthermore, the logic of substitution requires that employees are able to recover some of the cost of private coverage when switching to public insurance, perhaps in the form of higher wages or other improved benefits. While empirical evidence suggests that health insurance costs are passed back to workers (Gruber, 1994), it remains unclear whether recovery of these costs respond to individual or group choices of insurance. If the potential to recover group insurance expenditures through other forms of compensation responds largely to group choices of insurance, higher income individuals may not be able to recover any costs because employers have strong incentives to maintain their compensation costs in the form of private insurance coverage. These incentives include tax subsidies for employer spending on health insurance, and IRS nondiscrimination rules that require insurance to be offered to all full time employees or not at all.

Finally, advocates of public insurance argue that the entry of a new competitor in the high income insurance market has the potential to reduce private insurance prices. Faced with lower costs, children of some high income individ-

[1]Even if SCHIP and Medicaid offer comparable coverage to a private plan, it may still be a less attractive option than private coverage. Compensation to care providers is typically lower for Medicaid than private insurance, so providers are sometimes less willing to see publicly insured patients.

uals may choose to purchase a private health insurance plan for their children that provides better coverage than SCHIP. While the previously discussed mechanisms are only likely to mitigate crowd-out, the effect of lower private insurance rates may actually produce crowd-in — that is, an increase in private coverage that accompanies the increase in public coverage (Hacker, b,a).

Summarizing the discussion above, there are multiple reasons why we might expect to find a relationship between income and crowd-out. However, these reasons point to a different potential relationships, casting doubt on whether crowd-out is positively, negatively, or even non-monotonically related to income. While the posited mechanisms differ considerably, collectively they provide strong reasons that crowd-out rates calculated from low income populations may differ considerably from those in higher income populations. In the next section, I review the crowd-out literature as applied to low income populations. While more comprehensive literature reviews on the large crowd-out literature can be found elsewhere (mostly notably, Gruber and Simon (2008) and Congressional Budget Office (2007)), my review focuses primarily on methodology and discusses why the application of previous estimators to high income populations can produce biased estimates.

2.3 Past Research

Any review of past research on crowd-out necessarily begins with the canonical work of Cutler and Gruber (1996), whose paper inspired a sizable literature devoted to estimating crowd-out effects using a variety of data and estimation techniques. While Cutler and Gruber present a variety of different models in their paper, I focus here solely on the model that most closely approximates the method used by this paper.[2] Using individual-level data from the March Supplement of the Current Population Survey (CPS), Cutler and Gruber examine crowd-out during the initial Medicaid expansions of the 1987-1992 period.

Cutler and Gruber estimate two intermediate parameters: the rate at which public insurance is taken up as people become eligible, and the rate at which private insurance is dropped as people become eligible for public insurance. To obtain these estimates, Cutler and Gruber run two linear probability models of the form:

$$Private\ Coverage = \beta_1 Eligible_i + \beta X_i + \sum \alpha_s State_i + \sum \alpha_t Time_i + \epsilon_i \quad (2.1)$$

$$Public\ Coverage = \beta_2 Eligible_i + \beta X_i + \sum \alpha_s State_i + \sum \alpha_t Time_i + \epsilon_i \quad (2.2)$$

where *Public Coverage* and *Private Coverage* are dichotomous variables indicating coverage type, *Eligible* is a dichotomous variable indicating whether individual i is eligible for public health insurance, X is a set of demographic controls, and $State$ and $Time$ are state and year dummy variables. Since β_1 measures marginal take-up for private insurance (which is expected to be negatively signed) and β_2

[2] The most significant omission in this discussion of Cutler and Gruber's paper is their consideration of family-based measures of coverage. This approach captures spillovers from other family members — for example, when children are made eligible for public insurance, their parents may drop the entire family from coverage. While this is a substantively important problem, my paper solely focuses on estimating the direct crowd-out effect.

measures marginal take-up for public insurance as public insurance eligibility expands, one measure of crowd-out that can be derived from equations 1 and 2 is $\frac{-\beta_1}{\beta_2}$, which measures the fraction of individuals substituting public for private coverage as eligibility expands.

Cutler and Gruber also recognized that β_1 and β_2 were unbiased estimates of marginal insurance take-up only if eligibility for public health insurance was exogenous to the private and public coverage dependent variables. This assumption is clearly violated — low income families are more likely to be eligible for public health insurance and more likely to be uninsured, while states are highly likely to choose their level of public insurance coverage based in part on their preexisting insurance coverage profile. To address this issue, Cutler and Gruber use a "simulated instrument" for their measure of public insurance eligibility following the work of Currie and Gruber (1996a; 1996b). They begin with a nationally representative sample of children of each year and age, and apply each state's eligibility rules to calculate the fraction of children eligible for public insurance in each year/age sample. These values are then matched to each child's year of interview, state, and age to instrument for individual eligibility. Note that because the sample of children is nationally representative, it is unaffected by state-level demographic differences or differences in local economic conditions. Simulated eligibility therefore satisfies the need for an instrument correlated with individual eligibility for Medicaid that is not otherwise correlated with the demand for insurance.

Using their simulated eligibility variable, Cutler and Gruber obtain estimates of $\beta_1 = -0.074$ and $\beta_2 = 0.235$, which implies a crowd-out rate of 31%. Cutler and Gruber's estimate of β_2 furthermore implies that 23.5% of those made eligible by the public insurance expansions of 1987-1992 took up public insurance, and they note that 27% of the children who were newly made eligible for coverage in 1987 were uninsured. This implies that if the increase in coverage had only occurred among the uninsured population (i.e. there was no crowd-out), the take up rate among the uninsured would be almost 90%.

Following the publication of Cutler and Gruber's study, a host of papers challenging the original findings with different estimators and data sets emerged. One particularly popular approach was to estimate crowd-out using differences in differences (DD) estimators, an approach popularized by Ashenfelter and Card (1985). Under simple DD designs, outcomes are observed for two groups over two time periods, a pre-treatment and post-treatment period. While neither group is exposed to treatment during the pre-treatment period, the treatment group is exposed to treatment in the post-treatment period while the control group remains unexposed. Estimation of the treatment effect occurs by subtracting the average gain in the control group from the average gain in the treatment group. The objective of this design is to remove two potential sources of bias — biases from comparisons over time in the treatment group (i.e. insurance coverage is increasing/decreasing over time even in the absence of any treatment), and biases in comparing treatment and control groups resulting from permanent differences between those groups.

In contrast to the Cutler and Gruber approach, DD designs enjoy the advantage of relative simplicity in the sense that no instrument is required to obtain estimates. However, a critical assumption underlying DD analysis is that unmeasured, time-varying factors are assumed to have the same effect on treatment and control group members. When this assumption fails, the estimates are biased. This is particularly likely to be true in cases where the treatment and control groups are drawn from different populations — a situation that is true of *every* published DD design on crowd-out for children to date. Using the 1988 and 1993 Current Population Survey, Dubay and Kenney (1996) conduct a DD analysis comparing the change in insurance coverage of children relative to that of adult men. However, this approach assumes without providing justification that there are no other factors changing over time differentially for children and adult men. Notably, Dubay and Kenney obtain crowd-out estimates of 15% for children in poverty and 22% for those between 100-133% FPL. The trend of increasing crowd-out by income is consistent with the consensus of theoretical expectations discussed earlier, but the magnitude of crowd-out estimated by Dubay and Kenney are much lower than those obtained by Cutler and Gruber. Using data from the 1988 and 1992 National Longitudinal Survey of Youth, Yazici and Kaestner (2000) conducted a DD analysis, comparing change in insurance coverage of children becoming eligible to those not becoming eligible for public health insurance. They obtain a crowd-out estimate of 14.5%, consistent with the lower estimates of Dubay and Kenney. Blumberg et al. (2000) replicate this analysis using data from the 1990 Survey of Income and Program Participation, finding an even lower crowd-out rate of 5%. However, both the Yazici and Kaestner and Blumberg et al. DD analysis continue to make the assumption that time-varying factors do not differentially affect the eligible and ineligible populations. While their paper focuses on adult crowd-out rather than children, Kronick and Gilmer (2002) use a state-level DD design that compares coverage levels for four states in the early 1990s to other states in their region. Their design is admirable in using other states to construct a control unit with some affinity to the treated unit. In doing so, the control unit is less likely to be differentially affected by time-varying factors that may have contaminated earlier DD designs.

Recognizing the limitations of the original DD designs, Card and Shore-Sheppard (Card and Shore-Sheppard, 2004) measure crowd-out by exploiting discontinuous eligibility rules to identify the effect of Medicaid expansions on low income children. The 1991 Medicaid expansion (also known as the 100% expansion) extended eligibility to children born after September 30, 1983 living under the Federal Poverty Line, while the 1990 expansion (also known as the 133% expansion) extended Medicaid to children under 6 in families with income below 133 per cent of the poverty line. Using data from the 1992-93 Survey in Income and Program Participation, the 1991-92 Current Population Survey, and the 1992-93 Health Interview Survey, Card and Shore-Sheppard found that the 100% expansion increased Medicaid coverage by 10% for children born just after the cutoff date, while the 133% expansion had no effect on health insurance coverage. Under both expansions, Card and Shore-Sheppard did not find crowd-out rates that were statistically significant in any of their data sets.

Card and Shore-Sheppard's paper stands out in the crowd-out literature for several reasons. First, Card and Shore-Sheppard's crowd-out estimates continue the trend in the literature towards lower estimates — in fact, their null result is matched only by the Blumberg et al. (2000) study described earlier. Even more surprisingly, however, Card and Shore-Sheppard's finding of no effect on health insurance coverage between 100-133% FPL suggests that income can interact with coverage in highly unusual ways, for theoretical reasons discussed earlier. In short, Card and Shore-Sheppard's results provide additional empirical support for the kinds of unusual income dynamics that we observe later in analyzing the impact of Illinois' All Kids program.

While the Card and Shore-Sheppard method addresses many of the problems endemic to earlier DD studies, they are not well suited to the task of estimating crowd-out at higher income levels for two reasons. First, the Card and Shore-Sheppard method requires discontinuous eligibility for public insurance by age, something that is not true of the later SCHIP expansions. Secondly, Card and Shore-Sheppard's designs are difficult to generalize by income level because they estimate coverage effects for children born just after the cutoff date. This is unlikely to be representative of coverage effects for all children born after the cutoff date because the cost of coverage often varies drastically depending on the age of the child. Using the National Medical Expenditure Survey, Cutler and Gruber (1996) find that health care costs for infants average $2,486 per year, dropping to $399 per year for children between ages 6-9 before rising back to $930 per child between ages 15-18. These cost differentials suggest that crowd-out rates estimated using discontinuous eligibility may be highly sensitive to the choice of data set year, since each year's estimates will be conducted using children of a different age.

The lower crowd-out estimates obtained from the DD studies and Card and Shore-Sheppard's study subsequently led researchers to reexamine the original Cutler-Gruber type models with mixed results. Shore-Sheppard (2008) replicated the original Cutler-Gruber findings with Current Population Survey data, but found the results were highly sensitive to the choice of control variables. In particular, the inclusion of age*year interaction variables results in crowd-out estimates of zero and lower public insurance take-up rates than Cutler-Gruber, suggesting that omitted trends in insurance coverage by child age and state that are correlated with expansions in eligibility. This estimate of zero crowd-out is confirmed by the work of Ham and Shore-Sheppard (2005), who replicate the same analysis using data from the Survey of Income and Program Participation and find take-up and crowd-out rates of 13.6% and 0% respectively.

More recently, extensions of the original Cutler-Gruber models designed to examine the effect of the SCHIP expansions find significant levels of crowd-out. Using Current Population Survey data from 1996-2000, LoSasso and Buchmueller (2004) use the Cutler-Gruber design to evaluate the effect of SCHIP. They find a marginal take-up rate for public insurance of 8.1%, which is considerably lower than the rates found by Shore-Sheppard (2008) and Cutler and Gruber (1996) but in line with estimates from Ham and Shore-Sheppard (2005). Depending on the exact specification used, crowd-out was estimated to vary between 18 to 50%, consistent with the high crowd-out rates found by Cutler and Gruber (1996). More recently, Gruber and Simon (2008) applied Cutler-Gruber type models to the 1996 and 2001 Survey of Income and Program Participation and include the age*year interaction variables used by Shore-Sheppard (2008). Their estimates show a very low take-up rate of 5.5% for public insurance, accompanied by a direct crowd-out rate of 30% that is largely consistent with the LoSasso and Buchmueller (2004) estimates using CPS data.

Summarizing the general trends in the literature, Cutler and Gruber's initial estimates in 1996 suggest high crowd-out rates of approximately 30% for the Medicaid expansions of the early 1990s. However, estimates using alternative data sets and procedures, notably DD and discontinuous eligibility designs, find considerably less crowd-out, and a replication of the original Cutler-Gruber model finds that the original crowd-out estimates approach zero when age*year interaction variables are included. More recently however, researchers applying Cutler-Gruber models to the Medicaid expansions of the late 1990s find lower rates of insurance take-up than before accompanied by crowd-out rates around 30%. While this crowd-out rate is consistent with the original Cutler-Gruber estimates, the rates cannot be directly compared because the early 1990s expansions primarily targeted children living in poverty while the late 1990s expansions primarily targeted children between 100-200% FPL.[3]

Cutler-Gruber models provide an attractive means to estimate crowd-out for children below 200% FPL, but similar to DD and discontinuous eligibility designs, their application to the study of crowd-out in higher income populations is limited. While Medicaid will provide health insurance to some higher income

[3]Using the Survey of Income and Program Participation, Gruber and Simon (2008) find that 72.2% of the children who became eligible for public insurance between 1996 and 2002 were between 100-200% FPL.

children, almost universally these children are not representative of other children at their income level. A notable example of this selection bias appears in Tennessee's Medicaid program TennCare, which insures over 500,000 individuals who were uninsurable by private insurance due to pre-existing conditions (Chang, 2007). Estimation of crowd-out at higher income levels using Cutler-Gruber specifications is therefore likely to upwardly bias estimates of public insurance take-up and downwardly bias estimates of crowd-out because the eligible high income individuals receiving coverage are effectively forced to choose and stay in a public insurance plan. In the following section, I introduce a new method of estimating public insurance take-up and crowd-out that addresses the problems endemic to the application of difference-in-differences, discontinuous eligibility, and instrumental variables regression to higher income populations.

2.4 Methodology

In this section I present a new method to assess the effect of SCHIP using Illinois' All Kids program as our treatment unit of interest. This exposition closely follows the description in Abadie and Gardeazabal (2003) and Abadie, Diamond, and Heinmueller (2010). Supposed that we observe $J + 1$ regions, where the first region is exposed to an intervention of interest and the remaining J regions are not exposed. Then in region i and time t let Y_{it}^N represent the outcome that would be observed in the absence of the intervention for units $i = 1, \ldots, J + 1$ and periods $t = 1, \ldots, T$. Also, let Y_{it}^I represent the outcome that would be observed for unit i at time t if it is exposed to the intervention in periods $T_0 + 1$

to T, where T_0 is the number of pre-intervention periods and $1 \leq T_0 < T$.

Assuming the intervention has no effect on the outcome before the implementation period $t \in 1, \ldots, T_0$ for all units $i \in 1, \ldots, J+1$, the effect of exposure to the intervention on unit i at time t is $\alpha_{it} = Y_{it}^I - Y_{it}^N$. For untreated units (i.e. $i \neq 1$) or treated units before the intervention (i.e. $i = 1$ but $t > T_0$), the effect $\alpha_{it} = 0$. Then more generally, defining D_{it} to be an indicator value that takes the value of one if unit i is exposed to the intervention at time t (which is true if $i = 1$ and $t > T_0$) and zero otherwise, the observed outcome for any unit i at time t is $Y_{it} = Y_{it}^N + \alpha_{it} D_{it}$. Estimation of the effect on the exposed unit in each post-treatment time period $t > T_0$ is given $\alpha_{1t} = Y_{1t} = Y_{1t}^N$, with Y_{1t} observed for all time periods. The central issue of estimation therefore centers on how we estimate Y_{1t}^N, the outcome that would be observed in the treated region in the absence of the intervention of interest.

Following Abadie et al. (2010), I approach this problem by estimating Y_{1t}^N as a weighted combination of other states, chosen to resemble the demographic characteristics and health insurance profile of Illinois prior to the introduction of All Kids. This "synthetic" Illinois provides estimates of Y_{1t}^N, the outcome that would be observed in the absence of All Kids. Given J control states (i.e. states not treated with All Kids), I estimate $W = (w_1, \ldots, w_j)'$, a $(J \ x \ 1)$ vector of nonnegative weights that sum to 1. Each single weight w_j represents the weight of region j in the formulation of the synthetic Illinois, and each set of weights W produces a different synthetic Illinois.

To optimize the choice of W, let X_1 be a $(K \ x \ 1)$ vector of state level demographic and health coverage variables for Illinois, and let X_0 be a $(K\ x\ J)$ matrix containing the same variables for the J potential control regions. Also let V be a diagonal matrix of nonnegative components representing the relative importance of the different predictor variables. The vector of weights W^* that defines the combination of control regions best resembling Illinois during the pre-treatment period is chosen such that W^* minimizes $(X_1 - X_0W)'V(X_1 - X_0W)$ subject to $w_j \geq 0$ and $\sum_{j=1}^{J} w_j = 1$. While V can be selected subjectively based on previous knowledge about the relative importance of each predictor, in this application we choose V such that the path of the the synthetic outcome variable Y_{it}^N best approximates the path of the true observed outcome variable Y_{it}^I during the pre-intervention period $1 \leq T_0 < T$.

The synthetic control procedure enjoys many advantages over other methods used to estimate effects from comparative case studies. Abadie et al. (2010) note that the synthetic control procedure generalizes the DD model discussed previously. Even more generally, the procedure produces useful estimates in models with time-varying coefficients. This is not true for traditional DD designs, which allow for unobserved confounders but restricts the estimated effects of those confounders to be constant in time. In contrast, the synthetic control model allows the effects of confounding unobserved characteristics to vary with time, thus creating less model dependency.

Furthermore, the procedure directly addresses many of the issues raised in considering how earlier techniques used to estimate crowd-out could be applied to higher income populations. In contrast to the earlier DD designs where the treatment and control groups may not have been comparable (i.e. comparing children to adults), synthetic controls have the advantage of forcing the researcher to demonstrate an affinity between the treatment and control unit, in the sense that the predictor values of the synthetic control X_0W closely approximate those of the observed unit X_1 during the pre-treatment period. This affinity maximizes the chance the time-varying factors have the same effect on the treatment and control group. In effect, the procedure draws causal inference using exact matching, where the treatment unit is matched not only to the observed donor cases but also all possible convex combinations of those observed cases.

The synthetic control procedures also solves the selection issues found in the discontinuous eligibility and instrumental variables regression designs, where a direct application of those procedures to higher income populations will yield biased estimates. In both cases, bias occurs because the treatment units are not representative of the larger population at the stated income levels — discontinuous eligibility will disproportionately select a particular age group, while instrumental variables will disproportionately select individuals with pre-existing conditions in states such as Tennessee. Synthetic controls address this issue by allowing all of Illinois to be used as a treatment unit, because the All Kids treatment covers all children irrespective of their income level or preexisting conditions.

2.5 Results

In this section I present an application of the synthetic control method to Illinois' All Kids program. I begin by describing the application of this method to insurance coverage of the 400-500% FPL population in greater detail, focusing on the diagnostics used to determine model fit. I then apply the same procedure to coverage between 200-400% FPL to determine the net increase in child health insurance coverage resulting from All Kids. Next, I apply the same procedure to determine changes in public versus private coverage across the 200-500% FPL income level. These results then permit estimates of crowd-out at different income levels.

Data for these results was obtained from the annual March Supplement to the Current Population Survey from 1995-2009. CPS data is used for this paper because the synthetic control methodology requires samples designed to be representative within states, a characteristic that is notably not true of the Survey of Income and Program Participation and the Panel Study of Income Dynamics (Bureau, Bureau). I begin this analysis using 1995 data because it was the first year of the CPS following its redesign to improve the quality of its health insurance data (Swartz, 1997). All state level level estimates of the total, private, and public coverage variables are Kalman smoothed to reduce measurement error.

I begin by constructing a synthetic control for Illinois using CPS data from 1995-2006, the pre-treatment period before the enactment of All Kids. Recall that the objective is to approximate Illinois' predictors of child health insurance coverage between 400-500 %FPL using a convex combination of donor states. For control variables, I use the percentage of children covered with health insurance below FPL, from 100-200% FPL, 200-300% FPL, 300-400% FPL, and 400-500% FPL, along with a standard set of demographic controls including race, income,

	Illinois		Control States	
Variables	Real	Synthetic	Mean	Standard Deviation
Child Health Coverage, below FPL	77.52	77.49	75.99	6.12
Child Health Coverage, 1-200% FPL	78.26	78.41	77.62	5.28
Child Health Coverage, 2-300% FPL	86.34	85.60	84.16	3.82
Child Health Coverage, 3-400% FPL	89.61	89.50	88.19	2.69
Child Health Coverage, 4-500% FPL	90.80	90.78	89.93	2.25
Percent white	80.23	80.31	85.14	9.61
Mean income, U.S. dollars	36434.81	36368.34	33330.29	4037.54
Standard deviation, income	41722.77	41528.76	38049.24	4973.45
Unemployment Rate	5.63	5.62	5.21	.83
Percent College Educated	49.47	49.47	47.30	4.98

Table 2.1: Predictors of Child Health Insurance Coverage, 400-500% FPL: *All variables are averaged for the 1995-2006 period. 47 states were included in the donor state pool, excluding Alaska, Hawaii, Illinois, and the District of Columbia.*

income distribution, unemployment rate, and education. Donor states include all U.S. states except for Alaska, Hawaii, and the District of Columbia, though our results are robust to the inclusion of the those states as well. This control group allows estimates of children's health coverage in the post-intervention period between 2007-2009 because they did not enact public health expansions at that income level during the post-treatment period.

Table 2.1 summarizes the predictor values for the real Illinois and its synthetic control, averaged across the 1996-2006 pre-treatment period. First, in comparing the predictor values in columns 2 and 3, the synthetic control closely reproduces the values from Illinois during that period, suggesting a strong fit for the synthetic unit. Column 3 shows the same values calculated for the 47 states included in the donor pool, which largely represents the predictor values in the rest of the U.S. In comparing predictor values between Illinois and the U.S., we see some minor differences, most notably a \$3,105 difference in mean income. This is significant because it suggests that Illinois is not particularly unusual among U.S. states; hence the results of this study may potentially generalize to the U.S. at large.[4] Also note that for all variables, the synthetic control more closely approximates the predictor values of Illinois than the nation at large, especially with regard to income. This suggests that the synthetic unit probably serves as a better control unit than simply using all other states as a control, in the sense that the propensity of the synthetic unit to cover children between 400-500% FPL is closer to Illinois than the propensity of the nation at large. Finally, the standard deviation of the predictor variables for the 47 control states are shown on column 4. The relatively large deviations for all variables suggests that some combinations of U.S. control states will produce extremely poor synthetic controls that fail to reproduce the predictor values shown.

Table 2.2 explores the construction of the synthetic control more closely. While the estimated weights that define the combination of states used to construct the synthetic unit only approximately sum to one because of rounding, we see that 88.6% of the total weight is accounted for by four states: Maryland, Michigan, Missouri, and Oregon. These weights can be used to generate the synthetic control, which can then be used to estimate the effect of All Kids on

[4]Gruber (2008), for example, has argued that the health reform introduced by Governor Mitt Romney in Massachusetts would be unlikely to produce near-universal coverage if implemented in other states because of its unusually high income and coverage rates.

State	Weight	State	Weight
Alabama	0.016	Montana	0.001
Alaska	—	Nebraska	0
Arizona	0.007	Nevada	0.001
Arkansas	0.001	New Hampshire	0
California	0	New Jersey	0.002
Colorado	0.003	New Mexico	0.001
Connecticut	0.001	New York	0.001
Delaware	0.044	North Carolina	0.001
District of Columbia	—	North Dakota	0.003
Florida	0.001	Ohio	0.002
Georgia	0.002	Oklahoma	0.001
Hawaii	—	Oregon	0.253
Idaho	0.001	Pennsylvania	0.001
Indiana	0	Rhode Island	0
Iowa	0.012	South Carolina	0.001
Kansas	0	South Dakota	0
Kentucky	0.001	Tennessee	0.001
Louisiana	0.001	Texas	0.004
Maine	0.001	Utah	0
Maryland	0.308	Vermont	0
Massachusetts	0	Virginia	0.001
Michigan	0.180	Washington	0
Minnesota	0	West Virginia	0
Mississippi	0.001	Wisconsin	0
Missouri	0.145	Wyoming	0.001

Table 2.2: State Weights in Synthetic Illinois: *All weights rounded to three significant digits, so they may not exactly 1 as presented here. The results suggest that a synthetic Illinois is best fitted using largely a convex combination of Maryland, Michigan, Missouri, and Oregon.*

Illinois' child health insurances between 400-500% FPL.

Figure 2.1 displays the percentage of children between 400-500 FPL covered by either private or public health insurance in Illinois from 1995-2009, along with its synthetic counterpart. Notice that coverage in the synthetic Illinois very closely tracks the trajectory of the real Illinois for the entire pre-All Kids period. Combined with the high balance on all predictors as shown previously in Table 2.1, this suggests that synthetic Illinois provides a sensible approximation to the coverage that would have been observed between 2007-2009 in the absence of the All Kids SCHIP expansion. Also note that even in the absence of any expansion of public health insurance at that income level, the insurance rate rose considerably in the pre-treatment period between 1995-2006. Finally, note that after the 2006 policy intervention, coverage in the observed Illinois diverges from the synthetic. The gap between the true versus synthetic coverage represents the estimate of All Kids' effect on the total insurance rate at this income level. This gap is more dramatically shown in Figure 2.2, which plots the gap between the true and synthetic insurance rates over time. While the gap is negligible during the pre-treatment period, the large positive gap suggests that All Kids significantly increased health insurance coverage among children of this income group.

To evaluate the statistical significance of this gap, I conduct a permutation test to determine how likely a gap of this magnitude is likely to occur by chance under the null hypothesis. More specifically, I generate the distribution of mean squared prediction errors (MSPE) under the null hypothesis when no intervention has occurred, and compare this distribution to the observed MSPE of 7.5 observed

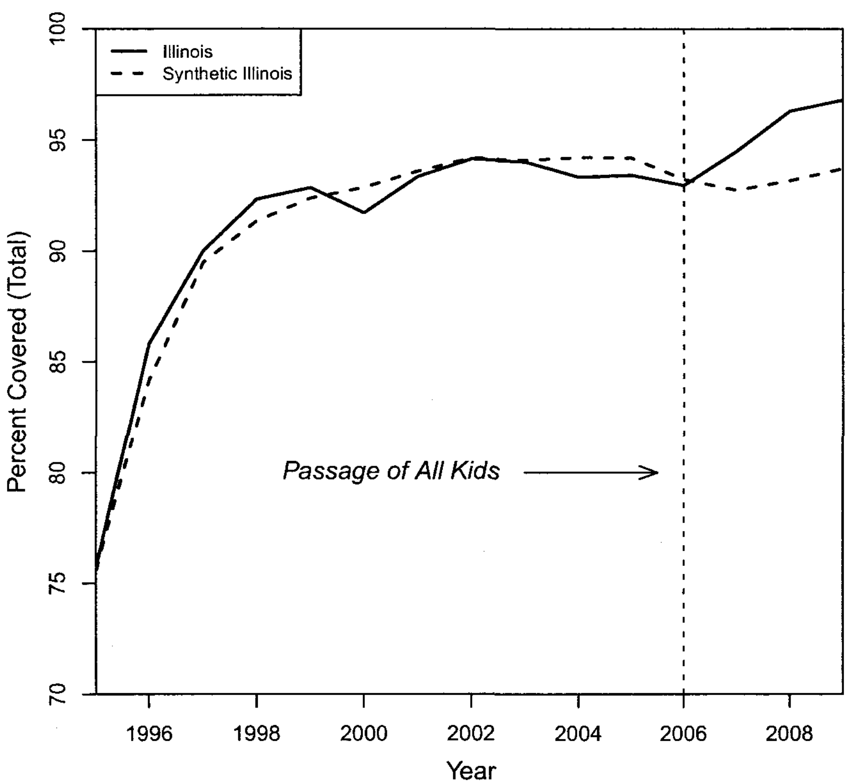


Figure 2.1: Coverage in True vs. Synthetic Illinois, 1995-2009: *Results shown for children between 400-500% FPL. Synthetic Illinois coverage was generated by taking a convex combination of coverage rates from different states using the weights shown in Table 2.*

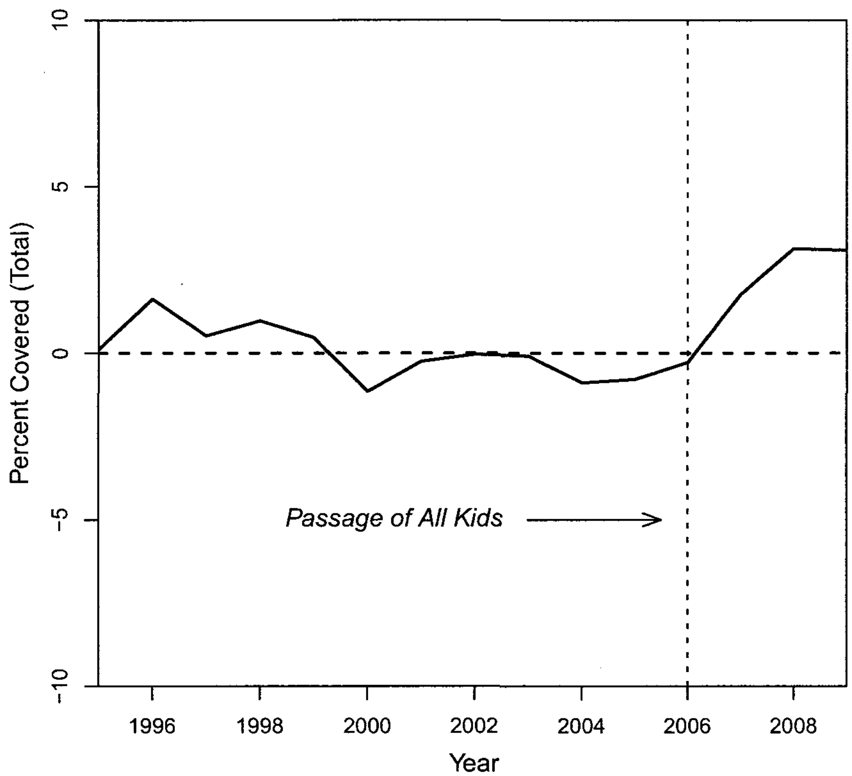


Figure 2.2: Coverage gap between True vs. Synthetic Illinois, 1995-2009: *Results shown for children between 400-500% FPL. The large gaps in the post-intervention period suggests a large effect on health insurance coverage.*

in the three post-treatment observations. To examine this distribution, note that in the pre-treatment period between 1995-2006 we have exactly 12 observations. Rather than fitting the model with all 12 observations as was done earlier, I fit the same model using only 9 training observations, saving the remaining 3 observations as testing data. Of the 220 possible permutations possible, I test 100 such permutations at random, calculating the MSPE of the remaining 3 observations for each trial. Figure 2.3 shows a histogram of the recorded MSPEs for all 100 trials, along with the MSPE we observe in Illinois. The figure shows that none of the control trials achieves such a large MSPE, and the inclusion of 100 trials in this sampling distribution suggests that we can reject the possibility of a null result at the $\alpha = 0.01$ level of significance.[5]

As an additional check on the sampling distribution shown in Figure 3, an additional test is to use a fake placebo treatment in a period before the actual treatment takes place. In effect, one fits the model using the first 6 observations, and measures the MSPE in the training period for the post-placebo period prior to the actual treatment. This procedure tests for the possibility that there is always divergence between the synthetic and observed results after even a fake treatment, however unlikely that may be. I conduct this test by using a placebo treatment in the year 2000, with results shown in Figure 2.4. As expected, the placebo treatment does not cause the true and synthetic coverage rates to diverge between 2001-2006, this providing additional evidence that the coverage gaps in the post-treatment period shown earlier in Figures 1 and 2 are unlikely to occur

[5]Our estimate of the statistical significance of this deviation is in fact likely understated, since the 100 control trials were generated with only 9 observations while the original synthetic Illinois was actually generated with 12 observations.

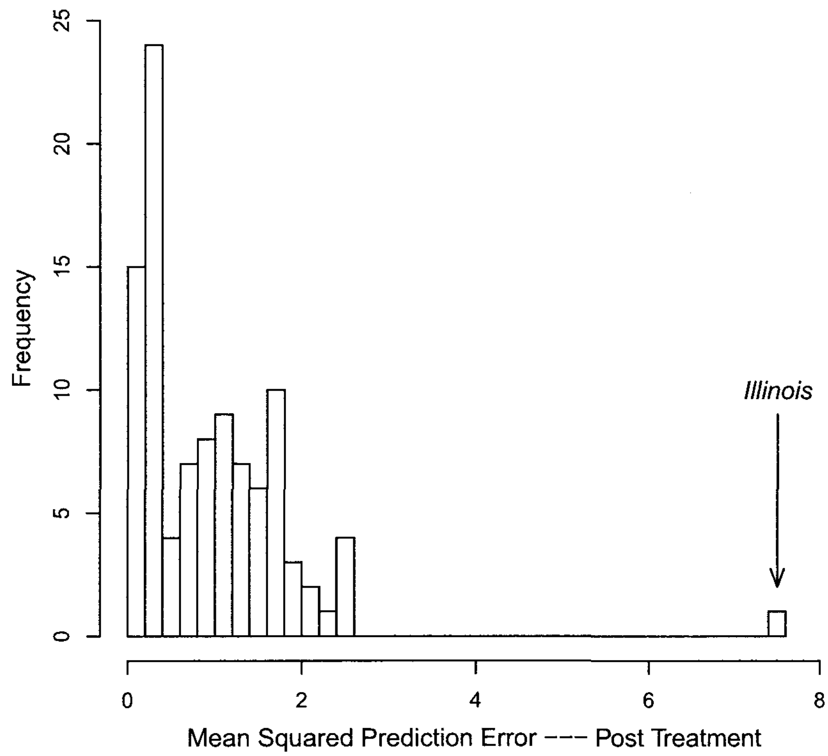


Figure 2.3: Distribution of Training MSPE, Illinois and 100 control trials: *Results shown for children between 400-500% FPL. The histogram shows that the gap observed from 2007-2009 in Figures 1 and 2 are unlikely to occur by chance, and can be rejected at the* $\alpha = 0.01$ *level of significance.*

by chance.

I repeat this analysis for the 200-300% and 300-400% income levels, but discard some control states that expanded SCHIP above 200% FPL. Results between 100-200% FPL cannot be obtained because there are no control states that did not enact an SCHIP expansion at that income level — however, coverage and crowd-out at this level have been analyzed using the methods discussed in the literature review. Figure 2.5 summarizes the results, showing the same coverage and gap plots as Figures 2.2 and 2.1 for different income levels. In both cases, the synthetic control accurately reproduces the observed coverage values during the pre-treatment period.[6] However, estimates in the post-treatment period differ substantially. Between 200-300% FPL, the synthetic control suggests that coverage would have increased in the post-treatment period even in the absence of All Kids. Nevertheless, All Kids appears to have increased health coverage even beyond the expected increase. Between 300-400% FPL however, the story is quite different. All Kids appears to have had no net effect on health insurance coverage at this income level, and the increase in coverage shown in the synthetic control between 200-300% resulting from factors other than All Kids also appears to be absent.

Table 2.5 summarizes the net effect of SCHIP on insurance coverage by income, tabulating the net increase in health insurance coverage by year and income. Between 200-300% FPL, total coverage increased by 1.61% after three years above what would have been expected in the absence of any policy inter-

[6]Predictor values are also well reproduced, however, to save space I omit weight and predictor value tables for the remaining models in the paper.

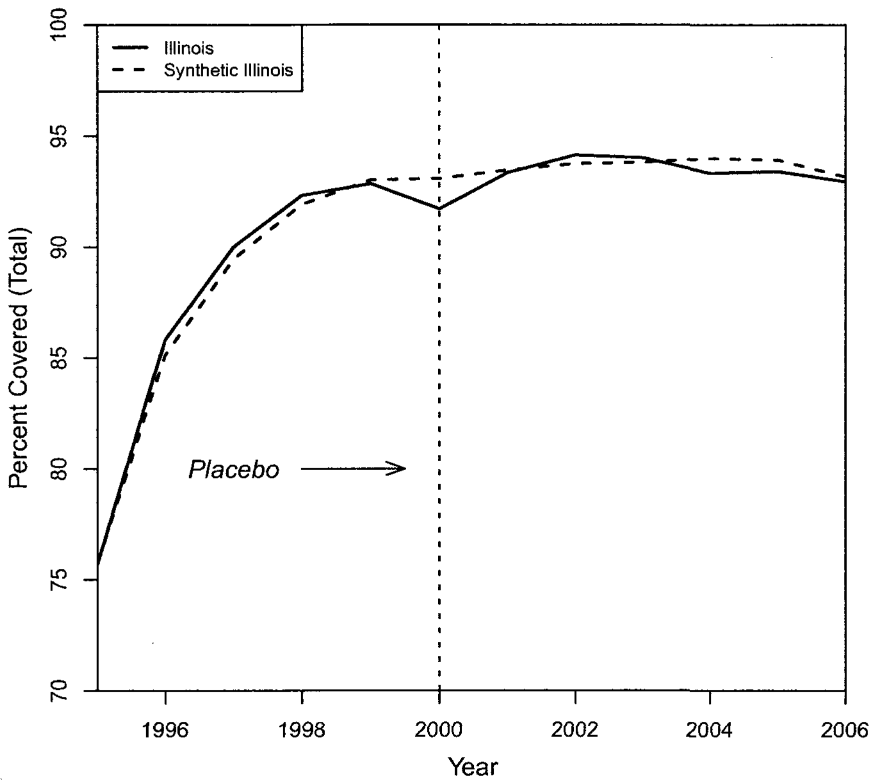
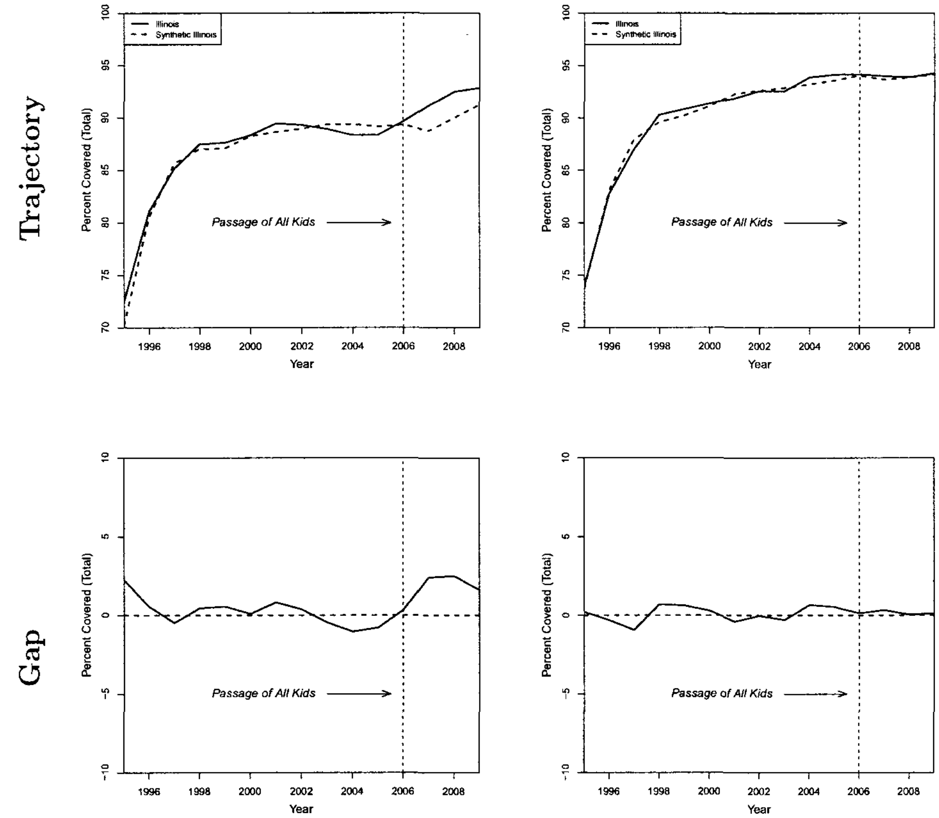


Figure 2.4: Placebo in Time — True vs. Synthetic Illinois, 1995-2006: *Results shown for children between 400-500% FPL. Here we falsely incorrectly assume that the All Kids treatment occurred in 2000 rather than 2007, and refit the model. The synthetic control still closely tracks the observed coverage rates in Illinois following the placebo treatment, providing further evidence that the coverage gaps in the post-treatment period shown earlier in Figures 1 and 2 are unlikely to occur by chance.*



200-300% FPL **300-400% FPL**

Figure 2.5: Summary of Insurance Coverage, 200-400% FPL: *Top panel shows observed vs. synthetic coverage for 200-300% and 300-400% FPL respectively, while the bottom panel shows the corresponding gap plots. At both income levels, the synthetic control accurately reproduces the observed coverage values. Coverage increases between 200-300% FPL, but the coverage increase between 300-400% FPL is statistically insignificant.*

FPL	2007	2008	2009
200-300%	2.40	2.49	1.61
300-400%	0.33	0.04	0.12
400-500%	1.76	3.14	3.10

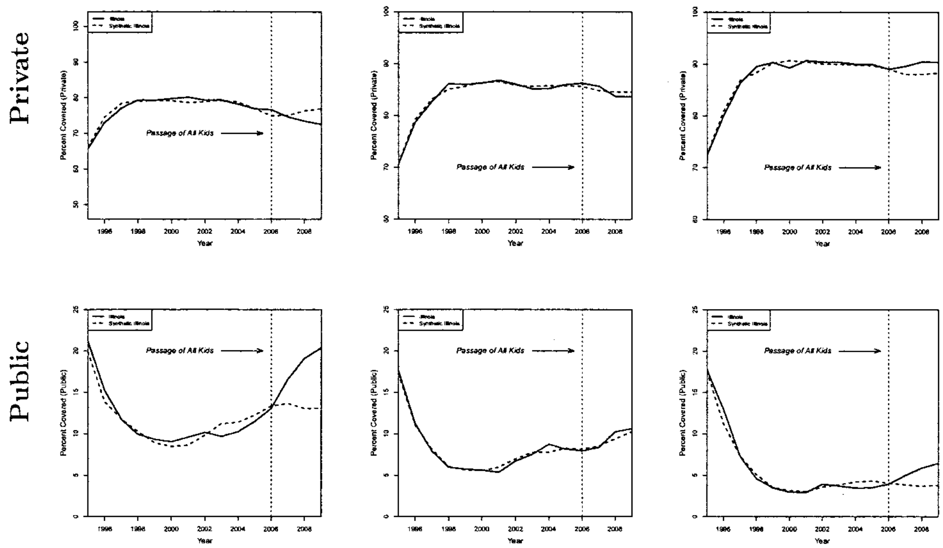
Table 2.3: Summary of Change in Total Insurance Coverage: *Numbers represent percent increases in net coverage by year, calculated as the difference in coverage between the observed and synthetic control. Differences for 300-400% FPL are not statistically significant at the 0.05 level of significance.*

vention. This increase covered 18.4% of the uninsured population at that that income level, so increase is substantively large. Between 300-400%, increases in coverage were not statistically significant. Between 400-500% FPL, the 3.1% increase in net coverage covered 49.2% of the uninsured. Health insurance coverage therefore does not appear to change monotonically with income.

I now extend the analysis presented above, estimating the same models separately for private and public health insurance coverage from 200% to 500% FPL with two changes. First, I replace the original coverage control variables for private or public versions of those same variables, as appropriate. Secondly, following Cutler and Gruber (1996) and Gruber and Simon (2008), I treat cases where an individual claims to have both private and public coverage together as cases where the individual is making a transition from private insurance to public insurance. While there are generally few cases where an individual claims both forms of insurance, this assumption will generally produce higher estimates of crowd-out than the alternative of discarding those cases altogether.

Figure 2.6 plots observed and synthetic coverage rates for private and public insurance between 1995-2009. Private insurance coverage is shown on the top panel, while public insurance coverage appears on the bottom panel. In all 6 cases, the synthetic control continues to closely reproduce the corresponding coverage rates observed in Illinois, providing evidence of good model fits. Also of interest is the increase in public insurance and the decrease in private insurance evident from at all income levels during the late 1990s. This suggests that there was some period of crowd-out prior to the implementation of SCHIP, whereby children substituted private insurance for public insurance. Figure 2.7 plots effect of All Kids on coverage in the same manner as Figure 2.6, calculating the effect as the difference between the observed and synthetic private/public insurance coverage. Notably, there exist large gaps in the post-treatment period for both private and public coverage from 200-300% FPL and 400-500% FPL.

Tables 2.5 and 2.5 present the effects shown in Figure 2.7 numerically for public and private insurance respectively during the post-treatment period. These values are needed to calculate crowd-out estimates in relative (percentage) terms, but they are interesting by themselves because they allow us to look at crowd-out in absolute terms as a proportion of the entire population. Between 200-300% FPL, note that private coverage drops by 4.37%. Substantively, this is a large amount — almost three times as much as the increase in overall coverage shown in table 2.5. In contrast, private coverage reduction between 300-400% FPL was a much more modest 0.96%, an important point that I return to shortly.



200-300% FPL **300-400% FPL** **400-500% FPL**

Figure 2.6: Private/Public Coverage in True vs. Synthetic Illinois: *From left to right, top panels show results for private coverage ranging from 200 to 500% FPL, while lower panels show results corresponding results for public coverage by income.*

FPL	2007	2008	2009
200-300%	2.95	6.05	7.25
300-400%	-0.17	0.83	0.40
400-500%	1.10	2.19	2.64

Table 2.4: Summary of Change in Public Insurance Coverage: *Numbers represent percent increases in public coverage by year, calculated as the difference in coverage between the observed and synthetic control.*

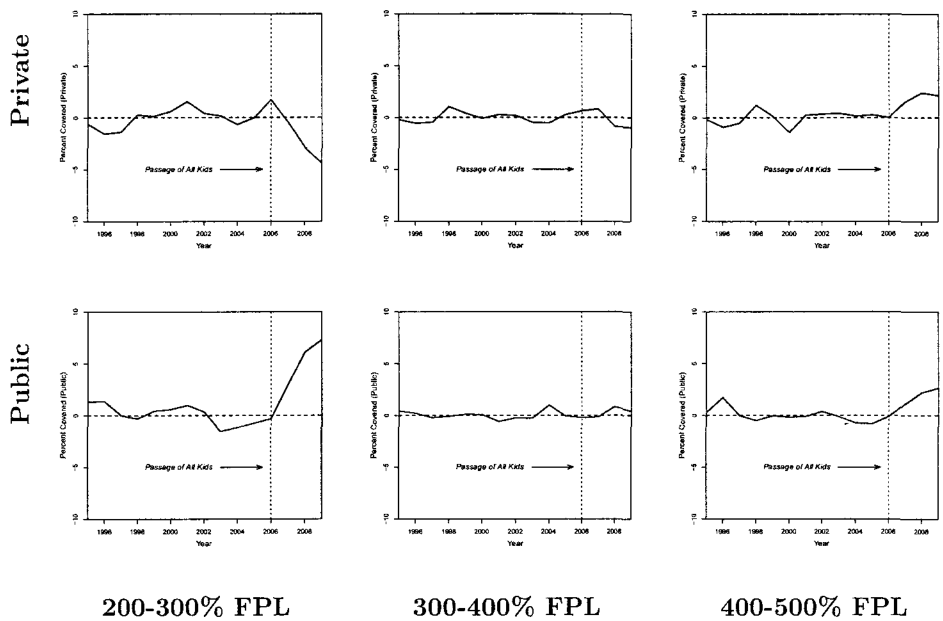


Figure 2.7: Private/Public Coverage Gaps in True vs. Synthetic Illinois: *From left to right, top panels show results for private coverage ranging from 200 to 500% FPL, while lower panels show results corresponding results for public coverage by income.*

FPL	2007	2008	2009
200-300%	-0.46	-2.86	-4.37
300-400%	0.83	-0.85	-0.96
400-500%	1.55	2.39	2.13

Table 2.5: Summary of Change in Private Insurance Coverage: *Numbers represent percent increase in private coverage by year, calculated as the difference in coverage between the observed and synthetic control.*

The crowd-out rate is defined as the fraction of children taking up public health insurance who, in the absence of a public health insurance option, would have taken up private health insurance instead. It therefore follows the formula:

$$Crowd - out = \frac{-\Delta private}{\Delta public}$$

where Δpublic and Δprivate are the entries found in Table 2.5 and 2.5 respectively. Applying the equation, we end up with the crowd-out rates shown in Table 2.5. Three results emerge from this analysis. Between 200-300% FPL, crowd-out reached 60% by 2009 — an estimate that is larger the magnitude of effect found by Cutler and Gruber (1996) and Gruber and Simon (2008) for Medicare below 200% FPL, but actually in line with the estimates speculated by the Kaiser report (on Medicaid and the Uninsured, 2007). Next, I obtain estimates of crowd-out between 300-400% that are absurdly high — a ratio of 2.4 implies that for every person taking up public insurance, 2.4 people are dropping private insurance. However, our discussion of absolute levels of crowd-out suggest that this ratio is driven by a numerator and denominator that are both very small, resulting in highly unstable estimates. Substantively, we therefore observe no crowd-out at this income level. Finally, between 400-500% FPL, we consistently observe negative ratios, implying crowd-in rather than crowd-out. This suggests that All Kids not only increased public coverage by 2.64%, it also simultaneously increased private coverage by 2.13%.

FPL	2007	2008	2009
200-300%	0.16	0.47	0.60
300-400%	4.94	1.03	2.40
400-500%	-1.40	-1.09	-0.81

Table 2.6: Estimates of crowd-out: *Calculated as -Δprivate/Δpublic from Table 4 and Table 5.*

2.6 Conclusion

Health economists widely believe that as eligibility levels for public health insurance programs are expanded to cover higher incomes, the possibility of substitution from private to public insurance increases because the expansions increasingly target those with access to private insurance. However, there exist many reasons why this relationship may not be so simple — a conclusion that previously drew support from the empirical work of Card and Shore-Sheppard (2004). In examining Illinois' All Kids program, I find additional evidence that the relationship between income and crowd-out is much more complex. More specifically, I find that for children between 200 to 300% of the Federal Poverty Line, the All Kids program produced an increase in health insurance coverage of 2% with 60% crowd-out. For children between 300-400% of FPL, All Kids produced no increase in overall coverage and no crowd-out. Finally, I find that for children between 400-500% FPL, SCHIP produced a 3-4% increase in coverage with crowd-in.

These findings are of substantive interest because there is clearly a desire by some members of Congress to continue expanding SCHIP eligibility to higher income populations, as seen in the recent 300% expansion passed with the 2009 SCHIP Reauthorization Act. My result suggests that this recent expansion should have a moderate effect on improving insurance coverage, at the cost of a crowd-out rate that is double the rate previously estimated for the 200-300% FPL population. Expansions beyond this income level however will not result in additional crowd-out, and may in fact lead to crowd-in between 400-500% FPL.

Despite the enormous problems posed by the threat of crowd-out, the presence of substantial crowd-out by itself does not necessarily answer the question of whether public expansions constitute sound or unsound public policy. Alternatives to public expansions also imperfectly target the uninsured — using a microsimulation model, Gruber (2008) finds that public expansions provide higher coverage at lower cost under SCHIP than alternatives such as tax credits. Furthermore, research suggests that public expansions improve health outcomes in cost-effective ways. For example, Currie and Gruber (1996b) find that the expansion of Medicaid eligibility to pregnant women raised Medicaid expenditures by $840,000 per infant life saved. This number is about half the value of life estimated by Manning et al. (1989), who use data from studies of willingness to pay for a small change in the probability of survival to estimate a value of life of $1.66 million.

In attempting to provide coverage and crowd-out estimates at higher income levels, my research poses several questions of a political nature. First, in considering the recent 200-300% SCHIP expansion, this paper questions whether a crowd-out rate of 60% is an acceptable price for 2% more coverage. Secondly, if crowd-out is a problem, is crowd-in a policy goal that is worth pursuing with public expansions? My research suggests that while substantial crowd-in occurs when public expansions target the 400-500% FPL, crowd-in does not result from expansions at lower incomes. Given the political infeasibility of extending coverage to such high income levels without first expanding coverage at lower levels, the path to this goal appears immensely challenging.

CHAPTER 3

The Structure of Utility in Spatial Models of Voting

3.1 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, 1997; McCarty et al., 2006). Recently ideal point models have also been applied to voting in non-legislative voting bodies such as the United Nations (Voeten, 2000), elections (for example, Herron and Lewis, 2007), 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 temporal 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.

In this paper, we seek to 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 point from their roll call vote (Kalandrakis, 2008). Moreover, it is not obvious that the data to which are 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 underlying 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 *no*) is governed by a systematic spatial voting component and an 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 their ideal policy than the alternative that is located farther from their ideal policy. The general form of the random utility function is:

$$U_j(O_j; X) = F(||O_j - X||) + \epsilon_j$$

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)o alternatives in the same policy space, F is a given monotonically decreasing function, 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 which 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 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., $F : x \to -x^2$), 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., $F : x \to \exp(-\frac{1}{2}x^2)$), then the probability of choosing the further away alternative goes to 1/2 as X is moved away from both O_j and O_n. As we discuss in greater detail in section

2, this difference has important implications for how actors respond to different alternatives. Typically, the central objective in fitting ideal point models is to estimate the ideal point (X) of each actor.[1] In this paper, we focus on the estimation of features of F.

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 to us 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 and as 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. Actors having more intense spatial preferences are relatively more likely to select the closer alternative in any situation 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

[1] In 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).

evidence that extremist legislators have higher-intensity utility functions than their moderate counterparts. This last result is substantively important to the formation and interpretation of law. It suggests that extremists are ideologically rigid whereas moderates are more likely to consider influences that arise outside liberal–conservative conflict.

Thirdly, we consider a model where the intensity of legislator preferences can vary asymmetrically. Legislator utility functions are permitted to skew to their left or right, but are also permitted to remain symmetric. Our results tentatively suggest that in the U.S. Congress, conditional on party, as legislators become more conservative their sensitivity to policy alternatives on the right increases. Correspondingly, liberal Democrats and moderate Republicans are the two groups who are most sensitive to policy alternatives to their left.

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. Our results here are based solely on a one-dimensional model, which accounts for most behavior in both the U.S. Congress and Supreme Court. Multiple dimensions introduce greater complexity and difficulty in comparing estimates across models, and we leave the comparison in higher dimensional spaces for future work.

We begin our paper by developing a model that nests Quadratic and Gaussian utility. This discussion begins with a substantive motivation of how the proposed change in the utility function can affect choice and then proceeds 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 to motivate the importance of the problem, we test the theory that extreme legislators will also have more intense preferences on four recent U.S. Senates and the U.S. Supreme Court. We conclude our paper with a discussion of of a model where the intensity of legislative preferences can vary in an asymmetric manner.

3.2 Gaussian vs. Quadratic Utility

3.2.1 Introduction

What does it mean for utility functions to be Gaussian or quadratic? Figure 3.1 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

[2]The 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.

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.

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.

In developing the NOMINATE model of ideal point estimation, Poole and Rosenthal were heavily influenced by the psychological experiments of Shepard, Nosofsky, and Ennis (Shepard, 1986; Nosofsky, 1986; Ennis, 1988), who examined how people judged similarity between stimuli such as light and sound intensity. Shepard (1987) found that when people judged the similarity between stimuli,

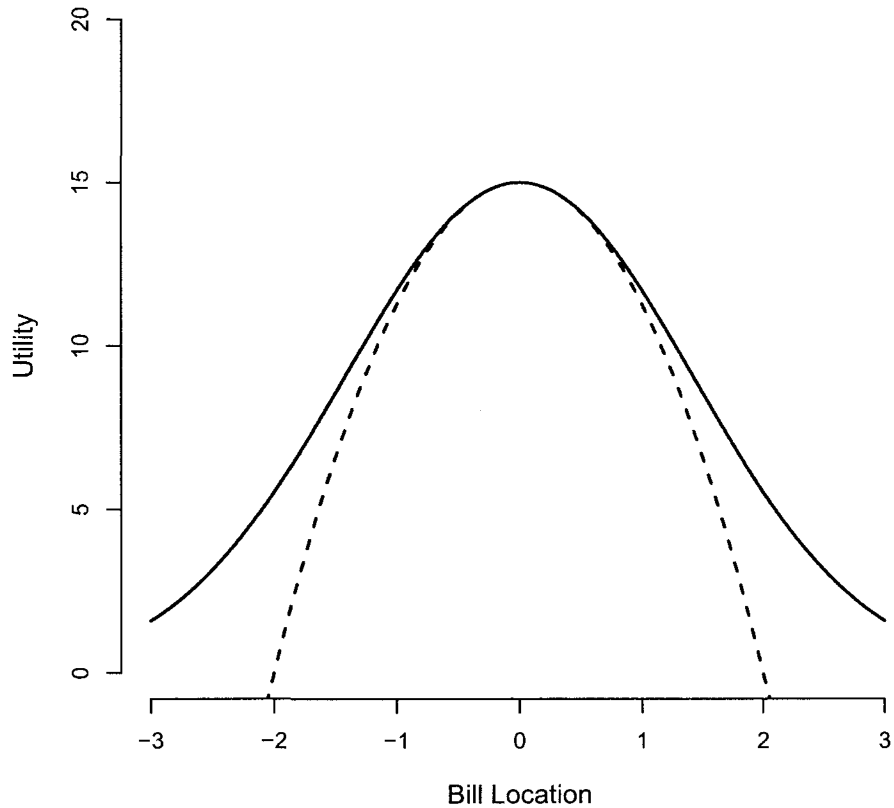


Figure 3.1: NOMINATE and IDEAL utility functions. *Lines show the deterministic utility functions assumed by NOMINATE (solid line) and IDEAL (dotted line) for a legislator with an ideal point of 0.*

they appeared to use an exponential response function. 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 $e^{-kA^2} = e^{-k(X_i - Oj)}$, where A is the distance measure between the ideal point and stimulus. Spatial models of voting can therefore be thought of as derivations of the Shepard–Nosofsky–Ennis similarity model.

3.2.2 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 ($i = 1, ..., p$), and q denote the number of roll call votes ($j = 1, ..., q$), and l represent the two possible choice 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 bill j. Then in the quadratic utility model, the utility that legislator i derives from voting yea or nay on bill j is:

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

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

In 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 are:

$$U_{ijy}^{Norm} = \beta exp\{-0.5 * w * (X_i - O_{jy})^2\} + \epsilon_{ijy}$$

$$U_{ijn}^{Norm} = \beta exp\{-0.5 * w * (X_i - O_{jn})^2\} + \epsilon_{ijn}$$

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

$$U_{ijy}^{Norm} = \beta \sum_{i=0}^{\infty} \frac{(-0.5 * w * (X_i - O_{jy})^2)^i}{i!} + \epsilon_{ijy}$$

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 jth 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 jth roll call is $Pr_{ijy} = \Phi[u_{ijn} - u_{ijy}]$.

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:

$$U_{ijy}^{Mix} = \beta \sum_{i=0}^{1} \frac{(-0.5 * w * (X_i - O_{jy})^2)^i}{i!} + \alpha\beta \sum_{i=2}^{\infty} \frac{(-0.5 * w * (X_i - O_{jy})^2)^i}{i!} + \epsilon_{ijy}$$

$$U_{ijn}^{Mix} = \beta \sum_{i=0}^{1} \frac{(-0.5 * w * (X_i - O_{jn})^2)^i}{i!} + \alpha\beta \sum_{i=2}^{\infty} \frac{(-0.5 * w * (X_i - O_{jn})^2)^i}{i!} + \epsilon_{ijn}$$

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. α can also be interpreted as the level of evidence supporting the Gaussian utility function in the data being estimated.

We use non-informative priors on the legislator and vote parameters. Given the p x 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:

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

where priors are uninformative and assumed to be distributed:

$$p(O_{jy}) \sim N(0, Inv - \chi^2)\forall j \in 1, ..., q$$

$$p(O_{jn}) \sim N(0, Inv - \chi^2)\forall j \in 1, ..., q$$

$$p(X_i) \sim N(0, Inv - \chi^2)\forall i \in 1, ..., p$$

$$p(\alpha) \sim Uniform(0, 1)$$

and 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.

3.2.3 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. 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 as a burn-in and thinning every 10th iteration. Posterior distributions from both data sets are shown in Figure 3.2. 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. These results are fully consistent with the theoretical expectations of the model. Under the quadratically generated Monte Carlo, α had a posterior mean of 0.265 and a posterior standard deviation of 0.087. While these numbers are not completely consistent with our expectation of an α of zero, the discrepancy can be explained by the fact that fitting a Gaussian utility model to quadratic data will still result in estimates that are somewhat reasonable given that they will have high similar utilities in all but the tail areas. The point to be emphasized from the Monte Carlo results is that the estimator is able to distinguish data from two similar but distinct utility functions.

We then applied the estimator to the 109th Senate, 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 3.3. 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

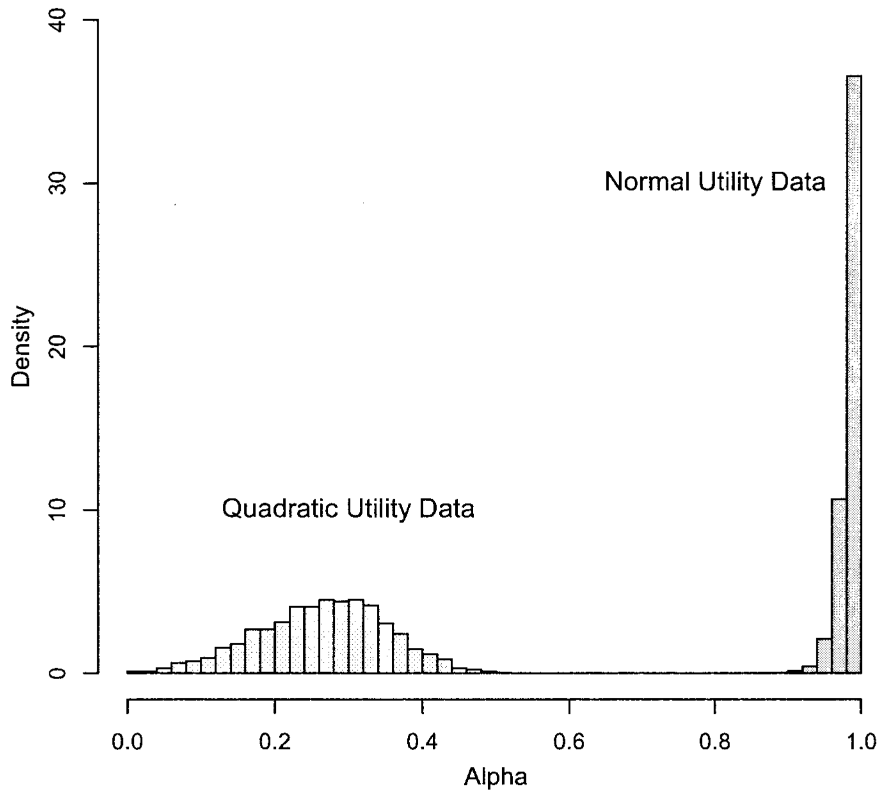
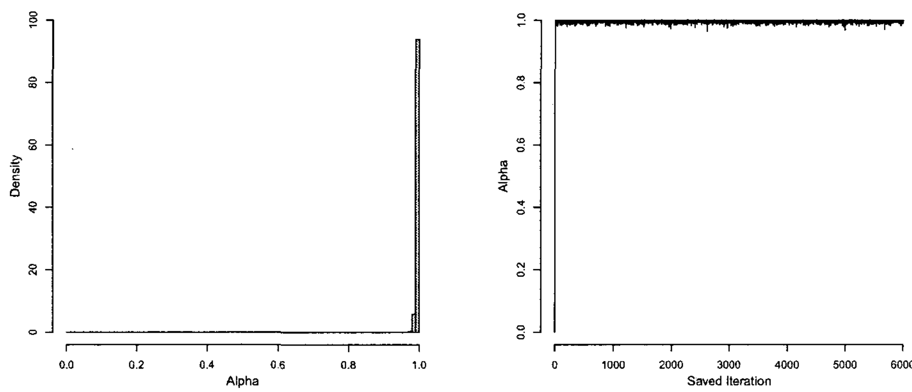


Figure 3.2: Posterior Distributions of Alpha Parameter. *Distributions are estimated Alpha parameters from two Monte Carlo data sets of 100 legislators and 500 roll calls. The distribution to the right is estimated from data generated from NOMINATE utilities, while the distribution to the left is estimated from data generated from quadratic utilities.*



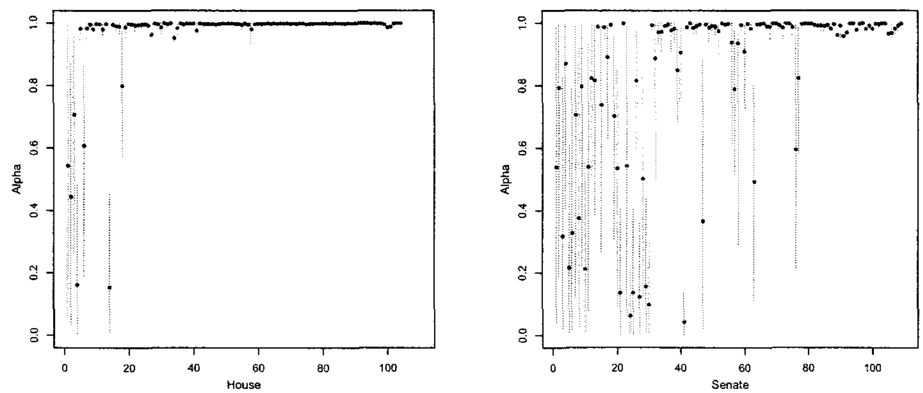
α **posterior** α **traceplot**

Figure 3.3: Posterior and Traceplot 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.*

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 however that while 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". In fact, there are substantively no differences in the legislator coordinates recovered between either the Gaussian, quadratic, or mixture models; ideal points recovered by the mixture model correlate at 0.958 with those obtained from the Gaussian model, and 0.950 with those obtained from the quadratic model.

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.[3] Figure 3.4 summarizes these results, plotting the posterior means and empirical 95 per cent confidence 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

[3]For these estimates, 6,000 iterations were used with a burnin of 1,000 iterations, and starting values for all parameters were drawn randomly.



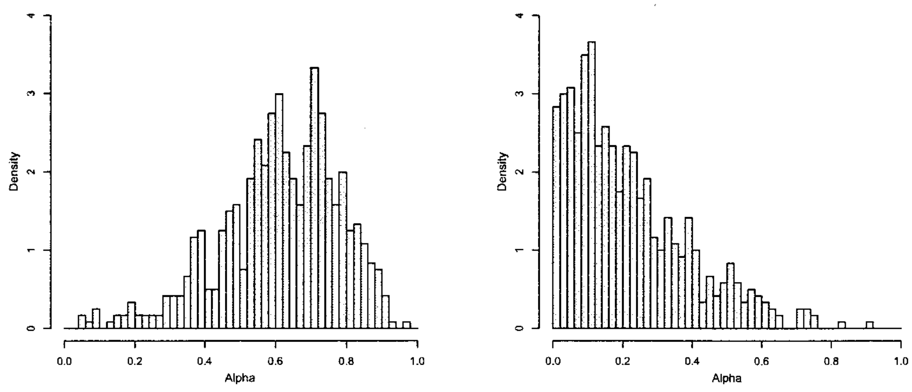
U.S. House **U.S. Senate**

Figure 3.4: 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 confidence intervals of α. $\alpha = 1$ is consistent with Gaussian utility while $\alpha = 0$ is consistent with quadratic utility.*

four times the size of the Senate. We plot the posterior densities of two early Congresses — the 6th House and the 5th Senate — in Figure 3.5. 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.

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. [4] Our results for the Supreme Court are consistent with

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



6th House **5th Senate**

Figure 3.5: 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.*

Gaussian utility, with a posterior mean of 0.998 for α and a posterior standard deviation of 0.002.[5]

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 across the world. To test this hypothesis, we conducted the same estimation for four different sources of choice data: the U.S. Supreme Court, the French Republic, the European Par-

[5]In 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 between quadratic and Gaussian utilities. This new result clearly suggests that a lower bound on the size of the roll call matrix needed before reasonable estimates of α can be estimated

liament, and the California Legislature. Our results, summarized in table 3.2.3, is consistent with our hypothesis that choice data outside the U.S. legislative context also appears to be fit well with normal utility functions.

3.3 Extremists and Intensity

3.3.1 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 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. 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

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
French First Republic	1.000	0.000
French Second Republic	1.000	0.000
French Third Republic	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, 2001-02	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, 2001-02	1.000	0.000

Table 3.1: 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.*

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 joining and dissenting from an opinion, as depicted in Figure 3.6. The deterministic components of Kennedy and Scalia's Gaussian utility functions are depicted on 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. The locations of the concurring and dissenting opinions are set at 0.4 and 0.6 respectively.

In the absence of stochastic utility, each justice joins the opinion closest to their ideal point — Kennedy is predicted to join in the concurrence, while both the low and high intensity Scalias are predicted to join the dissenting opinion.

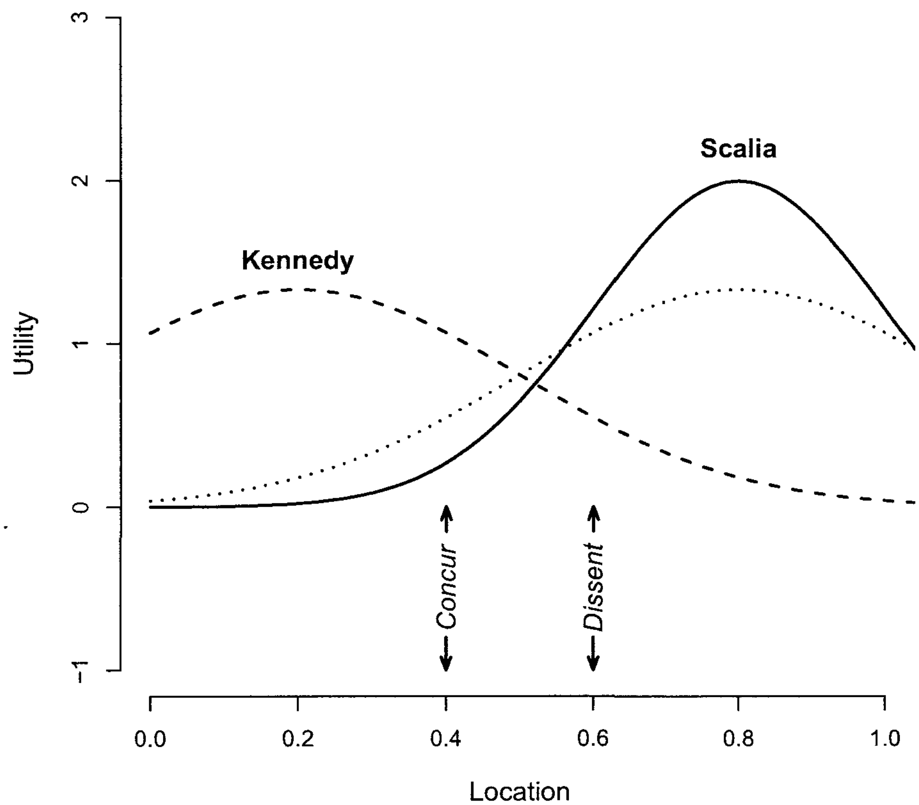


Figure 3.6: Utility of Functions of Kennedy and Two Scalias. *Kennedy's ideal point is set at 0.2, while both Scalias' ideal points are set at 0.8. While both the low and high-variance Scalias prefer to dissent, the low-variance Scalia experiences a much higher utility difference between concurrence and dissent, and is thus much less likely to err.*

But note that in the presence of random shocks to utility, the legislators have different propensities to commit 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 concurrence 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.

3.3.2 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. Again, let p denote the number of legislators ($i = 1, ..., p$), and q denote the number of roll call votes ($j = 1, ..., q$), and l represent the two possible choice 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 bill j. Then 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:

$$U_{ijy}^{Norm} = \beta exp\{-0.5 * w * (X_i - O_{jy})^2\} + \epsilon_{ijy}$$

$$U_{ijn}^{Norm} = \beta exp\{-0.5 * w * (X_i - O_{jn})^2\} + \epsilon_{ijn}$$

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:

$$U_{ijy}^{Intensity} = \beta exp\{-0.5 * w_i * (X_i - O_{jy})^2\} + \epsilon_{ijy}$$

$$U_{ijn}^{Intensity} = \beta exp\{-0.5 * w_i * (X_i - O_{jn})^2\} + \epsilon_{ijn}$$

This again leads to the standard probit formulation of the probability that legislator i votes Yea on the jth 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 jth roll call is $Pr_{ijy} = \Phi[u_{ijn} - u_{ijy}]$.

We use non-informative priors on the legislator and vote parameters. Given the p x 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:

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

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

We applied the intensity estimator separately to the 106th to 109th Senates, with the results of our estimation shown in Figure 3.7. 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 non-linear patterns in the distribution of weights.

Building on the social judgement 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 deviation from this trend at the conservative end of the 109th Senate.

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 weight parameter. This is largely due to Feingold's status in the Senate as an occasional ideological maverick who sometimes votes with Republicans against everyone else in his party. To understand the impact of this behavior on our 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 are attempting to understand why Feingold's estimates deviate

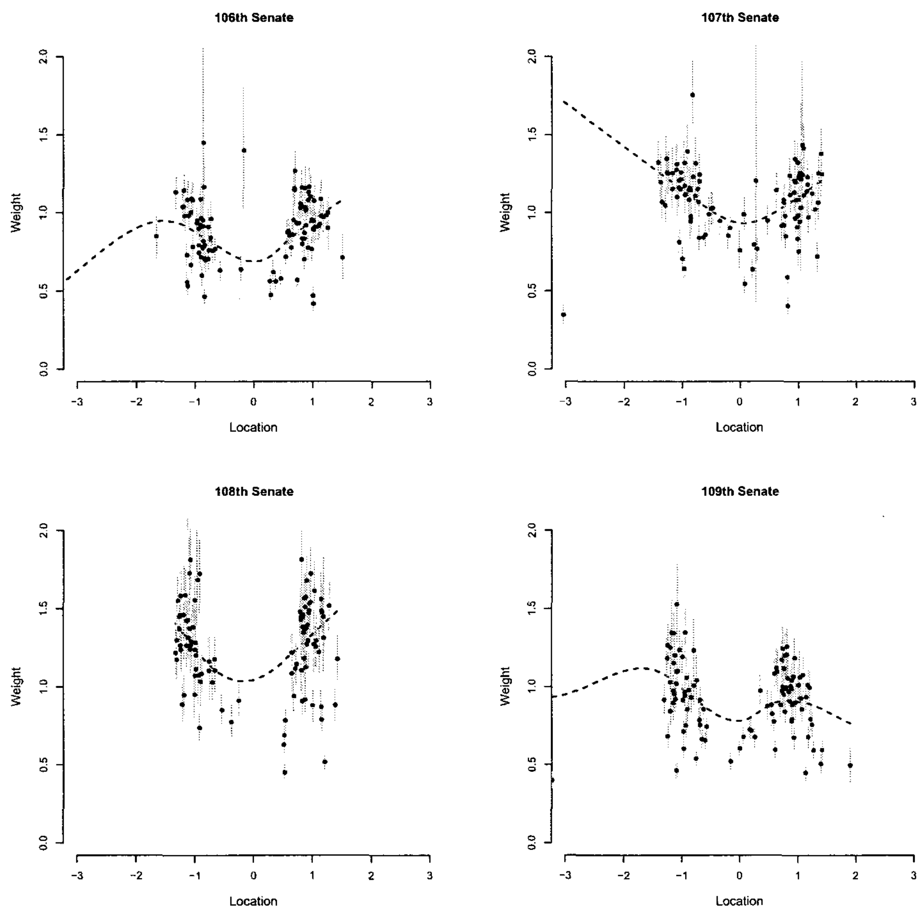


Figure 3.7: Ideal points vs. Estimated Weights: 106th to 109th Senate. *Lines reflect the 80 per cent confidence bands on each Senator's estimated weight. The dashed line is a lowess smoother of the points and provides evidence that extremists have larger weight parameters than their moderate counterparts. Russ Feingold (D–WI) is a notable outlier on three of the four graphs.*

	w=0.4	w=0.92
x=-1.31	-168.8	-245.80
x=-3.23	-159.5	-248.5

Table 3.2: Log-Likelihood for Feingold (D-WI), 109th Senate. *Feingold's estimated ideal point of x=-1.31 and weight of w=0.4 are notable outliers. We compute the log-likelihood for the same votes using Feingold and Harkin's (D-IA) ideal point and weight, and find that the difference in log-likelihood is largely the result of the different in weights.*

so substantially from those of Harkin.

Feingold's estimated ideal point of $x = -3.23$ and estimated weight of $w = 0.4$ contrast with an ideal point and weight of $x = -1.31$ and $w = 0.92$ for Harkin. We compute new log-likelihood values for each combination of weights and ideal points for Feingold's votes, and report them in table 3.3.2. As expected, Feingold's log-likelihood is maximized at Feingold's estimated weight and ideal point combination in the bottom left cell of the table. More importantly however, the table shows that as one moves from Feingold's coordinates to those of Harkin's in the top right corner, the change in weight is overwhelmingly the dominant factor explaining the discrepancy in fit.

To understand the impact of Feingold's maverick voting on his weight and ideal point estimates, consider one such vote in the 109th Senate. Here, we consider a Senate motion to concur in the House Amendment to the Senate Amendment to H.R 6111, a bill that amends the Internal Revenue Code of 1986.

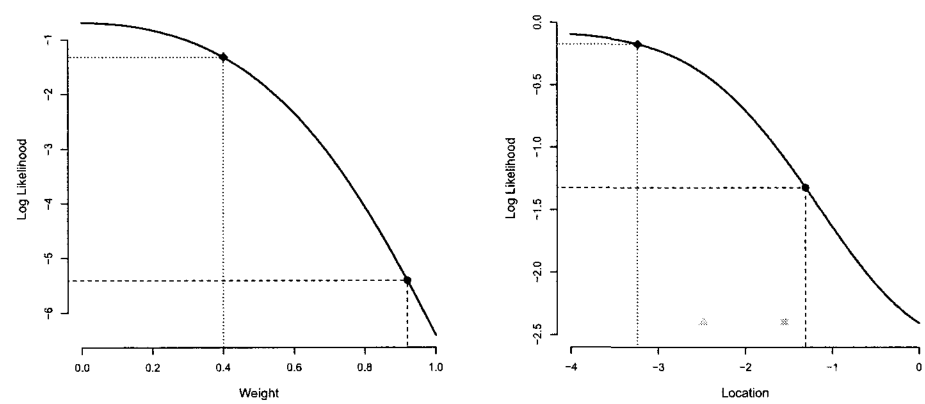


Figure 3.8: Log-Likelihood of Nay vote by Feingold (D–WI), H.R. 6111 concurrence with House. *The left panel shows the log-likelihood of a Nay vote for different values of w with $x = -3.23$, while the right panel shows the log-likelihood of a Nay vote for different values of x with $w = 0.92$. The diamond 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.*

The motion passed with a vote of 79-9, 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. Figure 3.8 on the left panel shows the log-likelihood of a Nay vote as Feingold's weight parameter changes, holding his ideal point fixed at his estimated value. 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.[6] On the figure to the right, we plot the same log-likelihood of a Nay vote as Feingold's ideal point changes, holding his weight fixed at his estimated value. This graphic suggests that 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.

Despite its smaller size, we also attempted to repeat this analysis for the 1953-2008 U.S. Supreme Court, as shown in Figure 3.9. Our results for the court suggest a pattern similar to the Senate, with the lowess-smoothed weights suggested that extremists have more sharply peaked utility functions than moderates.[7]

[6]It 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.

[7]An earlier version of this paper only used Supreme Court data from 1994-97, with inconclusive results.

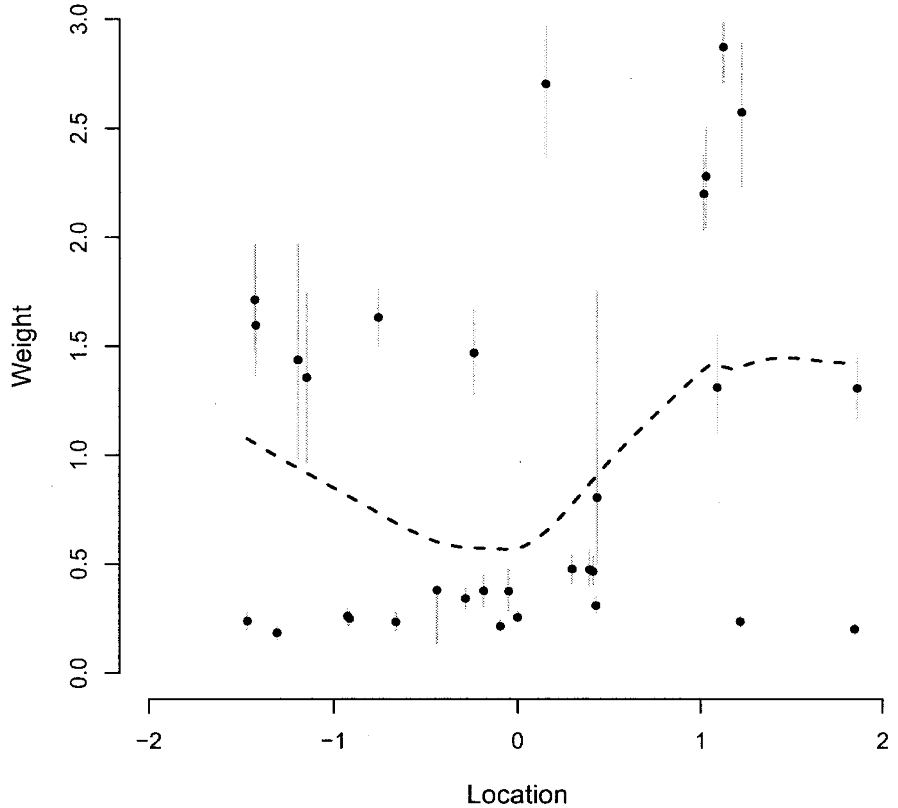


Figure 3.9: Ideal points vs. Estimated Weights: U.S. Supreme Court 1953-2008. *Lines reflect the 80 per cent confidence bands on each Justice's estimated weight. The dashed line is a lowess smoother of the points consistent with earlier results.*

3.4 Assymmetric Utility

3.4.1 What does Asymmetric Utility imply?

In the intensity model, legislator utility functions were permitted by their sensitivity to changes. By permitting variation in the weight parameter w_i for each legislator, we found that legislators at the extremes tended to exhibit greater sensitivity to policy shifts than moderates, implying that they are less likely to commit voting "errors" than their less ideological counterparts. While this model permits legislator utility functions to vary heterogeneously, it maintains the constraint of symmetric utility. Stated differently, the intensity models imply equal sensitivity to policy shifts to the left or right of the legislator, as long as the location of the choice O_j is equidistant to the alternative.

In this section, we consider the possibility of asymmetric utility, which implies that bill locations to the left or right of the legislator yield unequal levels of utility even if they are equidistant. Consider Figure ??, which shows the deterministic component of the utility function of current Senate Majority Leader Harry Reid (D–NV) in the 108th Senate as estimated by our estimator. Our estimates suggest that Reid is particularly sensitive to policy alternatives to his right, as shown by the skewed utility function. Under symmetric utility with an ideal point of -0.59, Reid should derive equal amounts of utility from an alternative located 1.0 units away from him. Instead, under the asymmetric utility function shown, Reid receives 0.636 utility from the left alternative and 0.285 from the right alternative, despite the fact that the two alternatives are equidistant from his ideal point.

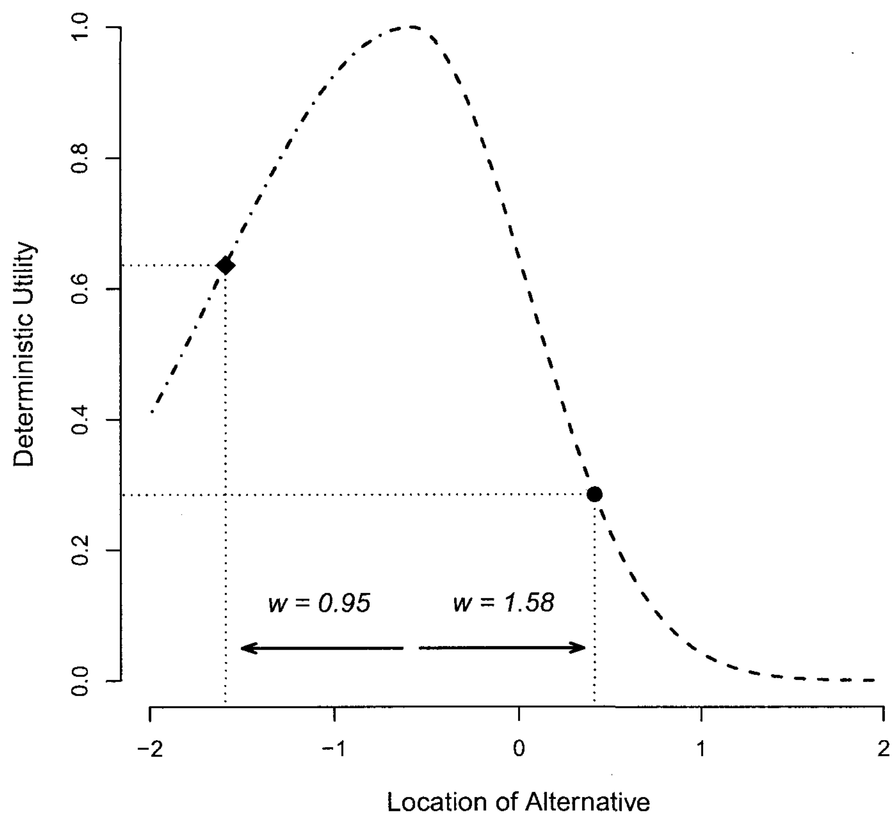


Figure 3.10: Impact of asymmetric weights for Harry Reid (D–NV), 108th Senate. *In the 108th Senate, Reid's estimated ideal point is -0.59, so utility from alternatives decreases as they move away from that point. Reid's right weight is 1.585, while his left weight is 0.951, for an estimated weight ratio of 1.67. The diamond and circle show Reid's utility from an alternative located 1.0 units to his left and right, respectively. Reid receives 2.23 times more utility from the left alternative, while the utilities would be constrained to be equal under a symmetric utility function.*

labelreid

In the formal theory literature, asymmetric utility functions have been considered in the work of Hammond, Bonneau, and Sheehan (2005). Hammond et al.'s book consider a range of different strategic models of Supreme Court decision making at different stages of the decision process. They note that the results of their models hold true whether utility is symmetric or asymmetric; however, they make no empirical claims whether Supreme Court justices exhibit asymmetric utility or not. A second application of asymmetric utility can be found in the work of Turunen-Red and Weymark (2008), who succeed in generalizing Harsanyi's social aggregation theorem for von Neumann–Morgenstern utility functions using asymmetric utility functions.

Empirically, the asymmetric utility model is prominent in the work of Narwa (2001), who attempted to determine the improvement in fit that would result from proximity models of voting with asymmetric utility. Narwa uses multinomial logit to test these models with 7 point scale survey data from the Netherlands Parliamentary Elections Survey. He found that that the asymmetric model performed slightly better than than the benchmark symmetric model and yielded 1.1 per cent more correct predictions — however, none of the improvements were statistically significant according to a χ^2 test. Our model differs from Narwa's study in several important ways. First, our model uses roll call data, which requires that ideal points and bill parameters must also be estimated, rather than reported on a survey. This complicates our analysis, but it also has the advantage of producing a continuous policy space, rather than one on a 7 point scale. Secondly, Narwa's study assume symmetry between liberals and conservatives, in the sense that if liberals are more sensitive to changes to their right, then conservatives are constrained to be equally sensitive to changes to their left and vice versa. We place no such restrictions in our model.

3.4.2 Estimation and Results

Our model builds on the intensity model developed previously, with one major addition. Rather than estimating a single weight parameter w_i for each legislator, we now estimate two weights for each legislator, a left and a right weight. When calculating the utility obtained from each alternative, the left weight is applied to all bill locations to the left of the legislator, while the right weight is applied to all bill locations to the right.

Again, let p denote the number of legislators ($i = 1, ..., p$), and q denote the number of roll call votes ($j = 1, ..., q$), and l represent the two possible choice 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 bill j. Then recall from the intensity model that given the signal to noise parameter β and individual weights w_i, the corresponding yea and nay utilities for legislator i on vote j were:

$$U_{ijy}^{Intensity} = \beta exp\{-0.5 * w_i * (X_i - O_{jy})^2\} + \epsilon_{ijy}$$

$$U_{ijn}^{Intensity} = \beta exp\{-0.5 * w_i * (X_i - O_{jn})^2\} + \epsilon_{ijn}$$

Rather than using a individual weight parameter w_i, the asymmetric model estimates left and right weights w_{il} and w_{ir} separately. Utilities are then calculated as follows. If the alternatives O_{jy} and O_{jn} are less than (i.e. to the left of) ideal point X_i, then:

$$U_{ijy}^{Asymmetric} = \beta exp\{-0.5 * w_{il} * (X_i - O_{jy})^2\} + \epsilon_{ijy}$$

$$U_{ijn}^{Asymmetric} = \beta exp\{-0.5 * w_{il} * (X_i - O_{jn})^2\} + \epsilon_{ijn}$$

Otherwise, utility is calculated with the right weight instead as:

$$U_{ijy}^{Asymmetric} = \beta exp\{-0.5 * w_{ir} * (X_i - O_{jy})^2\} + \epsilon_{ijy}$$

$$U_{ijn}^{Asymmetric} = \beta exp\{-0.5 * w_{ir} * (X_i - O_{jn})^2\} + \epsilon_{ijn}$$

This again leads to the standard probit formulation of the probability that legislator i votes Yea on the jth 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 jth roll call is $Pr_{ijy} = \Phi[u_{ijn} - u_{ijy}]$.

We use non-informative priors on the legislator and vote parameters. Given the p x 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:

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

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

We applied the asymmetric estimator separately to the 107th to 110th Senates, with the results of our estimation shown in Figure 3.11. On the X-axis, we plot ideal points of the senators, calculated as the posterior means of X_i. The posterior weight ration w_{ir}/w_{il} is plotted on the Y-axis, and an 80 per cent confidence band for each legislator's weight ratio is displayed. A weight ratio of 1 implies symmetric weights. Weight ratios above 1 imply greater sensitivity to choices on the right, while weight ratios below 1 imply greater sensitivity to choices on the left A lowess smoother is then applied to the points on each graph in an effort to detect nonlinear patterns in the distribution of weights.

While the trends are somewhat unclear, we note three things from our initial analysis. First, note that Russ Feingold (D–WI) is still an outlier, as shown by his extremely liberal ideal point estimates in the 107th and 110th Senate. This is consistent with our earlier analysis of how Feingold's maverick voting behavior can skew his estimates in the intensity model. Furthermore, note how the weight ratios for extremists can vary wildly — Feingold's 90 per cent bayesian confidence interval, for example, ranges from 0.10 all the way to 0.82. This is not surprising, because some of these weights are only identified by bill locations that are located

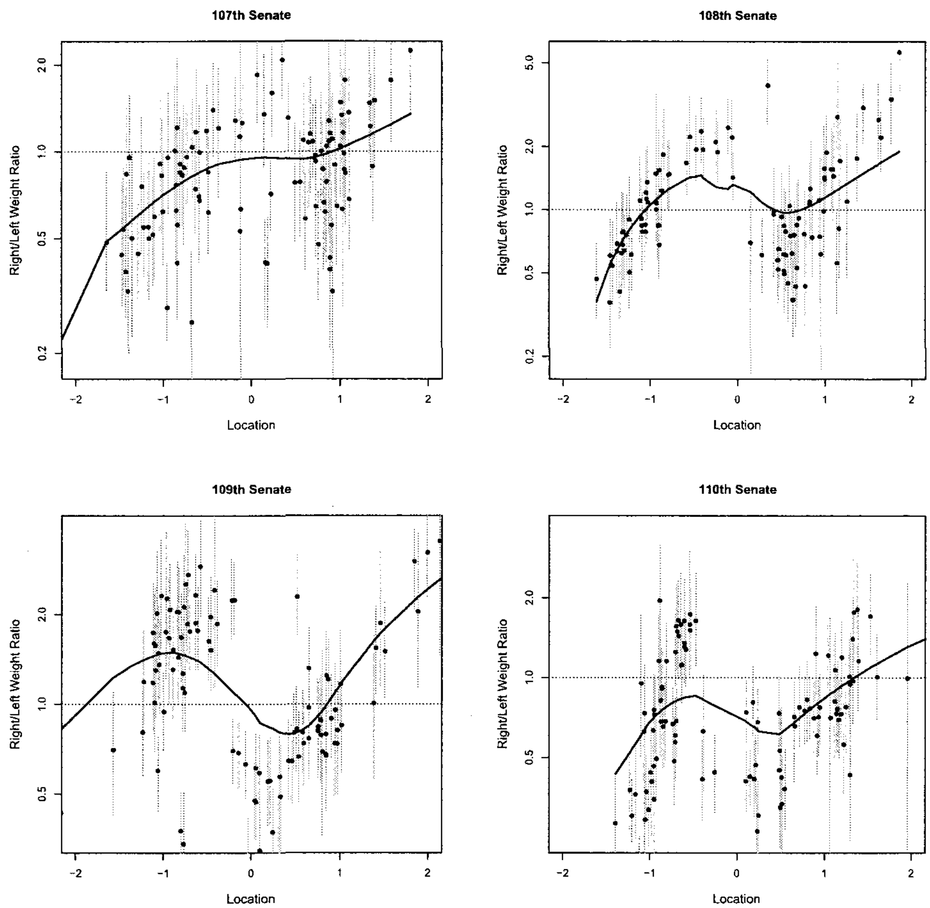


Figure 3.11: Ideal points vs. Estimated Weight Ratios: 107th to 110th Senate. *Lines reflect the 80 per cent confidence bands on each Senator's estimated weight. The dotted line shows the weight ratio at which utility functions are perfectly symmetric. The lowess smoother of the weight ratio suggests that conditional on party, as legislators become more conservative, their relative sensitivity to policy alternatives to the right increases.*

in even more extreme locations. Feingold's left weight, for example, are identified only by alternatives located to his left.

Secondly, the lowess smoother suggests that conditional on party, as legislators become more conservative their sensitivity to alternatives to the right also increases. This trend persists in all four Senates that we examined. In fact, it is the moderates of each party that are the most sensitive to policy alternatives on the left. This trend does not appear to be driven by party control of the legislature, as the 110th Senate is controlled by Democrats while the other three legislatures were controlled by Republicans.

Thirdly, consistent with the findings of Narwa (2001), we find few cases where the 90 per cent confidence interval does not cover the symmetric weight ratio of 1.0. Virtually every member of the 107th Senate has an estimated weight ratio that covers 1, though there are several moderates Democrats and conservative Republicans in the 108th through 110th that have weights clearly above 1.

We repeated this same analysis using data from the 1953-2008 U.S. Supreme Court, as shown in Figure 3.11. Here, we observe few systematic trends in weight ratios, although there appear to be a significant cluster of justices with extremely high ratios. More research here will be required to understand why these ratios appear so extreme.

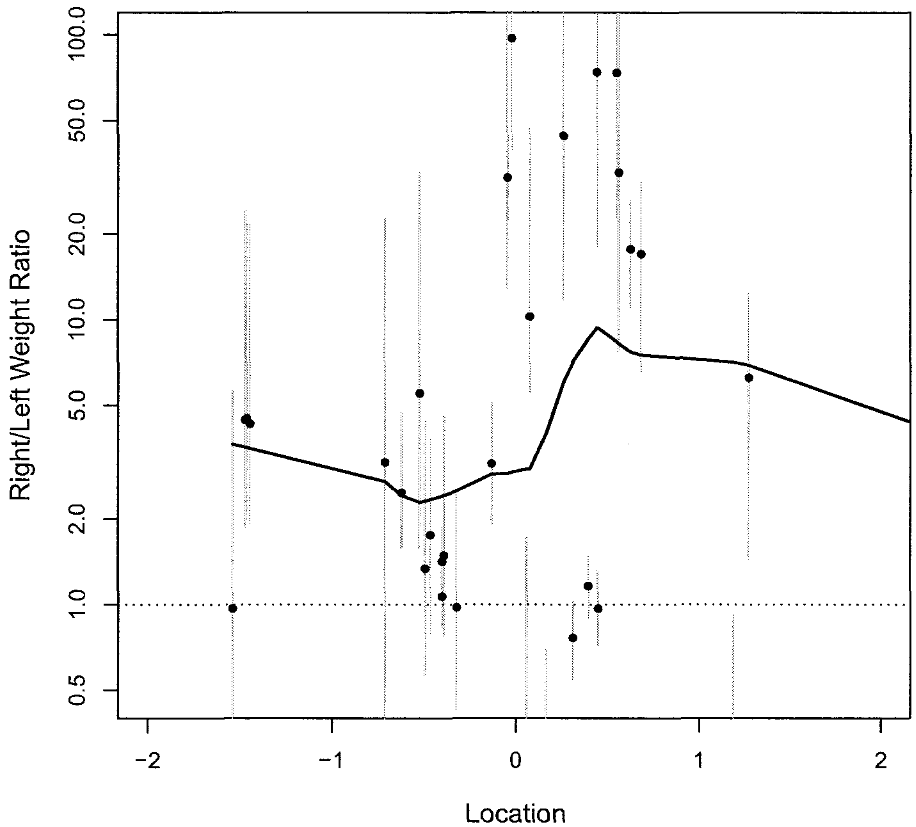


Figure 3.12: Ideal points vs. Estimated Weight Ratios: U.S. Supreme Court 1953-2008 *Lines reflect the 80 per cent confidence bands on each Justice's estimated weight. The dashed line is a lowess smoother of the points consistent with earlier results.*

3.5 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 (2008). 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.

Finally, we considered a third model extension examining the possibility that legislators have asymmetric utility functions. Here our results are somewhat inconclusive. Our results tentatively suggest that, conditional on party, as legislators become more conservative their sensitivity to policy alternatives on the right increases. Correspondingly, liberal Democrats and moderate Republicans are the two groups who are most sensitive to policy alternatives to their left. However, this trend does not appear to hold when the estimator is applied to U.S. Supreme Court data, and results from the Supreme Court are largely inconclusive.

CHAPTER 4

The Effect of Complex Voting on Ideal Point Estimates

4.1 Introduction

Ideal point models in political science have contributed much to our understanding of Congressional voting patterns. Building on the spatial theory of voting (Enelow and Hinich, 1984; Downs, 1957), these models posit that legislator ideal points can each be represented on a unidimensional or multidimensional space. Each vote is then associated with a Yea and Nay location in the same space. Utilities from a Yea or Nay vote on each roll call for each legislator are then calculated through the use of some utility function and a measure of distance between the legislator's bliss point and the bill locations. Legislators are then assumed to maximize utility given a random utility model with some random shock (McFadden, 1973). Scaling software is concerned with the estimation of these legislator and bill locations from roll call data, and popular static implementations of these procedures include NOMINATE (Poole and Rosenthal, 1985), Heckman-Snyder scores (Heckman and Snyder, 1997), and IDEAL (Clinton et al., 2004).

While theoretical differences between these different estimators exist, the ideal point estimates recovered by the majority of these estimators suggest that Congressional voting in the U.S. is remarkably stable. In particular, Congressional ideology is believed to exhibit higher levels of "constraint" versus the general electorate (Converse, 1964), is largely explained by one or two empirical dimensions (Poole and Rosenthal, 1991, 1997), has standard errors that approximate 4 per cent of the range of the actual coordinates (Carroll et al., 2009b), and produce virtually identical legislator coordinates despite different utility functions (Heckman and Snyder, 1997; Clinton et al., 2004; Carroll et al., 2009a). These results stand in stark contrast to the instability of coalitions that is predicted by the literature on social choice. With few exceptions (Plott, 1967), McKelvey (1986) has demonstrated that under many distributions of preferences and bill locations there exists some alternative that can defeat any bill. McKelvey's results thus suggest that the observed stability of roll call voting found by the ideal point literature may be somewhat overstated.

In this paper, I posit a potential reason why this discrepancy exists. I test the hypothesis that ideal point models are highly stable precisely because they ignore complex forms of voting that tend to produce more unstable ideal point estimates. In particular, I examine two sources of potential complexity that have largely been ignored in traditional ideal point models. First, ideal point models typically only examine Yea or Nay votes — however, they have largely failed to incorporate information from legislators who deliberately choose *not* to vote. Secondly, virtually all ideal point models make the assumption of independence across votes — that is, earlier votes have no effect on voting behavior on later votes, and expected voting behavior on later votes has no effect on voting behavior on earlier votes. However, this account of voting is inconsistent with "sophisticated" voting strategies in which legislators vote insincerely at early stage votes to obtain more favorable final passage outcomes than would otherwise be possible (Farquharson, 1969; Miller, 1995).

The canonical work on abstention in roll call voting is Poole and Rosenthal (1997), who find that abstention results largely from alienation. More specifically, abstentions are more likely on lopsided votes and less frequent on votes where the margin is close. In the modern era, Poole and Rosenthal also find that abstention is most prevalent on the majority side, a phenomenon they describe as "silent majorities." A key issue that arises in their work is one of pooling — different groups of legislators who may be abstaining for different reasons, either deliberately or involuntarily, cannot be separately identified. In this paper, I address this issue by focusing solely on a high-profile instance where a deliberate decision not to vote is permitted. Unlike most legislatures, the Illinois assembly allows legislators to vote Present – an procedure that became a campaign issue during 2008 Presidential campaign for Barack Obama, who voted present 129 times during his 6 years as a state senator. I examine the use of Present votes in Illinois, focusing on their potential to affect ideal point estimates.

The potential impact of sophisticated voting on ideal point estimates has also been studied by Poole and Rosenthal (1997), who find little improvement in fit using an ideal point model where some sophisticated voting is accounted for.[1]

[1] More specifically, Poole and Rosenthal estimate a model where each bill has two cutlines, permitting an ends-vs.-middle type coalition to be accounted for in a spatial framework.

Their work was subsequently challenged by Clinton and Meirowitz (2004), who conduct an analysis of the Compromise of 1790 through roll call analysis. The central insight of Clinton and Meirowitz is that amendments are not fully independent across votes on the same voting tree; early stage Yea or Nay locations should therefore be constrained to be identical to late stage vote locations based on the substantive ordering of the amendments and votes. After conducting their estimates, Clinton and Meirowitz found that their estimates of each legislator's induced preferences differed considerably from unconstrained estimates obtained through NOMINATE, a leading ideal point implementation by Poole and Rosenthal. In particular, correlations of induced preferences between the two procedures ranged from 0.72 on dimension 1 to 0.14 on dimension 2, providing strong evidence that estimates are highly sensitive to such model specifications. I consider a model similar to the one used by Clinton and Meirowitz and apply it to 109th House to test the stability of ideal point estimates.

I begin with a discussion of an ideal point estimator designed to allow a greater variety of choices — an estimator used in both the study of Present votes and vote dependencies. Next, I examine Obama's voting record in the Illinois Senate, focusing primarily on his use of the Present vote and the context in which they were used. I then switch to an analysis of the 109^{th} House, comparing ideal point estimates obtained from traditional procedures that assume independent votes to estimates that do not. I conclude with a discussion of my findings and its implications for future research.

4.2 Estimation

In this section, I introduce an ideal point estimator that allows more than two choices on each roll call. Let p denote the number of legislators (i=*1*,...,p); and q denote the number of roll call votes (j=*1*,...,q). Each roll call allows a set of at least $r \geq 2$, and k denotes the choice (k=*2*,...,r). In traditional ideal point models, the $r = 2$ choices that are permitted are Yea and Nay votes — this model, however, allows r to exceed two choices. For now, we remain agnostic about what constitutes a "choice".

Drawing on the spatial model of voting, there are three parameters of theoretical interest. The key parameter of interest is each legislator i's ideal point, denoted as x_i. In the same space, each choice k for each roll call j is located at θ_{jk}. Furthermore, each choice carries a valence parameter δ_{jk}, which measures the utility a legislator receives from selecting that choice independent of ideological concerns (Londregan, 1999). I assume a quadratic utility function for legislators. Following the random utility framework of McFadden (1973), legislator i's utility on roll call j from outcome k is:

$$U_{ijy} = u_{ijk} + \epsilon_{ijk}$$

where u_{ijk} represents the deterministic component of legislator utility, and ϵ_{ijk} represents the stochastic component. ϵ_{jk} is assumed to follow a type 1 extreme value distribution. The deterministic component of utility is composed of a roll call valence utility δ_{jk}, and a spatial component that declines as a quadratic function of the distance between the legislator i's ideal point and the location of outcome location θ_{jk}. We can further express the deterministic component of utility as follows:

$$u_{ijk} = \delta_{jk} - (x_i - \theta_{jk})^2 = -x_i^2 + \alpha_{jk} + \beta_{jk} * x_i$$

where $\alpha_{jk} = \delta_{jk} - \theta_{jk}^2$ and $\beta_{jk} = 2 * \theta_{jk}^2$. Decomposition of utility into these components simplifies the estimation of the desired parameters. Following Dhyrmes (1978), this allows us to express the probability that legislator i votes for outcome $m = k$ on roll call j as:

$$P_{m=k} = P_{ijk} = \frac{exp(u_{ijk})}{\sum_{m=1}^{r} exp(u_{ijm})}$$

The likelihood across all p votes and q legislators in roll call matrix V can then be expressed as:

$$p(V|\alpha, \beta, X) = \prod_{i=1}^{p} \prod_{j=1}^{q} \prod_{m=1}^{r} P_{ijm}^{C_{ijm}}$$

Note that when expressed in this manner, the $-x_i^2$ component of deterministic utility in each choice drops out from the numerator and denominator. In its currently expressed form, the choice equation is unidentified. To identify the model, I constrain α_{jk} and β_{jk} to equal 0 for the first outcome of all bills. Following Poole and Rosenthal (1997), I also discard any bills in which the losing side has less than 2.5 per cent of the vote from analysis. Finally, estimation is conducted using Markov Chain Monte Carlo software written in C in a Bayesian framework. The posterior that is simulated is:

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

where priors are uninformative and assumed to be distributed:

$$p(O_{jk}) \sim N(0, Inv - \chi^2) \forall j \in 1, ..., q$$

$$p(X_i) \sim N(0, Inv - \chi^2) \forall i \in 1, ..., p$$

This generalizes the Poole and Rosenthal NOMINATE model to any situation where legislators have $r > 1$ choices, with three differences. Rather than the multinomial logit form expressed here, NOMINATE calculates probabilities in probit form, which is justified by the assumption that ϵ_{jk} is distributed normal rather than Type 1 Extreme value. Poole and Rosenthal find that the choice function does not significantly affect estimates, but they originally choose the logistic choice function used here for computational simplicity. The model presented here also uses a quadratic utility function rather than NOMINATE's Normal utility function. While my earlier dissertation chapter suggests that normal utility functions tend to improve model fitting, estimates of ideal points are largely unaffected by the choice of utility function. Finally, the model included here includes a valence component δ_{jk} that measures the natural appeal of choice k, independent of spatial influence. In models with only two choices, valence cannot be separately identified, so it is omitted from NOMINATE (Lewis, 2001).

I test this model via Monte Carlo simulation. Using the data generating mechanism shown here, I generate a set of ideal point and bill parameters for $p = 100$ legislators, $q = 500$ roll call votes, and $r = 4$ choices for each roll call. I then use these parameters to generate a roll call matrix and attempt to recover the true ideal points using the estimator. The estimated ideal points X_i correlate with the true ideal points at $r = 0.998$, suggesting that the estimator successfully recovers ideal points generated under this model.

4.3 Voting Present: 91-93rd Illinois State Senate

In the 2008 U.S. Presidential election, Barack Obama's prior voting record in the Illinois State Senate came under scrutiny in both the Democratic primary and the general election. More specifically, Obama voted 'Present' 129 times — an unusual voting pattern permitted in Illinois. Criticism of Obama's qualifications to be President were articulated by former New York City Mayor Rudy Giuliani at the 2008 Republican Convention (Giuliani, 2008):

> And nearly 130 times, he couldn't make a decision. He couldn't figure out whether to vote "yes" or "no." It was too tough. He voted – he voted "present." I didn't know about this vote "present" when I was mayor of New York City. Sarah Palin didn't have this vote "present" when she was mayor or governor. You don't get "present." It doesn't work in an executive job. For president of the United States, it's not good enough to be present.

This analysis raises several questions. Substantively, what does voting Present mean, and why is it done? Even more specifically, was Obama's voting record in the Illinois State Senate unusual, and did it make him appear more moderate? Finally, as an issue in political methodology, Present votes are typically discarded in roll call analysis. How might this information be incorporated in an ideal point model, and how does this additional information affect our estimates?

In this section, I attempt to answer these questions by examining voting in the $91^{st} - 93^{rd}$ Illinois State Senate, which covers the 6 years that Obama served as a State Senator. The Illinois State Senate is comprised of 59 members, and was controlled by Republicans during the 91^{st} and 92^{nd} Congress before changing control to the Democrats in the 93^{rd} Congress. There is no supermajority required for legislation to pass, and the filibuster does not exist. In addition to standard Yea/Nay votes, Illinois' Senate allows members to vote "Present" on any legislation. Present votes are distinguished in the record from missing votes, where the senator is not available to vote, so they are unusual in that they indicate a *deliberate* decision not to vote Yea or Nay. However, Present Votes are not inconsequential because the Illinois Senate stipulates that all bills require the assent of a majority of the chamber (i.e. 30 votes), rather than a simple plurality (i.e. more Yea than Nay votes). Present votes therefore have the same legislative effect as a No vote.

Table 4.3 presents some descriptive statistics about general voting patterns in the Illinois Senate. These descriptive statistics illustrate two important characteristics of the data. First, many more roll calls were introduced in the 93^{rd} Senate after the Democratic takeover, a point that is important to emphasize here because it affects the comparability of the graphics I present later. The 93^{rd} Senate not only saw more roll calls, but the vast majority of those roll calls were also non-lopsided roll calls in which the losing side had more than 2.5 per cent of the vote. Lopsided roll calls are typically uninteresting in the sense that they

	91^{rd} Senate	92^{rd} Senate	93^{rd} Senate
Total Roll Calls	950	828	1,466
Non-lopsided Roll Calls	231	196	1,403
Roll Calls with $\geq$ 1 Present	96	68	800
Total Yea/Nay Votes	12,608	10,146	77,813
Total Present Votes	620	324	4,909
Present Vote Fraction	4.9%	3.2%	6.3%

Table 4.1: Descriptive Statistics of Voting Present, 91-93 Illinois Senate: *Lopsided roll calls include all roll calls where the losing side has less than 2.5 per cent of the vote. Yea/Nay/Present Votes only calculated from non-lopsided roll calls. Present Vote Fraction is Present Votes divided by Yea/Nay Votes. The Illinois Senate has 59 members at any one time.*

convey no spatial information, so the remainder of this analysis focuses solely on non-lopsided roll calls. Secondly, Present votes appear with some frequency in the data. Approximately 40 per cent of all non-lopsided roll calls have at least one person voting Present, and 3.2 to 6.3 per cent of all votes were Present votes. The key point here is that Present votes occur in sufficient frequency that they could, under certain situations, affect ideal point point estimates in a substantively significant manner.

Present votes have the same legislative effect as a No vote because bills that pass the Illinois Senate require a majority of the Assembly. Therefore, one potential reason why senators might vote Present is that it presents a way to kill legislation without having to vote No. Table 4.2 examines evidence for this hypothesis, and four key points emerge from this analysis. First, the vast majority of roll calls pass, consistent with our expectations of strong party discipline (Cox and McCubbins, 2005). Secondly, the hypothesis that Present votes are used to kill legislation appears weak. In examining the legislative impact of voting present, I define a "Pivotal Present Vote" as a bill that fails to pass, but which would have passed if all senators voting Present had switched their vote to Yea. The idea here is that pivotal present votes define the set of votes where Present votes actually "matter" in defeating a piece of legislation. Only 18 roll call votes are pivotal present votes, which suggests that voting Present was generally an ineffective tactic at killing legislation. Third, I tabulate the total number of Present votes by Obama on non-lopsided legislation, which total 55 Present votes across all three Congresses. Note that this figure is different from the 129 total Present votes attributed to Obama — this suggests that 74 of those Present votes occurred on non-controversial legislation. Finally, Table 4.2 also show how Obama's frequency of voting Present ranked in the legislature. During the two Republican-controlled Congresses, Obama's frequency of voting present ranked him in the upper tertile, while his frequency dropped to the median during the Democrat-controlled 93^{rd} Congress.

Table 4.3 continues this analysis by examining the 18 pivotal present votes in greater detail. With the potential exception of SB1704 (Pension Code Reform) and SB1963 (Consumer Protection Agency Act), none of the pivotal votes could be described as major pieces of legislation. Present votes therefore rarely appear

	91^{rd} Senate	92^{rd} Senate	93^{rd} Senate
Total Roll Calls	231	196	1,403
Total Passed	220	174	1,394
Total Failed	11	22	9
Total Pivotal Present Votes	4	5	9
Total Obama Present Votes	18	9	28
Obama Present Vote Frequency Ranking	17	13	28

Table 4.2: Potential Impact of Voting Present, 91-93 Illinois Senate: *Lopsided roll calls where the losing side has less than 2.5 per cent of the vote are excluded. Pivotal Present Votes are roll calls that failed to pass, but would have passed if members voting Present had instead voted Yea.*

to be decisive on major legislation. Of the 18 pivotal outcomes, Obama voted Present on 5 of those occasions, yet even on these 5 occasions the legislative impact of his Present votes appears to be minimal. Obama's Present votes in the 92^{nd} Congress occurred on two pieces of legislation where large numbers of senators (22 and 30 senators, respectively) joined him. In the 93^{rd} Senate, Obama voted Present on Election Code and Pension Code reform bills that narrowly failed, but the meaning of his votes here are unclear because they were votes against concurring with a House Amendment to a bill (typically a conference report), rather than a vote against the bill at third reading. The change to the Riverboat Gambling act appears to be the sole example of legislation where Obama's present vote clearly mattered, yet this appears to have been a fairly minor piece of legislation.

Senate	Bill	Yeas	Presents	Obama
91	SB688: Appropriations to Judicial Inquiry Board for ordinary and contingent expenses, Third Reading	26	19	Yea
91	SB748: Pre-Marital Education Requirement for Marriage without delay, Third Reading	24	14	Yea
91	SB786: Creation of Micro-Enterprise Assistance Council, Motion to Concur	26	4	Missing
91	SB897: Permits county sheriff to post Internet information about sex offenders, Third Reading	29	11	Yea
92	SB1107: Home Inspector License Act, Third Reading	26	7	Yea
92	SB445: Designates Qualified Non-Chicago Academic Medical Center Hospital, Third Reading	27	22	Present
92	SB657: Prevents employer from discharging employee for obtaining relief as victim of domestic violence, Concurrence	16	30	Present
92	SB2194: Amends Motor Fuel Law to increase Grade Crossing fund, Third Reading	29	3	Nay
92	SB609: Restricts proximity of adult entertainment establishments from schools, Concurrence	22	15	Yea
93	SB82: Amends Election Code to conform with Help America Vote Act, Concurrence	23	9	Present
93	SB100: Amends Compensation Review Act, allows judges have compensation increased by COL adjustments, Override	28	3	Nay
93	SB1704: Amends Pension Code for Chicago Police, Firefighters, Municipal, and Park District employees, Concurrence	28	13	Present
93	SB1960: Amends Illinois Government Ethics Act, Election Code, and University of Illinois Trustees Act, Concurrence	27	9	Nay
93	SB1963: Consumer Advocate Act, creates consumer protection agency, Third Reading	29	11	Yea
93	SB2228: Amends Criminal Code, makes technical change relating to applicability of common law, Third Reading	29	3	Yea
93	SB2230: Amends Criminal Code, makes technical change concerning definition of "conviction", Third Reading	29	4	Yea
93	SB2237: Amends Riverboat Gambling Act, makes technical change concerning concerning short title, Third Reading	27	5	Present
93	SB2249: Amends the Conveyances Act, makes technical change concerning Act's short title, Third Reading	29	2	Yea

Table 4.3: Obama and Party Line Votes, 91-93 Illinois Senate: *Lopsided roll calls where the losing side has less than 2.5 per cent of the vote are excluded. Pivotal Present Votes are roll calls that failed to pass, but would have passed if members voting Present had instead voted Yea.*

A second hypothesis justifying the use of Present votes is that is presents an indirect way for senators to oppose their party. If this is true, senators are likely to vote Present on party line roll calls that are supported by their party. I define a party line roll call as one where a majority of Democrats vote differently from a majority of Republicans, with Present votes excluded from the counts. Table 4.4 summarizes Obama's voting patterns on party line votes, with large differences in voting behavior between Congresses. Party line voting surged in Obama's final term, jumping from 34 party line votes in the 91^{st} and 92^{nd} to 535 party line votes in the Democrat-controlled 93^{rd} Congress. This increase was partly driven by the increase in total legislation in the 93^{rd}, but even accounting for the volume of legislation the 93^{rd} Congress was much more divided along partisan lines — 38 per cent of all roll calls in the 93^{rd} Congress were party line votes, compared to only 8 per cent before that. Across all three legislatures, Obama largely voted along party lines, siding with Democrats 92 per cent of the time. Notably, although there are no instances where Obama voted Present on party line roll calls supported by Democrats in either the 91^{st} or 91^{nd} Congress, this occurs 16 times in the 93^{rd} Congress. There is therefore some evidence that Obama used Present votes to indirectly oppose his party while the Democrats controlled the Senate.

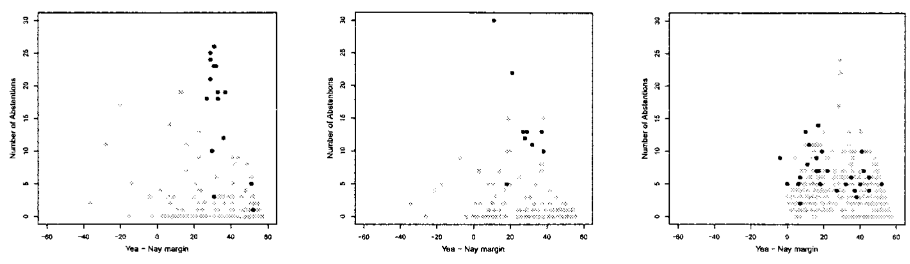
A third hypothesis for voting Present is that such votes are more likely when legislation is particularly controversial. If this is true, then abstentions are likely to occur when the Yea and Nay sides are closely matched on a vote. Table 4.1 examines the evidence for this claim by plotting the Yea - Nay margin on each

	91^{rd} Senate	92^{rd} Senate	93^{rd} Senate
Obama Votes with Democrats	23	4	497
Obama Votes with Republicans	2	2	9
Obama Votes Present, Democrats Support Bill	0	0	16
Obama Votes Present, Democrats Oppose Bill	0	0	1
Obama Misses Vote	1	2	12
Total Party Line Votes	26	8	535
Total Roll Calls	231	196	1,403

Table 4.4: Obama and Party Line Votes, 91-93 Illinois Senate: *Lopsided roll calls where the losing side has less than 2.5 per cent of the vote are excluded. Party line votes are defined as roll calls where a majority of Democrats vote differently from a majority of Republicans, with Present votes excluded in the counts.*

roll call against the number of Present votes on each roll call, with Obama's Present votes highlighted as darker points. While there are few votes that are close (i.e. have a Yea - Nay margin near 0), there is no evidence in any Congress that close votes are more likely to have high numbers of senators voting Present. Instead, Present votes appear to be most likely the Yea - Nay margin is around 35 votes. This trend is consistent with Obama's pattern of voting Present in the 91^{st} and 92^{nd} Senate — Obama largely votes Present only when a large number of other senators are voting likewise. Obama's propensity to vote Present in the 93^{rd} Senate, however, seems largely random.

While the previous graphic suggests that senators who vote Present frequently do so together, it provides little information about who votes Present together. To examine this issue, I first estimate ideal points for each Senate under the

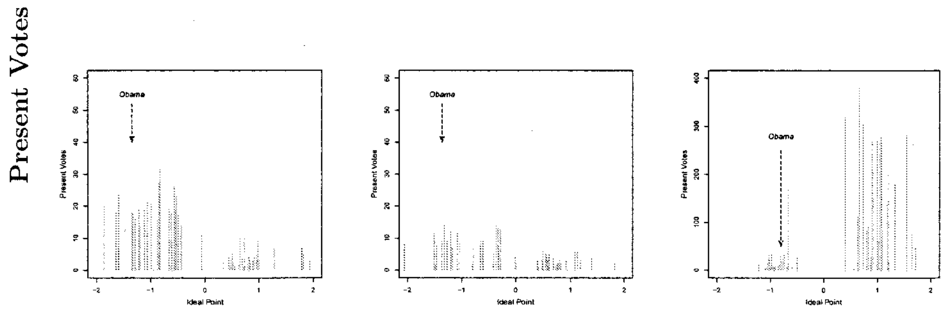


91st Senate **92nd Senate** **93rd Senate**

Figure 4.1: Frequency of Present Votes by Vote Margin, 91-93 Illinois Senate: *Vote margin on x-axis is defined as Yea minus Nay votes. Darker Points indicate the votes on which Obama Voted Present.*

traditional assumption that Present votes are missing data. Figure 4.2 then plots a histogram of present vote frequency by ideal point estimate. The histograms suggest that Present votes were largely used by the minority party — Democrats were particularly likely to use them in the 91^{st} and 91^{st} Senate, while Republicans were likely to use them in the 93^{rd}.

Ideal point models traditionally treat Present votes as missing data. However, Present votes clearly occur with moderate frequency and in a non-random manner. I therefore estimate ideal point models that incorporate Present votes in two ways. First, I estimate a model that treats Present votes as a third choice on each vote. I also estimate another model where I treat Present votes as No votes, a model motivated by the fact that Present and No votes have the same legislative effect. Ideal points estimated under these assumptions are then plotted against ideal points estimated under the traditional assumption of Present



91st Senate **92nd Senate** **93rd Senate**

Figure 4.2: Frequency of Present Votes by Ideal Point, 91-93 Illinois Senate: *Note that Y-axes are not on the same scale because many more votes occurred in the 93rd Senate.*

votes as missing data in Figure 4.3. Obama's specific ideal point estimates under different assumptions are presented in Table 4.5. The relationships are strongly linear, showing little variability in the recovered ideal point estimates regardless of the model estimated. Notably, Obama was criticized during the campaign for voting Present to appear moderate, yet there is no evidence of moderation in his ideal point estimates under any assumption. However, there is some evidence that suggests Obama may have become more moderate in the 93^{rd} senate, as his ideal point and ranking both moderate considerably. The evidence however is still inconclusive because estimates across legislatures are not comparably — in particular, because the 93^{rd} has a Democratic majority there are significantly more Democratic senators.

While little variability in ideal point estimates occurs when Present votes are incorporated into the model, we obtain large improvements in efficiency. Taking

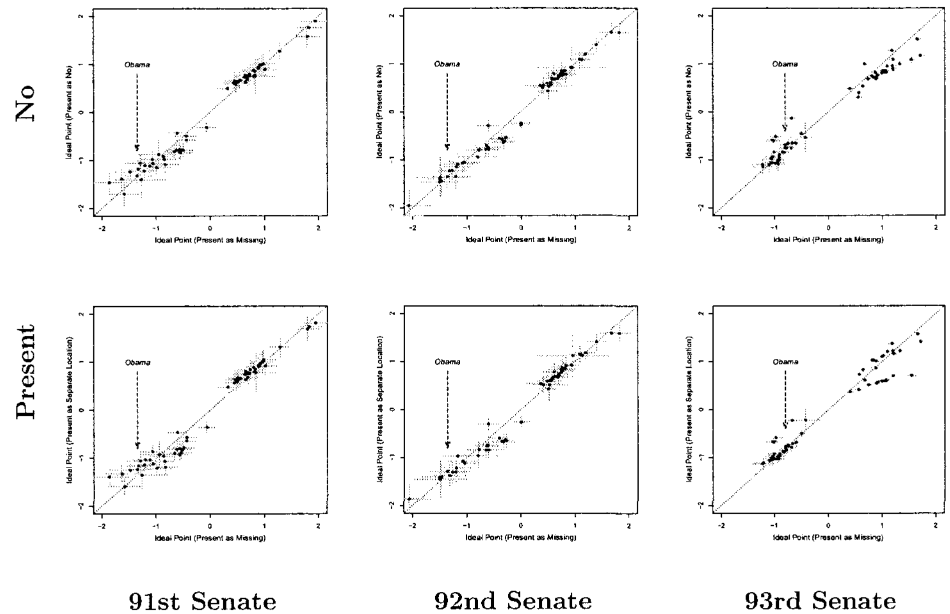


Figure 4.3: Private/Public Coverage in True vs. Synthetic Illinois: *From left to right, top panels show ideal point estimates comparing standard ideal point estimates to those derived from treating Present votes as No votes by Congress, while lower panels compare standard ideal point estimates to those derived from treating Present votes as a separate choice. Standard ideal point estimates are derived by discarding all Present votes. Bands represent 80 per cent confidence intervals.*

	91^{rd} Senate	92^{rd} Senate	93^{rd} Senate
Obama Ideal Point (Present Votes as Missing)	-1.34 (0.21)	-1.36 (0.25)	-0.80 (0.08)
Obama Ideal Point Rank	5	7	24
Obama Ideal Point (Present Votes as No Votes)	-1.32 (0.16)	-1.35 (0.19)	-0.69 (0.06)
Obama Ideal Point Rank	5	7	23
Obama Ideal Point (Present Votes as Choice)	-1.23 (0.12)	-1.29 (0.16)	-0.77 (0.06)
Obama Ideal Point Rank	6	10	23

Table 4.5: Obama Ideal Point Estimates Under Different Assumptions, 91-93 Illinois Senate: *Lopsided roll calls where the losing side has less than 2.5 per cent of the vote are excluded. Standard errors in parentheses.*

the 91^{st} Senate as an example, Democrats had a mean standard error of 0.19 in the traditional model where Present votes are omitted. However, this drops to 0.13 in the model where Presents are counted as No votes, and 0.12 in the case where Presents are counted as a third location. However, the increase in efficiency is marginal for Republicans, who have a mean standard error of 0.16 under the traditional model, but mean standard errors of 0.14 and 0.15 respectively under the models where Present votes are treated as no or separate votes respectively. An obvious hypothesis explaining this discrepancy is the large number of Present votes cast by the minority party — this hypothesis receives strong support from the fact that the pattern is reversed in the 93^{rd} Senate under a Democratic majority when Republican legislators vote Present much more frequently.

Summarizing the results presented here, the practice of voting Present largely appears to be a tactic employed by the minority party. When Present voting occurs, it is rarely decisive in the sense that the outcome of the vote would have changed regardless of how those who voted Present would otherwise have voted. In the few cases where Present votes are potentially decisive, the affected legislation is generally minor legislation. Present votes also do not appear to occur for particularly controversial legislation where the Yea minus Nay margin is small — in fact, they are most likely to occur when the margin is about 30 votes.

Obama's usage of Present votes was largely consistent with these patterns — his use of Present votes was in the upper tertile of legislators, but not particularly unusual. The majority of Obama's Present votes occurred on roll calls where large numbers of Democrats voted Present with him, particularly in the 91^{st} and 92^{nd} Congress. There is little evidence that Obama was more likely to vote Present when Present votes were potentially decisive, and little evidence that Obama was more likely to vote Present when votes were controversial. Some evidence suggests that Obama used Present votes to oppose his party on party line votes during the 93^{rd} Congress, but these votes were rarely decisive and had little impact on estimates of his ideal point. Notably, the incorporation of information from Present votes does not make Obama's estimated ideal point more moderate. Obama consistently appears as the 5^{th} or 6^{th} most liberal member of the Illinois Senate in the 91^{st} Senate, but he becomes the 23^{rd} or 24^{th} most liberal senator by the 93^{rd} Senate as he begins his run for the U.S. Senate.

4.4 The Impact of Bill Sequences: 109th Senate

Empirical models of spatial voting such as NOMINATE (1997) and IDEAL (2004) generally begin with the assumption that votes are conditionally independent across each other. Stated differently, in the context of the ideal point model presented in this paper, a legislator's probability for a particular vote on a bill is solely a function of her ideal point and the bill parameters for that bill, and is unaffected by bill parameters for other roll calls. However, this assumption is clearly violated in ideal point estimates from the U.S. Congress, which typically treat each amendment, motion, final passage vote, and concurrence with conference report as being fully independent even when they deal with the same bill. In fact, many votes in Congress are highly dependent on others, with early votes on amendment trees affecting later votes as one example of dependence across votes.

In this section, I examine the potential consequences of this assumption for ideal point estimation. By treating conditionally dependent votes as independent, three potential consequences may occur. First, the assumption of independence may overstate the true level of political polarization in the U.S. Legislature found in studies such as McCarty, Poole and Rosethal (2006). This occurs because as legislation moves from its initial stages to the final bill, multiple amendments and motions that are *de facto* identical and often polarizing occur. An example of this is the Motion to Recommit without Instructions, which is widely known as a motion that effectively kills legislation. This motion is therefore virtually identical to the final passage vote, which often occurs shortly if the motion to recommit fails. By treating the final passage vote and the motion to recommit as two separate votes on a partisan bill, traditional ideal point models place two cutlines separating the parties rather than one and thus overstates political polarization.

Building on the logic just described, the treatment of non-independent votes as fully independent may result in exaggerated precision of estimates. In treating near-identical roll call votes such as motions to recommit and final passage votes from the same bill as separate votes, traditional ideal point models assume there are more data points than may actually be true in reality. In a regression context, this would be analogous to randomly sampling a subset of the data and adding the subset as extra data before running the regression — the extra cases result in lower standard errors when the regression is run despite providing no additional independent information.

Finally, treating roll call votes as fully independent may discard potentially useful spatial information, resulting in differences in recovered ideal point estimates.[2] This argument is supported by the work of Clinton and Meirowitz (2004), who estimate an ideal point model where early stage Yea or Nay locations in a voting tree are constrained to be identical to late stage vote locations based on the substantive ordering of the amendments and votes. Using roll call data from the Compromise of 1790, Clinton and Meirowitz find that ideal point

[2]To see how the assumption of independence can discard spatial information, consider a sequence of k amendments on a bill, each permitting a Yea or Nay vote. Legislators therefore have 2^k different ways to vote on the k amendments, but only $k+1$ of those possibilities are spatially consistent. The assumption of independence thus assumes that 2^k-k-1 of the possible vote sequences are voting "errors" when in fact they may have useful spatial information under sophisticated voting strategies where some legislators do not vote sincerely.

estimates under their constrained model correlate with estimates from NOMINATE at $r = 0.72$. While the Clinton and Meirowitz model produces different estimates without assuming full independence across votes, their constraints on the bill locations assume sincere voting. Stated differently, consider an example where a legislator first votes between alternatives A_1 and A_2. If A_1 wins, legislators then choose between A_1 and the status quo SQ, where SQ is expected to win. If A_2 wins, legislators then choose between A_2 and the status quo SQ, but now A_2 is expected to win. Under the Clinton and Meirowitz procedure, the two A_2 are constrained to be the same, as are the two A_2 locations. However, this constraint is nonsensical under these conditions because the first choice between A_1 and A_2 is in reality a choice between SQ and A_2

In this section, I use the multinomial ideal point model developed earlier in this chapter to test these three potential outcomes. Rather than using Yea/Nay votes single roll calls as the choice data however, I use bill sequences as the unit of analysis. I define a bill sequence as the entire set of roll calls vote on a bill, including amendments, motions, concurrences, cloture, and final passage votes. I then consider each sequence of votes on each bill sequence as the unique choices within each bill sequence. Given a bill sequence of length k, there are 2^k possible "votes" within that sequence. In an effort to reduce computational complexity, I limit $k \leq 4$, permitting a maximum of $2^4 = 16$ different choices for each bill sequence. For bill sequences whose length exceeds $k = 4$, I truncate the sequence by only considering the last 4 votes in the sequence and discard all earlier votes.

In treating each bill sequence as a separate choice, this model implicitly assumes that each different vote sequence is an expression of a desired policy. Stated differently, legislators care about position-taking either in conjunction or in lieu of legislative outcomes. To see an example where this assumption is reasonable, consider a situation where legislators are voting on four budget amendments allocating money to a defense budget. Legislators have preferences over the total sum of money spent on the defense budget, which initially is \$0 before any amendments are introduced. In the first amendment legislators vote to increase the defense budget by \$1 billion to purchase additional rifles. In the second amendment, legislators vote to increase the defense budget by \$2 billion to purchase additional tanks. Next, legislators vote on an amendment to increase the defense budget by \$4 billion to purchase additional helicopters. Finally, legislators vote on an amendment to increase the defense budget by an additional \$9 billion to purchase additional bombers.

Given the choice to vote Yea or Nay on these four amendments, $2^4 = 16$ different vote sequences are possible, as shown in Table 4.6. In this table, four-letter codes designate how a legislator might vote on amendments 1 through 4 respectively — "YNNY", for example, indicates a legislator votes Yea on amendments 1 and 4 and Nay on amendments 3 and 4. Each vote sequence thus expresses a desire for a unique level of defense spending — the "YNNY"-voting legislator, for instance, wishes to spend \$1 billion + \$9 billion = \$10 billion on defense. The key point to note here is that each of the 16 possible vote sequences expresses a desire for a unique level of defense spending, also shown in Table 4.6.

I apply this estimator to data from the 109^{th} House, and contrast its results

Vote Sequence	Total Expenditure (bil.)	Vote Sequence	Total Expenditure (bil.)
YYYY	$16	NYYY	$15
YYYN	$7	NYYN	$6
YYNN	$3	NYNN	$2
YYNY	$12	NYNY	$11
YNNN	$1	NNNN	$0
YNNY	$10	NNNY	$9
YNYN	$5	NNYN	$4
YNYY	$14	NNYY	$13

Table 4.6: Expression of Desired Defense Spending Under Different Vote Sequences: *Designed as a motivating example for how different vote sequences imply different policy preferences. Votes represent Yea or Nay votes on four separate amendments adding $1 billion, $2 billion, $4 billion, and $9 billion respectively for extra rifles, tanks, helicopters, and bombers. We assume that legislators derive utility from position-taking, where their vote sequence expresses their desired for a certain level of total defense expenditure.The example demonstrates that each of the 16 different vote sequences map on to a different preferred level of defense expenditure.*

to estimates obtained under the traditional assumption of independent votes. The 109^{th} House consists of N=439 legislators voting on $Q_1 = 894$ non-lopsided votes.[3] After collapsing the votes into bill sequences however, the data contains $Q_2 = 346$ bill sequences. Results of this comparison appear as Figure 4.4. These figures suggest three conclusions. First, ideal point estimates do not appear noticeably changed when roll call votes are not assumed to be independent, as the ideal points recovered under the two assumptions line up tightly along the 45° line. Secondly, there does not appear to be any noticeable decrease in political polarization. The right panel however suggests modeling voting dependence appears to have a noticeable increase in uncertainty about most legislators' ideal point, as the majority of points lie above the 45° line.

4.5 Conclusion

In this paper, I examine the robustness of ideal point estimates to two potential sources of complexity. Separately, I consider the impact of deliberate abstentions in the Illinois state senate and the dependence across roll call votes that results from the amendment process in the U.S. Congress. In contrast to earlier work completed by Clinton and Meirowitz (2004), I find that ideal point estimates are highly stable even when these additional factors are accounted for. The primary implication of this finding is that previous applied research using ideal point estimates as a variable is largely robust to these additional complexities.

[3]There are 441 legislators in the roll call data set, but 2 of these members vote less than 20 times and have been discarded.This includes legislators who only served part of the term; hence, the number is greater than 435. Non-lopsided votes are defined as votes where the losing side has at least 2.5% of the vote.

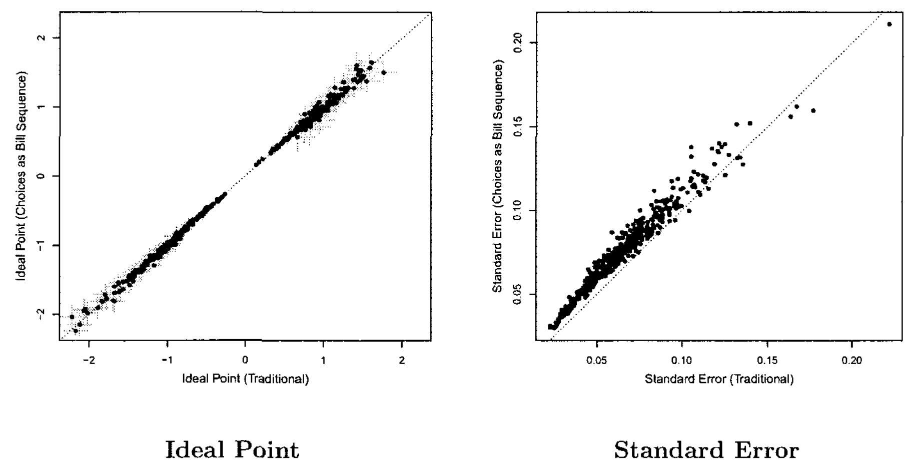


Figure 4.4: Effect of Bill Sequencing of Ideal Point Estimates, 109th House: *Left panel shows ideal points estimated from using traditional procedures vs. bill sequencing. Bands represent 80 per cent confidence intervals. The estimates show no significant change in the recovered coordinates and no decrease in political polarization. Right panel shows standard errors for each legislator's ideal point calculated from traditional procedures vs. bill sequencing. The results suggest that using bill sequences increases uncertainty about a legislator's ideal point for most legislators.*

My research however also suggests that the modeling vote dependence and abstentions can potentially affect the precision of the estimated ideal points. In examining Obama's voting record in the Illinois senate, I found that standard errors for Obama's ideal point decreased by approximately 25 per cent after information from Present votes was incorporated into the model. However, the net effect of vote dependence and abstentions remains unclear because they each have opposite effects. Modeling vote dependence suggests that traditional ideal point models overstate precision by counting near-identical votes as separate pieces of information, while modeling Present votes increases the precision of estimates. The answer to this question is important because studies attempting to answer questions such as the identity of the most liberal senator (Clinton and Jackman, 2004) are affected significantly by the standard errors of the estimates.

Bibliography

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. Working Paper.

Abadie, A. and J. Gardeazabal (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review 93*(1), 112–132.

Ashenfelter, O. and D. Card (1985). Using the longitudinal structure of earnings to estimate the effect of training programs. *The Review of Economics and Statistics 67*(4), 648–660.

Blumberg, L. J., L. Dubay, and S. A. Norton (2000). Did the medicaid expansions for children displace private insurance? an analysis using the sipp. *Journal of Health Economics 19*(1), 33–60.

Bureau, U. C. Comparison of sipp with other surveys. Retrieved 2010-02-01.

Card, D. and L. Shore-Sheppard (2004). Using discontinuous eligibility rules to identify the effects of the federal medicaid expansions on low income children. *Journal of Health Economics 86*(3), 752–766.

Carroll, R., J. Lewis, J. Lo, K. Poole, and H. Rosenthal (2009a). Comparing nominate and ideal: Points of difference and monte carlo tests. *Legislative Studies Quarterly 34*, check.

Carroll, R., J. Lewis, J. Lo, K. Poole, and H. Rosenthal (2009b). Measuring bias and uncertainty in dw-nominate ideal point estimates via the parametric bootstrap. *Political Analysis 17*, check.

Chang, C. F. (2007). Evolution of tenncare yields valuable lessons. *Managed Care 11*, 45–49.

Clinton, J. and S. Jackman (2004). The most liberal senator? analyzing and interpreting congressional roll calls. *PS: Political Science and Politics 37*, 805–811.

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.

Clinton, J. and A. Meirowitz (2004). Testing explanations of strategic voting in legislatures: A reexamination of the compromise of 1790. *American Journal of Political Science 48*, 675–689.

Converse, P. (1964). The nature of belief systems in mass publics. In *Ideology and Discontent*. New York: Free Press.

Cox, G. and M. McCubbins (2005). *Spatial Models of Parliamentary Voting*. Berkeley: University of California Press.

Currie, J. and J. Gruber (1996a). Health insurance eligibility, utilitzation of medical care, and child health. *Quarterly Journal of Economics 111*(2), 431–466.

Currie, J. and J. Gruber (1996b). Saving babies: The efficacy and cost of recent changes in the medicaid eligbility of pregnant women. *Journal of Political Economy 104*(6), 1263–1296.

Cutler, D. and J. Gruber (1996). Does public health insurance crowd-out private insurance? *Quarterly Journal of Economics 111*(2), 391–430.

Dhrymes, P. (1978). *Introductory Econometrics.* New York: Springer.

Downs, A. (1957). *An Economic Theory of Democracy.* New York: Harper.

Dubay, L. and G. Kenney (1996). Revisiting the issues: The effects of medicaid expansions on insurance coverge of children. *The Future of Children 6*(1), 152–161.

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.

Farquharson, R. (1969). *Theory of Voting.* New Haven: Yale University Press.

Giuliani, R. (2008). Giulianis speech at the republican national convention.

Gruber, J. (1994). The incidence of mandated maternity benefits. *American Economic Review 84*(3), 622–641.

Gruber, J. (2008). Covering the uninsured in the united states. *Journal of Economic Literature 46*(3), 571–606.

Gruber, J. and K. Simon (2008). Crowd-out 10 years later: Have recent public insurance expansions crowded out private health insurance? *Journal of Health Economics 27*(3), 201–217.

Hacker, J. The case for public plan choice in national health reform. Retrieved 2010-02-01.

Hacker, J. Health care for america. January 11, 2007.

Haley, J. and G. Kenney (1996). Coverage gaps for medicaid-eligible children in the wake of federal welfare reform. *Inquiry 40*(2), 158–168.

Ham, J. and L. Shore-Sheppard (2005). The effect of medicaid expansions for low-income children on medicaid participation and private insurance coverage: Evidence from the sipp. *Journal of Public Economics 89*(1), 57–83.

Hammond, T. H., C. W. Bonneau, and R. S. Sheehan (2005). *Strategic behavior and policy choice on the U.S. Supreme Court.* Palo Alto: Stanford University Press.

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.

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.

Jackman, S. (2008). pscl: Political science computational laboratory.

Kalandrakis, T. (2008). Rationalizable voting. Working Paper.

Keisler, C., B. Collins, and N. Miller (1969). *Attitude Change: A Critical Analysis of Theoretical Approaches*. New York: Wiley Press.

Kronick, R. and T. Gilmer (2002). Insuring low-income adults: Does public coverage crowd out private? *Health Affairs 21*(1), 225–239.

Lewis, J. B. (2001). Estimating voter preference distributions from individual-level voting data. *Political Analysis 9*, 275–297.

Londregan, J. (1999). Estimating legislator's preferred points. *American Journal of Political Science 8*, 35–56.

Londregan, J. B. (2000). Estimating legislators' preferred points. *Political Analysis 8*, 35–56.

LoSasso, A. and T. Buchmueller (2004). The effect of the state children's health insurance program on health insurance coverage. *Journal of Health Economics 23*(5), 1059–1082.

Manning, W. G., E. B. Keeler, J. P. Newhouse, E. M. Sloss, and J. Wasserman (1989). The taxes of sin: Do smokers and drinkers pay their way? *Journal of the American Medical Association 261*(3), 1604–1609.

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 (2006). *Polarized America: The Dance of Ideology and Unequal Riches.* Cambridge: MIT Press.

McFadden, D. (1973). Conditional logit analysis of qualitatative choice behavior. In *Frontiers of Economics.* New York: Academic Press.

McKelvey, R. (1986). Covering, dominance, and institution free properties of social choice. *American Journal of Political Science 30*, 283–314.

Miller, N. (1995). *Committees, Agendas, and Voting.* Poststrasse: Harwood.

Narwa, D. (2001). How general should the proximity model be? *European Journal of Political Economy 17*, 53–74.

Nosofsky, R. M. (1986). Attention, similarity, and the context theory of classification. *Journal of Experimental Psychology 115*, 39–57.

Office, C. B. (2007). The state children's health insurance program. Retrieved 2010-02-01.

on Medicaid, K. C. and the Uninsured (2007). Schip reauthorization: Key questions in the debate–a description of new administrative guidance and the house and senate proposals. Issue Brief publication number 7675.

Peress, M. (2009). Small chamber ideal point estimation. *Political Analysis 17*, 276–290.

Plott, C. (1967). A notion of equilibrium and its possibility under majority rule. *American Economic Reivew 57*, 787–806.

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.

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.

Shore-Sheppard, L. (2008). Stemming the tide? the effect of expanding medicaid eligibility on health insurance coverage. *The B.E. Journal of Economic Analysis and Policy 8*(2), 45–49.

Swartz, K. (1997). Changes in the 1995 current population survey and estimates of health insurance coverage. *Inquiry 34*(1), 70–79.

Turunen-Red, A. and J. A. Weymark (2008). Paretian aggregation of asymmetric utility functionals. Working Paper.

Voeten, E. (2000). Clashes in the assembly. *International Organization 54*(2), 185–215.

Yazici, E. and R. Kaestner (2000). Medicaid expansions and the crowding out of private health insurance. *Inquiry 37*(1), 23–32.