Spring 5-15-2013

Essays on Quantitative Methods for Consequences of Political Institutions

Tsung-han Tsai
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Political Science

Dissertation Examination Committee:
Jeff Gill, Chair
Nan Lin
Tse-min Lin
Jacob Montgomery
Guillermo Rosas
Itai Sened

Essays on Quantitative Methods for Consequences of Political Institutions

by

Tsung-han Tsai

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2013
St. Louis, Missouri

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

It is the most wonderful experience of my life that I complete the doctoral degree at Washington University in St. Louis, where I received support and encouragement from a great number of individuals over the past five years. I am deeply grateful to Jeff Gill, who is my professional mentor and dissertation advisor. Working along side Jeff in my graduate career, I benefited greatly from his mentoring in terms of professional development and integration. As one of the leading scholars in the discipline of political methodology, Jeff gave me the most important and helpful advice and guidance in the whole process of this dissertation project. I am also thankful for his countless remarks, suggestions, and comments on many revisions of this manuscript. I would also like to thank Guillermo Rosas for his enthusiasm, his encouragement, and his insightful advice to develop my research ideas and help me to make them better. Guillermo has supported and inspired me since the early years of my doctoral program at WashU. I am very appreciative of his help to me to get through some tough times. Itai Sened's willingness to share his knowledge, perspective, and experience were crucial to my understanding of positive political theory and the execution of this research. I am eternally obliged to his pointed and pragmatic professional advice. I am also indebted to my committee members, Jacob Montgomery, Nan Lin, and Tse-min Lin, whose constructive guidance and feedback has shaped my intellectual development. In addition, many faculty members of Washington University have been, at one time or another, pivotal to my academic career development. I would be remiss if I did not mention Matthew Gabel, Ryan Moore, Randall Calvert, Andrew Martin, Brian Crisp, Margit Tavits, John Patty, Norman Schofield, Dawn Brancati, and Siddhartha Chib. I also appreciate the support and help from Dan O'Neill, Shu-shan Lee, Chung-ying Lin, Xun Pang, Noel Pereyra Johnston, Santiago Olivella, Adriana Crespo-Tenorio, Christopher Claassen, Joshua Potter, Keith Schnakenberg, Abdulhadi Sahin, Constanza Schibber, and many other fellow students and friends. Last, but certainly not least, I thank my family and Chia-yi Lee, whose support, patience, and love have allowed me to complete this dissertation, and dedicate this dissertation to them.

ABSTRACT OF THE DISSERTATION

Essays on Quantitative Methods for Consequences of Political Institutions

by

Tsung-han Tsai

Doctor of Philosophy in Political Science,

Washington University in St. Louis, 2013.

Professor Jeff Gill, Chairperson

In this dissertation, I develop and apply sophisticated Bayesian models to the analysis of institutional effects on electoral and legislative behavior in the policy making process. Leveraging the flexibility of Bayesian methods for statistical modeling, I deal with several methodological problems encountered by political scientists, and social scientists in general, in some established research agenda. This dissertation shows the improvement of the ability to evaluate the success of conflicting theories when these methodological issues are properly dealt with.

The consequences of political institutions are investigated at three different levels in this dissertation: countries, political parties, and individual legislators. First of all, at the country level, I investigate whether there is a difference between the performances of democratic and nondemocratic regimes in social provision policy in

18 Latin American countries by focusing on the rarely changing property of political regimes. An appropriate model for the dynamic nature of rarely changing variables is built to thoroughly explore how democratic institutions improve social welfare. Second, at the party level, I develop a Bayesian structural equation model to examine the interdependence between party policy strategies and party support in multiparty systems, in an effort to illustrate the endogenous dynamics of multiparty systems. The results show that party manifestos do not provide clear-cut division of party policy positions. Instead, party labels are more important information than changes in party manifestos to the electorate. Finally, at the level of legislators, I focus on the role of the president and political parties in Brazilian legislative process, in which political exchanges between the government and legislature is an essential feature. By recognizing the existence of the non-ideological effect on voting behavior, I develop a random item-difficulty ideal-point model implied by the spatial voting model to analyze the relationship between coalition dynamics and party-based voting behavior of legislators.

Chapter 1

Introduction

The study of political institutions is central to the discipline of political science, and the social sciences in general. Political institutions are, as Douglass North states, "the rules of the game in a society or, more formally, ... the humanly devised constraints that shape human interaction" (1990, p. 3). Consequently, institutions structure the preferences of players and the distribution of resources, and reduce uncertainty (or increase predicability) by establishing a stable framework of human interaction, whether political, economic, or social, in a society (March and Olsen, 2006; North, 1990). Moreover, in modern societies institutions play a major role in linking public preferences to political outcomes, and these outcomes in turn induce public policy decisions (Powell Jr., 2004, 2007; Persson and Tabellini, 2000).

"Political institutions," broadly defined, refer to a variety of forms: from formal rules like written constitutions, through organizations like political parties, to existent social norms. In this dissertation, I investigate formal rules and political organizations. Specifically, I concentrate on the types of political regime, forms of government, and party systems. First, regime types, such as democratic and nondemocratic regimes, determine the extent to which the citizens have the powers to make policy decisions and the degree to which the government is credibly committed to or bound by institutional arrangements (Acemoglu and Robinson, 2006; Bueno de Mesquita et al., 1999). Second, the forms of government, which refers to presidential and parliamentary systems, determine how these powers can be exercised by political representatives and how conflicts among these representatives can be solved (Lijphart, 1999; Shugart, 2006; Shugart and Carey, 1992). Finally, party systems, the formation of which is structured by social cleavages (Lipset and Rokkan, 1990), electoral rules (Duverger, 1954; Taagepera and Shugart, 1989), or both (Amorim Neto and Cox, 1997; Boix, 2007), reflect interactions resulting from inter-party relations and dimensions on which party competition takes place in the electoral and legislative arena (Sartori, 1976).

Focusing on the types of political regime, forms of government, and party systems, this dissertation investigates institutional effects in the policy making process, including a general comparison between the performances of democratic and nondemocratic regimes, multiparty competition in parliamentary systems, and legislative behavior in multiparty presidential systems. Interpreting institutions as *exogenous constraints*, or as an exogenously given *game form* (Shepsle, 1979), this dissertation attempts to answer the following questions: (1) do political regimes make a significant difference in social provision policy? (2) what are the determinants and results of party policy shifts in multiparty parliamentary systems? and (3) does the president construct a parliamentary-style coalition and rule through it or construct a series of policy coalitions for different issues when facing a multiparty legislature? These are important questions in the study of comparative politics and comparative political economy, and there is still disagreement among leading scholars about the answers to these questions. To understand the political consequences, and the resulting public policy decisions, of these institutions, this dissertation takes rational choice approaches to institutional analysis, or so-called "rational choice institutionalism," (Shepsle, 2006; Shepsle and Bonchek, 1997; Weingast, 1996, 2002), contrasting the other two approaches within the "new institutionalism:" historical institutionalism and sociological institutionalism (Hall and Taylor, 1996; March and Olsen, 1984). This approach provides the microfoundations for macropolitical phenomena. From the perspective of rational choice institutionalists, I provide theoretical explanations to these questions.

Leading comparativists and political economists have made a great effort to develop theories for explaining political and economic phenomena. This dissertation contributes to the literature by focusing on three methodological issues that are not properly dealt with in previous empirical analyses but are important to the improvement of the ability to evaluate the success of conflicting theories. The first issue concerns the characteristics of political regimes, and political institutions in general—being rarely changing or invariant over time. The second issue is about endogenous dynamic components in multiparty, electoral competition. Regarding the third issue, I deal with the violation of the core assumption of ideal-point models, namely that the association between observed items is explained only by the latent trait variable.

This dissertation leverages the philosophy and flexibility of Bayesian methods on statistical modeling to deal with these methodological issues. The Bayesian approach to statistical modeling has been increasingly applied to a variety of statistical problems and has shown its advantages in many respects including model specification, model estimation, model selection, and statistical inference (Carlin and Louis, 2000; Gelman et al., 2004; Gill, 2008*a*; Greenberg, 2007). The Bayesian approach is also advantageous to handle complexed data structure such as time-series cross-sectional (TSCS) data and hierarchical data structure (Gelman and Hill, 2007; Shor et al., 2007). This dissertation seeks to examine the effects of institutions on electoral and legislative behavior, and policy outcomes as well as the applicability of Bayesian methods for methodological issues encountered by political scientists in empirical research.

In Chapter 2, *A Bayesian Approach to Endogenous Rarely Changing Variables in Time-series Cross-sectional Analyses*, I investigate whether there is a difference between the performances of democratic and nondemocratic regime in social provision policy. Since political regime is rarely changing, their dynamic effects on the outcome are of concern to researchers who evaluate how political regimes affect public provision policy. However, estimating the dynamic effects of rarely changing variables in the analysis of time-series cross-sectional (TSCS) data by conventional estimators may be problematic when unit effects are included in the model. Building on the structure of simultaneous equation modeling and on error-component formulations, this chapter proposes a model to explicitly account for the correlation between unit effects and covariates and shows that the proposed model performs as well, or better than alternative estimators in modeling TSCS data. Applying the proposed model to 18 Latin American countries, I find weak evidence of the effects of political regimes on social security and welfare spending. One explanation to this result is that the variation between democracies is not taken into account.

Chapter 3, *Party Policy Strategies in Multiparty Systems: Bayesian Structural Equation Modeling for Dynamic Party Competition*, examines the interdependence between party policy strategies and party support in multiparty systems, in an effort to illustrate the endogenous dynamics of multiparty systems. To evaluate theoretical arguments, I propose a Bayesian structural equation model (SEM) to analyze the Comparative Manifesto Project dataset for Britain and Israel. The results show that the interdependence between party policy strategies and party support is weak since party manifestos do not provide clear-cut division of party policy positions. The results also show that party labels are more important information than changes in party manifestos to the electorate in Britain. These findings present important implications for party competition and for democratic representation.

In Chapter 4, *The President, Political Parties, and Legislative Behavior in Brazil: An Application of Bayesian Multilevel IRT Modeling*, I focus on the role of the president and political parties in Brazilian legislative process at the national level, in which political exchanges between the government and legislature is an essential feature. The existence of the non-ideological effect on voting behavior makes the estimation of legislators' policy preferences based on roll calls questionable. In this chapter, I derive a spatial model of voting in which voting behavior is induced by both ideological motivations and coalition dynamics, and develop a random item-difficulty ideal-point model implied by the spatial voting model. Applying the proposed model to the analysis of roll-call votes in the Brazilian Chamber of Deputies between 2003 and 2006, I find that the coalition dynamic is influential but not definitive on party-based voting behavior of legislators. I also show that the estimated positions of Brazilian parties from the proposed model correspond approximately to their perceived ones on the ideological dimension.

Finally, Chapter 5 concludes with a discussion of contributions made by this dissertation to the literature.

Chapter 2

A Bayesian Approach to Endogenous Rarely Changing Variables in Time-series Cross-sectional Analyses

In national-level studies of public social spending, many efforts have been made to explain why some countries consistently pursue more welfare-enhancing policies than others. From the perspective of institutional scholars, democratic institutions play an important role in linking public preferences to policy decisions. In theory, democratic regimes enable politicians to make credible policy commitments to the electorate through competitive elections, so people who prefer redistributive policies are able to influence policy decisions by electing their representatives, compared to nondemocratic ones (North, 1990; Olson, 1993). However, the realization of redistributive policies in democracies may be observed in the long term because it takes time for politicians to establish their credibility of providing public goods (Keefer, 2007; Keefer and Vlaicu, 2008) or represent the interests of the underprivileged (Huber et al., 2006; Huber, Mustillo and Stephens, 2008). This raises two empirical questions: Do democracies pursue more redistributive policies than nondemocracies? Does democracy matter in the long term for social welfare programs?

This chapter addresses these questions by focusing on the estimation issue of dynamics of rarely changing variables such as political regimes in time-series cross-sectional (TSCS) data analyses. In the analysis of TSCS data, which have both intertemporal and cross-sectional variations, researchers are able to control for unobserved heterogeneity across units to eliminate omitted variable bias in estimation by including unit-specific effects. However, it is problematic to estimate time-invariant and/or slowly changing variables with unit-specific effects by two conventional approaches: fixed-effects (FE) models and random-effects (RE) models (Baltagi, 2005; Hsiao, 2003; Wooldridge, 2002). This problem becomes more troublesome when the lagged dependent variable (LDV) is included into a panel data model, which is usually used to account for dynamics, since the assumption of strict exogeneity of covariates is violated (Baltagi, 2005; Hsiao, 2003; Wooldridge, 2002).

Building on the structure of simultaneous equation modeling and on error-component formulations, a Bayesian simultaneous equation model is developed here with hierarchical features that accommodate the correlation between unit effects and explanatory variables. The complexity of this specification requires estimation with Markov chain Monte Carlo (MCMC) methods. A Bayesian approach offers flexibility for complex model specifications and resolves the inferential problems that arise in non-Bayesian multilevel models (Carlin and Louis, 2000; Gelman and Hill, 2007; Gill, 2008*a*).

To assess the performance of the Bayesian simultaneous equation model with hierarchical features presented in this chapter, I employ a Monte Carlo study, in which I compare the proposed model with alternative estimators in estimating the coefficients of correlated time-variant and rarely changing variables. The simulation results show that the proposed model not only performs as well, or better than alternative estimators in terms of bias and efficiency, but also provides additional information on the degree of the correlation between covariates and unit effects. The proposed model is applied to analyzing the effects of political regimes on social spending in Latin America, where countries have different democratic experiences and different social welfare systems. Consistent to the results in some previous studies (Avelino, Brown and Hunter, 2005; Kaufman and Segura-Ubiergo, 2001) but contrary to others (Huber, Mustillo and Stephens, 2008), this chapter finds little evidence that democracy has an effect on social security and welfare spending in Latin America.

This chapter has two primary contributions. First, methodologically, I propose a model to deal with the problem of the correlation between unit effects and covariates in the framework of RE models, which provide a more efficient and less biased estimation. Second, substantively, this chapter contributes to our understanding of the effects of political regimes on social spending in general. The results of democratic effects on social spending in Latin American countries provide researchers opportunities to re-examine theoretical arguments.

2.1 Dynamic Models for Rarely Changing Variables

I start with a discussion of a dynamic panel data model and show how the effects of slowly changing variables are represented by this model specification.[1] The definition of time-invariant and slowly changing variables in this chapter is based on Plümper and Troeger (2007), which defines two categories of time-invariant variables. In the first category, variables are time invariant by definition such as gender and race. In the second category, variables are time invariant only for the period under analysis. A time-invariant variable can be turned into a rarely changing variable due to the selection of time periods. The time-invariant variables considered in this chapter belong to the second category. In TSCS data, it is more often that these variables are time invariant in some units and rarely changing in others. Hereafter, I use rarely changing or slowly changing variables to refer to variables that are not time invariant for all units and time-invariant variables to those that are time invariant for all units.

2.1.1 Dynamics of Rarely Changing Variables

Suppose that there exist TSCS data with a continuous outcome variable, y_{jt}, for unit $j = 1, \cdots, J$ measured at time $t = 1 \cdots, T_j$, where the number of time points for individual units may differ. A major advantage of TSCS data is that researchers are able to control for unobserved heterogeneity across units and/or through time to

[1]The effects of time-invariant or rarely changing variables on the outcome can be modeled within the potential outcome framework (Rubin, 1974) and can be estimated by matching methods (Rubin, 1973; Rosenbaum and Rubin, 1983) or by Bayesian approaches (Chib, 2007; Chib and Jacobi, 2007). However, since group-level effects are ignored in matching methods and the dynamics of explanatory variables are the main concern, these methods are not considered in this chapter.

avoid omitted variable bias. A standard approach to deal with this problem is to allow intercepts to vary across units and over time in a panel data model given by

$$y_{jt} = \boldsymbol{x}'_{jt}\boldsymbol{\alpha} + \boldsymbol{w}'_{jt}\boldsymbol{\beta} + \boldsymbol{z}'_{j}\boldsymbol{\gamma} + \delta_j + \lambda_t + \varepsilon_{jt}, \tag{2.1}$$

where $\boldsymbol{x}_{jt}$ is a vector of time-variant explanatory variables, $\boldsymbol{w}_{jt}$ is a vector of slowly changing explanatory variables, and $\boldsymbol{z}_j$ is a vector of time-invariant explanatory variables with corresponding unknown coefficients $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$, respectively, δ_j and λ_t denote unit-specific and time-specific effects, respectively, and ε_{jt} denotes the error term. It is assumed that the error term is independently, identically distributed with mean zero and finite variance σ^2_ε.

To model dynamics, one standard approach is to include a lagged dependent variable into Equation (2.1). Thus, the model is specified as:

$$y_{jt} = \phi y_{j(t-1)} + \boldsymbol{x}'_{jt}\boldsymbol{\alpha} + \boldsymbol{w}'_{jt}\boldsymbol{\beta} + \boldsymbol{z}'_{j}\boldsymbol{\gamma} + \delta_j + \lambda_t + \varepsilon_{jt}, \tag{2.2}$$

where $y_{j(t-1)}$ is the LDV with an autoregressive coefficient ϕ.[2] It is argued that when including the LDV appropriately captures dynamics of explanatory variables, autocorrelation can be eliminated (Beck and Katz, 1996, 2009; Keele and Kelly, 2006). This works particularly well when covariates are rarely changing and their effects decay over time. For simplicity, I consider a balanced TSCS data structure (that is,

[2]This specification for TSCS models has an identical counterpart in time series models (Enders, 2004; Hamilton, 1994). A debate on this dynamic specification in political science can be found in Achen (2000), Beck (2001), Beck and Katz (1996, 2009), De Boef and Keele (2008), Keele and Kelly (2006), and Plümper, Troeger and Manow (2005).

$T_j = T$ for all j) and a dynamic panel data model with an LDV, unit-specific effects, and one explanatory variable which is slowly changing. Thus, the model has the form:

$$y_{jt} = \phi y_{j(t-1)} + \beta w_{jt} + \delta_j + \varepsilon_{jt}. \tag{2.3}$$

Moreover, I assume that the process $\{y_{jt}\}$ is stationary for individual units, that is, formally, $|\phi| < 1$, which means that there is no unit roots.

To explain how the inclusion of an LDV can capture the dynamic effects of the slowly changing variable w_{jt} on the outcome, by recursive substitution, Equation (2.3) can be rewritten as:

$$y_{jt} = \beta \sum_{q=0}^{\infty} \phi^q w_{j(t-q)} + \frac{\delta_j}{1-\phi} + \frac{\varepsilon_{jt}}{1-\phi}, \tag{2.4}$$

where q refers to the number of lags of w_{jt}. By calculating of the marginal effect of w_{jt} on $y_{j(t+p)}$, one can obtain the dynamic multiplier[3]:

$$\frac{\partial y_{j(t+p)}}{\partial w_{jt}} = \phi^p \beta, \tag{2.5}$$

where p denotes the length of time between the input w_{jt} and the outcome $y_{j(t+p)}$ and t denotes the dates of the observations. For a given unit j, the dynamic multiplier depends only on p, not on t. Therefore, $\frac{\partial y_{jt}}{\partial w_{j(t-1)}} = \frac{\partial y_{j(t+1)}}{\partial w_{jt}} = \phi\beta$.

To derive the dynamic effects of the slowly changing variable on the outcome, suppose that, for a given unit j, w_{jt} increases one unit at time t, that is, $w_{jt} - w_{j(t-1)} = 1$, and persists to time $t+p$. Since $\frac{\partial y_{jt}}{\partial w_{j(t-1)}} = \frac{\partial y_{j(t+1)}}{\partial w_{jt}} = \phi\beta$ and $w_{jt} = \cdots = w_{j(t+p)}$, the impact on $y_{j(t+p)}$ of a permanent change in w_{jt} is given by

$$\frac{\partial y_{j(t+p)}}{\partial w_{jt}} + \frac{\partial y_{j(t+p)}}{\partial w_{j(t+1)}} + \cdots + \frac{\partial y_{j(t+p)}}{\partial w_{j(t+p)}} = \beta(\phi^p + \cdots + \phi^2 + \phi + 1). \tag{2.6}$$

[3]The dynamic multiplier is also referred to as the impulse response function in time series analysis.

For a stationary process $\{y_{jt}\}$, that is, when $|\phi_j| < 1$, the limit of Equation (2.6) as p goes to infinity is the long-run effect of w_j on y_j given by the following[4]:

$$\lim_{p\to\infty}(\frac{\partial y_{j(t+p)}}{\partial w_{jt}} + \cdots + \frac{\partial y_{j(t+p)}}{\partial w_{j(t+p)}}) = \frac{\beta}{1-\phi}. \tag{2.7}$$

An example is presented in Figure 2.1, which shows a permanent one-unit change in the input variable at time period $t = 1$ and the corresponding changes in the outcome, as presented in Equation (2.6), assuming $\phi = 0.9$ and $\beta = 1$. As can be seen in Figure 2.1, the effect of the input on the outcome variable is distributed across several extended time periods. Figure 2.1 also shows that the change in the outcome is increasing and approaches to the long-run effect of the input, that is, $\frac{1}{1-0.9} = 10$, when the time period t goes to infinity.

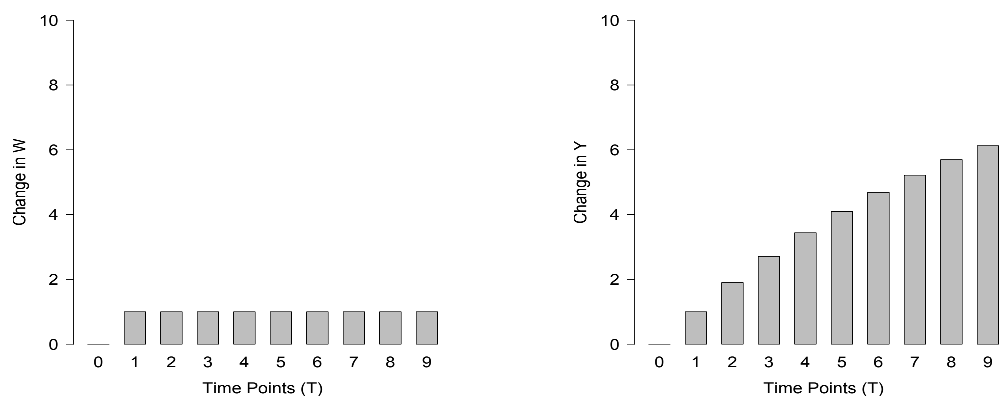


Figure 2.1.: Dynamics of the Input W and the Outcome Y.

[4]As the counterpart in time series models, in TSCS models the cumulative effect on y of a transitory change in w in the limit is the same with the long-run effect, that is, $\sum_{p=0}^{\infty} \frac{\partial y_{j(t+p)}}{\partial w_{jt}} = \frac{\beta}{1-\phi}$ (Enders, 2004; Hamilton, 1994).

2.1.2 Estimation Problems

Two primary issues exist in estimating panel data models: extremely low within panel variation in explanatory variables, and the potential endogeneity of explanatory variables. One case of endogeneity is the correlation between unit effects and covariates, which is considered here. These two issues may lead to inefficient or biased coefficient estimates. It is well known that applied researchers face a quandary when estimating a static panel data model, e.g., Equation (2.1), by two commonly used approaches: FE models and RE models (Baltagi, 2005; Cameron and Trivedi, 2005; Hsiao, 2003; Wooldridge, 2002). Assuming that unobserved unit effects (and/or time effects) are "fixed" across units (and/or over time), the FE model computes unbiased coefficients for time-varying variables even though covariates are correlated with unit effects. However, the FE model is unable to estimate time-invariant variables and produces inefficient estimated coefficients of rarely changing variables since it utilizes only the variation within each unit (Plümper and Troeger, 2007). To estimate time-invariant and rarely changing variables, researchers may employ the RE model, which assumes that the unobserved unit effects are random variables that are distributed independently of the covariates and are independent of covariates. Nevertheless, the RE model is biased if in fact unit effects are correlated with covariates (Baltagi, 2005; Hsiao, 2003; Wooldridge, 2002).

When researchers are interested in dynamics of slowly changing variables as my motivation stated in the first section, the FE and RE models provide consistent parameter estimates under certain conditions. For a dynamic panel data model, e.g., Equation (2.2), on the one hand, since the inclusion of the LDV violates the assumption of strict exogeneity of the regressors, estimates of the time-varying variables under the FE model are biased and inconsistent when T is small, given $J \to \infty$ (Nickell, 1981). On the other hand, the consistency of parameters estimates of the RE model depends on the assumptions about the initial conditions and on infinite sample size T or J, given that the correlation between the covariates and unit effects is zero, or at least, is explicitly modeled (Anderson and Hsiao, 1981, 1982; Hsiao, 2003).

Many scholars have proposed methods to deal with the problem of the correlation between unit effects and covariates and/or to estimate time-invariant and slowly changing variables. One approach is in the form of instrumental variables under the framework of transformed FE models. The major issue for instrumental variable estimators is the choice of instruments and corresponding assumptions. For example, to estimate static panel data models, Hausman and Taylor (1981) specify the exogeneity status of the explanatory variables and use the information on exogenous covariates for endogenous variables. For dynamic panel data models, Anderson and Hsiao (1982) work on instrumental variables for first-differenced equations, which eliminate the problem of correlation between covariates and unit effects. The subsequent studies adopt the generalized method of moments (GMM) framework to derive estimators (e.g., Arellano and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998). However, the GMM estimators may suffer from the problem of instrument proliferation (Roodman, 2009), especially when J is small, which is often the case in studies of political economy.

Also working under the framework of FE models, Plümper and Troeger (2007) propose a three-step procedure for the estimation of time-invariant and rarely changing variables in FE models with unit effects, which is called "fixed-effects vector decomposition" (FEVD).[5] In their Monte Carlo simulations for a static panel data model, they investigate the finite sample properties of several estimators under the situation that time-invariant or time-variant variables are correlated with the unit effects, and find that the FEVD procedure is superior to alternative estimators no matter what the size of correlation between regressors and the unit effects is. Furthermore, they claim that the FEVD procedure also works for dynamic structures in the way that one can simply include the LDV in stages 1 and 3 in the FEVD procedure. However, since the FEVD procedure estimates a standard FE model in the first stage, the FEVD estimator for time-varying variables and rarely changing variables is biased in a dynamic panel data model when T is finite.

Another approach is to explicitly account for the linear dependence between unit effects and explanatory variables under the structure of multilevel modeling, which is a general term for RE models. To do so, when there is potential correlation between time-varying and rarely changing variables and unit effects, one can control for possible correlation by including the within-group means of the endogenous covariates

[5] In a recent symposium on the FEVD procedure, Breusch et al. (2011) and Greene (2011) show that the FEVD estimator can be expressed as an instrumental variable estimator for a particular set of instruments.

as group-level predictors (Bafumi and Gelman, 2006; Gelman, 2006; Mundlak, 1978). When the correlation between unit effects and covariates is explicitly accounted for, the asymptotic properties of RE models hold. However, there is little discussion on finite sample properties of this approach in dynamic multilevel models in estimating time-invariant and rarely changing variables when there is potential correlation between covariates and unit effects although some literature implies that this approach works fine asymptotically (Anderson and Hsiao, 1982; Hsiao, 2003).

2.1.3 A Simultaneous Equation Model for Panel Data

It has been shown that RE estimators for dynamic models are consistent when units $J \to \infty$ and T is finite, assuming that the initial values of outcomes are treated as fixed constants (Anderson and Hsiao, 1982; Sevestre and Trognon, 1985). This property holds only if explanatory variables are not correlated with unit effects or, at least, the correlation is explicitly modeled. In other words, if we are able to model the correlation between covariates and unit effects, RE models would be less biased than those without modeling the correlation and would still be more efficient than FE models in estimating coefficients for rarely changing variables since FE models are biased and inconsistent when $J \to \infty$ and T is small (Hsiao, 2003, pp. 78-90), and are inefficient for rarely changing variables (Plümper and Troeger, 2007).

Building on the structure of simultaneous equation models and multilevel modeling (Chamberlain and Griliches, 1975; Chib, 2008; Zellner, 1970), I develop a model to explicitly account for the dependence between unit effects and endogenous rarely changing variables. For illustration, I consider the model presented in Equation (2.3), which contains an LDV, a rarely changing variable, and unit effects. The resulting analysis also holds for endogenous time-varying variables and can be generalized to models containing both exogenous and endogenous explanatory variables.

I start with the assumption that the error term of the rarely changing variables can be decomposed into "within effects" and "between effects". Within effects are changes over time within the specific individual unit and between effects are effects specific to the individual unit which does not vary over time. Thus, the slowly changing variable can be expressed as

$$w_{jt} = \zeta_0 + \eta_j + \xi_{jt}, \tag{2.8}$$

where ζ_0 denotes the grand mean, η_j denotes between effects, which can be treated as "unit effects" for w_{jt}, and ξ_{jt} denotes within effects (or what is left over). Combining Equation (2.8) with Equation (2.3), we have a simultaneous equation model as follows:

$$w_{jt} = \zeta_0 + \eta_j + \xi_{jt}, \tag{2.9}$$

$$y_{jt} = \phi y_{j(t-1)} + \beta_0 + \beta_1 w_{jt} + \delta_j + \varepsilon_{jt}, \tag{2.10}$$

where I assume that

$$\xi_{jt} \overset{\text{iid}}{\sim} N(0, \sigma_\xi^2), \tag{2.11}$$

$$\varepsilon_{jt} \overset{\text{iid}}{\sim} N(0, \sigma_\varepsilon^2). \tag{2.12}$$

Next, to model the correlation across equations, that is, the dependence between δ_j and w_{jt}, I assume that there exists an unobserved common factor influencing both

y_{jt} and w_{jt} through δ_j and η_j, respectively. Consequently, to model the dependence between δ_j and w_{jt}, I work on the unit effects terms in Equation (2.9) and (2.10), that is, η_j and δ_j. I assume that the joint distribution for the vector of unit effects $\boldsymbol{\psi}_j = (\eta_j, \delta_j)$ is Gaussian. In particular, I assume that

$$(\eta_j, \delta_j) \sim N_2(\mathbf{0}, \mathbf{\Omega}), \forall j, \tag{2.13}$$

where the variance-covariance matrix is of the form

$$\mathbf{\Omega} = \begin{pmatrix} \sigma_\eta^2 & \sigma_{\delta\eta} \\ \sigma_{\delta\eta} & \sigma_\delta^2 \end{pmatrix}. \tag{2.14}$$

Since our interest is the correlation between δ_j and w_{jt}, given $\mathbf{\Omega}$ and σ_ξ^2, we can simply derive this value by the definition of correlation given by

$$\begin{aligned} \operatorname{corr}(\delta_j, w_{jt}) &= \frac{\operatorname{Cov}(\delta_j, w_{jt})}{\sqrt{\operatorname{Var}(\delta_j)\operatorname{Var}(w_{jt})}}, \\ &= \frac{\operatorname{Cov}(\delta_j, \eta_j)}{\sqrt{\operatorname{Var}(\delta_j)\operatorname{Var}(w_{jt})}}, \\ &= \frac{\sigma_{\delta\eta}}{\sqrt{\sigma_\delta^2(\sigma_\xi^2 + \sigma_\eta^2)}}. \end{aligned} \tag{2.15}$$

The second line is derived because the correlation between w_{jt} and δ_j in fact results from the correlation between η_j and δ_j.

I estimate this model by a Bayesian approach, so I complete the model specification by defining the prior distribution. Let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_\varepsilon, \zeta_0, \sigma_\xi, \mathbf{\Omega})$ denote the model parameters where $\boldsymbol{\beta} = (\phi, \beta_0, \beta_1)'$. Following the conventional approach (Carlin and

Louis, 2000; Gelman et al., 2004; Robert, 2001), I use conjugate prior distributions for these parameters as follows:

$$\sigma_\varepsilon^2 \sim IG(v_0/2, d_0/2), \tag{2.16}$$

$$\boldsymbol{\beta}|\sigma_\varepsilon^2 \sim N_3(\boldsymbol{b}_0, \sigma_\varepsilon^2 \boldsymbol{B}_0), \tag{2.17}$$

$$\sigma_\xi^2 \sim IG(g_0/2, h_0/2), \tag{2.18}$$

$$\zeta_0|\sigma_\xi^2 \sim N(m_0, \sigma_\xi^2/a_0), \tag{2.19}$$

$$\boldsymbol{\Omega} \sim \text{Inv-Wishart}(\nu_0, \boldsymbol{\Lambda}_0^{-1}), \tag{2.20}$$

where $v_0, d_0, \boldsymbol{b}_0, \boldsymbol{B}_0, g_0, h_0, m_0, a_0, \nu_0$, and $\boldsymbol{\Lambda}_0$ are hyperparameters, which can be assigned values to reflect prior information about the corresponding parameters or to give diffuse forms.[6] Furthermore, I assume priori independence of $(\boldsymbol{\beta}, \sigma_\varepsilon^2), (\zeta_0, \sigma_\xi^2)$, and $\boldsymbol{\Omega}$ and, therefore, our prior distribution of $\boldsymbol{\theta}$ is given by

$$\begin{aligned}\pi(\boldsymbol{\theta}) = IG&\left(\sigma_\varepsilon^2 | \frac{v_0}{2}, \frac{d_0}{2}\right) N_3\left(\boldsymbol{\beta}|\sigma_\varepsilon^2, \boldsymbol{b}_0, \boldsymbol{B}_0\right) \\ &\times IG\left(\sigma_\xi^2 | \frac{g_0}{2}, \frac{h_0}{2}\right) N\left(\zeta_0|\sigma_\xi^2, m_0, a_0\right) IW\left(\boldsymbol{\Omega}|\nu_0, \boldsymbol{\Lambda}_0^{-1}\right).\end{aligned} \tag{2.21}$$

Let $\boldsymbol{y}_j = (y_{j1}, y_{j2}, \cdots, y_{jT})'$, $\boldsymbol{y}_{j,-1} = (y_{j0}, y_{j1}, \cdots, y_{j(T-1)})'$, $\boldsymbol{w}_j = (w_{j1}, w_{j2}, \cdots, w_{jT})'$, $\boldsymbol{\xi}_j = (\xi_{j1}, \xi_{j2}, \cdots, \xi_{jt})'$, and $\boldsymbol{\varepsilon}_j = (\varepsilon_{j1}, \varepsilon_{j2}, \cdots, \varepsilon_{jt})'$ be $T \times 1$ vectors of observations for unit $j = 1, 2, \cdots J$. The simultaneous equation model has the form

$$\begin{pmatrix} \boldsymbol{w}_j \\ \boldsymbol{y}_j \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \beta_1 \mathbf{I}_T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{w}_j \\ \boldsymbol{y}_j \end{pmatrix} + \begin{pmatrix} \zeta_0 \mathbf{I}_T & \mathbf{0} \\ \beta_0 \mathbf{I}_T & \phi \mathbf{I}_T \end{pmatrix} \begin{pmatrix} \boldsymbol{e} \\ \boldsymbol{y}_{j,-1} \end{pmatrix} + \begin{pmatrix} \boldsymbol{e} & \tilde{0} \\ \tilde{0} & \boldsymbol{e} \end{pmatrix} \begin{pmatrix} \eta_j \\ \delta_j \end{pmatrix} + \begin{pmatrix} \boldsymbol{\xi}_j \\ \boldsymbol{\varepsilon}_j \end{pmatrix}, \tag{2.22}$$

[6]In the Bayesian approach, proper prior distributions can help in identification of the sampling model (Lindley, 1972). In this case, the simultaneous equation is identified since ζ_0 is assumed to be normally distributed with mean known.

where $\mathbf{0}$ is a $T \times T$ matrix, $\boldsymbol{e} = (1, \cdots, 1)'$ is a $T \times 1$ vector, $\tilde{0} = (0, \cdots, 0)'$ is a $T \times 1$ vector. The multivariate regression representation of the structural form is given by

$$\boldsymbol{Y}_j = \boldsymbol{B}\boldsymbol{Y}_j + \boldsymbol{\Gamma}\boldsymbol{X}_j + \boldsymbol{Z}\boldsymbol{\psi}_j + \boldsymbol{U}_j, \tag{2.23}$$

where $\boldsymbol{\psi}_j = (\eta_j, \delta_j)'$, $\boldsymbol{U}_j \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, and

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_\xi^2 \mathbf{I}_T & \mathbf{0} \\ \mathbf{0} & \sigma_\varepsilon^2 \mathbf{I}_T \end{pmatrix}. \tag{2.24}$$

Also, let $\boldsymbol{y} = (\boldsymbol{y}_1, \cdots, \boldsymbol{y}_J)$, $\boldsymbol{w} = (\boldsymbol{w}_1, \cdots, \boldsymbol{w}_J)$, and $\boldsymbol{\psi} = (\boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_J)$. The likelihood function for the simultaneous equation model denoted by $p(\boldsymbol{w}, \boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\theta})$ has the form

$$p(\boldsymbol{w}, \boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\theta}) = \prod_{j=1}^{J} p(\boldsymbol{w}_j, \boldsymbol{y}_j|\boldsymbol{\psi}, \boldsymbol{\theta}), \tag{2.25}$$

where $p(\boldsymbol{w}_j, \boldsymbol{y}_j|\boldsymbol{\psi}, \boldsymbol{\theta}) = N_K(\boldsymbol{B}\boldsymbol{Y}_j + \boldsymbol{\Gamma}\boldsymbol{X}_j + \boldsymbol{Z}\boldsymbol{\psi}_j, \boldsymbol{\Sigma}) N_2(\boldsymbol{\psi}_j|\mathbf{0}, \boldsymbol{\Omega})$, where $K = 2T$.

From the Bayes theorem, the joint posterior distribution of interest, $\pi(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{w}, \boldsymbol{\psi})$, is as follows:

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{w}, \boldsymbol{\psi}) \propto p(\boldsymbol{y}, \boldsymbol{w}|\boldsymbol{\psi}, \boldsymbol{\theta}) p(\boldsymbol{\psi}|\boldsymbol{\theta}) \pi(\boldsymbol{\theta}). \tag{2.26}$$

This joint distribution is of a type that can be efficiently processed by MCMC methods (see Casella and George, 1992; Chib and Greenberg, 1995; Gelfand and Smith, 1990), which can be implemented in programs such as WinBUGS (Lunn et al., 2000) and JAGS (Plummer, 2003).[7]

The drawback of the proposed model is that the correlation between the LDV and unit effects is not explicitly accounted for. Without accounting for the correlation,

[7]Under the assumed conjugate prior distributions, the full conditional distributions required for implementing the Gibbs sampler have been derived and can be found in the Appendix.

the proposed model shares the same properties with the RE model. That is, the consistency of parameter estimates depends on the assumptions about the initial conditions and on infinite sample size T or J (Hsiao, 2003).

The statistical model presented in this section has several advantages in the analysis of TSCS data, in which the standard assumptions underlying classical linear regression models are violated (see Stimson, 1985; Beck and Katz, 1995). First of all, the structure of simultaneous equation modeling allows researchers to explicitly model correlations across equations. With regard to the correlation between unit effects and covariates, I connect time-invariant components of the error terms in two equations, by which we are able to investigate the degree of the correlation.

Second, the structure of multilevel modeling appropriately captures the error structure in TSCS data and uses the information available from different levels. Multilevel models allow the variation within and between units to be estimated conditional on the information at all levels rather than assuming no variability (e.g., the pooled OLS model) or maximal variance (e.g., the FE model) at the unit level (Gelman and Hill, 2007; Raudenbush and Bryk, 2002; Shor et al., 2007). With this "borrowing strength" from various levels, multilevel modeling provides partial pooling estimates in each unit, appropriately accounting for uncertainty, and allows researchers to estimate rarely changing variables along with the unit effects.

Third, Bayesian inferences is easier for complex model specifications. The combination of simultaneous equation modeling and multilevel modeling produce complicated model specifications. The conventional estimators such as maximum likelihood (ML) estimator may provide invalid inferences when the information in the data may not be enough to precisely estimate the elements of the covariance matrix (Gelman, 2006). With regard to hierarchical data like TSCS data, in which the number of higher-level units is often small and the number of observations varies across groups, large-sample properties do not hold (Raudenbush and Bryk, 2002). In contrast, Bayesian inferences is based on describing the posterior distribution given the data and the prior distribution and are thus not only flexible for complex model specifications but also able to resolve inferential problems that rely on large-sample theory.

Finally, Bayesian approaches allow researchers to describe all unknown quantities with probability statements that reflects levels of uncertainty (Gill, 2008*a*). Sometimes researchers are not only interested in the parameters specified in the model but also interested in a function of these parameters. For example, for a dynamic panel data model, the immediate effect (β_1), the dynamic effects presented in Equation (2.6), and the long-run effect of a slowly changing variable presented in Equation (2.7) are of main interest. Using Bayesian approaches, we are able to obtain the posterior distributions of these quantities and, thus, any information on these quantities.

2.2 Monte Carlo Simulations

In this section, I employ Monte Carlo simulations to assess the finite sample properties of the proposed Bayesian simultaneous equation model (BSEM) model, compared to alternative estimators that have been proposed or suggested to estimate correlated rarely changing variables.

2.2.1 Simulation Design

The data generation process (DGP) for the simulations is as follows:

$$y_{jt} = \mu + \phi y_{j(t-1)} + \beta_1 x_{1jt} + \beta_2 x_{2jt} + \beta_3 x_{3jt} + \beta_4 w_{1jt} + \beta_5 w_{2jt} + \beta_6 w_{3jt} + \delta_j + \varepsilon_{jt}, \quad (2.27)$$

where x_1, x_2, and x_3 are time-varying variables and w_1, w_2, and w_3 are rarely changing variables, δ_j denotes the unit effects, and ε_{jt} denotes the error term. The two groups of variables are drawn from a normal distribution with different means and variances; δ_j is normally distributed with mean zero and variance one; the error term ε_{jt} is assumed to be white noise and drawn from a standard normal distribution. Among the covariates, x_3 and w_3 are correlated with the unit effects δ_j and, thus, x_3, w_3, δ, and ε vary for each replication while x_1, x_2, w_1, and w_2 are fixed across all experiments. In the experiments, the correlation between x_3 and the unit effects $\text{corr}(x_3, \delta_j) = \{0, 0.1, 0.2, \cdots, 0.9\}$ and the correlation between w_3 and the unit effects $\text{corr}(w_3, \delta_j) = \{0, 0.1, 0.2, \cdots, 0.9\}$ are varied. The true values of the coefficients are held constant throughout all experiments as follows:

$$\mu = 1, \phi = 0.8, \beta_1 = 0.5, \beta_2 = 2, \beta_3 = -1.5, \beta_4 = -2.5, \beta_5 = 1.8, \beta_6 = 3.$$

I simulate the data with different numbers of time periods $T = \{15, 30, 60\}$ and fixed number of units $J = 20$. The number of replications is 50 for each experiment.

The DGP presented in Equation (2.27) follows the simulation design in Plümper and Troeger (2007) but differs in three fundamental respects. First, the unit effects vary across replications for each experiment, which account for random variation in the unit effects (Breusch et al., 2011). Second, the rarely changing variables I consider are persistent for the rest of periods under analysis once they change at a certain (randomly assigned) time point. Finally, the dynamics of explanatory variables are taken into account by the inclusion of the LDV in the data generation process.[8]

The BSEM employed here contains two endogenous covariates, which is an extension of the one discussed in Section 2.1.3. This Bayesian model with vague priors is estimated with MCMC techniques and implemented in `JAGS` 3.1.0 called from `R` version 2.14.2 (`R2jags` Su and Yajima, 2012). The estimation was performed with three parallel chains of 50,000 iterations each to be conservative. The first half of the iterations were discarded as a burn-in and 5 as thinning, and thus 15,000 samples were generated. The convergence of Markov chains was tested by standard diagnostic tools such as Geweke, Gelman-Rubin, Raftery-Louis, and Heidelberger-Welch (Gill, 2008*b*) and was conducted by an easy to use `R` function `superdiag` that integrates all of the standard empirical diagnostics (Tsai and Gill, 2012). The results show no evidence of non-convergence.

The alternative estimators considered include the pooled OLS estimator, the FE estimator, the FEVD estimator, the RE model (without modeling the correlation),

[8]I assume that the initial value $\{y_{j0}\}$ is random and is correlated with unit effects (see Anderson and Hsiao, 1981, 1982). Thus, the initial value $\{y_{j0}\}$ is allowed to affect the equilibrium level and is generated from a normal distribution with equilibrium mean and equilibrium variance. However, $\{y_{j0}\}$ is assumed to be fixed for all estimators in the estimation process.

and a multilevel model (MLM) accounting for the correlation. The three stages for the FEVD estimator is implemented as follows: (1) fit a standard FE model for Equation (2.27); (2) regress the estimated unit effects from stage 1 on rarely changing variable only; (3) rerun Equation (2.27) but include the residuals from stage 2.[9] This procedure is implemented in R. For the MLM, I include $\bar{x}_{3j}$ and $\bar{w}_{3j}$ as the group-level predictors to model the correlation as suggested by Mundlak (1978). The RE and MLM are estimated by maximum likelihood estimation (MLE) through the `lme4` R package (Bates, Maechler and Ben, 2012).[10]

To compare competing estimators, I follow the literature in reporting the average bias and the root mean squared error (RMSE) of estimates. To avoid the bias to be cancelled out, I calculate the average absolute value of bias. The RMSE captures both the bias and the efficiency of the estimators, which is calculated based on the formula $\sqrt{\frac{\sum_{s=1}^{50}\left(\hat{\beta}^{(s)}-\beta_{\text{true}}\right)^2}{50}}$, where $\hat{\beta}$ denotes the estimate, s in the superscript denotes the sth replication, β_{true} is the true value, and 50 is the total number of replications for each experiment.

---

[9]Formally, in stage 2, I fit the model $\hat{u}_j = \beta_4 w_{1jt} + \beta_5 w_{2jt} + \beta_6 w_{3jt} + h_j$, where $\hat{u}_j = \bar{y}_j - \phi^{FE}\bar{y}_{j(-1)}\beta_1^{FE}\bar{x}_{1j} + \beta_2^{FE}\bar{x}_{2j} + \beta_3^{FE}\bar{x}_{3j}$. In stage 3, I fit the model $y_{jt} = \mu + \phi y_{j(t-1)} + \sum_{k=1}^{3}\beta_k x_{kjt} + \sum_{k=4}^{6}\beta_k w_{kjt} + \varphi\hat{h}_j + \varepsilon_{jt}$.

[10]I am especially interested in the performance of the MLM and FEVD in a dynamic setting since they are increasingly used in empirical research, but there are few discussions on their performance. I do not consider GMM estimators here and leaves for future research because GMM estimators explicitly account for the correlation between the LDV and unit effects while the estimators compared here do not.

2.2.2 Simulation Results

In Table 2.1, I report the average absolute value of bias and RMSE over 10 experiments for three different designs: $T = \{15, 30, 60\}$. The results of experiments in which the correlation between the time-varying variable x_3 and unit effects δ is varied are displayed from the second column to the seventh column; those in which the correlation between the rarely changing variable w_3 and unit effects δ is varied are displayed in the eighth column to the thirteenth column. Notice that in the former $\mathrm{corr}(w_3, \delta_j)$ is fixed at the value of 0.3 while in the latter $\mathrm{corr}(x_3, \delta_j)$ is fixed at the value of 0.3.

These results in Table 2.1 are summarized as follows. First, on average, the pooled OLS produces the poorest estimates in terms of the absolute bias and RMSE. As we know and the results confirm, the correlation between the LDV and unit effects would seriously bias the pooled OLS estimator. Furthermore, the correlation between covariates and unit effects exacerbates the problem. The magnitude of bias and RMSE does not decrease as T increases.

Second, among these six estimators, the FEVD estimator produces estimates with the largest bias and RMSE for the rarely changing variable in all cases while its estimates for the LDV and the time-varying variable are exactly the same with those under the FE model. The results indicate that the FEVD does not perform well for estimating rarely changing variables in dynamic panel data models as Plümper and Troeger (2007) claim.

Table 2.1: Average Absolute Bias and RMSE

	corr(x_3, δ) = {0, ⋯ , 0.9}						corr(w_3, δ) = {0, ⋯ , 0.9}					
	Bias			RMSE			Bias			RMSE		
	y_{t-1}	x_3	w_3	y_{t-1}	x_3	w_3	y_{t-1}	x_3	w_3	y_{t-1}	x_3	w_3
OLS	0.015	0.228	0.095	0.017	0.241	0.119	0.016	0.131	0.160	0.018	0.149	0.185
FE	0.006	0.063	0.126	0.007	0.081	0.159	0.006	0.052	0.129	0.007	0.065	0.166
FEVD	0.006	0.063	0.353	0.007	0.081	0.393	0.006	0.052	0.387	0.007	0.065	0.421
RE	0.004	0.099	0.103	0.005	0.113	0.126	0.005	0.050	0.159	0.006	0.063	0.183
MLM	0.005	0.063	0.121	0.006	0.081	0.153	0.005	0.052	0.124	0.006	0.065	0.160
BSEM	0.005	0.063	0.113	0.006	0.080	0.142	0.005	0.052	0.119	0.006	0.065	0.151
	corr(w_3, δ) = 0.3, T = 15						corr(x_3, δ) = 0.3, T = 15					
	y_{t-1}	x_3	w_3	y_{t-1}	x_3	w_3	y_{t-1}	x_3	w_3	y_{t-1}	x_3	w_3
OLS	0.018	0.251	0.112	0.019	0.261	0.140	0.019	0.146	0.155	0.021	0.159	0.180
FE	0.004	0.044	0.083	0.005	0.054	0.104	0.004	0.034	0.080	0.005	0.042	0.100
FEVD	0.004	0.044	0.311	0.005	0.054	0.345	0.004	0.034	0.351	0.005	0.042	0.380
RE	0.003	0.070	0.079	0.004	0.080	0.098	0.004	0.033	0.108	0.005	0.041	0.127
MLM	0.003	0.044	0.081	0.004	0.054	0.103	0.004	0.034	0.078	0.004	0.042	0.098
BSEM	0.003	0.043	0.080	0.004	0.054	0.100	0.004	0.034	0.077	0.004	0.042	0.096
	corr(w_3, δ) = 0.3, T = 30						corr(x_3, δ) = 0.3, T = 30					
	y_{t-1}	x_3	w_3	y_{t-1}	x_3	w_3	y_{t-1}	x_3	w_3	y_{t-1}	x_3	w_3
OLS	0.021	0.265	0.156	0.023	0.276	0.189	0.023	0.159	0.185	0.024	0.171	0.220
FE	0.004	0.030	0.065	0.005	0.037	0.083	0.003	0.023	0.064	0.004	0.029	0.080
FEVD	0.004	0.030	0.338	0.005	0.037	0.366	0.003	0.023	0.351	0.004	0.029	0.376
RE	0.003	0.044	0.065	0.004	0.052	0.083	0.003	0.023	0.081	0.004	0.029	0.098
MLM	0.003	0.030	0.064	0.004	0.037	0.081	0.003	0.023	0.063	0.004	0.029	0.079
BSEM	0.003	0.030	0.063	0.004	0.037	0.081	0.003	0.023	0.063	0.004	0.029	0.078
	corr(w_3, δ) = 0.3, T = 60						corr(x_3, δ) = 0.3, T = 60					

Third, the RE is more biased than the FE in estimating correlated covariates x_3 and w_3 while the RE is more efficient than the FE in estimating slightly correlated rarely changing variable w_3. However, as the correlation is modeled under the MLM, the MLM and the FE perform equally well in estimating coefficients for the LDV and correlated covariates in terms of bias and RMSE. In other words, when the correlation between unit effects and covariates is modeled under the random-effects framework in the way suggested by Mundlak (1978), the MLM is less biased at the expense of efficiency.

Last but not least, the BSEM performs as well, or better than the FE and MLM in terms of bias and RMSE. We can see that the BSEM is as less biased as the FE in all cases and is more efficient than the FE and MLM when T is less than 30. In other words, the proposed model, which explicitly accounts for the correlation between unit effects and covariates, is less biased than the RE and more efficient than the FE. These properties of the proposed model are remarkable especially when the sample size is small.

I then show the effects of the correlation on the absolute bias and RMSE of the estimates for the LDV, time-varying variable x_3, and rarely changing variable w_3. To save space, I only present the results from the design where $T = 15$, but the results remain the same when $T = \{30, 60\}$. Figure 2.2 presents the absolute value of bias and the RMSE of the six estimators for the LDV, x_3 and w_3 when $\text{corr}(x_3, \delta_j)$ is varied and $\text{corr}(w_3, \delta_j)$ is fixed at the value of 0.3. The three panels on the left-hand side ((a), (c), and (e)) show the absolute values of bias and, as can be seen, the pooled OLS produces estimates of the LDV and x_3 with the largest bias. The RE does not perform well for correlated covariate x_3 and gets worse as the size of the correlation between unit effects and the time-varying variable increases. The FEVD estimator has the poorest estimates of rarely changing variables and the FE, MLM, and BSEM

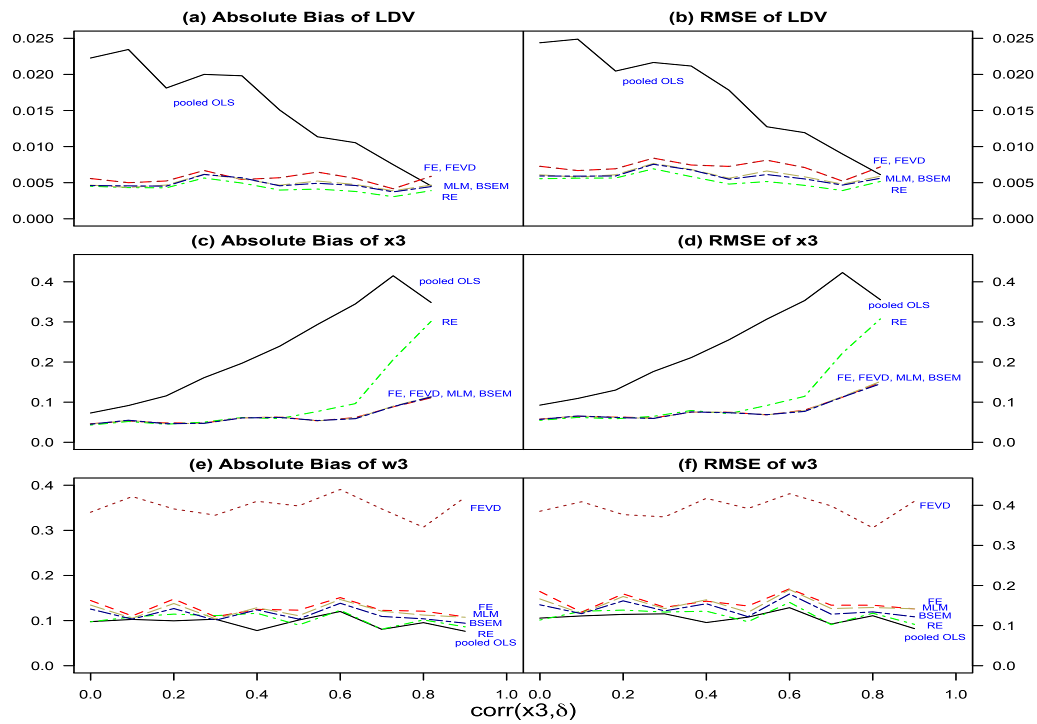


Figure 2.2.: Change in the Absolute Value of Bias and RMSE for corr$(x_3, \delta_j) = 0.3$.

perform more or less equally well in estimating all of these three variables no matter what the size of the correlation is.

As can be seen, the same pattern appears in the RMSE of these estimators resented in the three panels on the right-hand side of Figure 2.2. Simply put, the pooled OLS has the largest RMSE for the LDV and time-varying variable x_3 while the FEVD estimator performs poorly for the estimate of rarely changing variable w_3. The BSEM performs a little bit better than the FE and MLM in terms of RMSE. These results hold no matter what the size of the correlation between unit effects and the rarely changing variable is.

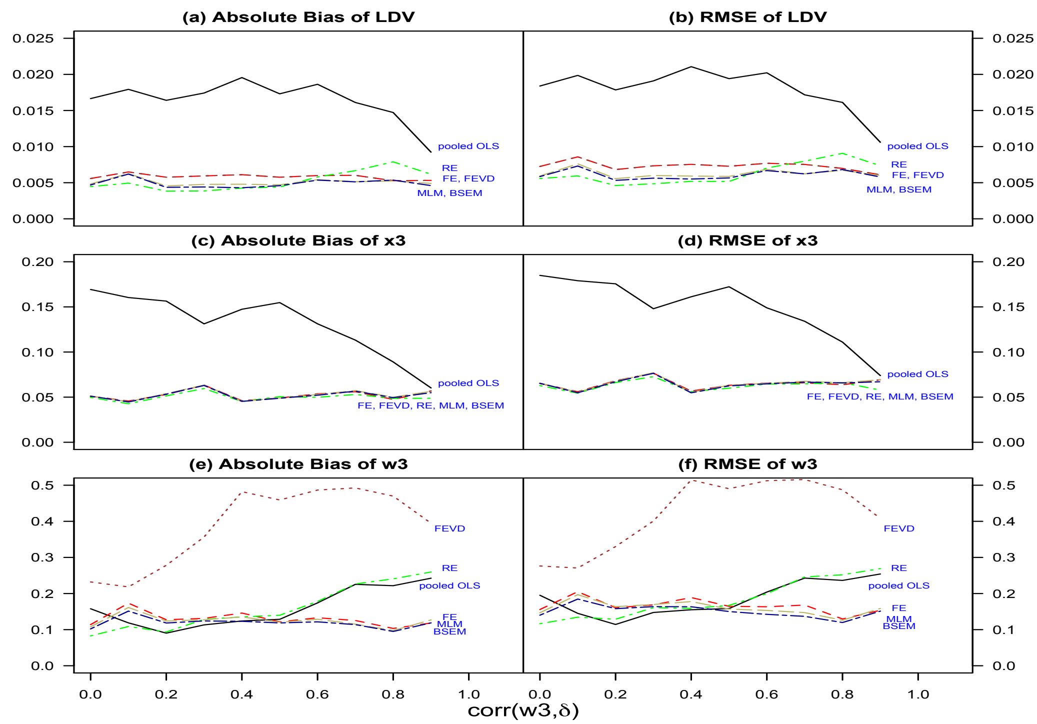


Figure 2.3.: Change in the Absolute Value of Bias and RMSE for $\text{corr}(w_3, \delta_j) = 0.3$.

Figure 2.3 presents the absolute value of bias and the RMSE of the six estimators for the LDV, time-varying variable x_3 and rarely changing variable w_3 when $\text{corr}(x_3, \delta_j)$ is fixed at the value of 0.3 and $\text{corr}(w_3, \delta_j)$ is varied. Figure 2.3 shows that, in general, the pooled OLS has largest biased and inefficient estimates of the LDV and x_3; the FEVD perform poorly in terms of the bias and RMSE in estimating the coefficient of w_3. The FE, MLM, and BSEM perform equally well in estimating the coefficients of the LDV and the correlated time-varying variable x_3 and rarely changing variable w_3.

Finally, I present the estimates of correlation between unit effects and the time-varying variable and rarely changing variable from the proposed model. To save the

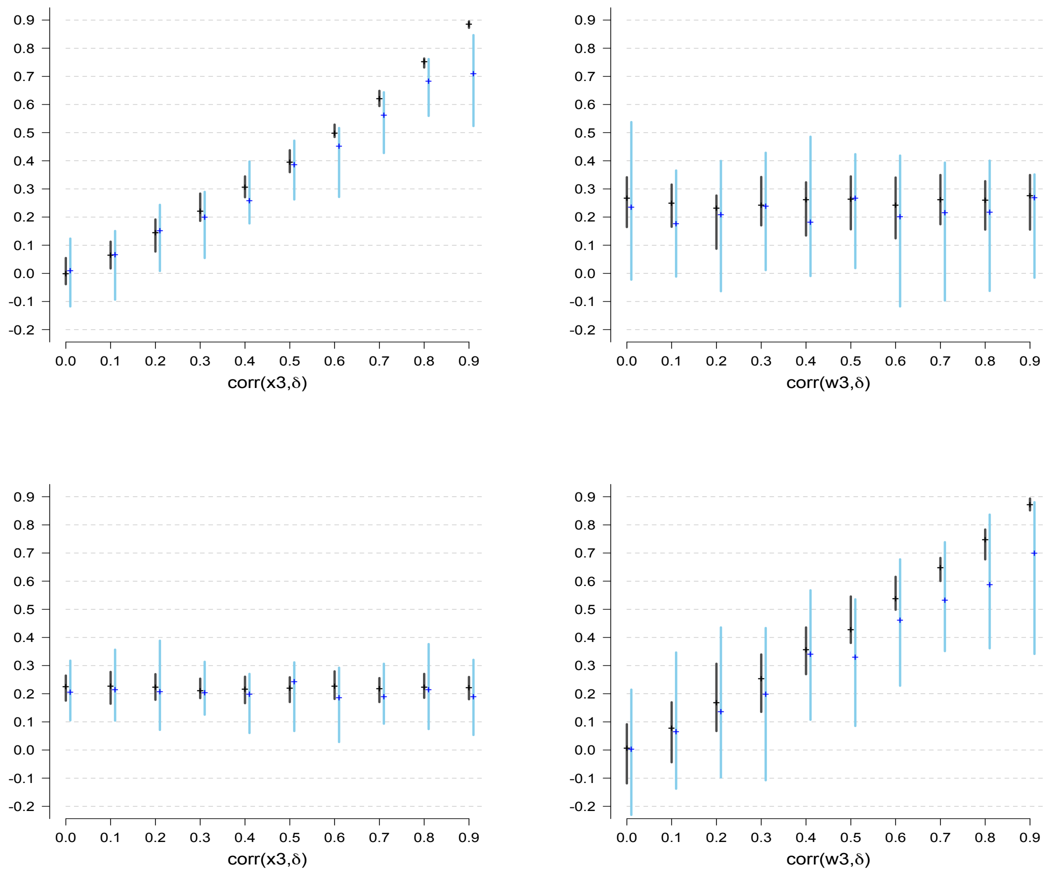


Figure 2.4.: The 80% Highest Posterior Density Intervals of the Estimates of the Correlation between Unit Effects and Covariates.

space, I only show the results from the design where $T = 15$ in Figure 2.4. The two panels on the top are from the design where corr(x_3, δ_j) is varied while corr(w_3, δ_j) is fixed at the value of 0.3; the two panels at the bottom are from the design where corr(w_3, δ_j) is varied while corr(x_3, δ_j) is fixed at the value of 0.3. In this figure, black lines represent the 80% highest posterior density (HPD) intervals of the correlation between unit effects and covariates from generated data while blue lines represent the

80% HPD intervals of the correlation estimates. As can be seen in Figure 2.4, the estimates of correlation between unit effects and the time-varying variable and rarely changing variable are approximately close to those in the generated data.

In short, the proposed model not only perform as well, or better than classical estimators in estimating coefficients for correlated covariates in terms of bias and efficiency but also provide approximately accurate estimates for correlation between unit effects and covariates. Therefore, the proposed model is preferred when the sample size is small and the degree of the correlation between unit effects and covariates is of interest rather than eliminating unit effects such as the FE or assuming no correlation like the RE.

2.3 Application: Social Spending in Latin America

Many studies on democracy and social welfare policy suggest that democracies redistribute more than nondemocracies and thus can improve the welfare of the poor because democratic institutions enable those who prefer redistributive policies to influence policy decisions through electoral competition and representation (Acemoglu and Robinson, 2006; Amenta and Poulsen, 1996; Boix, 2003; Meltzer and Richard, 1981; Bueno de Mesquita et al., 2003). However, the realization of redistributive policies in democracies may be observed in the long term because it takes time for politicians to establish their credibility of providing public goods (Keefer, 2007; Keefer and Vlaicu, 2008) or represent the interests of the underprivileged (Huber et al., 2006;

Huber, Mustillo and Stephens, 2008). Therefore, there is a disagreement on whether we would observe immediate or long-run effects of democracy on social policy.

Some researchers address this issue by focusing on Latin American countries at least for two reasons. First, there is great variation in social welfare systems among Latin American countries. Second, most Latin American countries experienced regime change between democracy and authoritarianism, which provide observations to investigating the effects of democratic regimes between the short run and the long run. However, the empirical results in previous studies are not consistent. For example, some studies show that democratic regimes tend to spend more on overall social programs than authoritarian regimes (Avelino, Brown and Hunter, 2005; Brown and Hunter, 1999) while others find no robust evidence that democracy has an impact on aggregate social expenditures (Kaufman and Segura-Ubiergo, 2001). Moreover, for disaggregate spending, some find that democracies have a positive effect on social security spending in the long term (Huber, Mustillo and Stephens, 2008) while others do not find strong evidence both in the short and long term (Avelino, Brown and Hunter, 2005; Kaufman and Segura-Ubiergo, 2001).

There are at least two methodological issues in the previous quantitative research on the effects of political regimes on social spending. First, the slowly changing property of political regimes are not taken into account or are incorrectly modeled. Some of these studies recognize the problem of estimating rarely change variables in TSCS analysis and, thus, conduct a cross-sectional analysis and do not model persistence of political regimes (e.g., Persson and Tabellini, 2003). Another studies consider the long-term properties of political regimes, but incorrectly model the long-term effects of democracy by cumulating yearly values of democracy variable (Huber, Mustillo and Stephens, 2008).

Second, the absence of country-specific effects in statistical models is one of the problems in past empirical research (Ross, 2006). One of the reasons to exclude unit effects is that when the model is estimated by the OLS, these unit effects eliminate any variation in the outcome explained by time-invariant variables (Plümper, Troeger and Manow, 2005). As discussed before, excluding the unit effects may be at the risk of having omitted variable bias and provide invalid inference. Therefore, as discussed in Section 2, it is better to include an LDV and unit effects into model specification to evaluate the effects of democracy on social welfare spending.

2.3.1 Data and Measurements

I apply the Bayesian simultaneous equation model presented in Section 2 to analyzing social spending in Latin America, which handles the above issues. The dataset I used was collected by Huber et al. (2008).[11] This data set covers a number of political and economic variables in 18 Latin American countries from 1970 to 2000. In this application, the outcome variable is social security and welfare spending, and the main explanatory variable is the type of political regimes.

[11]The data on social security and welfare spending were collected by Huber et al. (2008) from the IMF and those on education and health spending were from the Economic Commission for Latin American and the Caribbean (ECLAC), Cominetti's dataset (1996), and the IMF. The data set can be downloaded from the website `http://www.unc.edu/~jdsteph/common/data-common.html`.

Before discussing model results, two important points need to be observed. First, social security and welfare spending is measured as a percentage of GDP (`SSW`). Since it is bounded between 0% and 100%, we can argue that the proportion of GDP spent on social security is stationary (Beck and Katz, 2009). Second, the distribution of social security spending is right-skewed, so I used the logarithm transformation, which is better described by a normal distribution.[12]

The type of political regimes (`REGIME`), which is rarely changing, has three categories : nondemocratic regime as 1, restricted democracy as 2, and full democracy as 3, which is based on the classification of regimes in Rueschemeyer, Stephens and Stephens (1992). I differentiate restricted democracies from full democracies to represent the difference in the level of credibility that politicians commit structured by political institutions (Keefer, 2007). I control for gross domestic production per capita (million US dollars) adjusted for purchasing power parities (`GDPPC`), the percentage of population that lives in the urban area (`UBNPOP`), the percentage of aged population (`POP65`), export and import as a percentage of GDP as a measurement of trade openness (`TRADE`), and foreign direct investment as a percentage of GDP (`FDI`).[13]

[12]Six observations (1981-1986 in Peru) have values of zero in the measurement of social security/welfare spending, which makes logarithm transformation produce negative infinity. For these six observations, I treat them as missing rather than replace them with small values. Looking at the data carefully, we observe that the measure of spending on social security/welfare is missing in 1979, 1980 and from 1987 to 1989. Consequently, it is reasonable to treat them as missing. It turns out that this setup does not affect the results.

[13]I updated the observations on `FDI` for missingness from World Development Indicators (WDI 2011) from the World Bank. This updated data can be downloaded in `http://data.worldbank.org/data-catalog/world-development-indicators`. For the remaining missing values in `TRADE` and `FDI`, I employed multiple imputation (Rubin, 1987) by the `R` package `mice` (Van Buuren and Groothuis-Oudshoorn, 2011).

Since I do not know which explanatory variable is correlated with unit effects, I model all of them. Therefore, the simultaneous equation model is given by Equation (2.28):

$$\begin{aligned}
\ln(\texttt{SSW}_{jt}) &= \beta_0 + \phi \ln(\texttt{SSW}_{j(t-1)}) + \beta_1 \texttt{REGIME}_{jt} + \beta_2 \texttt{GDPPC}_{jt} + \beta_3 \texttt{UBNPOP}_{jt} \\
&\quad + \beta_4 \texttt{POP65}_{jt} + \beta_5 \texttt{TRADE}_{jt} + \beta_6 \texttt{FDI}_{jt} + \delta_t + \varepsilon_{jt}, \\
\texttt{REGIME}_{jt} &= \zeta_1 + \eta_{1j} + \xi_{1jt}, \\
\texttt{GDPPC}_{jt} &= \zeta_2 + \eta_{2j} + \xi_{2jt}, \\
\texttt{UBNPOP}_{jt} &= \zeta_3 + \eta_{3j} + \xi_{3jt}, \\
\texttt{POP65}_{jt} &= \zeta_4 + \eta_{4j} + \xi_{4jt}, \\
\texttt{TRADE}_{jt} &= \zeta_5 + \eta_{5j} + \xi_{5jt}, \\
\texttt{FDI}_{jt} &= \zeta_6 + \eta_{6j} + \xi_{6jt}.
\end{aligned} \tag{2.28}$$

This model is estimated with MCMC techniques and implemented in `JAGS` 3.1.0 called from `R` (`R2jags`).[14]

2.3.2 Results of Analysis

The results of the determinants of social spending are displayed in Table 2.2. The second column presents the estimates under the proposed Bayesian simultane-

[14]Vague priors are used for parameters. I check the sensitivity to this specification and find that it does not influence the results. The estimation was performed with three parallel chains of 50,000 iterations each to be conservative. The first half of the iterations were discarded as a burn-in period and 5 as thinning and thus 15,000 samples were generated. The convergence of MCMC chains is conducted by using an easy to use `R` function `superdiag` that integrates all of the standard empirical diagnostics (e.g., Geweke, Gelman-Rubin, Raftery-Louis, and Heidelberger-Welch) (Tsai and Gill, 2012), and there is no evidence of non-convergence in these chains.

ous equation model. From the results, we find little evidence that political regimes have an immediate effect on social security and welfare spending. However, as many researchers expect, the effect of urban population on social spending is positive and significant at conventional statistical levels. It indicates that countries that have a 1% of urban population greater than the average level have a 2.3% ($e^{0.023} \simeq 1.023$) increase in social security spending.

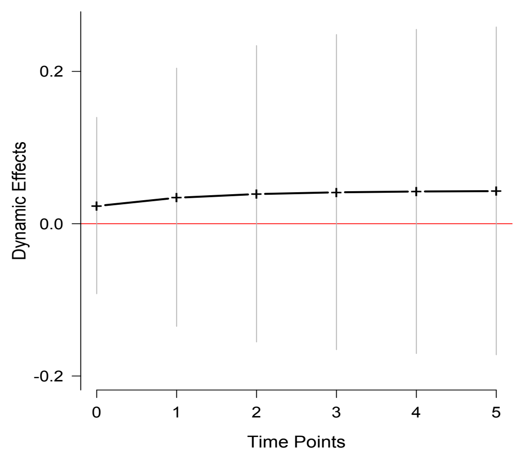


Figure 2.5.: Dynamic Effects of Political Regimes on Social Spending.

Based on Equation 2.6, Figure 2.5 presents the dynamic effects of political regimes over six time periods. As can be seen, the 95% credible intervals cover the value of zero at each time point. Therefore, there is little evidence that political regimes have an impact on social security and welfare spending after these countries change from one type of regimes to another and the regime type persists. In other words, there is little evidence that political regimes influence social security spending both in the short and long term.

Table 2.2: Determinants of Social Security Spending, 1971-2000

Variables	BSEM[a]	Pooled OLS[b]	FE	FEVD	RE	MLM[c]
Intercept	−0.846	−0.772	−.−−	0.001	−1.029	−1.652
	(0.455)	(0.187)	−.−−	(0.170)	(0.312)	(0.693)
$\ln(\texttt{SSW}_{j(t-1)})$	0.459	0.684	0.422	0.422	0.457	0.447
	(0.037)	(0.051)	(0.037)	(0.036)	(0.036)	(0.036)
REGIME	0.023	−0.011	0.025	−0.382	0.009	0.026
	(0.059)	(0.033)	(0.060)	(0.050)	(0.053)	(0.060)
GDPPC	−0.069	0.006	−0.076	−0.076	−0.062	−0.073
	(0.039)	(0.023)	(0.039)	(0.024)	(0.035)	(0.039)
UBNPOP	0.023	0.010	0.025	0.025	0.022	0.024
	(0.009)	(0.004)	(0.009)	(0.004)	(0.007)	(0.009)
POP65	0.042	0.065	0.046	0.046	0.092	0.042
	(0.070)	(0.014)	(0.072)	(0.019)	(0.051)	(0.072)
TRADE	−0.002	0.002	−0.002	−0.002	−0.001	−0.002
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
FDI	0.001	0.002	0.001	0.001	−0.003	−0.001
	(0.015)	(0.018)	(0.016)	(0.014)	(0.015)	(0.016)

Note: Standard errors in parentheses; *a* posterior means reported; *b* PCSEs reported; *c* within-group means included as group-level predictors.

Table 2.2 also presents the estimates from alternative estimators that are widely used in political economy. The third column, first, presents the results from the pooled OLS with panel-corrected standard errors (PCSEs). It shows that the estimate of the LDV is biased upward under the pooled OLS, which is expected (Baltagi, 2005; Hsiao, 2003; Wooldridge, 2002). We also see that the potential biased estimate of POP65 leads to very a different inference from those based on the BSEM, FE, and MLM, which are less biased. Second, comparing the results of the FE estimator represented

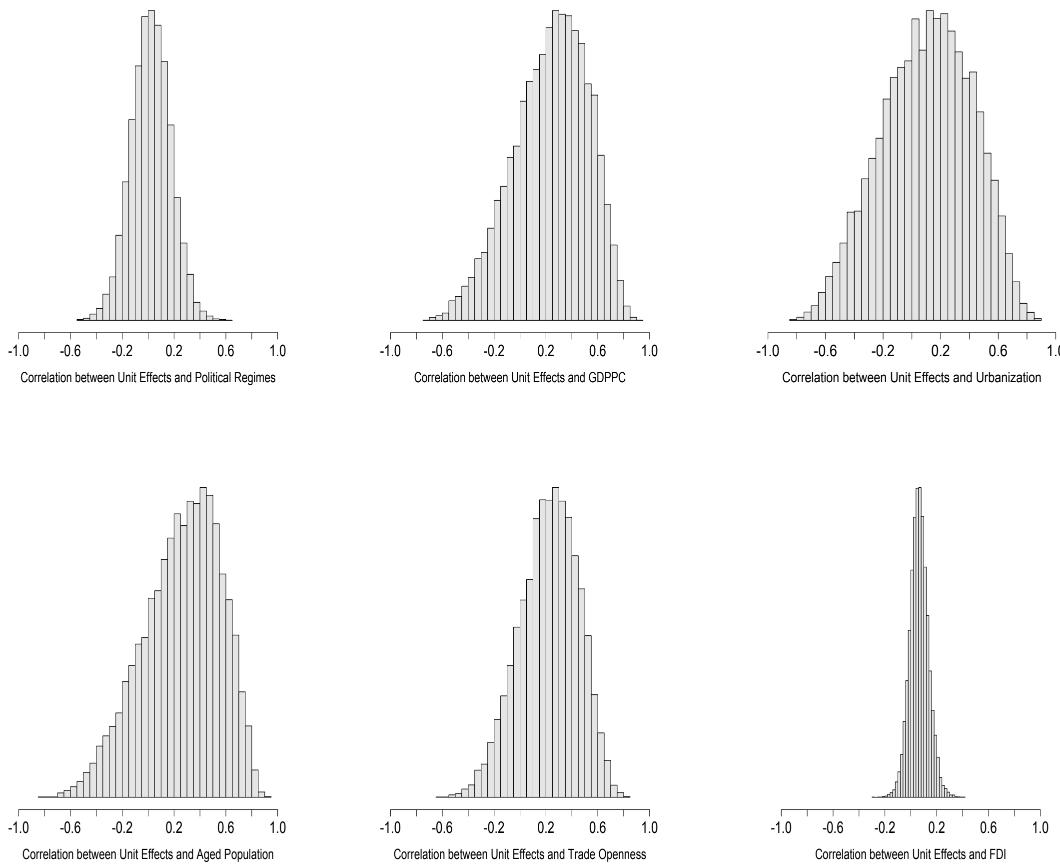


Figure 2.6.: Densities of Estimated Correlation between Unit Effects and Covariates.

in the fourth column with those of the BSEM in the second column, we see that the two estimators produce similar estimates, which means that the BSEM is as less biased as the FE. Moreover, the BSEM has smaller uncertainty than the FE and this uncertainty can be decreases by using informative priors. Third, the estimates under the FEVD reported in the fifth column are the same with those under the FE, except the intercept and political regimes. It shows that the estimate for political regimes under the FEVD estimator is negative and significant at conventional levels.

According to the finding in the simulations, this is incorrect and misleading. We also observe that the FEVD estimator provides smaller standard errors than the FE estimator (Breusch et al., 2011), which may also lead to incorrect inferences, e.g, the effect of `POP65`. Fourth, the results from the RE model are reported in the sixth column, which shows that the RE is more efficient than the BSEM, FE, and MLM. However, it is likely that the estimates are biased when the assumption that the correlation between unit effects and covariates is violated. The discrepancy between estimates under the FE and RE shows the possibility. Finally, I include within-group means of all six covariates as group-level predictors for a multilevel model. The results in the seventh column show that the MLM is as less biased as the FE, which indicates that including within-group means into the model accounts for the correlation between unit effects and covariates. However, under the MLM, we have no information on the degree of the correlation, which is provided by the BSEM.

Finally, I show the estimated correlation parameters provided by the BSEM in Figure 2.6. As can be seen, only the distribution of the estimated correlation between unit effect and political regimes centers around zero, which implies the weak evidence of correlation between unit effects and political regimes. In contrast, Figure 2.6 shows that `GDPPC`, `UBNPOP`, `POP65`, `TRADE`, and `FDI` are moderately and positively correlated with unit effects.

2.4 Discussion

Existing literature argues that democracies pursue more welfare-enhancing policies than nondemocracies. Recently, some studies argue that the realization of redistributive policies in democracies may be observed in the long term (Keefer and Khemani, 2005; Keefer and Vlaicu, 2008). However, the empirical evidence of how political regimes affect redistributive policies is inconsistent. Following studies of political regimes and redistributive policies in Latin America, this chapter examines the dynamic effects of political regimes on social programs by focusing on estimating the effects of rarely changing variables within a Bayesian framework.

The dynamic effects of rarely changing variables such as political regimes are often of great interest in studies of political economy. However, the classical estimators are problematic to estimate the effects of time-invariant and rarely changing variables along with unit effects, the control of which is important in the analysis of welfare states (Kaufman and Segura-Ubiergo, 2001; Ross, 2006). This chapter shows the flexibility and advantages of the Bayesian approach to estimating the coefficients of endogenous explanatory variables along with the parameters of the correlation between unit effects and covariates. The finite sample properties of the proposed model and alternative estimators that are widely applied to panel data analyses are explored in Monte Carlo simulations. The results show that the proposed model not only performs as well, or better than alternative estimators in terms of bias and efficiency, but also provides additional information on the degree of the correlation between unit effects and covariates, which will not be discovered without the proposed model. Applying the Bayesian simultaneous equation model to the study on social spending in Latin America, this chapter finds weak evidence that political regimes affect social welfare spending both in the short and long term.

2.5 Appendix: The Description of the Gibbs Sampler

Suppose that we have TSCS data with variables y_{jt} and w_{jt} for unit $j = 1, \cdots, J$ measured at time $t = 1, \cdots, T$. Let $\boldsymbol{y}_j = (y_{j1}, y_{j2}, \cdots, y_{jT})'$, $\boldsymbol{y}_{j,-1} = (y_{j0}, y_{j1}, \cdots, y_{j(T-1)})'$, and $\boldsymbol{w}_j = (w_{j1}, w_{j2}, \cdots, w_{jT})'$ be $T \times 1$ vectors of observations for unit $j = 1, 2, \cdots J$. Also, let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_\varepsilon, \zeta_0, \sigma_\xi, \boldsymbol{\Omega})$, be our parameters of interest, where $\boldsymbol{\beta} = (\phi, \beta_0, \beta_1)$ and $\boldsymbol{\Omega} = \begin{pmatrix} \sigma_\eta^2 & \sigma_{\eta\delta} \\ \sigma_{\eta\delta} & \sigma_\delta^2 \end{pmatrix}$. The simultaneous equation model has the form

$$\begin{pmatrix} \boldsymbol{w}_j \\ \boldsymbol{y}_j \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \beta_1 \mathbf{I}_T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{w}_j \\ \boldsymbol{y}_j \end{pmatrix} + \begin{pmatrix} \zeta_0 \mathbf{I}_T & \mathbf{0} \\ \beta_0 \mathbf{I}_T & \phi \mathbf{I}_T \end{pmatrix} \begin{pmatrix} \boldsymbol{e} \\ \boldsymbol{y}_{j,-1} \end{pmatrix} + \begin{pmatrix} \boldsymbol{e} & \tilde{0} \\ \tilde{0} & \boldsymbol{e} \end{pmatrix} \begin{pmatrix} \eta_j \\ \delta_j \end{pmatrix} + \begin{pmatrix} \boldsymbol{\xi}_j \\ \boldsymbol{\varepsilon}_j \end{pmatrix}, \tag{2.29}$$

where $\mathbf{0}$ is a $T \times T$ matrix, $\boldsymbol{e} = (1, \cdots, 1)'$ is a $T \times 1$ vector, $\tilde{0} = (0, \cdots, 0)'$ is a $T \times 1$ vector. The multivariate regression representation of the structural form for Equation (2.29) is given by

$$\boldsymbol{Y}_j = \boldsymbol{B}\boldsymbol{Y}_j + \boldsymbol{\Gamma}\boldsymbol{X}_j + \boldsymbol{Z}\boldsymbol{\psi}_j + \boldsymbol{U}_j, \tag{2.30}$$

where $\boldsymbol{U}_j \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ and

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_\xi^2 \mathbf{I}_T & \mathbf{0} \\ \mathbf{0} & \sigma_\varepsilon^2 \mathbf{I}_T \end{pmatrix}. \tag{2.31}$$

Also, let $\boldsymbol{\psi} = (\boldsymbol{\psi}_1, \cdots, \boldsymbol{\psi}_J)$, where $\boldsymbol{\psi}_j = (\eta_j, \delta_j)'$ for $j = 1, \cdots, J$. The likelihood function for the simultaneous equation model denoted by $p(\boldsymbol{w}, \boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\theta})$ has the form

$$p(\boldsymbol{w}, \boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\theta}) = \prod_{j=1}^{J} p(\boldsymbol{w}_j, \boldsymbol{y}_j|\boldsymbol{\psi}, \boldsymbol{\theta}), \tag{2.32}$$

where $p(\boldsymbol{w}_j, \boldsymbol{y}_j|\boldsymbol{\psi}, \boldsymbol{\theta}) = N_K(\boldsymbol{B}\boldsymbol{Y}_j + \boldsymbol{\Gamma}\boldsymbol{X}_j + \boldsymbol{Z}\boldsymbol{\psi}_j, \boldsymbol{\Sigma}) N_2(\boldsymbol{\psi}_j|\mathbf{0}, \boldsymbol{\Omega})$, where $K = 2T$. By the properties of conditional distributions of normal distributions (Greenberg, 2007), we can write down:

$$p(\boldsymbol{w}|\boldsymbol{\psi}, \boldsymbol{\theta}) p(\eta|\boldsymbol{\theta}) = N_T\left(\boldsymbol{w}_j|\zeta_0 \mathbf{I}_T + \eta_j \mathbf{I}_T, \sigma_\xi^2 \mathbf{I}_T\right) N\left(\eta|\frac{\sigma_{\eta\delta}\delta}{\sigma_\delta^2}, \sigma_\eta^2 - \frac{\sigma_{\eta\delta}^2}{\sigma_\delta^2}\right), \tag{2.33}$$

$$p(\boldsymbol{y}|\boldsymbol{\psi}, \boldsymbol{\theta}) p(\delta|\boldsymbol{\theta}) = N_T\left(\boldsymbol{w}_j|\phi \boldsymbol{y}_{j,-1} + \beta_0 \mathbf{I}_T + \beta_1 \boldsymbol{w}_j + \delta_j \mathbf{I}_T, \sigma_\varepsilon^2 \mathbf{I}_T\right) N\left(\delta|\frac{\sigma_{\eta\delta}\eta}{\sigma_\eta^2}, \sigma_\delta^2 - \frac{\sigma_{\eta\delta}^2}{\sigma_\eta^2}\right). \tag{2.34}$$

Then, by using the lemma in Lindley and Smith (1972), we have the unconditional distribution of $\boldsymbol{w}$ and $\boldsymbol{y}$ as follows:

$$p(\boldsymbol{w}|\boldsymbol{\theta}) = N_T\left(\boldsymbol{w}_j|(\zeta_0 + \frac{\sigma_{\eta\delta}\delta}{\sigma_\delta^2})\mathbf{I}_T, (\sigma_\xi^2 + \sigma_\eta^2 - \frac{\sigma_{\eta\delta}^2}{\sigma_\delta^2})\mathbf{I}_T\right), \tag{2.35}$$

$$p(\boldsymbol{y}|\boldsymbol{\theta}) = N_T\left(\boldsymbol{y}_j|\phi \boldsymbol{y}_{j,-1} + \beta_0 \mathbf{I}_T + \beta_1 \boldsymbol{w}_j + \frac{\sigma_{\eta\delta}\eta}{\sigma_\eta^2}\mathbf{I}_T, (\sigma_\varepsilon^2 + \sigma_\delta^2 - \frac{\sigma_{\eta\delta}^2}{\sigma_\eta^2})\mathbf{I}_T\right). \tag{2.36}$$

Chapter 3

Party Policy Strategies in Multiparty Systems: Bayesian Structural Equation Modeling for Dynamic Party Competition

Do politicians change their parties' policy orientations over time in pursuit of votes? Do party policy shifts lead to the electoral consequence desired by politicians? These are central questions in the study of party policy platforms. Since Anthony Downs' seminal work (1957), scholars have used the framework of spatial models to investigate the determinants and results of party policy shifts (e.g., Adams, Merrill and Grofman, 2005; Budge, 1994). However, some of these studies pay little attention to the interdependence between party positioning and electoral results and analyze these two questions separately. Therefore, endogenous dynamic components in electoral competition is less well-understood.

Investigating the dynamics of party policy positioning in the electoral arena is important not only in studying political competition but also in studying democratic representation. Some scholars emphasize the dynamic nature of political process, which is ignored in mainstream models of party competition (Kollman, Miller and Page, 1992, 1998; Laver, 2005). More importantly, in multiparty systems, the number and identity of competing parties are changing endogenously (Laver and Schilperoord, 2007). This phenomenon suggests that politicians may remain or change their parties' ideological positions and associated policies depending on what rival parties they compete with (Schofield and Sened, 2006). Moreover, in multiparty parliamentary systems, party policy positions determine party composition of the government, which in turn makes policy decisions implemented (Austen-Smith and Banks, 1988; Axelrod, 1970; De Swaan, 1973). These implemented policies indicate what citizen preferences are represented. Thus, studying the dynamics of party policy positioning also contributes to research on democratic representation (Budge and McDonald, 2006; McDonald, Mendes and Budge, 2004; Stimson, MacKuen and Erikson, 1995) and responsiveness (Powell Jr., 2004).

In the study of party shifts, little empirical research is conducted to investigate party policy strategies in a dynamic model of party competition which considers the interdependence between party policy positioning and party support in elections. Some existing empirical studies of party policy shifts give a test of Budge's (1994) theory, but ignore the possibility of changing party systems and measurement errors in estimating the unobserved policy positions (e.g., Adams et al., 2004; Adams and Somer-Topcu, 2009*b*). Other studies do not provide systematic empirical results although they analyze the interdependence between party support and party positions in a dynamic two-party system (Kollman, Miller and Page, 1992, 1998), or a dynamic multiparty system (Laver, 2005; Laver and Schilperoord, 2007) through simulations.

To fill this gap, this chapter develops a hierarchical structural equation model analyzed through a Bayesian approach using Markov chain Monte Carlo (MCMC) methods for estimation. Structural equation modeling (SEM) is employed here because, first, party policy orientations are unobserved and can be modeled by the measurement equation of SEM as well as measurement errors. Second, the simultaneous equation of SEM can examine the relationship between two endogenous variables: party policy orientations and electoral outcomes. Finally, a hierarchical framework is incorporated into the structural equation model to account for heterogeneous effects due to the changing sets of parties. Moreover, a Bayesian approach with MCMC computational methods offers flexibility for the complex model specifications of SEM (Lee, 2007; Song et al., 2011), and resolves the inferential problems that arise in non-Bayesian approaches (Carlin and Louis, 2000; Gelman and Hill, 2007; Gill, 2008*a*).

This chapter evaluates theoretical arguments using the Comparative Manifesto Project (CMP) database (Klingemann et al., 2006).[1] The CMP dataset is the most important data on party manifestos across countries and over time, and is widely used to estimate party policy positions (Gabel and Huber, 2000; Klingemann, Hofferbert and Budge, 1994; Laver and Budge, 1992). The dynamics of party competition in two multiparty systems are examined: Britain and Israel. Based on the data and the model specified, I find little evidence that party policy positions influence election

[1]The CMP dataset is publicly available and can be downloaded from the Manifesto Project Database website: `https://manifesto-project.wzb.eu/`.

results, and weak effects of past elections on party policy repositioning. In British politics, party labels are more important than small changes in party manifestos to the electorate. In contrast, the distinction between Israeli parties on the left-right, ideological dimension is not always clear to voters.

3.1 Party Policy Strategies and Multiparty Systems

In representative democracies, organizations like political parties are established by politicians to achieve their multiple goals and to solve a variety of problems in the policy making process (Aldrich, 1995; Müller and Strøm, 1999; Strøm, 1990), in which the public's preferences are indirectly transformed into policy decisions constrained by political institutions. In the spatial theory literature, party policy stances provide important information in the electoral competition process based on which voters can express their preferences about representation (Downs, 1957) and in the government formation process based on which politicians can choose their coalition partners that are policy-connected to their parties (Axelrod, 1970; De Swaan, 1973). For these purposes, politicians can maintain or change their parties' policy orientations over time.

3.1.1 Party Shifts and Election Results

In his influential study of party policy shifts, Budge (1994) argues that politicians are uncertain about election results and voters are uncertain which party better represents their preferences. In such a political environment, party elites provide voters distinctly ideological alternatives and associated policies as information about party policy positions (Budge, 1994; Enelow and Hinich, 1984; Wittman, 1977, 1983) and thus voters would be disappointed about party movement since they may see that as a lack of commitments. Therefore, parties that are clearly marked out by their ideology such as leftist parties (Adams, Haupt and Stoll, 2009) and niche parties (Adams et al., 2006) or that heavily rely on resources from activists (Schofield and Sened, 2006) are less likely to change their policy positions in response to past election results.[2]

On the other hand, past elections serve as one important source of information on the public's preferences over policy alternatives in an uncertain environment. Relying on this information, politicians can adjust their parties' policy programmes in response to past vote gains or losses. When a party loses votes, party elites would change their party's policy position (Janda et al., 1995; Somer-Topcu, 2009), which might be in a reversed direction of their previous shift (Budge, 1994). When a party gains votes, politicians would continue moving in the same direction (Budge, 1994) or stay put (Janda et al., 1995; Somer-Topcu, 2009) since they presume that shifting in the same direction or their position is favored by voters.

[2]Ideological constraints on party policy shifts can also be understood from the perspective of the organizational theory. For example, as organizations, the change process in parties is restricted by their cognitive and institutional friction and thus politicians are unable to immediately alter their parties' platforms (Jones, Sulkin and Larsen, 2003; Walgrave and Nuytemans, 2009), where cognitive limitations refer to ideological constraints and institutional friction refers to the rigidity of internal decision rules.

With regard to electoral effects of party positioning, the existing literature provides two conflicting arguments and inconsistent evidence.[3] On the one hand, some studies argue that policy moderation leads to vote gains (Calvert, 1985; Downs, 1957; Lin, Enelow and Dorussen, 1999) and find empirical evidence supporting this argument (Adams, Merrill and Grofman, 2005; Adams and Somer-Topcu, 2009*a*; Alvarez, Nagler and Bowler, 2000; Ezrow, 2005). On the other hand, parties may enhance their vote gains by presenting distinctly noncentrist positions (Cox, 1990; Schofield and Sened, 2006; Wittman, 1983) and is also supported by empirical evidence (Adams and Merrill, 1999; Adams, Merrill and Grofman, 2005; Schofield and Sened, 2006). In a general analysis, however, Adams et al. (2006) find that neither centripetal nor centrifugal policy shifts have an impact on party support.

Many empirical studies of party shifts rely on the data on party manifestos. Although considerable efforts are devoted to examining electoral effects of party positioning and the effects of election results on party shifts, the studies discussed above pay little attention to endogenous dynamic components in party competition in their empirical analyses. For example, James Adams and his coauthors consider that past party shifts (rather than current party movement) may affect current election results (Adams and Somer-Topcu, 2009*a*) or that past election results may influence current party shifts (Adams et al., 2004, 2006; Somer-Topcu, 2009), but they do not consider

[3]Instead of discussing centripetal and centrifugal incentives in electoral competition, Tavits (2007) shows that parties lose votes when they shift in principled issues, but not in pragmatic issues. In addition, Adams et al. (2006) focus on niche parties and find that niche parties are penalized electorally for moderating and shifting their policy positions.

that both past party shifts and past election results are themselves endogenous to the political process.

3.1.2 Endogenous Dynamics of Multiparty Competition

To better understand the interdependence between parties' electoral support and policy strategies, some scholars emphasize the dynamic nature of political process and point out the inability of mainstream models of party competition for modeling dynamic systems (e.g., Fowler and Smirnov, 2007; Laver and Sergenti, 2011). These scholars assume that political actors such as voters and politicians are not perfectly informed and, thus, both voters and politicians make decisions by looking backward and learning from the past rather than looking forward strategically. In other words, voters continually support or switch to parties that better represent their preferences; politicians constantly adapt their parties' policy platforms to the change in voters' affiliation. This dynamic nature of political process cannot be explained by traditional models of political competition which search for equilibrium solutions because dynamic models of multiparty competition are analytically intractable and, as Laver (2005, p.263) observes, "party competition may never achieve equilibrium" (cf. Schofield and Sened, 2006).

To appropriately model a system with an endogenous dynamic component, more and more scholars rely on the techniques of agent-based modeling (ABM), which is suited to investigate outcomes resulting from continuous interactions between a number of adaptive agents in an evolving dynamic setting (Fowler and Smirnov, 2007;

Laver and Sergenti, 2011). For example, assuming a multidimensional policy space, Kollman, Miller and Page (1992, 1998) and Laver (2005) illustrate the consequences of politicians' selection of different adaptive strategies in two-party competition and multiparty competition, respectively.[4] Fowler and Smirnov (2005) and Smirnov and Fowler (2007) focus on two-party competition in a one-dimensional policy space and investigate how different electoral outcomes affect party policy strategies.

In addition to the interdependence between voters and politicians, an important and unrealistic assumption in traditional spatial models of political competition is relaxed by the techniques of ABM, that is, politicians and their parties are implicitly or explicitly assumed to confront the same competitors in political competition. In most of the multiparty systems, or proportional representation (PR) electoral systems which encourage multiparty systems (Cox, 1997; Duverger, 1954), party entry and party exit from political competition is highly likely, which implies that the set of competing parties may change in different elections. To address this issue, Laver and Schilperoord (2007) extend Laver's (2005) ABM of dynamic party competition and treat the number and identity of parties as endogenous to the political process. One of the implications from this study is that the changing set of competing parties in elections may affect the interdependence between party support and party policy strategies.

[4] Kollman, Miller and Page (1992, 1998) consider three different search algorithms: random adaptive parties, climbing adaptive parties, and generic adaptive parties. Laver (2005) evaluates four different adaptive strategies: (1) stickers, (2) aggregators, (3) hunters, and (4) predators. Budge, Ezrow and McDonald (2010) provide a comparison between two sets of parties' policy strategies discussed in Budge (1994) and Laver (2005) and show their similarities.

However, from the settings of ABMs, it is not clear how voters determine which party better represents their preferences if they make decisions only by looking backward. Since there is only one party or a subset of parties forming the government prior to elections, voters need additional information to recognize policy positions of parties that are not in the government. One source of this information is party programmes and/or manifestos. If voters also look at "instantaneous" information to make decisions, then party manifestos also serve as one important source of information on current party policy positions. Moreover, although the techniques of ABM are proposed to model dynamic party competition, the interdependence between elections and party positioning developed by the techniques of ABM is not tested empirically.

In this chapter I extend Budge's (1994) theory to dynamic multiparty competition. In his article, Budge argues that parties are clearly differentiated from each other in terms of ideology and, therefore, party policy orientations are quite stable in the sense that parties move within their ideological space in a unidimensional policy space. He also argues that past elections are important information on the public's preferences over policy alternatives based on which politicians may move their party. In dynamic party systems, these arguments imply that party elites may change their party's policy strategies for the current election in response to past election results, but the change of party policy positions depends on how their party's position differs from rival parties' in the previous election and how "informative" past election results are. For example, a party may gain votes in the previous election. If this party is significantly distinct from its rivals in terms of policy positions, it may stay put or maintain the distinction because party elites presume that the distinction is the reason why they gain votes. But if this party's vote share is extremely lower than other parties' even though they gain votes in the previous election, party elites may change their party's position because they are far away from where the majority of the voters are.

Furthermore, the repositioning of this party in the current election in turn affects current election results. However, the electoral effects of party repositioning are contingent on previous election results because the decision rules for politicians to set party policy are influenced by the previous election. In other words, the previous election partially influences how party repositioning affects the current election. For example, a party may have a significant change in their position in the current election, compared to that in the previous election. But if this party has a high vote share in the previous election, the change in party position may not lead to vote gains in the current election. I now turn to modeling these complex effects.

3.2 The Bayesian Structural Equation Model

In this section, a longitudinal structural equation model is proposed to deal with the interdependence between party support and party policy orientations as well as measurement issues of party policy positions. The path diagram of the model is presented in Figure 3.1. Consider a set of observations for party policy orientations and party support, respectively: for party $j = 1, \cdots, J$ measured at election time point $t = 1, \cdots, T_j$, we have a $p_1 \times 1$ random vector $\boldsymbol{y}_{jt}$ for the former and a $p_2 \times 1$

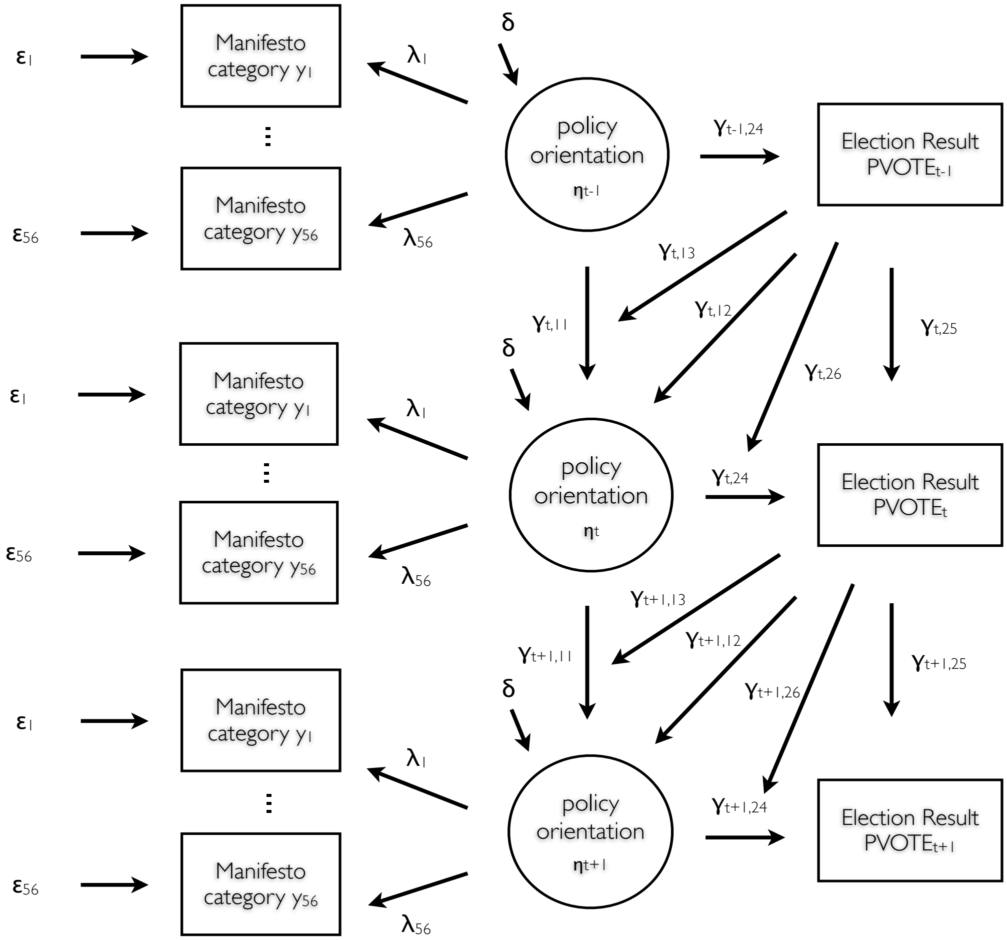


Figure 3.1.: Path Diagram of the Structural Equation Model for Dynamics of Party Competition.

random vector $\boldsymbol{x}_{jt}$ for the latter. The number of elections parties participate, T_j, may differ from party to party.

The framework of SEM consists of two components: a measurement equation and a structural (or simultaneous) equation. The measurement equation, which relates the latent variables to corresponding indicators and takes the measurement errors into account, is basically equivalent to a confirmatory factor analysis (CFA) given by

$$\boldsymbol{u}_{jt} = \boldsymbol{\alpha} + \boldsymbol{\Lambda}\boldsymbol{\eta}_{jt} + \boldsymbol{\varepsilon}_{jt}, \tag{3.1}$$

where $\boldsymbol{u}_{jt} = (\boldsymbol{y}'_{jt}, \boldsymbol{x}'_{jt})'$ is a $p \times 1$ vector of observations and $p = p_1 + p_2$, $\boldsymbol{\alpha}$ is a $p \times 1$ vector of intercepts, $\boldsymbol{\Lambda}$ is a $p \times q$ factor loading matrix, $\boldsymbol{\eta}_{jt}$ is a $q \times 1$ vector of latent common factors representing party policy orientations and party support, and $\boldsymbol{\varepsilon}_{jt}$ is a $p \times 1$ vector of unique factors. The vector of unique factors $\boldsymbol{\varepsilon}_{jt}$ is assumed to be independent of latent factors $\boldsymbol{\eta}_{jt}$ and identically and independently distributed (i.i.d.) with a normal distribution having mean zero vector and a covariance matrix $\boldsymbol{\Psi}$, specifically, $\boldsymbol{\varepsilon}_{jt} \sim N_p(\mathbf{0}, \boldsymbol{\Psi})$, where $\boldsymbol{\Psi}$ is a diagonal matrix with elements ψ_h^2 for $h = 1, \cdots, p$.

Simultaneous equation models are widely used in the field of econometrics to deal with the problem of endogeneity (e.g., Hamilton, 1994; Hsiao, 2003). The simultaneous equation representing relationships between a set of endogenous latent variables is specified in Equation (3.2):

$$\boldsymbol{\eta}_{jt} = \boldsymbol{\beta}_j + \boldsymbol{\Gamma}_t \mathbf{F}(\boldsymbol{\eta}_{j1}, \cdots, \boldsymbol{\eta}_{j(t-1)}) + \boldsymbol{\delta}_{jt}, \tag{3.2}$$

where $\boldsymbol{\beta}_j$ is a $q \times 1$ vector of unit-specific effects, $\boldsymbol{\Gamma}_t$ is a matrix of coefficients that captures election-to-election effects, $\mathbf{F}(\boldsymbol{\eta}_{j1}, \cdots, \boldsymbol{\eta}_{j(t-1)})$ is a vector-valued function which allows flexible autoregressive structures and accounts for dynamics, and $\boldsymbol{\delta}_{jt}$ is a $q \times 1$ vector of residuals. It is assumed that the vector of residuals $\boldsymbol{\delta}_{jt}$ is independent of $\boldsymbol{\eta}_{j1}, \cdots, \boldsymbol{\eta}_{j(t-1)}$ and $\boldsymbol{\varepsilon}_{jt}$, and identically and independently normally distributed with means zero and covariance matrix $\boldsymbol{\Sigma}$, that is, $\boldsymbol{\delta}_{jt} \sim N_q(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a diagonal matrix with elements σ_l^2 for $l = 1, \cdots, q$. The longitudinal structural equation model proposed here is not identified without imposing the identification restrictions. Identifiability constraints on either factor loadings or factor variances are two common approaches.[5] I discuss the identification strategies for data analysis in the next section.

This is a Bayesian specification, so I complete the model specification by defining prior distributions. Let $\boldsymbol{\theta}_u = (\boldsymbol{\alpha}, \boldsymbol{\Lambda}, \boldsymbol{\Psi})$ denote the model parameters in Equation (3.1) and $\boldsymbol{\theta}_\eta = (\boldsymbol{\beta}, \boldsymbol{\Gamma}, \boldsymbol{\Sigma})$ denote the model parameters in Equation (3.2) where $\boldsymbol{\beta} = \{\boldsymbol{\beta}_j; j = 1, \cdots, J\}$ and $\boldsymbol{\Gamma} = \{\boldsymbol{\Gamma}_t; t = 1, \cdots, T_j\}$. Following the conventional approach (Carlin and Louis, 2000; Gelman et al., 2004; Robert, 2001), I use conjugate prior distributions for these parameters. The parameters that are involved in the measurement equation given in Equation (3.1) have the following prior distributions:

$$\boldsymbol{\alpha} \sim N_p(\boldsymbol{\alpha}_0, \boldsymbol{A}_0), \tag{3.3}$$

$$\psi_h^2 \sim IG(a_0/2, b_0/2), \tag{3.4}$$

$$\boldsymbol{\Lambda}_h | \psi_h^2 \sim N_q(\boldsymbol{\Lambda}_0, \psi_h^2 \boldsymbol{H}_0), \tag{3.5}$$

where $\boldsymbol{\Lambda}_h'$ is the hth row of $\boldsymbol{\Lambda}$ for $h = 1, \cdots, p$. The parameters that are involved in the simultaneous equation given in Equation (3.2) have the following prior distributions:

[5]The former involves setting each factor's scale to the scale of one of its indicators, also called unit loading identification (ULI) constraint; the latter is to standardize the latent factors with zero means and variances of unity, also called unit variance identification (UVI) constraint (Kline, 2010). Moreover, in a Bayesian analysis, informative priors that limit parameters to an appropriate range can also achieve an identified model (Congdon, 2006; Lee, 2007).

$$\boldsymbol{\beta}_j \sim N_q(\boldsymbol{\beta}_0, \boldsymbol{B}_0), \tag{3.6}$$

$$\sigma_l^2 \sim IG(c_0/2, d_0/2), \tag{3.7}$$

$$\boldsymbol{\Gamma}_{tl}|\sigma_l^2 \sim N(\boldsymbol{\Gamma}_0, \sigma_l^2 \boldsymbol{G}_0), \tag{3.8}$$

where $\boldsymbol{\Gamma}'_{tl}$ is the lth row of $\boldsymbol{\Gamma}_t$ for $l = 1, \cdots, q$.[6] We can give hyperparameters, $\boldsymbol{\alpha}_0, \boldsymbol{A}_0, a_0, b_0, \boldsymbol{\Lambda}_0, \boldsymbol{H}_0,\ \boldsymbol{\beta}_0, \boldsymbol{B}_0, c_0, d_0, \boldsymbol{\Gamma}_0$, and $\boldsymbol{G}_0$, actual values to reflect prior information about the corresponding parameters or to give diffuse forms. Furthermore, I assume priori independence of $\boldsymbol{\theta}_u$, and $\boldsymbol{\theta}_\eta$ and, therefore, the prior distribution of is given by

$$\begin{aligned}\pi(\boldsymbol{\theta}_y, \boldsymbol{\theta}_\eta) = N_p(\boldsymbol{\alpha}|\boldsymbol{\alpha}_0, \boldsymbol{A}_0) \prod_{h=1}^{p} IG\left(\psi_h^2 | \frac{a_0}{2}, \frac{b_0}{2}\right) N_q(\boldsymbol{\Lambda}_h | \boldsymbol{\Lambda}_0, \psi_h^2 \boldsymbol{H}_0) \\ \times \prod_{j=1}^{J} N_q(\boldsymbol{\beta}_j | \boldsymbol{b}_0, \boldsymbol{B}_0) \prod_{l=1}^{q} IG\left(\sigma_l^2 | \frac{c_0}{2}, \frac{d_0}{2}\right) \prod_{l=1}^{q} \prod_{t=1}^{T_j} N_q(\boldsymbol{\Gamma}_{tl} | \boldsymbol{\Gamma}_0, \sigma_l^2 \boldsymbol{G}_0).\end{aligned} \tag{3.9}$$

Let $\boldsymbol{u} = \{\boldsymbol{u}_{jt} : j = 1, \cdots, J; t = 1, \cdots, T\}$ and $\boldsymbol{\eta} = \{\boldsymbol{\eta}_{jt} : j = 1, \cdots, J; t = 1, \cdots, T_j\}$. The complete data likelihood function for the SEM denoted by $p(\boldsymbol{u}, \boldsymbol{\eta}|\boldsymbol{\theta}_y, \boldsymbol{\theta}_\eta)$ has the form

$$\begin{aligned}p(\boldsymbol{u}, \boldsymbol{\eta}|\boldsymbol{\theta}_y, \boldsymbol{\theta}_\eta) &= p(\boldsymbol{u}|\boldsymbol{\eta}, \boldsymbol{\theta}_u) p(\boldsymbol{\eta}|\boldsymbol{\theta}_\eta) \\ &= \prod_{j=1}^{J} \prod_{t=1}^{T_j} N(\boldsymbol{u}_{jt}|\boldsymbol{\alpha} + \boldsymbol{\Lambda}\boldsymbol{\eta}_{jt}, \boldsymbol{\Psi}) N(\boldsymbol{\eta}_{jt}|\boldsymbol{\beta}_j + \boldsymbol{\Gamma}_t \mathbf{F}(\boldsymbol{\eta}_{j1}, \cdots, \boldsymbol{\eta}_{j(t-1)}), \boldsymbol{\Sigma})\end{aligned} \tag{3.10}$$

[6] $\alpha_0, \boldsymbol{A}_0, a_0, b_0, \boldsymbol{\Lambda}_0, \boldsymbol{H}_0, \boldsymbol{\beta}_0, \boldsymbol{B}_{al}, c_0, d_0, \boldsymbol{\Gamma}_0, \boldsymbol{G}_0$ are hyperparameters, which can be assigned values to reflect prior information about the corresponding parameters.

From the Bayes theorem, the joint posterior distribution of interest, $\pi(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta|\boldsymbol{u}, \boldsymbol{\eta})$, is as follows:

$$\begin{aligned}\pi(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta|\boldsymbol{u}, \boldsymbol{\eta}) &\propto p(\boldsymbol{u}, \boldsymbol{\eta}|\boldsymbol{\theta}_y, \boldsymbol{\theta}_\eta)\pi(\boldsymbol{\theta}_y, \boldsymbol{\theta}_\eta), \\ &= p(\boldsymbol{u}|\boldsymbol{\eta}, \boldsymbol{\theta}_u)p(\boldsymbol{\eta}|\boldsymbol{\theta}_\eta)\pi(\boldsymbol{\theta}_u)\pi(\boldsymbol{\theta}_\eta). \qquad (3.11)\end{aligned}$$

Utilizing the idea of data augmentation (Albert and Chib, 1993; Tanner and Wong, 1987), I augment the data by including the unobserved latent variables $\boldsymbol{\eta}$ and operate with the conditional densities

$$\pi(\boldsymbol{\eta}|\boldsymbol{u}, \boldsymbol{\theta}_\eta) \propto p(\boldsymbol{\eta}|\boldsymbol{\theta}_\eta)p(\boldsymbol{u}|\boldsymbol{\eta}, \boldsymbol{\theta}_u), \qquad (3.12)$$

$$\pi(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta|\boldsymbol{u}, \boldsymbol{\eta}) \propto p(\boldsymbol{u}|\boldsymbol{\eta}, \boldsymbol{\theta}_u)p(\boldsymbol{\eta}|\boldsymbol{\theta}_\eta)\pi(\boldsymbol{\theta}_u)\pi(\boldsymbol{\theta}_\eta). \qquad (3.13)$$

The Bayesian estimates of the parameters and latent variables can be obtained by generating a sufficiently large number of observations from the joint posterior distribution through MCMC methods (see Casella and George, 1992; Chib and Greenberg, 1995; Gelfand and Smith, 1990), which can be implemented in a variety of existing programs such as `WinBUGS` (Lunn et al., 2000) and `JAGS` (Plummer, 2003).[7]

The Bayesian structural equation model has several advantages to the application of dynamic party competition. First, the measurement equation of SEM captures the relationships between unobserved latent variables and observed indicators as well as measurement errors. Party policy orientations cannot be directly observed and can only be measured based on information such as party electoral manifestos and

[7]Under the assumed conjugate prior distributions, the full conditional distributions required for implementing the Gibbs sampler can be found in the Appendix.

survey data. The estimation of unobserved party orientations is modeled by the CFA model.[8]

Second, the simultaneous equation of SEM assesses the relationships among latent variables. In the case of dynamic multiparty competition, I treat both party support and party policy orientations as endogenous and I study the interdependence between party support and party policy orientations through a function of lagged endogenous variables. This function allows flexible autoregressive structures and accounts for dynamics, as I show in the analysis that follows.

Third, a multilevel framework is incorporated into the structural equation model to examine the possibility that the changing sets of competing parties influence the relationships between party policy orientations and party support. The proposed model allows the coefficients matrix $\mathbf{\Gamma}$ in the structural equation to vary over elections. Along with the inclusion of unit effects, the setting of varying coefficients explicitly account for past behavior and time-invariant party-specific effects, which enables us to understand better the determinants of both party support and party positions over time.

Finally, a Bayesian approach with MCMC methods offers flexibility for complex model specifications of SEM (Congdon, 2006; Lee, 2007; Song et al., 2011) and resolves the inferential problems that arise in non-Bayesian approaches (Carlin and Louis, 2000; Gelman and Hill, 2007; Gill, 2008*a*). What makes SEM different from

[8]The factor analytic techniques are also used in several methods of manifesto-based left-right placement of political parties (see Gabel and Huber, 2000, for an evaluation for different methods). However, these methods do not provide any uncertainty of estimated party positions (Benoit, Laver and Mikhaylov, 2009).

the common regression model is the consideration of measurement errors and the existence of latent random variables. Bayesian approaches to SEM directly model the relationship between the latent factors and their indicators rather than the sample covariance matrix (Lee, 2007). The inferences on latent factors given observed indicators are appropriately reflected by Bayesian inferences, which are based on the posterior distribution given the data and the prior distribution rather than relying on large-sample theory.

3.3 Data and Model Specifications

The proposed model is applied to analyzing the interdependence between party support and party policy orientations in two very different multiparty parliamentary systems: Britain and Israel. Britain has a heavily institutionalized party system among Western European countries in the sense that the three major parties in the UK have survived intact and dominated since World War II. Moreover, some scholars have considered Britain as a typical example expressing the unidimensional, left-right political environment (e.g., Adams, Merrill and Grofman, 2005; Budge, 1994). In Israel, on the contrary, most of the parties, including both large and small parties, merge or split in elections for the Knesset over the post 1948 period, and the semantic content of left-right dimension in Israel may not be similar to the semantic content of left-right dimension in many other countries, where these labels (left and right) tend to be applied to different positions on an underlying economic-distributive divide.[9]

[9]Some scholars claim that the left-right, unidimensional political environment is good enough to describe party competition in modern democracies (Budge, 1994). However, it is not true in some

For instance, the two largest parties (Labour and Likud) were losing their fortunes and formed alliances with small parties during the second half of the 1990s due to the direct election of the prime minister and the rising importance of peace and terrorism issues (Arian and Shamir, 1995, 1999, 2002). Because of the changing sets of competing parties and a non-left-right policy dimension underlying party competition, the Israeli multiparty system provides an important case to investigate the dynamics of party competition rather than a difficult case that has way too often been dropped from research (e.g., Adams and Somer-Topcu, 2009*b*; Somer-Topcu, 2009). Thus the dramatic differences in these two cases test the robustness of the model specification.

Following the literature on party policy positioning, this chapter utilizes the CMP data base, which classifies 56 policy categories from party manifestos across countries and over time (Klingemann et al., 2006). The dataset covers party manifestos and election results in general elections of 1992, 1997, 2002, and 2005 in Britain and those of 1992, 1996, and 1999 in Israel. The unit of analysis is a parliamentary party at a given election time-point. In both countries, parties either do not gain seats or do not appear in every election in some districts so that the number of observations within each party differs, which constitutes an unbalanced time-series cross-sectional (TSCS) data structure.

This chapter focuses on the interdependence between party support and party policy orientations. For party $j = 1, \cdots, J$ and election $t = 1, \cdots, T_j$, party support is measured by party j's vote share at election t ($\texttt{PVOTE}_{jt}$). In this case, party support

countries such as Israel. For example, Schofield and Sened (2006) argue that Israeli electoral competition is better explained by a two-dimensional (security and religion) policy space.

is assumed to be measured without measurement errors. With regard to party policy orientations, it is assumed that a collection of p_1-variate vector of indicators $\boldsymbol{y}_{jt}$ measures the unobserved party policy orientations. Following most of the empirical studies of party policy positions, I assume a single latent factor underlying parties' policy concerns. Formally, the measurement equation given in Equation (3.1) can be simplified as

$$\boldsymbol{y}_{jt} = \boldsymbol{\alpha} + \boldsymbol{\Lambda}\eta_{jt} + \boldsymbol{\varepsilon}_{jt}, \tag{3.14}$$

where $\boldsymbol{\alpha}$ is a $p_1 \times 1$ vector of intercepts, $\boldsymbol{\Lambda}$ is a $p_1 \times 1$ vector of factor loadings, η_{jt} is the single latent factor, and $\boldsymbol{\varepsilon}_{jt}$ is a $p_1 \times 1$ vector of unique factors.

The dynamics of party competition are studied through the simultaneous equation given in Equation (3.2). The outcome variables in the simultaneous equation is party shifts (the change in party j's position at election t compared with its position in the previous election $t-1$) and vote gains or losses (the change in party j's vote share at election t compared with its vote share in the previous election $t-1$). I consider different operationalizations of covariates for party shifts and the change in vote share. First, for the regression of party shifts, as discussed in Section 3.1, parties that are clearly marked out by their ideology are less likely to change their policy positions. In unidimensional policy space, these parties are often referred to those located at positions away from the center. I measure party j's distance from the center at the previous election by computing the absolute difference between party j's factor score and the mean factor score of all parties at election $t-1$. Moreover, party elites may adjust their party's policy programmes in response to past election results, but only when this information is informative. To measure this, I calculate the difference between party j's vote share and the maximum vote share gained by a party at election $t-1$. Second, for the regression of vote gains or losses, I examine the effects of centrifugal policy shifts on the change in vote share. I do so by computing the difference between party j's factor score at election t and the mean factor score of all parties at election $t-1$. To evaluate the effects of past elections on the current election, I include party j's vote share at election $t-1$. I also include interaction terms to capture conditional effects in both regression equations. In specific, the relationships between party support and party policy orientations is specified by the following structural equations. For $t = 1$,

$$\begin{pmatrix} \eta_{jt} \\ \mathtt{PVOTE}_{jt} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma_{t,24} & 0 & 0 \end{pmatrix} \begin{pmatrix} \widetilde{\eta}_{j(t-1)} \\ \widetilde{\mathtt{PVOTE}}_{j(t-1)} \\ \widetilde{\eta}_{j(t-1)} * \widetilde{\mathtt{PVOTE}}_{j(t-1)} \\ \eta^*_{jt} \\ \mathtt{PVOTE}_{j(t-1)} \\ \eta^*_{jt} * \mathtt{PVOTE}_{j(t-1)} \end{pmatrix} + \begin{pmatrix} \beta_{j,1} \\ \beta_{j,2} \end{pmatrix} + \begin{pmatrix} \delta_{jt,1} \\ \delta_{jt,2} \end{pmatrix}, \tag{3.15}$$

and for $t = 2 \cdots, T_j$,

$$
\begin{pmatrix} \eta_{jt} - \eta_{j(t-1)} \\ \texttt{PVOTE}_{jt} - \texttt{PVOTE}_{j(t-1)} \end{pmatrix} = \begin{pmatrix} \gamma_{t,11} & \gamma_{t,12} & \gamma_{t,13} & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma_{t,24} & \gamma_{t,25} & \gamma_{t,26} \end{pmatrix} \begin{pmatrix} \widetilde{\eta}_{j(t-1)} \\ \widetilde{\texttt{PVOTE}}_{j(t-1)} \\ \widetilde{\eta}_{j(t-1)} * \widetilde{\texttt{PVOTE}}_{j(t-1)} \\ \eta^*_{jt} \\ \texttt{PVOTE}_{j(t-1)} \\ \eta^*_{jt} * \texttt{PVOTE}_{j(t-1)} \end{pmatrix}
$$

$$
+ \begin{pmatrix} \beta_{j,1} \\ \beta_{j,2} \end{pmatrix} + \begin{pmatrix} \delta_{jt,1} \\ \delta_{jt,2} \end{pmatrix}, \tag{3.16}
$$

where $\beta_{j,1}$ and $\beta_{j,2}$ denote party-specific effects and $\gamma_{t,11}, \gamma_{t,12}, \gamma_{t,13}, \gamma_{t,24}, \gamma_{t,25}$ and $\gamma_{t,26}$ are unknown varying-coefficients, $\widetilde{\eta}_{j(t-1)} = |\eta_{j(t-1)} - \bar{\eta}_{t-1}|$, $\widetilde{\texttt{PVOTE}}_{j(t-1)} = \max(\texttt{PVOTE}_{j(t-1)}) - \texttt{PVOTE}_{j(t-1)}$, and $\eta^*_{jt} = |\eta_{jt} - \bar{\eta}_{t-1}|$.

Different identifiability constraints on the latent factor are used for the two countries based on substantive reasons. For the UK, firstly, following most of the studies on party platforms, I assume unidimensional, left-right policy space underlying these four elections. Second, according to the way that the left-right positions of parties are derived in the CMP data set, I fix the factor loading $(\lambda_1, \cdots, \lambda_{56})$ shown in Figure 3.1 to 1 for "rightist" indicators, -1 for "leftist" indicators, and 0 for others, with higher factor scores denoting a more right-wing emphasis.[10]

[10]The CMP dataset also provides estimates of party left-right positions calculated as the difference between the sum of the references of one group (identified as right) and another group (identified as left), with larger values denoting a more right-wing emphasis (Laver and Budge, 1992). The two groups include 26 out of 56 policy indicators. Most of the empirical studies assume unidimensional policy space and, thus, the left-right positions are used. However, the left-right positions are calculated assuming no measurement errors or no measures of uncertainty are provided.

Regarding the Israeli elections, I do not impose any assumption about the relationship between the manifesto indicators and the latent factor, which reflects the lack of knowledge about what the single latent factor would be. In this case, identifiability constraints are given by standardizing the latent factor with unit variance ($Var(\delta) = 1$ in Figure 3.1). With standardized factors, all the loadings can be taken as free parameters. This setting allows the measurement model to assign weights to the different indicators based on the covariance between the latent factor and indicators.[11] The higher the loadings, the higher the covariance, which provides information on what the underlying latent factor might be.

The values of the hyperparameters in the structural equation model are chosen to reflect vague priors.[12] One exception is the prior for the factor loadings for the Israeli case. Since the latent factor is standardized and the indicators, which are positive, summarize the information of the latent factor, relatively informative prior $N(1,1)$ is used for the factor loadings (Congdon, 2006; Johnson and Albert, 1999). Moreover, the factor loadings are constrained to be positive to avoid label switching (Congdon, 2006, p. 430).[13]

[11] This approach is also used in the "vanilla" method proposed by Gabel and Huber (2000). In their article, Gabel and Huber (2000) test whether measurement errors can be accounted for by country-specific effects and certain covariates such as manifesto's length, party size, and government participation. This examination can be easily done in SEM by directly including a random effect term and fixed covariates in measurement equation. Including fixed covariates also provides a more precise relationship between the latent factor and its indicators.

[12] The sensitivity of posterior distributions to the chosen inputs of the hyperparameters is checked by using informative priors. In specific, according to substantive reasons, $\gamma_{t,11}$ is expected to be negative, and $\gamma_{t,12}$ and $\gamma_{t,25}$ are expected to be positive. These prior expectations are conducted by assuming truncated distributions with corresponding support and small variances. The results are not significantly different.

[13] The implementation of latent trait models via MCMC techniques may raise label switching and identification issues (Skrondal and Rabe-Hesketh, 2004). For latent traits, a labelling issue occurs

The Bayesian structural equation model is estimated with MCMC techniques and implemented in `JAGS` 3.1.0 called from `R` version 2.14.2 (`R2jags`, Su and Yajima, 2012). The estimation was performed with three parallel chains of 50,000 iterations each to be conservative. The first half of the iterations were discarded as a burn-in period and 5 as thinning and thus 15,000 samples were generated.[14]

The estimation of policy orientations in this chapter differs from existing literature in the following way. Measurement errors in the CMP dataset are taken into account in estimating party policy positions and in evaluating the relationships between latent variables. It is claimed that the CMP manifesto indicators contain measurement errors and this problem is not well dealt with (Benoit, Laver and Mikhaylov, 2009). To do so, existing studies of party policy shifts provide "reliable" policy change by considering larger changes in the CMP scale, e.g., ±4-point change (Budge, Ezrow and McDonald, 2010; Tavits, 2007). Considering measurement errors in the structural equation model not only reflects the fact that party manifestos do not precisely provide positions of parties, but also improves the estimates of the regression coefficients since measurement errors can cause bias and inconsistency in estimating causal effects (King, Keohane and Verba, 1994).

because the direction in which the most latent variables are measured is arbitrary. This can be resolved by using formally constrained priors (Congdon, 2006, pp. 425-427).

[14]The convergence of Markov chains was tested by standard diagnostic tools such as Geweke, Gelman-Rubin, Raftery-Louis, and Heidelberger-Welch (Gill, 2008*b*) and was conducted by an easy to use `R` function `superdiag` that integrates all of the standard empirical diagnostics (Tsai and Gill, 2012). The results show no evidence of non-convergence in these chains.

3.4 Results of Analysis

The proposed structural equation model involves a large number of parameters (more than 200). To save space, I report only the results of our main theoretical interests (e.g., factor scores, factor loadings, and regression coefficients).

3.4.1 The British Case

Figure 3.2 shows the estimates of factor scores with uncertainty for each party at different electoral time periods. In the 1992 and 1997 elections (upper panels), the Labour Party (Labour) is the most leftist-oriented party among the set of competing parties and it is not as close to the center as other parties. Compared to the factor scores estimated by the structural equation model, the CMP estimates of left-right party positions, represented by red dots in Figure 3.2, display more extreme positions of the Conservative Party (CONS) and dramatic change of the Labour Party's position (from left to right of the center). Moreover, the results show that in the 2001 election (lower-left panel) the Liberal Democrats (Liberal) and the UK Independence Party (UKIP) are at two extreme positions on the left-right dimension and are significantly different from each other. The estimated positions in 2005 election from the structural equation model shows the general impression of where the three main parties are.

In Figure 3.2, we do not see substantial differences among these parties in terms of left-right positions. Traditional spatial theory predicts convergence of party positions (Downs, 1957) while many scholars have argued that parties do not converge in policy

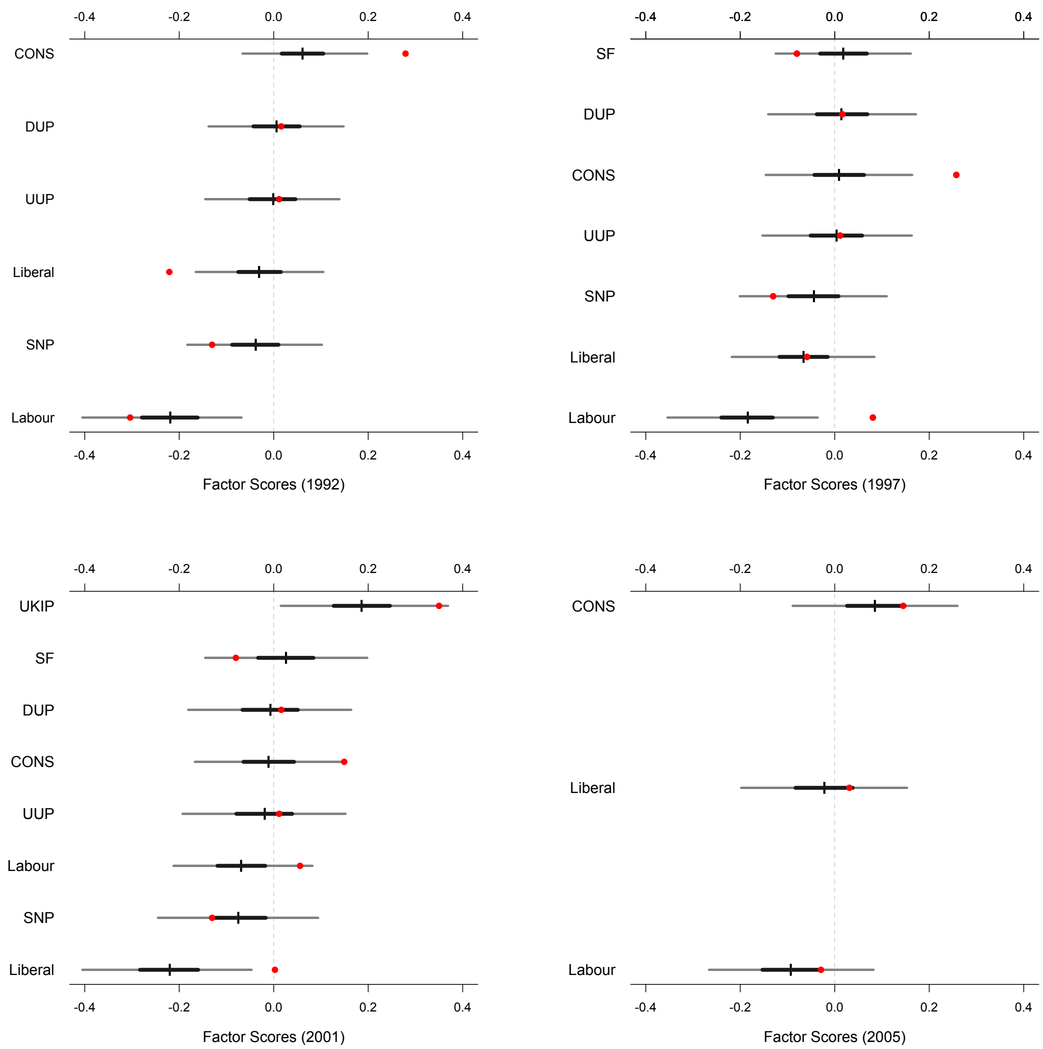
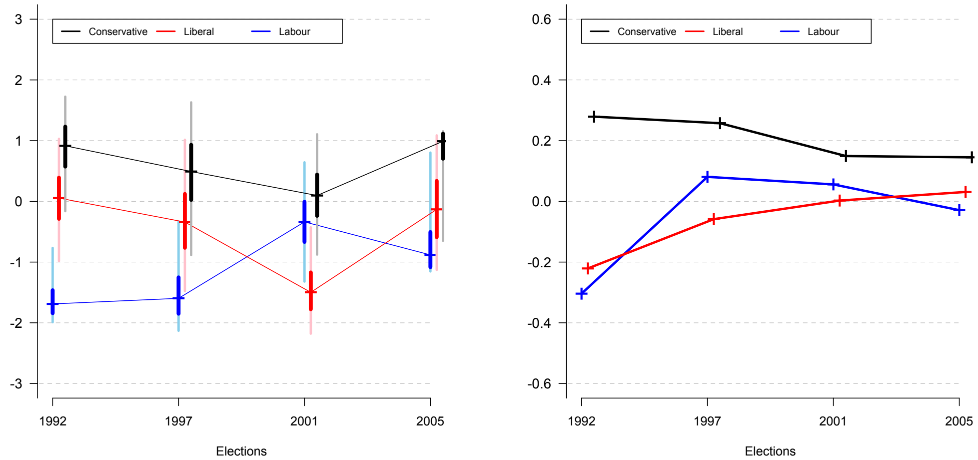


Figure 3.2.: Factor Scores across Parties over time for the UK.

terms for different reasons, especially in multiparty systems (e.g., Adams, Merrill and Grofman, 2005; Aldrich, 1983; Budge, Ezrow and McDonald, 2010; Stokes, 1963). To investigate whether the indifference of party left-right positions results from party movement toward the same position, I provide party movement along the left-right policy dimension over the four elections in Figure 3.3. Panel (a) in Figure 3.3 shows



(a) Standardized factor scores

(b) CMP left-right estimates

Figure 3.3.: Left-right Movement of UK Parties, 1992–2005.

the standardized factor scores over time estimated by the structural equation model, which suggests that the UK parties, at least for the three main parties, do not continuously move toward the center of the unidimensional policy space. Panel (b) in Figure 3.3 shows the CMP left-right estimates over time, from which we observe a certain degree of convergence of party positions.

The result of indifferent party positions is helpful for us to acknowledge that we are uncertain about the true positions and to re-examine Budge's argument that parties are clearly differentiated from each other in terms of ideological positions. In other words, we find that party positions overlap, given the information from party manifestos. As can be seen in Panel (a) of Figure 3.3, although the movements of the Conservative Party and the Labour Party, as the rightist and leftist party, respectively,

Table 3.1: Coefficients of the Structural Equation: Britain, 1992-2005

Outcome	Policy Shifts		Vote Share Change	
Variable	Coefficient	Std. Dev.	Coefficient	Std. Dev.
$\tilde{\eta}_{1992}$	0.000	0.011	$-.--$	$-.--$
$\tilde{\eta}_{1997}$	0.000	0.011	$-.--$	$-.--$
$\tilde{\eta}_{2001}$	0.000	0.011	$-.--$	$-.--$
$\widetilde{\texttt{PVOTE}}_{1992}$	0.000	0.003	$-.--$	$-.--$
$\widetilde{\texttt{PVOTE}}_{1997}$	-0.001	0.003	$-.--$	$-.--$
$\widetilde{\texttt{PVOTE}}_{2001}$	0.008	0.007	$-.--$	$-.--$
$\tilde{\eta}_{1992} \times \widetilde{\texttt{PVOTE}}_{1992}$	-0.001	0.011	$-.--$	$-.--$
$\tilde{\eta}_{1997} \times \widetilde{\texttt{PVOTE}}_{1997}$	-0.001	0.011	$-.--$	$-.--$
$\tilde{\eta}_{2001} \times \widetilde{\texttt{PVOTE}}_{2001}$	0.002	0.011	$-.--$	$-.--$
η^*_{1992}	$-.--$	$-.--$	0.001	0.576
η^*_{1997}	$-.--$	$-.--$	0.000	0.578
η^*_{2001}	$-.--$	$-.--$	-0.002	0.571
η^*_{2005}	$-.--$	$-.--$	0.012	0.570
$\texttt{PVOTE}_{1992}$	$-.--$	$-.--$	-0.851	0.159
$\texttt{PVOTE}_{1997}$	$-.--$	$-.--$	-0.825	0.161
$\texttt{PVOTE}_{2001}$	$-.--$	$-.--$	-0.861	0.166
$\eta^*_{1997} \times \texttt{PVOTE}_{1992}$	$-.--$	$-.--$	0.066	0.563
$\eta^*_{2001} \times \texttt{PVOTE}_{1997}$	$-.--$	$-.--$	-0.066	0.556
$\eta^*_{2005} \times \texttt{PVOTE}_{2001}$	$-.--$	$-.--$	-0.092	0.555
SF	0.011	0.079	0.516	3.870
LABOUR	-0.014	0.053	32.795	5.204
LIBERAL	-0.031	0.064	16.295	3.407
CONS	0.006	0.055	29.445	4.753
UUP	0.009	0.074	0.671	3.212
SNP	-0.004	0.073	1.637	3.204
DUP	0.013	0.073	0.369	3.232
UKIP	0.086	0.099	-0.093	5.246

do not cross each other, those of the Liberal Democrat Party do. In other words, given the party manifestos and the competing parties, the left-right positions of the Liberal Democrats and the Labour Party Party are not clearly differentiated from each other when the Labour Party moves toward the center and so could be those of the Liberal

Democrats and the Conservative Party when the Conservative Party is close to the center.

The estimates of coefficients in the simultaneous equation in the structural equation model are presented in Table 3.1. First, considering the determinants of party shifts, the result suggests that a party's previous position, whether being away from the center or not, does not influence the change in party position in current election. Likewise, we do not find that past election results affect current party shifts. Second, for election results, the result suggests that neither centripetal nor centrifugal party shifts influence a party's vote gains or losses. The weak relationship between party shifts and party support is consistent with the findings in previous empirical studies (Adams et al., 2004, 2006). However, we do observe that the previous election influences the current election in the UK, which indicates that a party with a 1 percent higher vote share in the previous election is likely to lose 0.8 percent of vote in the current election.

We observe that party-specific effects of the three main parties do have an effect on their vote gains and losses. This result means that the vote shares are determined by what these parties are, which do not change over time. In other words, the existing party reputation is more important than changes in party manifestos for the electorate.

3.4.2 The Israeli Case

To understand the underlying single latent factor in Israeli electoral competition, I first look at factor loadings, which indicate the covariance between the latent factor and indicators. Table 3.2 displays the CMP policy indicators and their factor loadings that are reliably larger than 1 at 80 percent highest probability density (HPD) intervals.[15] As can be seen, the underlying latent factor is correlated with policy categories such as `Military`, `Peace`, `Free Enterprise`, `Incentives`, `Economic Goals`, `Social Justice`, `Welfare State Expansion`, `Education Expansion`, `National Way of Life`, and `Multiculturalism`.[16] We can see that some of these policy categories are included as indicators to estimate the CMP left-right party positions, but others are not. The factor loadings in Table 3.2 show that those policy categories included in the CMP estimates are not correlated to each other in the way that is assumed by the method computing the CMP estimates. For example, increasing military expenditure (`Military`) is considered to be opposed to peaceful means of solving crises (`Peace`) in the CMP estimates. However, these two policy categories are positively correlated to the latent factor in Israel. Therefore, assuming the semantic content of left-right dimension in Israeli electoral politics is not appropriate.

Figure 3.4 shows the standardized factor scores for each party at different electoral time periods. The higher the scores are, the more concern over the issues listed in

[15] I report HPD because the factor loadings may be not symmetric and report the level of 80 percent rather than conventional statistical levels (e.g., 95 percent) because low probability reflects our uncertainty over the underlying latent factor.

[16] Descriptions of these policy categories can be found in the codebook of the CMP dataset.

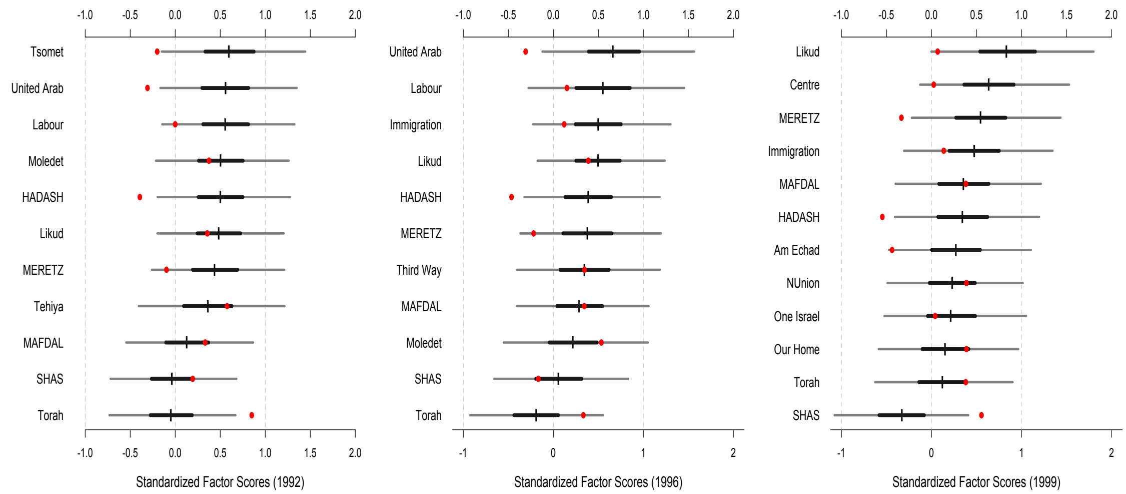
Table 3.2: Factor Loadings and Corresponding Policy Category: Israel, 1992-1999

CMP Policy Category	Mean Factor Loadings	Std. Dev.	80% HPD Interval	
External Relations				
Military: Positive[a] (per104)	2.214	0.898	[1.516	3.794]
Peace: Positive[b] (per106)	2.214	0.898	[1.684	3.683]
Economy				
Free Enterprise: Positive[a] (per401)	2.174	0.846	[1.741	3.153]
Incentives: Positive[a] (per402)	2.174	0.846	[1.119	2.259]
Economic Goals (per408)	2.187	0.816	[1.618	2.993]
Welfare and Quality of Life				
Social Justice: Positive (per503)	2.307	0.860	[1.622	3.146]
Welfare State Expansion: Positive[b] (per504)	2.347	0.823	[1.516	2.777]
Education Expansion: Positive[b] (per506)	2.347	0.823	[3.118	5.173]
Fabric of Society				
National Way of Life: Positive[a] (per601)	2.161	0.838	[1.859	3.666]
Multiculturalism: Positive (per607)	2.161	0.838	[1.593	3.311]

Note: The superscript *a* and *b* denote indicators that are identified by Laver and Budge (1992) as right and left, respectively.

Table 3.2 the parties are. This result suggests little evidence that, overall, party policy positions in Israel are significantly different from each other along the unidimensional policy space. This is because the common factor covers several important policy issues in these elections (Arian and Shamir, 1995, 1999, 2002). Although some parties care more about certain policy issues than others, these parties cannot be differentiated from each other if all of these policy issues condense into one criterion.

Comparing the estimates of party positions from the structural equation model (represented by the line) and the CMP dataset (represented by the red dots), we



(a) (b) (c)

Figure 3.4.: Standardized Factor Scores across Parties over Time for Israel.

observe different, even opposite, patterns.[17] For example, Panel (a) in Figure 3.4 shows that United Torah Judaism (Torah) is ranked as the party (the alliance of parties) that shows least concern over the listed policy issues in the 1992 election. The CMP estimates, however, present that Torah is the most rightist oriented, which emphasizes policy categories such as `Military`, `Free Enterprise`, and `Incentives`. Similar conflicting results appear in the 1996 and 1999 elections.

The estimates of coefficients in the simultaneous equation for Israel are presented in Table 3.3.[18] Like the British case, the result suggests that past election results or past party policy positions do not influence current party policy orientations. Also,

[17]The values from the two estimates are not comparable. I compare these two sets of estimates by considering relative party positions.

[18]The estimates of party-specific effects are not shown here to save space since there are 18 parties. Moreover, we do not find strong party-specific effects in Israel.

we do not find that centripetal or centrifugal party shifts influence a party's vote gains or losses. Unlike the results in the British elections, which indicates the effects of past elections on the current election, the previous election does not have an effect on the current election in Israel. One explanation to this result is that the sets of competing parties are different during these three elections.

Table 3.3: Coefficients of the Structural Equation: Israel, 1992-1999

Outcome	Policy Shifts		Vote Share Change	
Variable	Coefficient	Std. Dev.	Coefficient	Std. Dev.
$\tilde{\eta}_{1992}$	0.000	0.101	$-.--$	$-.--$
$\tilde{\eta}_{1996}$	0.001	0.099	$-.--$	$-.--$
$\widetilde{\texttt{PVOTE}}_{1992}$	-0.014	0.024	$-.--$	$-.--$
$\widetilde{\texttt{PVOTE}}_{1996}$	-0.018	0.037	$-.--$	$-.--$
$\tilde{\eta}_{1992} \times \widetilde{\texttt{PVOTE}}_{1992}$	0.002	0.057	$-.--$	$-.--$
$\tilde{\eta}_{1996} \times \widetilde{\texttt{PVOTE}}_{1997}$	0.011	0.076	$-.--$	$-.--$
η^*_{1992}	$-.--$	$-.--$	0.750	0.758
η^*_{1996}	$-.--$	$-.--$	0.012	0.625
η^*_{1999}	$-.--$	$-.--$	0.026	0.635
$\texttt{PVOTE}_{1992}$	$-.--$	$-.--$	-0.499	0.355
$\texttt{PVOTE}_{1996}$	$-.--$	$-.--$	-0.580	0.405
$\eta^*_{1996} \times \texttt{PVOTE}_{1992}$	$-.--$	$-.--$	-0.138	0.569
$\eta^*_{1999} \times \texttt{PVOTE}_{1997}$	$-.--$	$-.--$	-0.180	0.583

In sum, employing an appropriate technique for dynamic multiparty competition, which considers the uncertainty of unobserved party positions and endogenous dynamics in political process, I find little evidence of the interdependence between party support and party policy orientations in multiparty systems, given the specified model and the data. But I find that party brands matter in a system without frequent party exit and entry.

3.5 Conclusion

This chapter aims at investigating the interdependence between party policy strategies and party support in multiparty systems, in an effort to illustrate the endogenous dynamics of multiparty systems. To do so, a Bayesian structural equation model is proposed to analyze dynamic party competition in the UK and Israel. These are important cases because they represent two very different types of multiparty competition: Britain is an institutionalized party system while the party system in Israel changes over time. I first present an analysis of party competition in the UK, which has rigidly institutionalized party system and is widely recognized as a typical example of a unidimensional, left-right political environment. The results show that, although the Liberal Democrats are generally located between the the Conservative Party (a centre-right party) and the Labour Party (a centre-left party), party manifestos do not provide clear-cut division of party policy positions, which in turn weakens the magnitude of the interdependence between party support and party policy positions. However, I find that party labels do influence electoral outcomes. This is a revealing finding because party labels rather than party manifestos are more important information on party positions for voters.

Second, I explore the policy dimension underlying Israeli electoral competition and report the results of the relationships between party policy strategies and party support. The findings suggest that Israeli parties do not compete with each other in a traditional left-right policy space and that Israeli electoral competition is inappropriately explained by unidimensional policy space. Moreover, the results show that the birth and death of parties deteriorate the distinction between party policy alternatives in Israel. Frequent party system changes and indistinct party policy positions diminish the interdependence between party support and party policy orientations. In other words, the findings show little evidence that party policy positions influence election results and weak effects of past election results on party policy positioning.

The application of SEM to these two very different multiparty parliamentary systems shows its flexibility. I set up the measurement equation to reflect what we know about the relationships between the observed variables and the latent factors in these two countries. While I assume a unidimensional policy space in Israel but have no knowledge of what the single factor is, I assume a unidimensional, left-right policy space in the UK. This model can be extend a multidimensional policy space which may better explain the Israeli electoral competition.

Political parties are what induces democracies to be responsive. Based on this point of view, the ideological positions of parties play two major roles: the representation of public preferences and the political commitments of politicians to voters. Generally speaking, election results are not influenced by party movement but are determined by the reputation of party brands. This implies that the general perceived ideological positions are more important than the changes in party manifestos in specific elections to voters.

3.6 Appendix: Estimation of the Model with Markov Chain Monte Carlo

Let $\boldsymbol{\theta}_u = (\boldsymbol{\alpha}, \boldsymbol{\Lambda}, \boldsymbol{\Psi})$ denote the model parameters in the measurement equation and $\boldsymbol{\theta}_\eta = (\boldsymbol{\beta}, \boldsymbol{\Gamma}, \boldsymbol{\Sigma})$ denote the model parameters in the simultaneous equation where $\boldsymbol{\beta} = (\boldsymbol{\beta}_j; j = 1, \cdots, J\}$ and $\boldsymbol{\Gamma} = \{\boldsymbol{\Gamma}_t; t = 1, \cdots, T_j\}$. From the Bayes theorem, the joint posterior distribution of interest, $\pi(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta | \boldsymbol{u}, \boldsymbol{\eta})$, is as follows:

$$\begin{aligned}\pi(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta | \boldsymbol{u}, \boldsymbol{\eta}) &\propto p(\boldsymbol{u}, \boldsymbol{\eta} | \boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta)\pi(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta), \\ &= p(\boldsymbol{u}|\boldsymbol{\eta}, \boldsymbol{\theta}_u)p(\boldsymbol{\eta}|\boldsymbol{\theta}_\eta)\pi(\boldsymbol{\theta}_u)\pi(\boldsymbol{\theta}_\eta). \end{aligned} \tag{3.17}$$

Utilizing the idea of data augmentation (Albert and Chib, 1993; Tanner and Wong, 1987), we augment the data by including the unobserved latent variables $\boldsymbol{\eta}$ and operate with the conditional densities

$$\pi(\boldsymbol{\eta}|\boldsymbol{u}, \boldsymbol{\theta}_\eta) \propto p(\boldsymbol{\eta}|\boldsymbol{\theta}_\eta)p(\boldsymbol{u}|\boldsymbol{\eta}, \boldsymbol{\theta}_u), \tag{3.18}$$

$$\pi(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta | \boldsymbol{u}, \boldsymbol{\eta}) \propto p(\boldsymbol{u}|\boldsymbol{\eta}, \boldsymbol{\theta}_u)p(\boldsymbol{\eta}|\boldsymbol{\theta}_\eta)\pi(\boldsymbol{\theta}_u)\pi(\boldsymbol{\theta}_\eta). \tag{3.19}$$

The grouped Gibbs sampler is implemented as follows:

1. Generate $\boldsymbol{\eta}$ from $\pi(\boldsymbol{\eta}|\boldsymbol{u}, \boldsymbol{\theta}_\eta)$.

2. Generate $\boldsymbol{\theta} = (\boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta)$ from $\pi(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta | \boldsymbol{u}, \boldsymbol{\eta})$.

3. Go to 1.

We first derive the distribution of latent factors $\boldsymbol{\eta}$ given in Equation (3.18). Given

$$\boldsymbol{u}_{jt} = \boldsymbol{\alpha} + \boldsymbol{\Lambda}\boldsymbol{\eta}_{jt} + \boldsymbol{\varepsilon}_{jt}, \tag{3.20}$$

$$\boldsymbol{\varepsilon}_{jt} \sim N_p(\mathbf{0}, \boldsymbol{\Psi}), \tag{3.21}$$

and

$$\boldsymbol{\eta}_{jt} = \boldsymbol{\beta}_j + \boldsymbol{\Gamma}_t\mathbf{F}(\boldsymbol{\eta}_{j1}, \cdots, \boldsymbol{\eta}_{j(t-1)}) + \boldsymbol{\delta}_{jt}, \tag{3.22}$$

$$\boldsymbol{\delta}_{jt} \sim N_q(\mathbf{0}, \boldsymbol{\Sigma}), \tag{3.23}$$

we have $p(\boldsymbol{\eta}_{jt}|\boldsymbol{\theta}_\eta) = N(\boldsymbol{\eta}_{jt}|\boldsymbol{\beta}_j + \boldsymbol{\Gamma}_t\mathbf{F}(\boldsymbol{\eta}_{j1}, \cdots, \boldsymbol{\eta}_{j(t-1)}), \boldsymbol{\Sigma})$ and we also know that $p(\boldsymbol{u}_{jt}|\boldsymbol{\eta}, \boldsymbol{\theta}_u) = N(\boldsymbol{u}_{jt}|\boldsymbol{\alpha} + \boldsymbol{\Lambda}\boldsymbol{\eta}_{jt}, \boldsymbol{\Psi})$. It can be shown (Lindley and Smith, 1972) that

$$\pi(\boldsymbol{\eta}_{jt}|\boldsymbol{u}_{jt}, \boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta) = N\left(\boldsymbol{\eta}_1, \boldsymbol{\Sigma}_1\right), \tag{3.24}$$

where $\boldsymbol{\eta}_1 = \boldsymbol{\Sigma}_1[\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_j + \boldsymbol{\Gamma}_t\mathbf{F}(\cdot)) + \boldsymbol{\Lambda}^T\boldsymbol{\Psi}^{-1}(\boldsymbol{u}_{jt} - \boldsymbol{\alpha})]$ and $\boldsymbol{\Sigma}_1 = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Lambda}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\Lambda})^{-1}$. Hence, we can obtain $\pi(\boldsymbol{\eta}|\boldsymbol{u}, \boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta) = \prod_{j=1}^{J}\prod_{t=1}^{T_j}\pi(\boldsymbol{\eta}_{jt}|\boldsymbol{u}_{jt}, \boldsymbol{\theta}_u, \boldsymbol{\theta}_\eta)$.

We then derive the distribution of $\boldsymbol{\theta}_u$ and $\boldsymbol{\theta}_\eta$. Since we assume priori independence of $\boldsymbol{\theta}_u$ and $\boldsymbol{\theta}_\eta$, their marginal conditional densities are as follows:

$$\pi(\boldsymbol{\theta}_u|\boldsymbol{u}) \propto p(\boldsymbol{u}|\boldsymbol{\eta}, \boldsymbol{\theta}_u)\pi(\boldsymbol{\theta}_u), \tag{3.25}$$

$$\pi(\boldsymbol{\theta}_\eta|\boldsymbol{\eta}) \propto p(\boldsymbol{\eta}|\boldsymbol{\theta}_\eta)\pi(\boldsymbol{\theta}_\eta). \tag{3.26}$$

Let $u_{jt,h}$ and α_h be the hth element of $\boldsymbol{u}_{jt}$ and $\boldsymbol{\alpha}$, respectively, for $h = 1, \cdots, p$. Under the assumed conjugate prior distributions, the marginal conditional density of $\boldsymbol{\theta}_u = (\boldsymbol{\alpha}, \boldsymbol{\Lambda}, \boldsymbol{\Psi})$ is

$$\boldsymbol{\alpha} \sim N_p(\boldsymbol{\alpha}_1, \boldsymbol{A}_1), \tag{3.27}$$

$$\psi_h^2 \sim IG(a_1/2, b_1/2), \tag{3.28}$$

$$\boldsymbol{\Lambda}_h|\psi_h^2 \sim N_q(\boldsymbol{\Lambda}_1, \boldsymbol{H}_1), \tag{3.29}$$

where

(1) $\boldsymbol{\alpha}_1 = \boldsymbol{A}_1 \left[\boldsymbol{A}_0^{-1}\boldsymbol{\alpha}_0 + \boldsymbol{\Psi}^{-1} \sum_{j=1}^{J} \sum_{t=1}^{T_j} (\boldsymbol{u}_{jt} - \boldsymbol{\Lambda}\boldsymbol{\eta}_{jt}) \right]$

(2) $\boldsymbol{A}_1 = (\boldsymbol{A}_0^{-1} + \boldsymbol{\Psi}^{-1})^{-1}$

(3) $a_1 = a_0 + \sum_{j=1}^{J} T_j$

(4) $b_1 = b_0 + \sum_{j=1}^{J} \sum_{t=1}^{T_j} (u_{jt,h} - \alpha_h - \boldsymbol{\Lambda}'_h \boldsymbol{\eta}_{jt})^2$

(5) $\boldsymbol{\Lambda}_1 = \boldsymbol{H}_1 [\psi_h^{-2} (\boldsymbol{H}_0^{-1} \boldsymbol{\Lambda}_0 + \sum_{j=1}^{J} \sum_{t=1}^{T_j} \boldsymbol{\eta}_{jt} (u_{jt,h} - \alpha_h))]$

(6) $\boldsymbol{H}_1 = \psi_h^2 (\boldsymbol{H}_0^{-1} + \sum_{j=1}^{J} \sum_{t=1}^{T_j} \boldsymbol{\eta}_{jt} \boldsymbol{\eta}'_{jt})^{-1}$.

Now consider the conditional distribution of $\boldsymbol{\theta}_\eta = (\boldsymbol{\beta}, \boldsymbol{\Gamma}, \boldsymbol{\Sigma})$ that is proportional to $p(\boldsymbol{\eta}|\boldsymbol{\theta}_\eta)\pi(\boldsymbol{\theta}_\eta)$. Let $\eta_{jt,l}$ and β_{jl} be the lth element of $\boldsymbol{\eta}_{jt}$ and $\boldsymbol{\beta}_j$, respectively, for $l = 1, \cdots, q$. Following the same reasoning as before, it can be shown that:

$$\boldsymbol{\beta}_j \sim N_q(\boldsymbol{\beta}_1, \boldsymbol{B}_1), \tag{3.30}$$

$$\sigma_l^2 \sim IG(c_1/2, d_1/2), \tag{3.31}$$

$$\boldsymbol{\Gamma}_{t,l} | \sigma_l^2 \sim N(\boldsymbol{\Gamma}_1, \sigma_l^2 \boldsymbol{G}_1), \tag{3.32}$$

where

(1) $\boldsymbol{\beta}_1 = \boldsymbol{B}_1 [\boldsymbol{B}_0^{-1} \boldsymbol{\beta}_0 + \boldsymbol{\Sigma}^{-1} \sum_{j=1}^{J} \sum_{t=1}^{T_j} (\boldsymbol{\eta}_{jt} - \boldsymbol{\Gamma}_t \mathbf{F}(\cdot))]$

(2) $\boldsymbol{B}_1 = (\boldsymbol{B}_0^{-1} + \boldsymbol{\Sigma}^{-1})^{-1}$

(3) $c_1 = c_0 + \sum_{j=1}^{J} T_j$

(4) $d_1 = d_0 + \sum_{j=1}^{J} \sum_{t=1}^{T_j} (\eta_{jt,l} - \beta_{j,l} - \boldsymbol{\Gamma}_{t,l} \mathbf{F}(\cdot))^2$

(5) $\boldsymbol{\Gamma}_1 = \boldsymbol{G}_1 [\sigma_l^{-2} (\boldsymbol{G}_0^{-1} \boldsymbol{\Gamma}_1 + \mathbf{F}(\cdot) \sum_{j=1}^{J} \sum_{t=1}^{T_j} (\eta_{jt,l} - \beta_{j,l})]$

(6) $\boldsymbol{G}_1 = \sigma_l^2(\boldsymbol{G}_0^{-1} + \mathbf{F}(\cdot)\mathbf{F}(\cdot)')^{-1}$.

Chapter 4

The President, Political Parties, and Legislative Behavior in Brazil: An Application of Bayesian Multilevel IRT Modeling

The framework of spatial models has been widely applied to the explanation of many aspects of the political process. One of the most important achievements is theories of lawmaking initially developed from the American Congress (e.g., Cox and McCubbins, 1993, 2005; Krehbiel, 1998) and applied to cross-national studies (e.g., Carey, 2009). To test empirical implications of these models, one essential requirement for operationalization is the development of measurements of political actors' positions in policy spaces or ideological spaces. For example, in the study of legislative politics, many analysts utilize roll-call votes—the recorded votes in legislatures—to measure policy preferences of legislators, which are called ideal points, by using some statistical techniques (e.g., Clinton, Jackman and Rivers, 2004; Poole and Rosenthal, 1997). Despite its merits, however, the principal problem of using roll-call data is that voting records are usually a consequence of political bargaining and, thus, they may not accurately reveal legislators' preferences (Krehbiel, 2000).

Legislative politics in Brazilian Chamber of Deputies is a typical case that legislative behavior can be as much a product of political negotiation as it is of ideological or policy concerns. The legislative powers of the executive authorized by the 1988 constitution allow the president to control much of the distribution of political resources upon which legislators depend for their political survival. With these powers, the president is able to exchange political favors for legislative support by strategically forming coalitions (Alston and Mueller, 2006; Amorim Neto, 2006; Pereira and Mueller, 2004*a*). Although they hold different points of view, leading scholars of Brazil are in agreement on that legislative behavior is structured by coalition dynamics (Ames, 2002; Amorim Neto, 2002*a*; Figueiredo and Limongi, 2000). When, besides ideology, the government-opposition dynamic drives party-based voting behavior in the legislature, standard ideal point estimates do not distinguish between these two influences (Zucco, 2009; Zucco and Lauderdale, 2011).

In this chapter, I show how the setting of random item parameters under the framework of item response theory (IRT) models (Clinton, Jackman and Rivers, 2004) is able to recover estimates of the policy preferences of legislators and to provide tests of the government-opposition effect on legislative behavior. The logic behind this modeling strategy is as follows: Assuming a one-dimensional, ideological content of the underlying politics, legislators make different voting decisions even if they hold the same ideological position. The difference between voting decisions is systematically explained by the government or opposition status of parties through changing the "difficulty" of roll calls confronted by legislators. The proposed model is applied to the analysis of roll-call votes in the Brazilian Chamber of Deputies between 2003 and 2006 to estimate ideal points of parties as well as to investigate the government-opposition conflict. The results show that the estimated positions of Brazilian parties correspond approximately to their perceived ones on the ideological dimension. Moreover, I find that the coalition dynamic is influential but not definitive on party-based voting behavior of legislators.

Uncovering whether Brazilian presidents construct parliamentary-style coalitions in the Congress and rule through them or construct a series of ad hoc coalitions for different issues is important not only as it enhances our knowledge of executive-legislative relations but also from the perspective of democratic representation. This chapter illustrates party-based voting and the role of the president in the legislative process and the representation of voters' preferences. Therefore, studying the role of parties on executive-legislative relations, this chapter contributes to several themes in this field, including party influence in legislatures (Aldrich, 1995; Cox and McCubbins, 1993; Krehbiel, 1999; McCarty, Poole and Rosenthal, 2001), patterns of governance in presidential systems (Linz and Valenzuela, 1994; Shugart and Carey, 1992), and the role of parties as the primary vehicle of representation (Kitschelt, 2000; Sartori, 1976).

4.1 Presidential Powers and Legislative Behavior in Brazil

There is a debate over the role of political parties in structuring legislative voting in the study of Brazilian politics.[1] One side of the debate focuses on the incentives structured by electoral rules and argues that an open-list proportional representation (OLPR) system, which is used to elect members of the Chamber of Deputies, leads to multiple, undisciplined parties in the legislature (Ames, 1995*a*; Mainwaring, 1991). Elected under a system of OLPR in multimember districts, legislators face severe internal party competition and, thus, seek personal votes (Carey and Shugart, 1995; Shugart, Valdini and Suominen, 2005). The constituent-centered nature of legislators' electoral bases weakens party leaders' control over their members in the legislature so that parties are less disciplined and unable to serve as the primary vehicle for executive-legislative cooperation (Mainwaring, 1993, 1999; Mainwaring and Shugart, 1997). Consequently, the nationally minded president has to bargain with individualistic, pork-oriented deputies on a case by case basis to obtain support for legislative agendas (Ames, 1995*b*, 2001).

The other side argues that Brazilian presidents have had success in enacting their legislative agendas supported by disciplined parties due to the centralized decision-making processes inside Congress (Figueiredo and Limongi, 2000; Cheibub, Figueiredo and Limongi, 2009). This argument emphasizes institutions within Congress that empower party leaders to influence the legislative agenda, thus rendering legislative behavior in line with party leaders' indication (Lyne, 2008; Pereira and Mueller, 2004*b*).

[1] See Amorim Neto (2002*b*) for a brief review on this debate.

In other words, incentives from the electoral arena that push toward individualistic behavior are countered by the incentives within Congress that induce party-based behavior. As a result, the president, who cares about policies affecting national interests, are able to count on reliable support from the political parties included in the presidential coalition through bargaining with party leaders, who concern policy reputation of their parties.

Although these two schools of thought draw attention to different political institutions that structure legislators' incentives, both agree that strong presidential powers are the key component to mold a coalition in the Brazilian Congress. The legislative powers of the executive authorized by the Constitution of 1988 lead to a relatively powerful president in executive-legislative relations since the president controls the legislative agenda and resources upon which legislators depend for their political survival (Ames, 2002; Figueiredo and Limongi, 2000; Pereira and Mueller, 2004*a*). The president is able to successfully get their proposals approved through their control over the legislative agenda along with the strategic use of patronage for political support by allocating resources either directly to individual legislators, or through political parties (Alston and Mueller, 2006). Alternatively, the existence of strong legislative powers of the president allows the executive to rule through its prerogatives such as decree without constructing a stable majority coalition (Amorim Neto, 2002*a*, 2006). Therefore, the Brazilian Congress has the potential to oscillate between "atomistic" (per Ames) and "parliamentary" (per Figueiredo and Limongi) modes, depending on the president's strategic choices (Amorim Neto, Cox and McCubbins, 2003).

Even though the president can rely on various legislative powers to rule, I argue that the president would rather form a stable, party-based coalition and rule through it than construct a legislative policy coalition on a case by case basis or rely on executive prerogatives for several reasons. First, as discussed above, the internal rules of Congress enhance party leaders' influence on rank-and-file behavior and, thus, ensure party unity (Figueiredo and Limongi, 2000; Lyne, 2008). Bargaining with party leaders allows the president to achieve a high rate of success in the legislative arena at relatively low cost. Second, the policy reputation of parties provides important information based on which the president (and party leaders) can choose their coalition partners in a multiparty system (Lyne, 2008), just like the process of government formation in parliamentary systems (Axelrod, 1970; De Swaan, 1973). With this information, the presidential calculus of the trade-off between policy outcomes and patronage can be more precise.

Third, parties are the major agents in the electoral arena although an OLPR system encourages candidate-centered competition (Morgenstern, 2004). In Brazil, the ballot structure only indicates the partisanship of candidates and allows voters to choose candidates among parties rather than among factions or coalitions. Fourth, parliamentary-style executive cabinets could be formed for electoral purposes. A party-list electoral system and concurrent electoral cycles in presidentialism make the national policy campaign possible (Shugart and Carey, 1992) and, therefore, a party-based, electoral alliance is likely to become a "parliamentary agenda cartel" (Amorim Neto, Cox and McCubbins, 2003). Finally, in Brazil parties have been becoming more programmatic and politicians have been becoming more party oriented since 1989, compared to those in the earlier period (1945–1964) (Hagopian, Gervasoni and Moraes, 2009; Lyne, 2005). This raises the importance of a party's policy reputation as a commitment of politicians to the electorate, which changes legislators' incentives to preserve their autonomy from party leaders in both the electoral and legislative arenas (Aldrich, 1995; Levy, 2004; Snyder and Ting, 2002).

According to the above discussion, the voting behavior of legislators in Brazil is a consequence of the combination of electoral rules, the decision-making process within the assembly, and presidential powers. The effects induced by these institutions make political parties meaningful players in the assembly and in executive-legislative relations. Indeed, evidence has shown that at the collective level political parties have served to organize Brazilian legislative politics and that the unity of the main parties has stabilized in recent years (Ames, 2002; Cheibub, Figueiredo and Limongi, 2009; Desposato, 2006; Morgenstern, 2004). Moreover, there is also evidence of clear government-opposition cleavage in legislative voting on a partisan basis recently (Lyne, 2005; Zucco, 2009). Not only parties in the government but also those in the opposition show consistent records of support for their camps' legislative program.

At the individual level, I argue that legislators belonging to parties included in the presidential coalition are willing to support government proposals further away from their ideal points, compared to the status quo, since they receive patronage and remain just as well off as they did at the status quo. In contrast, those in the opposition have less incentives to vote with the government since they receive nothing even if they do so. The point is that the voting behavior of legislators is not only explained by preferences but also influenced by factors induced by the coalition dynamic. This phenomenon corresponds to concerns for the inconsistency between behavior and preferences generally (Krehbiel, 1993, 2000) and in Brazil in particular (Ames, 2002; Saiegh, 2009; Zucco, 2009). Therefore, this suggests the need for an item response and ideal point model that can sort out the non-ideological effect.

4.2 The Ideal Point Model

In this section, I derive an ideal-point model with random item parameters from a simple spatial model of voting. The government-opposition effect on legislative voting behavior is captured by random item-difficulty parameters under the framework of multilevel modeling. The proposed model can be used to recover estimates of the policy preferences of legislators and parties on a left-right, ideological dimension and to evaluate non-ideological effects on voting behavior.

4.2.1 The Spatial Model of Legislative Voting

For each proposal $k = 1, \cdots, K$, legislator $i = 1, \cdots, N$ from party $j[i] = 1, \cdots, J$ makes a choice between a "Yea" decision and a "Nay" decision, where $j[i]$ denotes index variables for party affiliation of legislator i. My purpose is to model the decisions made in a unidimensional Euclidean proposal space. I assume that each legislator's decision depends on the value the legislator attaches to the policy positions of the status quo and the alternative, and to the patronage (cost) of voting with (against) the government. In other words, it is assumed that a legislator is rational in the sense that the legislator will vote for the proposal if the utility the legislator attaches to the alternative is greater than the utility the legislator attaches to the status quo, regardless of the expected actions of the other legislators.

To operationalize this model, I start with random utility functions. Let $U_{i,k}^{(Y)}$ be the utility for legislator i of voting Yea on proposal k, and $U_{i,k}^{(N)}$ be the utility for legislator i of voting Nay on proposal k. Also, let $K_g \subseteq \{1, 2, \cdots, K\}$ denote the set of policy proposals issued by the government. Note that the complement of K_g is the set of policy proposals issued by the opposition, denoted by K_o. Similarly, let $J_g \subset \{1, 2, \cdots, J\}$ denote the set of parties in the government coalition, and J_o the set of parties in the opposition. The utility of legislator $i = 1, \cdots, N$ on proposal $k = 1, \cdots, K$ is assumed to be given by

$$\begin{cases} U_{i,k}^{(Y)} & = -\|\theta_i - x_k^{(Y)}\|^2 + \delta_{j[i],k} \boldsymbol{I}(j \in J_g) \boldsymbol{I}(k \in K_g) \\ & \qquad - \lambda_{j[i],k} \boldsymbol{I}(j \in J_g) \boldsymbol{I}(k \in K_o) + \eta_{i,k} \\ U_{i,k}^{(N)} & = -\|\theta_i - x_k^{(N)}\|^2 + \zeta_{i,k}, \end{cases} \tag{4.1}$$

where $\|\cdot\|$ represents the Euclidean norm, $\theta_i \in \mathbb{R}$ is legislator i's ideal point, $x_k^{(Y)} \in \mathbb{R}$ is the location of the proposal k under a Yea vote, $x_k^{(N)} \in \mathbb{R}$ is the location of the proposal k under a Nay vote, $\delta_{j[i],k} > 0$ is the value of the patronage, $\lambda_{j[i],k} > 0$ is the cost, $\eta_{i,k}$ and $\zeta_{i,k}$ represent the stochastic elements of utility, and $\boldsymbol{I}(\cdot)$ denotes the indicator function. I assume that $\eta_{i,k}$ and $\zeta_{i,k}$ are normally distributed with zero means and fixed variances.

I assume that legislator i will vote Yea on proposal k when $y^*_{i,k} = U^{(Y)}_{i,k} - U^{(N)}_{i,k} > 0$. We can write and simplify this utility difference as follows (see Clinton, Jackman and Rivers, 2004):

$$
\begin{aligned}
y^*_{i,k} &= U^{(Y)}_{i,k} - U^{(N)}_{i,k} \\
&= -\Big[(x^{(Y)}_k)^2 - (x^{(N)}_k)^2 - \delta_{j[i],k}\boldsymbol{I}(j \in J_g)\boldsymbol{I}(k \in K_g) + \lambda_{j[i],k}\boldsymbol{I}(j \in J_g)\boldsymbol{I}(k \in K_o)\Big] \\
&\qquad + 2(x^{(Y)}_k - x^{(N)}_k)\theta_i + (\eta_{i,k} - \zeta_{i,k}) \\
&= \beta_k(-\alpha_{j[i],k} + \theta_i) + \varepsilon_{i,k},
\end{aligned} \tag{4.2}
$$

where $\alpha_{j[i],k} = \frac{1}{\beta_k}\Big[(x^{(Y)}_k)^2 - (x^{(N)}_k)^2 - \delta_{j[i],k}\boldsymbol{I}(j \in J_g)\boldsymbol{I}(k \in K_g) + \lambda_{j[i],k}\boldsymbol{I}(j \in J_g)\boldsymbol{I}(k \in K_o)\Big]$, $\beta_k = 2(x^{(Y)}_k - x^{(N)}_k)$, and $\varepsilon_{i,k} = (\eta_{i,k} - \zeta_{i,k})$. We can see that legislator i belonging to a governing party ($j \in J_g$) is rewarded (with an additional term $\delta_{j[i],k}$) when supporting for a government proposal ($k \in K_g$) and is punished (by subtracting an additional term $\lambda_{j[i],k}$) when supporting for the opposition ($k \in K_o$). Therefore, this implies that, when $\beta_k > 0$,

$$
\begin{cases}
\alpha_{j[i],k} < \alpha_{j'[i],k} \quad \text{for} \quad j \in J_g,\ j' \in J_o, k \in K_g \\
\alpha_{j[i],k} > \alpha_{j'[i],k} \quad \text{for} \quad j \in J_g,\ j' \in J_o, k \in K_o,
\end{cases} \tag{4.3}
$$

and, when $\beta_k < 0$,

$$
\begin{cases}
\alpha_{j[i],k} > \alpha_{j'[i],k} \quad \text{for} \quad j \in J_g,\ j' \in J_o, k \in K_g \\
\alpha_{j[i],k} < \alpha_{j'[i],k} \quad \text{for} \quad j \in J_g,\ j' \in J_o, k \in K_o.
\end{cases} \tag{4.4}
$$

Basically, it means that the effect of coalition dynamics influences a legislator's voting decisions through changing the values of the party-specific parameters $\alpha_{j,k}$.

4.2.2 The Random Difficulty IRT Model

Let $y_{i,k} = 1$ if legislator i from party j votes Yea on the proposal k and $y_{i,k} = 0$ otherwise. The spatial model of voting can be translated into a statistical model by noting the relationship between the theoretical utility difference and observed votes. I assume that

$$\begin{cases} y_{i,k} = 1 \quad \text{if} \quad y^*_{i,k} > 0 \\ y_{i,k} = 0 \quad \text{if} \quad y^*_{i,k} \leq 0, \end{cases} \tag{4.5}$$

where

$$y^*_{i,k} = \beta_k(\theta_i - \alpha_{j[i],k}) + \varepsilon_{i,k}, \quad \varepsilon_{i,k} \overset{\text{iid}}{\sim} N(0,1). \tag{4.6}$$

This results in a standard two-parameter item response model; the only difference is that item difficulty parameters $\alpha_{j[i],k}$ vary across parties. Therefore, we have a response function given by

$$\Pr(y_{i,k} = 1) = \Phi(\beta_k(\theta_i - \alpha_{j[i],k})), \tag{4.7}$$

where $\Phi(\cdot)$ denotes the standard normal distribution function, and the slope parameters β_k and intercepts $\alpha_{j,k}$ are equivalent to item discrimination parameters and item difficulty parameters, respectively, in the item-response modeling literature (Albert and Chib, 1993; Embretson and Reise, 2000; Rasch, 1960).

According to the spatial model of voting discussed above, the values of item difficulty depend on the government or opposition status of parties. This suggests that item difficulty parameters can be modeled through a framework of multilevel modeling (Gelman and Hill, 2007; Gill, 2008*a*; Raudenbush and Bryk, 2002). Formally, the party-specific item difficulty parameters $\alpha_{j,k}$ can be modeled as a function of coalition dynamics as follows:

$$\alpha_{j,k} = \gamma_{0,k} + \gamma_{1,k} z_j + \nu_{j,k}, \quad \nu_{j,k} \overset{\text{iid}}{\sim} N(0, \sigma^2_{\nu_k}), \tag{4.8}$$

where z_j contains the information on the government or opposition status of party j, $\gamma_{0,k}$ and $\gamma_{1,k}$ are coefficient parameters, and $\nu_{j,k}$ is the error term. Suppose that $z_j = 1$ if $j \in J_g$ and $z_j = 0$ otherwise. From Equation (4.3) and Equation (4.4), given that $\beta_k > 0$, we would expect that

$$\begin{cases} \gamma_{1,k} < 0 \quad \text{for} \quad k \in K_g \\ \gamma_{1,k} > 0 \quad \text{for} \quad k \in K_o, \end{cases} \tag{4.9}$$

and, given that $\beta_k < 0$, we would expect that

$$\begin{cases} \gamma_{1,k} > 0 \quad \text{for} \quad k \in K_g \\ \gamma_{1,k} < 0 \quad \text{for} \quad k \in K_o. \end{cases} \tag{4.10}$$

Substantively, this means that, for any two legislators with exactly the same ideal point on an ideological scale but from different camps in terms of government-opposition conflict, the legislator belonging to governing parties is more likely to vote with the government and the other legislator who is in the opposition is more likely to support the proposals of the opposition.

In typical item-response and ideal point models, ideal points are assumed to be sampled from a single distribution. However, legislators from different parties may have distinct, or even opposing, preferences. To account for the between-party variation, a multilevel model is also applied to the ideal points clustered by parties (Bafumi et al., 2005; Fox, 2010; Fox and Glas, 2001). I assume that each legislator i is drawn from party-specific distributions, centered on the party means $\mu_{\theta_{j[i]}}$ for $j = 1, \cdots, J$, which is given by

$$\theta_i \sim N(\mu_{\theta_{j[i]}}, \sigma_\theta^2), \tag{4.11}$$

where σ_θ^2 is the variance. Although the specification of a prior distribution for ideal points follows that of Zucco and Lauderdale (2011), the difference is that I assume hyper priors for these party means rather than fix them at certain values estimated from survey data as Zucco and Lauderdale do. I explain the specification of priors in the next section.

Let $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_K)'$, $\boldsymbol{A}$ be a $J \times K$ matrix with kth column $\boldsymbol{\alpha}_k = (\alpha_{1,k}, \cdots, \alpha_{J,k})'$, $\boldsymbol{\gamma}$ be a $K \times 2$ matrix with kth column $\boldsymbol{\gamma}_k = (\gamma_{0,k}, \gamma_{1,k})'$, $\boldsymbol{\sigma}_\nu^2 = (\sigma_{\nu_1}^2, \cdots, \sigma_{\nu_K}^2)'$, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_N)'$, $\boldsymbol{\mu}_\theta = (\mu_{\theta_1}, \cdots, \mu_{\theta_J})'$, $\boldsymbol{Y}$ be the $N \times K$ matrix of observed votes with (i, k)th element $y_{i,k}$, and $\boldsymbol{Z}$ be a $J \times 2$ matrix with $j$the row $\boldsymbol{z}_j = (1, z_j)$. The random item-difficulty ideal-point model is estimated via a simulation-based Bayesian approach processed by MCMC methods (see Casella and George, 1992; Chib and Greenberg, 1995; Gelfand and Smith, 1990) and, thus, the inference is based on the joint posterior distribution given by

$$\pi(\boldsymbol{\theta}, \boldsymbol{A}, \boldsymbol{\beta}, \boldsymbol{\mu}_\theta, \sigma_\theta^2, \boldsymbol{\gamma}, \boldsymbol{\sigma}_\nu^2 | \boldsymbol{Y}, \boldsymbol{Z}) \propto p(\boldsymbol{Y} | \boldsymbol{\theta}, \boldsymbol{A}, \boldsymbol{\beta}, \boldsymbol{\mu}_\theta, \sigma_\theta^2, \boldsymbol{\gamma}, \boldsymbol{\sigma}_\nu^2; \boldsymbol{Z}) \pi(\boldsymbol{\beta}, \boldsymbol{\mu}_\theta, \sigma_\theta^2, \boldsymbol{\gamma}, \boldsymbol{\sigma}_\nu^2), \tag{4.12}$$

where $\pi(\boldsymbol{\beta}, \boldsymbol{\mu}_\theta, \sigma_\theta^2, \boldsymbol{\gamma}, \boldsymbol{\sigma}_\nu^2)$ is the prior distribution of corresponding parameters and

$$p(\boldsymbol{Y}|\boldsymbol{\theta}, \boldsymbol{A}, \boldsymbol{\beta}, \boldsymbol{\mu}_\theta, \sigma^2_\theta, \boldsymbol{\gamma}, \boldsymbol{\sigma}^2_\nu; \boldsymbol{Z}) = \prod_{i=1}^{N}\prod_{k=1}^{K}\prod_{j=1}^{J} \Phi(\beta_k(\theta_i - \alpha_{j[i],k}))^{y_{i,k}} \times [1 - \Phi(\beta_k(\theta_i - \alpha_{j[i],k}))]^{1-y_{i,k}}. \quad (4.13)$$

is the likelihood function, given the assumptions of independence across legislators and roll calls.

4.2.3 Prior Distribution and Identification

It is well known that item response models suffer from two identification problems: *scale invariance* and *rotational invariance* (see, e.g., Albert, 1992; Johnson and Albert, 1999). The problem of scale invariance occurs because the metric (location and scale) of the latent traits is only known up to a linear transformation. Therefore, one must anchor the metric of the latent traits. Additionally, the problem of rotational invariance refers to the fact that, for the unidimensional case, multiplying all of the model parameters by -1 would not change the value of the likelihood function. Substantively, the model cannot determine which direction is left or right in terms of ideology.

In the Bayesian context, the use of informative prior distributions resolves these two identification problems (Johnson and Albert, 1999; Martin and Quinn, 2002). I begin by identifying relative party positions on the left-right, ideological dimension. For the specification of prior distributions for party means $\boldsymbol{\mu}_\theta$, I utilize the knowledge of the perceived ideological ordering of parties. In specific, I select (at least) two parties from the same camp in terms of government-opposition status (from either the government or the opposition, or from both). Moreover, these parties must be considered as confronting to each other in terms of ideological positions, say, the left-right cleavage.[2] The remaining parties that are not selected are treated as "centre" ones. Let $L \subset \{1, 2, \cdots, J\}$, $C \subset \{1, 2, \cdots, J\}$, and $R \subset \{1, 2, \cdots, J\}$ denote the set of leftist, centre, and rightist parties, respectively. Note that L, C, and R partition the set of parties and are mutually exclusive. Following the conventional approach (Carlin and Louis, 2000; Gelman et al., 2004; Robert, 2001), I assume conjugate prior distributions for unknown parameters. That is, for parameters μ_{θ_j} for $j = 1 \cdots, J$ and σ_θ^2 from Equation (4.11), I assume that

$$\mu_{\theta_j} \sim N(\mu_{\theta_L}, h_0\sigma_\theta^2)\boldsymbol{I}(\mu_{\theta_j} < 0), \quad \text{for} \quad j \in L, \tag{4.14}$$

$$\mu_{\theta_j} \sim N(\mu_{\theta_C}, h_0\sigma_\theta^2), \quad \text{for} \quad j \in C, \tag{4.15}$$

$$\mu_{\theta_j} \sim N(\mu_{\theta_R}, h_0\sigma_\theta^2)\boldsymbol{I}(\mu_{\theta_j} > 0), \quad \text{for} \quad j \in R, \tag{4.16}$$

$$\sigma_\theta^2 \sim IG(c_0/2, d_0/2), \tag{4.17}$$

where $\mu_{\theta_L}, \mu_{\theta_C}, \mu_{\theta_R}, c_0, d_0$, and h_0 are hyperprior parameters.

This specification strategy achieves two goals. First, the problem of rotational invariance is resolved by restricting the mean of the pre-identified leftist party to be negative and that of the pre-identified rightist party to be positive.[3] As a result, leftist

[2]Basically, I assume that the government coalition is not purely ideological. This assumption excludes a complete collinearity between government-opposition cleavage and ideological conflict because they would be difficult to identify without further information (Zucco and Lauderdale, 2011).

[3]In the psychometrics literature, the rotation invariance problem is usually solved by restricting item discrimination parameters to be positive (e.g., $\beta_k \sim N(\mu_\beta, \sigma_\beta^2)\boldsymbol{I}(\beta_k > 0)$). This is because respondents are assumed to answer test items correctly if they have higher ability. However, for

parties are on the left and rightist parties are on the right of the underlying scale. This form of constraint is sometimes known as 'anchoring' (Skrondal and Rabe-Hesketh, 2004, p. 66). Second, since the two selected parties are from the same camp in terms of government-opposition status, it ensures that the assumed underlying ideological dimension and the government-opposition conflict do not completely overlap.

To deal with the problem of scale invariance, the metric of ideal points is rescaled via a linear transformation in each MCMC iteration (Bafumi et al., 2005; Fox, 2010; Jackman, 2009). By imposing identification of normalization $\tilde{\theta}_i^{(t)} = (\theta_i^{(t)} - \bar{\theta}^{(t)})/s_\theta^{(t)}$, where $\bar{\theta}^{(t)} = \frac{1}{N}\sum_{i=1}^N \theta_i^{(t)}$, $s_\theta^{(t)} = \sqrt{\frac{1}{N}\sum_{i=1}^N (\theta_i^{(t)} - \bar{\theta}^{(t)})^2}$, and t indexes MCMC iterations, we have ideal points $\tilde{\theta}_i^{(t)}$ with an identified metric. All of the other parameters are transformed correspondingly.

For item parameters $\boldsymbol{\beta}$ and $\boldsymbol{A}$ from Equation (4.7) and Equation (4.8), I assume that

$$\beta_k \sim N(\mu_\beta, \sigma_\beta^2), \quad \text{for} \quad k = 1, \cdots, K, \tag{4.18}$$

$$\sigma_{\nu_k}^2 \sim IG(a_0/2, b_0/2), \quad \text{for} \quad k = 1, \cdots, K, \tag{4.19}$$

$$\boldsymbol{\gamma}_k \sim N_2(\boldsymbol{\gamma}_0, \sigma_{\nu_k}^2 \boldsymbol{G}_0), \quad \text{for} \quad k = 1, \cdots, K, \tag{4.20}$$

where $\mu_\beta, \sigma_\beta^2, a_0, b_0, \boldsymbol{\gamma}_0$, and $\boldsymbol{G}_0$ are hyperparameters. Furthermore, let $\boldsymbol{\Omega} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\sigma}_\nu^2)$ and $\boldsymbol{\Psi} = (\boldsymbol{\mu_\theta}, \sigma_\theta^2, \sigma_{\mu_\theta}^2)$. The prior distribution is given by

roll-call voting, there is no "correct answer" to each roll call. In fact, rightist legislators are more likely to vote for rightist proposals and against leftist proposals. Thus, the identification strategy of assuming the item discrimination parameters to be positive is not used.

$$
\begin{aligned}
\pi(\mathbf{\Omega}, \mathbf{\Psi}) &= \pi(\mathbf{\Omega})\pi(\mathbf{\Psi}), \\
&= \pi(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\sigma}_{\nu}^2)\pi(\boldsymbol{\mu_{\theta}}, \sigma_{\theta}^2), \\
&= p(\boldsymbol{\beta})p(\boldsymbol{\gamma}|\boldsymbol{\sigma}_{\nu}^2)p(\boldsymbol{\sigma}_{\nu}^2)p(\boldsymbol{\mu_{\theta}}|\sigma_{\theta}^2)p(\sigma_{\theta}^2),
\end{aligned}
\tag{4.21}
$$

where I assume priori independence of $\mathbf{\Omega} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\sigma}_{\nu}^2)$ and $\mathbf{\Psi} = (\boldsymbol{\mu_{\theta}}, \sigma_{\theta}^2, \sigma_{\mu_{\theta}}^2)$.

4.2.4 Illustration of Random Item Difficulties

Roll-call data are widely used to estimate ideal points of legislators, which in turn are used to test theories of legislative politics in the U.S. (e.g., Clinton, 2007; McCarty, Poole and Rosenthal, 2006; Poole and Rosenthal, 1985, 1997) and Latin American countries (e.g., Morgenstern and Nacif, 2002; Morgenstern, 2004). Although several approaches have been applied to the estimation of ideal points of legislators, one of the primary concerns about ideal point estimates based on roll-call data is that these voting records do not differentiate voting behavior induced by preferences and non-ideological effects such as party affiliation (Krehbiel, 1993, 2000). Brazil's assembly is a typical case that voting behavior of legislators is a consequence of political negotiation. As discussed in the previous section, legislative voting in the Brazilian Congress reflects the result of the president's distribution of resources combined with underlying preferences (Alston and Mueller, 2006; Pereira and Mueller, 2004*a*; Zucco, 2009). Especially, political exchanges between the government and Congress is an essential feature of Brazilian politics. Therefore, it is important to consider the impact of non-ideological factors when ideal points are estimated based on the similarity of the legislators' voting records.

I have shown that the effect of government-opposition conflict on voting behavior can be captured by the proposed random item-difficulty ideal-point model. In what follows, I show how this model provides more accurate ideal point estimates than a standard IRT model does. Figure 4.1 presents an item characteristic curve (ICC) for a government proposal in two scenarios: The one on the left has party-specific item difficulties and the one on the right has a common item difficulty across parties. Suppose that Party 1 is included in the government and Party 2 is in the opposition, ignoring the remaining parties. The left panel of Figure 4.1 shows that Party 1 has lower value of item difficulty (the black dotted curve) than Party 2 does (the red dotted curve). In other words, for a legislator whose ideal point is at $d1$, this legislator has a 70% probability of voting Yea when being in a governing party but has a 30% probability of voting Yea when being in a party in the opposition. The difference between these two probabilities reflect the influence of government-opposition conflict on the voting decision.

However, if we assume a common item difficulty parameter for Party 1 and Party 2, then we will obtain incorrect estimates of ideal points and, in turn, incorrect estimated party positions. As the right panel of Figure 4.1 shows, the purple curve represents the ICC for a proposal that has the same item difficulty parameters across parties. Given this assumption, we end up with $b1$ as the legislator's ideal point who belongs to Party 1 and $r1$ as the legislator's who belongs to Party 2. We can see that

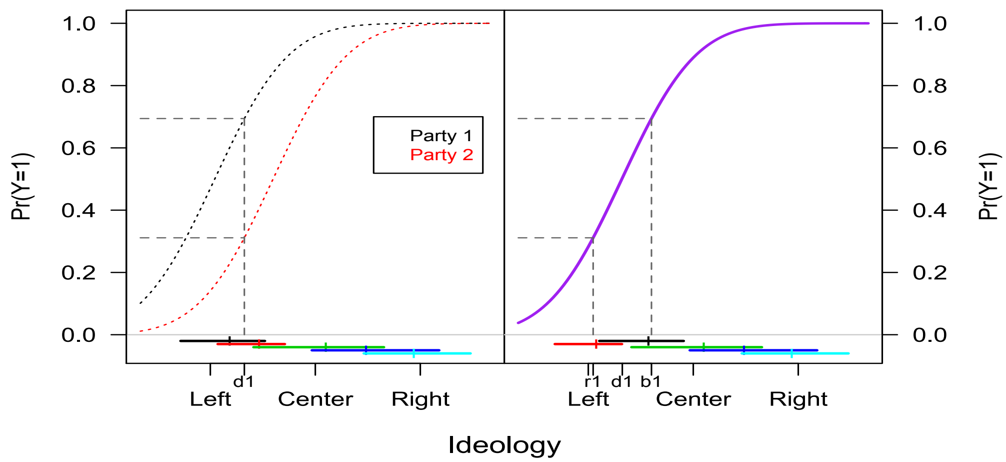


Figure 4.1.: Random Item-Difficulty Ideal-Point Model.

$b1$ is at the right of the true ideal point $d1$ and $r1$ is at the left of $d1$. As a result, at the aggregate level, we have estimates of party positions showing that Party 2 is more leftist-oriented than Party 1, which is not accurate.

The Bayesian random item-difficulty ideal-point model proposed here has several advantages. First, this model allows us to detect whether legislators' voting behavior is influenced by non-ideological effects such as government-opposition conflict while estimating ideal points of legislators and parties. Second, related to the first point, since institution-structured behavior is differentiated from preference-induced behavior, we are more confident to claim that the estimated ideal points reflect the preferences of legislators. Third, the random item-difficulty ideal-point model deals with the problem of differential item functioning (DIF). Substantially, the relationship between roll calls and ideal points varies across parties. In a more general sense, this means that the response probability depends on group membership given the level of the latent traits and can be formally represented by

$$\Pr(y_{i,k} = 1|\theta_i) \neq \Pr(y_{i,k} = 1|\theta_i, J), \tag{4.22}$$

where J characterizes the group respondent i belongs to.

4.3 Monte Carlo Experiments

In what follows, I conduct a series of Monte Carlo analyses to assess the proposed random item-difficulty ideal-point model, compared to the standard IRT model. In particular, I am interested in the estimates of party positions and the effect of government-opposition status on item difficulty parameters and on voting behavior of legislators.

4.3.1 Simulation Design

The simulated data are generated in the following steps. First, the ideal points of legislators are drawn from party-specific distributions given by

$$\theta_i \sim N(\mu_{\theta_{j[i]}}, \sigma^2_{\theta_{j[i]}}), \quad \text{for} \quad i = 1, \cdots, N_j; j = 1, \cdots, J, \tag{4.23}$$

in which

$$|\{1, \cdots, J\}| = 5, J_g = \{1, 3, 4\}, J_o = \{2, 5\},$$

$$(N_1, N_2, N_3, N_4, N_5) = (14, 23, 28, 22, 13),$$

$$(\mu_{\theta_1}, \mu_{\theta_2}, \mu_{\theta_3}, \mu_{\theta_4}, \mu_{\theta_5}) = (-1.6, -1, 0.2, 0.8, 1.5),$$

$$(\sigma_{\theta_1}, \sigma_{\theta_2}, \sigma_{\theta_3}, \sigma_{\theta_4}, \sigma_{\theta_5}) = (0.6, 0.2, 0.5, 0.5, 1.5),$$

where $|\{1, \cdots, J\}|$ is the cardinality of the set of parties denoting the number of parties. The true values of parameters are chosen arbitrarily and should not influence the results.

Second, for item discrimination and item difficulty parameters, I hold that

$$K = 40, K_g = \{1, 2, \cdots, 30\}, K_o = \{31, 32, \cdots, 40\}.$$

Moreover, due to the purpose of this chapter, item discrimination parameters β_k are fixed at the value of one for $k = 1 \cdots, K$. Regarding item difficulty parameters, I first draw item-difficulty parameters α_k from a normal distribution given by

$$\alpha_k \sim N(0, 1), \quad \text{for} \quad k = 1, \cdots, K, \tag{4.24}$$

and then I create party-specific item difficulties as follows:

$$\begin{cases} \alpha_{j,k} = \alpha_k - 0.5 \quad \text{for} \quad k \in K_g; j \in J_g \\ \alpha_{j,k} = \alpha_k + 0.5 \quad \text{for} \quad k \in K_o; j \in J_g, \end{cases} \tag{4.25}$$

and

$$\begin{cases} \alpha_{j,k} = \alpha_k + 0.5 \quad \text{for} \quad k \in K_g; j \in J_o \\ \alpha_{j,k} = \alpha_k - 0.5 \quad \text{for} \quad k \in K_o; j \in J_o. \end{cases} \tag{4.26}$$

Finally, for $i = 1, \cdots, N$, $k = 1, \cdots, K$, and $j = 1 \cdots, J$, I generate roll-call data by the following response function

$$\begin{cases} y_{i,k} = 1 \quad \text{if} \quad \Phi(\beta_k \theta_i - \alpha_{j[i],k}) \geq 0.5 \\ y_{i,k} = 0 \quad \text{if} \quad \Phi(\beta_k \theta_i - \alpha_{j[i],k}) < 0.5. \end{cases} \tag{4.27}$$

A simulation-baed Bayesian approach to the random item-difficulty ideal-point model is estimated with MCMC techniques and implemented in `JAGS` 3.1.0 (Plummer, 2003, 2012) called from `R` version 2.15.2 (`R2jags` Su and Yajima, 2012). The estimation is performed with two parallel chains of 10,000 iterations each. The first half of the iterations are discarded as a burn-in and 2 as thinning, and thus 5,000 samples are generated. This simulation is repeated for 10 times.

4.3.2 Simulation Results

In the estimation process, I select Party 1 as the pre-identified leftist party and Party 4 as the rightist party, both of which are included in the government coalition. Formally, $L = \{1\}, C = \{2, 3, 5\}$, and $R = \{4\}$. Hyperparameters are assigned values to represent vague priors.

The results of simulations are presented in two parts. First, Table 4.1 shows the simulated and estimated party positions. The upper panel of Table 4.1 provides information from the simulated data and the lower panel represents estimates from the proposed ideal-point model. Regarding the simulation data, the values in the parentheses are the true values of $\boldsymbol{\mu}_\theta$. From the second column to the fifth column, the values represent estimated party means, standard errors, and the 2.5th and 97.5th percentiles based on the simulated ideal points of legislators across the number of simulations. For example, the values in the second column are calculated by taking the mean of clustered legislators' ideal points by parties by the following procedure: $\bar{\theta}_j = \frac{1}{10}\sum_{m=1}^{10}(\frac{1}{N_j}\sum_{i=1}^{N_j}\theta_i^{(m)})$ for $j = 1, \cdots, 5$.

Table 4.1: Simulated and Estimated Party Means

	Mean	s.d.	2.5%	97.5%
Simulation				
`Party 1` (−1.6)	−1.600	0.161	−1.781	−1.291
`Party 2` (−1)	−1.025	0.062	−1.090	−0.914
`Party 3` (0.2)	0.179	0.094	0.059	0.340
`Party 4` (0.8)	0.805	0.082	0.692	0.903
`Party 5` (1.5)	1.514	0.095	1.377	1.669
Estimation				
`Party 1` (−1.17)	−1.226	0.107	−1.382	−1.038
`Party 2` (−1.02)	−1.079	0.062	−1.175	−0.985
`Party 3` (0.26)	0.340	0.069	0.256	0.454
`Party 4` (0.93)	0.943	0.064	0.855	1.036
`Party 5` (0.94)	0.900	0.066	0.823	1.000

The lower panel of Table 4.1 shows the party means calculated based on both the aggregate-level and individual-level estimates from the random item-difficulty ideal-point model. We can see that the estimated party means displayed in the parentheses are close to the true values, except for the estimated position of Party 5. Furthermore, the left-right ordering of parties is the same as the setup of the experiments. However, the distance between Party 4 and 5 is indistinguishable. This can also be seen from the estimates of party positions calculated from individual ideal point estimates. Fourth of the five parties have estimated positions quite similar to the simulated data, except Party 5. In other words, Party 1 is in the far left of the scale; Party 2 is recognized as a leftist party and at the right of Party 1; Party 3 is at the center among these five parties. But the distinction between Party 4 and Party 5 is not clear.[4]

Table 4.2: Hypothesized and Estimated Effects of Government Status

Coef.	Hypothesized impact	Estimated impact
$\gamma_{1,k}$ for $k = 1, \cdots, 30$	–	–
$\gamma_{1,k}$ for $k = 31, \cdots, 40$	+	+

Second, Table 4.2 compares the hypothesized and estimated effects of government-opposition status. According to Equation (4.9), we expect a negative sign of coefficient $\gamma_{1,k}$ for $k = 1, \cdots, 30$ and a positive sign of coefficient $\gamma_{1,k}$ for $k = 31, \cdots, 40$. As can be seen from Table 4.2, all of the coefficient estimates are in the same direction as we expect based on the spatial model of voting.

4.4 Roll-Call Data Analysis

Returning to legislative behavior in Brazil. In the study of Brazilian legislative politics, legislators' voting records have been widely used to evaluate party discipline, party strength, and, most broadly, party government (Ames, 2002; Amorim Neto, Cox and McCubbins, 2003; Figueiredo and Limongi, 2000, just to name a few). Re-

[4]When a standard two-parameter ideal point model is employed, the estimated party means (not shown) provide incorrect ordering of parties, in which Party 2 is in the far left of the scale, in addition to putting Party 4 at the right of Party 5.

cently, Zucco (2009) has shown that the voting behavior of Brazilian deputies does not solely depend on their ideology. Instead, their voting behavior is strongly influenced by the government-opposition conflict. Thus, ideal points estimated by standard unidimensional IRT models mostly reflect a government-opposition dimension of conflict rather than a left-right, ideological dimension. To distinguish between ideological motivations and political inducements from the executive, Zucco and Lauderdale (2011) employ a two-dimensional item response model and utilize survey of legislators to identify party positions on a left-right, ideological dimension and explore the content of the second dimension. They show that the second dimension corresponds to the government-opposition conflict and has become the dominant dimension of conflict in recent years.

The analysis of roll calls in this chapter follows Zucco and Lauderdale (2011) in the sense that I need to pre-identify party positions on the left-right, ideological dimension. However, the procedure differs in three fundamental respects. First, I assume a unidimensional policy space underlying Brazilian legislative politics, which follows many studies of Brazilian politics (e.g., Figueiredo and Limongi, 2000; Pereira and Mueller, 2004*b*; Rosas, 2005). Second, I rely on information on the ideological ordering of parties to identify party positions. In specific, I constrain (at least) one perceived leftist party on the left and one perceived rightist party on the right of the ideological scale by assuming proper prior distributions. In such case, party positions are estimated in the same model. Finally, the effect of government-opposition conflict on legislative voting is evaluated by a regression model under the framework of multilevel modeling.

4.4.1 Data and Model Specification

I apply the Bayesian random item-difficulty ideal point model developed in Section 3 to analyzing legislative behavior in the Brazilian Congress. I measure voting behavior using roll calls taken on the floor of the *Câmara de Deputados* between 2003 and 2006.[5] This period is selected because the cabinet in this period is considered as one with non-overlapping between the ideological cleavage and government-opposition divide, and the results are comparable to those in previous studies (Zucco, 2009; Zucco and Lauderdale, 2011).

In the analysis, I exclude roll calls in which one side obtained less than 2.5% of the votes after excluding legislators who participated in less than 95% roll calls. This results in 294 roll calls and 704 legislators.[6] To evaluate the effect of the government-opposition dynamic, I include a dummy variable for government and opposition status of parties as the covariate z_j in Equation (4.8), that is, $z_j = 1$ if party j is in the government and $z_j = 0$ otherwise.[7]

To identify the left-right, ideological scale, I select Partido Popular Socialista (PPS, Socialist People's Party) and Partido Democratico Trabalhista (PDT, Demo-

[5]The data set I use was collected by Zucco (2013) and can be found on his website.

[6]The missingness is not considered as a Nay vote.

[7]During the period between 2003 and 2006, some parties changed cabinet status. To simplify the analysis, I treat parties as governing parties if there were more than half of roll calls were proposed when they were in the government.

cratic Labor Party) as pre-identified leftist parties, and Partido Liberal (PL, the Liberal Party) and Pardito da Frente Liberal (PFL, the Liberal Front Party) as rightist parties. Among these pre-identified parties, PPS and PL are from the government coalition and PDT and PFL from the opposition. The estimation was performed with two parallel chains of 10,000 iterations each. The first half of the iterations were discarded as a burn-in and 2 as thinning, and thus 5,000 samples were generated.

4.4.2 Empirical Results

The first set of results are shown in Figure 4.2, which displays the estimates of party positions $\boldsymbol{\mu}_\theta$. As cab be seen, these estimated party positions correspond approximately to what would be the perceived ideological positions of Brazilian parties during 2003 and 2006, except for the PP, PMDB, and PSDB. The estimated position of the PP locates at the left of the scale but it is considered as a centre-right or right-wing party. The reason why the position of the PP deviates from its perceived position is that the party has supported the government before it joined the cabinet (see Zucco and Lauderdale, 2011). Regarding the PMDB and PSBD, their estimated positions are at the right of their perceived positions at the center. The PMDB suffers from conflict between its factions and that might be the reason why it is misplaced (Zucco and Lauderdale, 2011). Although these three parties are misplaced, the proposed model in general perform better than a standard IRT model in capturing the ideological dimension.[8]

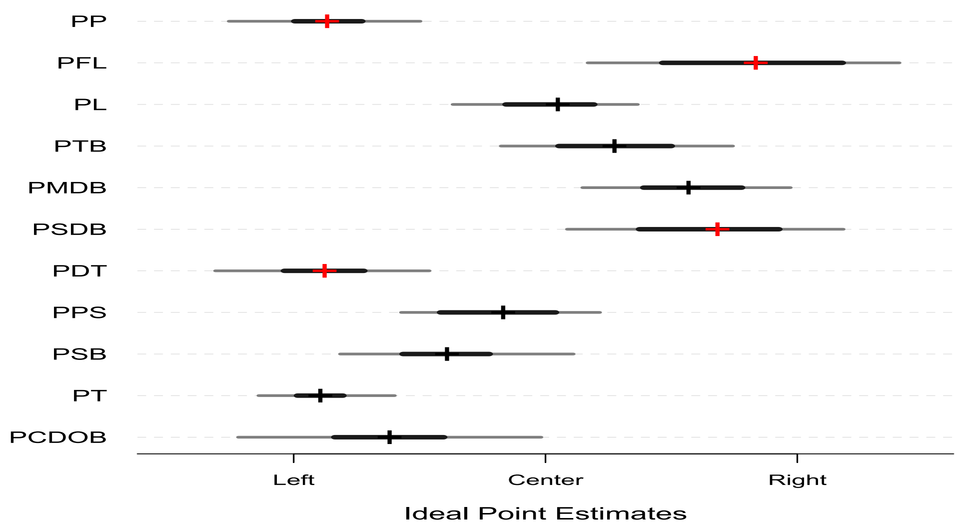


Figure 4.2.: Estimates of Brazilian Party Positions.

Table 4.3 presents the results concerning the investigation of government-opposition conflict in legislative behavior. First, consider the third row of Table 4.3 where $\beta_k > 0$. There are 142 roll calls for which legislators in the government are more likely to vote Yea than those in the opposition, and the effect of government-opposition status on voting behavior is significant at the conventional statistical level (p-value < 0.05) in 14 of 142 roll calls. Moreover, on 65 roll calls legislators in the governing parties are less likely to vote Yea and the effect of government-opposition status is significantly different from zero in 31 of these 65 roll calls. Second, the fourth row of Table 4.3

[8]The ideal point estimates of parties by a standard IRT model can be found in Zucco (2009), which shows that parties are distinguished by the government-opposition conflict rather than by ideology.

Table 4.3: Analysis of Roll Calls, 2003–2006

	No. of Roll Calls		
+/− of β_k	$\gamma_{1,k} < 0$	$\gamma_{1,k} > 0$	Total
+	142 (14)	65 (31)	207 (45)
−	69 (5)	18 (7)	87 (12)

Note: The values in parentheses are the numbers of Roll Calls with the credible interval of $\gamma_{1,k}$ not covering zero.

shows that the effect of government-opposition status on voting behavior is significantly less than zero at the conventional statistical level on 5 of 69 roll calls and significantly larger than zero on 7 of 18 roll calls.

In sum, the results of analysis displayed here indicate that the government-opposition conflict indeed influences legislative voting behavior in the Brazilian Congress although it is observed in about only one-fifth (57/294 = 19.4%) of the roll calls in the 2003–2006 period. Most importantly, legislative voting is somewhat party-based in the sense that political parties play an major role in the government-opposition conflict. These results suggest that the legislative behavior during 2003 and 2006 is neither a purely atomistic nor a purely parliamentary mode. In other words, party membership in the cabinet indeed influences voting behavior of legislators but is not definitive.

4.5 Conclusion

Legislative voting records are not always a good measure for legislators' ideology or preferences. This is especially true when legislative voting behavior is a consequence of political negotiation such as the executive-legislative relation in Brazil. This chapter develops a random item-difficulty ideal-point model, in which item parameters differ across parties and the differences can be systematically explained by party membership in the cabinet. In the process, I first derive a spatial model of voting in which voting behavior is induced by both ideological motivations and coalition dynamics. Next I show that the statistical model implied by the spatial voting model is a two-parameter item-response model with random difficulty parameters. This model can be employed to estimate ideal points of legislators and parties and to detect the effect of government-opposition dynamic at the same time.

Applying the Bayesian random item-difficulty ideal-point model proposed here to analyzing roll-call votes in the Brazilian Chamber of Deputies between 2003 and 2006, I find that the estimates of party positions correspond approximately to their perceived ideological positions, except for a few unusual cases. Moreover, the evidence shows that the government-opposition conflict matters for party-based legislative voting but it is not definitive all the time.

This chapter makes a contribution that is both theoretical and methodological. More and more studies of legislative voting in the Brazilian Congress stress the role of the president and political parties. I derive a spatial model of voting to illustrate how legislative behavior is influenced by party membership in the cabinet. This chapter is relevant to the debate over the power of legislative parties (Aldrich, 1995; Aldrich and Rohde, 1997-98; Krehbiel, 1999, 2000). Also the fact that legislative behavior is a consequence of government-opposition conflict suggests the need for an ideal point model that can sort out the non-ideological effect. It is shown that the proposed Bayesian random item-difficulty ideal-point model performs well in differentiating between the ideological preferences and non-ideological effects. Although the statistical model is derived from the Brazilian legislative process, it can be easily adapted to dealing with general cases.

Chapter 5

Conclusion

In this dissertation, I investigate the effects of political institutions on electoral and legislative behavior in the policy making process. I provide a detailed effort in the applicability of Bayesian methods to methodological issues encountered by political scientists in empirical research. Specifically, I first investigate the difference between democratic and nondemocratic regimes in social provision policy and deal with the problem of rarely changing variables in longitudinal data analysis. Second, I examine the interdependence between party policy shifts and election results in multiparty systems and handle the endogenous dynamic components in electoral competition. Finally, I explore ideological and non-ideological effects on legislative behavior in a multiparty presidential system and deal with the violation of the core assumption of item-response and ideal point models. This dissertation enhances our knowledge of the role of political institutions in explaining political and economic phenomena.

The Bayesian models proposed in this dissertation allow us to recognize the critical components of the data and specify them in the model in a theoretically-informed manner. This dissertation improves our ability to successfully evaluate the factors that affect the relationship under investigation and makes several contributions to the literature. First of all, unlike previous studies that exploit indirect institutional effects (e.g., Persson and Tabellini, 2003) or consider only variations between countries (e.g., Ross, 2006), the Bayesian model for TSCS data proposed in Chapter 2 is able to detect the short- and long-term effects of rarely changing variables. This kind of models can be applied to the study of the consequences of policies enacted by politicians such as regulations and the structure of the relevant political institutions, where we would expect the short-term impacts of policies and long-term impacts of institutions.

Second, contrary to a literature that looks at CMP data as the be-all end-all of party positions, the Bayesian structural equation model proposed in Chapter 3 allows us to be more serious about uncertainty, about the estimation of a latent position, and about the reciprocal effects of party positions and electoral performance. By estimating a Bayesian SEM, we know that the general perceived ideological positions informed by party labels are more important information than specific party manifestos in electoral campaigns to voters. This discussion improves our understanding of the relationship between party positions and electoral performance and democratic representation.

Finally, the estimation of legislators' policy preferences are of central interest to scholars of legislative politics. However, the ideal point models rely on the assumption that it is primarily policy preferences that drive legislative voting behavior. By recognizing the violation of this assumption in Brazilian legislative politics, I develop a Bayesian random item-difficulty ideal-point model to differentiate between ideological and government-opposition effects on party-based voting behavior of legislators. This model is useful to distinguish between influences on legislative behavior in most of multiparty systems, where voting decisions are affected by ideological motivations and non-ideological inducements from political parties or the executive.

Bibliography

Acemoglu, Daron and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.

Achen, Christopher H. 2000. Why lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables. The Annual Meeting of the Midwest Political Science Association (Chicago).

Adams, James, Andrea B. Haupt and Heather Stoll. 2009. "What Moves Parties?: The Role of Public Opinion and Global Economic Conditions in Western Europe." *Comparative Political Studies* 42(5):611–639.

Adams, James, Michael Clark, Lawrence Ezrow and Garrett Glasgow. 2004. "Understanding Change and Stability in Party Ideologies: Do Parties Respond to Public Opinion or to Past Election Results?" *British Journal of Political Science* 34(04):589–610.

Adams, James, Michael Clark, Lawrence Ezrow and Garrett Glasgow. 2006. "Are Niche Parties Fundamentally Different from Mainstream Parties? The Causes and the Electoral Consequences of Western European Parties' Policy Shifts, 1976-1998." *American Journal of Political Science* 50(3):513–529.

Adams, James and Samuel Merrill. 1999. "Modeling Party Strategies and Policy Representation in Multiparty Elections: Why are Strategies So Extreme?" *American Journal of Political Science* 43(3):765–791.

Adams, James, Samuel Merrill and Bernard Grofman. 2005. *A Unified Theory of Party Competition: A Cross-National Analysis Integrating Spatial and Behavioral Factors*. New York: Cambridge University Press.

Adams, James and Zeynep Somer-Topcu. 2009*a*. "Moderate Now, Win Votes Later: The Electoral Consequences of Parties' Policy Shifts in 25 Postwar Democracies." *Journal of Politics* 71(2):678–692.

Adams, James and Zeynep Somer-Topcu. 2009*b*. "Policy Adjustment by Parties in Response to Rival Parties' Policy Shifts: Spatial Theory and the Dynamics of Party Competition in Twenty-Five Post-War Democracies." *British Journal of Political Science* 39(4):825–846.

Albert, James H. 1992. "Bayesian Estimation of Normal Ogive Item Response Curves Using Gibbs Sampling." *Journal of Educational and Behavioral Statistics* 17(3):251–269.

Albert, James H. and Siddhartha Chib. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data." *Journal of the American Statistical Association* 88(442):669–679.

Aldrich, John H. 1983. “A Downsian Spatial Model with Party Activism.” *The American Political Science Review* 77(4):974–990.

Aldrich, John H. 1995. *Why Parties?: The Origin and Transformation of Political Parties in America*. Chicago: University Of Chicago Press.

Aldrich, John H. and David W. Rohde. 1997-98. “The Transition to Republican Tule in the House: Implications for Theories of Congressional Politics.” *Political Science Quarterly* 112(4):541–567.

Alston, Lee J. and Bernardo Mueller. 2006. “Pork for Policy: Executive and Legislative Exchange in Brazil.” *Journal of Law, Economics, and Organization* 22(1):87–114.

Alvarez, R. Michael, Jonathan Nagler and Shaun Bowler. 2000. “Issues, Economics, and the Dynamics of Multiparty Elections: The British 1987 General Election.” *American Political Science Review* 94(1):131–149.

Amenta, Edward and Jane D. Poulsen. 1996. “Social Politics in Context: The Institutional Politics Theory and Social Spending at the End of the New Deal.” *Social Forces* 75(1):33–60.

Ames, Barry. 1995*a*. “Electoral Rules, Constituency Pressures, and Pork Barrel: Bases of Voting in the Brazilian Congress.” *Journal of Politics* 57(2):324–343.

Ames, Barry. 1995*b*. “Electoral Strategy under Open-List Proportional Representation.” *American Journal of Political Science* 39(2):406–433.

Ames, Barry. 2001. *The Deadlock of Democracy in Brazil.* Ann Arbor: University of Michigan Press.

Ames, Barry. 2002. Party Discipline in Brazil's Chamber of Deputies. In *Legislative Politics in Latin America*, ed. S. Morgenstern and Benito Nacif. New York: Cambridge University Press chapter 7, pp. 185–221.

Amorim Neto, Octavio. 2002*a*. Presidential Cabinets, Electoral Cycles, and Coalition Discipline in Brazil. In *Legislative Politics in Latin America*, ed. S. Morgenstern and Benito Nacif. New York: Cambridge University Press chapter 3, pp. 48–78.

Amorim Neto, Octavio. 2002*b*. "The Puzzle of Party Discipline in Brazil." *Latin American Politics and Society* 44(1):127–144.

Amorim Neto, Octavio. 2006. "The Presidential Calculus: Executive Policy Making and Cabinet Formation in the Americas." *Comparative Political Studies* 39(4):415–440.

Amorim Neto, Octavio, Garry W. Cox and Mathew D. McCubbins. 2003. "Agenda Power in Brazil's Camara dos Deputados, 1989-98." *World Politics* 55(4):550–578.

Amorim Neto, Octavio and Gary W. Cox. 1997. "Electoral Institutions, Cleavage Structures, and the Number of Parties." *American Journal of Political Science* 41(1):149–74.

Anderson, T.W. and Cheng Hsiao. 1981. "Estimation of Dynamic Models with Error Components." *Journal of the American Statistical Association* 76(375):598–606.

Anderson, T.W. and Cheng Hsiao. 1982. "Formulation and Estimation of Dynamic Models Using Panel Data." *Journal of Econometrics* 18(1):47–82.

Arellano, M. and Olympia Bover. 1995. "Another Look at the Instrumental Variable Estimation of Error-Components Models." *Journal of Econometrics* 68(1):29–51.

Arellano, Manuel and Stephen Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *The Review of Economic Studies* 58(2):277–297.

Arian, Asher and Michal Shamir, eds. 1995. *The Elections in Israel, 1992*. Albany: State University of New York Press.

Arian, Asher and Michal Shamir, eds. 1999. *The Elections in Israel, 1996*. Albany: State University of New York Press.

Arian, Asher and Michal Shamir, eds. 2002. *The Elections in Israel, 1999*. Albany: State University of New York Press.

Austen-Smith, David and Jeffrey S. Banks. 1988. "Elections, Coalitions, and Legislative Outcomes." *The American Political Science Review* 82(2):405–422.

Avelino, George, David S. Brown and Wendy Hunter. 2005. "The Effects of Capital Mobility, Trade Openness, and Democracy on Social Spending in Latin America, 1980-1999." *American Journal of Political Science* 49(3):625–641.

Axelrod, Robert M. 1970. *Conflict of Interest: A Theory of Divergent Goals with Applications to Politics*. Chicago: Markham Pub. Co.

Bafumi, Joseph and Andrew Gelman. 2006. Fitting Multilevel Models When Predictors and Group Effects Correlate. The Annual Meeting of the Midwest Political Science Association (Chicago).

Bafumi, Joseph, Andrew Gelman, David K. Park and Noah Kaplan. 2005. "Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation." *Political Analysis* 13(2):171–187.

Baltagi, Badi H. 2005. *Econometric Analysis of Panel Data*. 3 ed. New York: John Wiley & Sons Inc.

Bates, Douglas, Martin Maechler and Bolker Ben. 2012. *lme4 ver. 0.999999-0 (R package)*.

URL: *http://lme4.r-forge.r-project.org/*

Beck, Nathaniel. 2001. "Time-Series-Cross-Section Data: What Have We Learned in the Past Few Years?" *Annual Review of Political Science* 4(1):271–293.

Beck, Nathaniel and Jonathan N. Katz. 1995. "What To Do (and Not To Do) With Time-series Cross-section Data." *American Political Science Review* 89(3):634–647.

Beck, Nathaniel and Jonathan N. Katz. 1996. "Nuisance vs. Substance: Specifying and Estimating Time-Series-Cross-Section Models." *Political Analysis* 6(1):1–36.

Beck, Nathaniel and Jonathan N. Katz. 2009. "Modeling Dynamics in Time-Series-Cross-Section Political Economy Data." *California Institute of Technology Social Science Working Paper* 1304:1–31.

Benoit, Kenneth, Michael Laver and Slava Mikhaylov. 2009. "Treating Words as Data with Error: Uncertainty in Text Statements of Policy Positions." *American Journal of Political Science* 53(2):495–513.

Blundell, Richard and Stephen Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models." *Journal of Econometrics* 87(1):115–143.

Boix, Carles. 2003. *Democracy and Redistribution*. New York: Cambridge University Press.

Boix, Carles. 2007. The Emergence of Parties and Party Systems. In *The Oxford Handbook of Comparative Politics*, ed. Carles Boix and Susan C. Stokes. Oxford: Oxford University Press chapter 21, pp. 499–521.

Breusch, Trevor, Michael B. Ward, Hoa Thi Minh Nguyen and Tom Kompas. 2011. "On the Fixed-Effects Vector Decomposition." *Political Analysis* 19(2):123.

Brown, David S. and Wendy Hunter. 1999. "Democracy and Social Spending in Latin America, 1980-92." *American Political Science Review* 93(4):779–790.

Budge, Ian. 1994. "A New Spatial Theory of Party Competition: Uncertainty, Ideology and Policy Equilibria Viewed Comparatively and Temporally." *British Journal of Political Science* 24(4):443–467.

Budge, Ian, Lawrence Ezrow and Michael D. McDonald. 2010. "Ideology, Party Factionalism and Policy Change: An Integrated Dynamic Theory." *British Journal of Political Science* 40(4):781–804.

Budge, Ian and Michael D. McDonald. 2006. "Choices Parties Define: Policy Alternatives in Representative Elections, 17 Countries 1945V1998." *Party Politics* 12(4):451.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge: MIT Press.

Bueno de Mesquita, Bruce, James D. Morrow, Randolph M. Siverson and Alastair Smith. 1999. "Policy Failure and Political Survival: The Contribution of Political Institutions." *Journal of Conflict Resolution* 43(2):147–161.

Calvert, Randall. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science* 29(1):69–95.

Cameron, A. Colin and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Carey, John M. 2009. *Legislative Voting and Accountability*. New York: Cambridge University Press.

Carey, John M. and Matthew Søberg Shugart. 1995. "Incentives to Cultivate a Personal Vote: A Rank Ordering of Electoral Formulas." *Electoral Studies* 14(4):417–439.

Carlin, Bradley P. and Thomas A. Louis. 2000. *Bayes and Empirical Bayes Methods for Data Analysis*. New York: Chapman & Hall/CRC.

Casella, George and Edward I. George. 1992. “Explaining the Gibbs Sampler.” *American Statistician* 46(3):167–174.

Chamberlain, Gary and Zvi Griliches. 1975. “Unobservables with a Variance-components Structure: Ability, Schooling, and the Economic Success of Brothers.” *International Economic Review* 16(2):422–449.

Cheibub, José Antonio, Argelina Figueiredo and Fernando Limongi. 2009. “Political Parties and Governors as Determinants of Legislative Behavior in Brazil’s Chamber of Deputies, 1988–2006.” *Latin American Politics and Society* 51(1):1–30.

Chib, Siddhartha. 2007. “Analysis of Treatment Response Data Without the Joint Distribution of Potential Outcomes.” *Journal of Econometrics* 140(2):401–412.

Chib, Siddhartha. 2008. Panel Data Modeling and Inference: A Bayesian Primer. In *The Econometrics of Panel Data*, ed. L. Mátyás and Patrick Sevestre. Springer chapter 15, pp. 479–515.

Chib, Siddhartha and Edward Greenberg. 1995. “Understanding the Metropolis-Hastings Algorithm.” *American Statistician* 49(4):327–335.

Chib, Siddhartha and Liana Jacobi. 2007. “Modeling and Calculating the Effect of Treatment at Baseline from Panel Outcomes.” *Journal of Econometrics* 140(2):781–801.

Clinton, Joshua D. 2007. “Lawmaking and Roll Calls.” *Journal of Politics* 69(2):457–469.

Clinton, Joshua D., Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2):355–70.

Cominetti, Rossella. 1996. Social Expenditure in Latin America: An Update. Technical report Santiago, Chile: CEPAL Technical Department.

Congdon, Peter. 2006. *Bayesian Statistical Modelling*. Wiley.

Cox, Gary W. 1990. "Centripetal and Centrifugal Incentives in Electoral Systems." *American Journal of Political Science* 34(4):903–935.

Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. New York: Cambridge University Press.

Cox, Gary W. and Mathew D. McCubbins. 1993. *Legislative Leviathan: Party Government in the House*. Berkeley: University of California Press.

Cox, Gary W. and Mathew D. McCubbins. 2005. *Setting the Agenda: Responsible Party Government in the US House of Representatives*. New York: Cambridge University Press.

De Boef, Suzanna and Luke Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52(1):184–200.

De Swaan, Abram. 1973. *Coalition Theories and Cabinet Formations: A Study of Formal Theories of Coalition Formation Applied to Nine European Parliaments After 1918*. San Francisco: Jossey-Bass.

Desposato, Scott W. 2006. "The Impact of Electoral Rules on Legislative Parties: Lessons from the Brazilian Senate and Chamber of Deputies." *The Journal of Politics* 68(4):1018–1030.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York, Harper.

Duverger, Maurice. 1954. *Political Parties: Their Organization and Activity in the Modern State.* New York: Wiley.

Embretson, Susan E. and Steven P. Reise. 2000. *Item Response Theory for Psychologists.* Mahwah: Lawrence Erlbaum.

Enders, Walter. 2004. *Applied Econometric Time Series.* New York: John Wiley & Sons Inc.

Enelow, James M. and Melvin J. Hinich. 1984. *The Spatial Theory of Voting: An Introduction.* New York: Cambridge University Press.

Ezrow, Lawrence. 2005. "Are Moderate Parties Rewarded in Multiparty Systems? A Pooled Analysis of Western European Elections, 1984–1998." *European Journal of Political Research* 44(6):881–898.

Figueiredo, Argelina Cheibub and Fernando Limongi. 2000. "Presidential Power, Legislative Organization, and Party Behavior in Brazil." *Comparative Politics* 32(2):151–170.

Fowler, James H. and Oleg Smirnov. 2005. "Dynamic Parties and Social Turnout: An Agent-Based Model." *American Journal of Sociology* 110(4):1070–1094.

Fowler, James H. and Oleg Smirnov. 2007. *Mandates, Parties, and Voters: How Elections Shape the Future*. Philadelphia: Temple University Press.

Fox, Jean-Paul. 2010. *Bayesian Item Response Modeling: Theory and Applications*. New York: Springer.

Fox, Jean-Paul and Cees A.W. Glas. 2001. "Bayesian Estimation of A Lultilevel IRT Model Using Gibbs Sampling." *Psychometrika* 66(2):271–288.

Gabel, Matthew J. and John D. Huber. 2000. "Putting Parties in their Place: Inferring Party left-right Ideological Positions from Party Manifestos Data." *American Journal of Political Science* 44(1):94–103.

Gelfand, Alan E. and Adrian F. M. Smith. 1990. "Sampling-Based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association* 85(410):398–409.

Gelman, Andrew. 2006. "Multilevel (Hierarchical) Modeling: What It Can and Cannot Do." *Technometrics* 48(3):432–435.

Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Gelman, Andrew, John B. Carlin, Hal S. Stern and Donald B. Rubin. 2004. *Bayesian Data Analysis*. 2 ed. New York: Chapman & Hall/CRC.

Gill, Jeff. 2008*a*. *Bayesian Methods: A Social and Behavioral Sciences Approach*. 2 ed. Boca Raton: Chapman & Hall/CRC press.

Gill, Jeff. 2008*b*. "Is Partial-Dimension Convergence a Problem for Inferences from MCMC Algorithms?" *Political Analysis* 16(2):153–178.

Greenberg, Edward. 2007. *Introduction to Bayesian Econometrics*. New York: Cambridge University Press.

Greene, William. 2011. "Fixed Effects vector Decomposition: A Magical Solution to the Problem of Time-Invariant Variables in Fixed Effects Models?" *Political Analysis* 19(2):135.

Hagopian, Frances, Carlos Gervasoni and Juan A. Moraes. 2009. "From Patronage to Program The Emergence of Party-Oriented Legislators in Brazil." *Comparative Political Studies* 42(3):360–391.

Hall, Peter A. and Rosemary C.R. Taylor. 1996. "Political Science and the Three New Institutionalisms." *Political Studies* 44(5):936–957.

Hamilton, James D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.

Hausman, Jerry A. and William E. Taylor. 1981. "Panel Data and Unobservable Individual Effects." *Econometrica: Journal of the Econometric Society* 49(6):1377–1398.

Hsiao, Cheng. 2003. *Analysis of Panel Data*. Cambridge University Pr.

Huber, Evelyne, Francois Nielsen, Jenny Pribble and John D. Stephens. 2006. "Politics and Inequality in Latin America and the Caribbean." *American Sociological Review* 71(6):943–963.

Huber, Evelyne, John D. Stephens, Thomas Mustillo and Jenny Pribble. 2008. *Social Policy in Latin America and the Caribbean Dataset, 1960-2006*. University of North Carolina.

Huber, Evelyne, Thomas Mustillo and John D. Stephens. 2008. "Politics and Social Spending in Latin America." *The Journal of Politics* 70(2):420–436.

Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. Chichester: Wiley.

Janda, Kenneth, Robert Harmel, Christine Edens and Patricia Goff. 1995. "Changes in Party Identity: Evidence from Party Manifestos." *Party Politics* 1(2):171–196.

Johnson, Valen E. and James H. Albert. 1999. *Ordinal Data Modeling*. New York: Springer-Verlag.

Jones, Bryan D., Tracy Sulkin and Heather A. Larsen. 2003. "Policy Punctuations in American Political Institutions." *American Political Science Review* 97(1):151–169.

Kaufman, Robert R. and Alex Segura-Ubiergo. 2001. "Globalization, Domestic Politics, and Social Spending in Latin America: A Time-series Cross-section Analysis, 1973-97." *World Politics* 53(4):553–587.

Keefer, Philip. 2007. "Clientelism, Credibility, and the Policy Choices of Young Democracies." *American Journal of Political Science* 51(4):804–821.

Keefer, Philip and Razvan Vlaicu. 2008. "Democracy, Credibility, and Clientelism." *Journal of Law, Economics, and Organization* 24(2):371–406.

Keefer, Philip and Stuti Khemani. 2005. "Democracy, Public Expenditures, and the Poor: Understanding Political Incentives for Providing Public Services." *The World Bank Research Observer* 20(1):1–27.

Keele, Luke and Nathan J. Kelly. 2006. "Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables." *Political Analysis* 14(2):186–205.

King, Gary, Robert O. Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.

Kitschelt, Herbert. 2000. "Linkages between Citizens and Politicians in Democratic Polities." *Comparative Political Studies* 33(6-7):845.

Kline, Rex B. 2010. *Principles and Practice of Structural Equation Modeling*. New York: Guilford Press.

Klingemann, Hans-Dieter, Andrea Volkens, Judith Bara, Ian Budge and Michael McDonald. 2006. *Mapping Policy Preferences II: Estimates for Parties, Electors, and Governments in Eastern Europe, the European Union and the OECD, 1990-2003*. Oxford: Oxford University Press.

Klingemann, Hans-Dieter, Richard I. Hofferbert and Ian Budge, eds. 1994. *Parties, Policies, and Democracy*. Oxford: Westview.

Kollman, Ken, John H. Miller and Scott E. Page. 1992. “Adaptive Parties in Spatial Elections.” *American Political Science Review* 86(4):929–937.

Kollman, Ken, John H. Miller and Scott E. Page. 1998. “Political Parties and Electoral Landscapes.” *British Journal of Political Science* 28(1):139–158.

Krehbiel, Keith. 1993. “Where's the Party?” *British Journal of Political Science* 23(2):235–266.

Krehbiel, Keith. 1998. *Pivotal Politics: A Theory of US Lawmaking*. Chicago: University of Chicago Press.

Krehbiel, Keith. 1999. “Paradoxes of Parties in Congress.” *Legislative Studies Quarterly* 24(1):31–64.

Krehbiel, Keith. 2000. “Party Discipline and Measures of Partisanship.” *American Journal of Political Science* 44(2):206–221.

Laver, Michael. 2005. “Policy and The Dynamics of Political Competition.” *American Political Science Review* 99(2):263–281.

Laver, Michael and Ernest Sergenti. 2011. *Party Competition: An Agent-based Model*. Princeton: Princeton University Press.

Laver, Michael and Ian Budge, eds. 1992. *Party Policy and Government Coalitions*. St. Martin's Press.

Laver, Michael and Michel Schilperoord. 2007. "Spatial Models of Political Competition with Endogenous Political Parties." *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1485):1711–1721.

Lee, Sik-Yum. 2007. *Structural Equation Modeling: A Bayesian Approach*. Hoboken: John Wiley & Sons Inc.

Levy, Gilat. 2004. "A Model of Political Parties." *Journal of Economic Theory* 115(2):250–277.

Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven: Yale University Press.

Lin, Tse-Min, James M. Enelow and Han Dorussen. 1999. "Equilibrium in Multicandidate Probabilistic Spatial Voting." *Public Choice* 98(1):59–82.

Lindley, D.V. 1972. *Bayesian Statistics: A Review*. Philadelphia : Society for Industrial and Applied Mathematics.

Lindley, D.V. and A.F.M. Smith. 1972. "Bayes Estimates for the Linear Model." *Journal of the Royal Statistical Society. Series B (Methodological)* 34(1):1–41.

Linz, Juan J. and Arturo Valenzuela. 1994. *The Failure of Presidential Democracy: Comparative Perspectives*. Baltimore: Johns Hopkins University Press.

Lipset, Seymour M. and Stein Rokkan. 1990. Cleavages Structures, Party Systems, and Voter Alignments. In *The West European Party System*, ed. P. Mair. Oxford: Oxford University Press chapter 9, pp. 91–138.

Lunn, David J., Andrew Thomas, Nicky Best and David Spiegelhalter. 2000. "WinBUGS-A Bayesian Modelling Framework: Concepts, Structure, and Extensibility." *Statistics and Computing* 10(4):325–337.

Lyne, Mona M. 2005. "Parties as Programmatic Agents: A Test of Institutional Theory in Brazil." *Party politics* 11(2):193–216.

Lyne, Mona M. 2008. "Proffering Pork: How Party Leaders Build Party Reputations in Brazil." *American Journal of Political Science* 52(2):290–303.

Mainwaring, Scott. 1991. "Politicians, Parties, and Electoral Systems: Brazil in Comparative Perspective." *Comparative Politics* 24(1):21–43.

Mainwaring, Scott. 1993. "Presidentialism, Multiparty Systems, and Democracy: The Difficult Equation." *Comparative Political Studies* 26(2):198–228.

Mainwaring, Scott. 1999. *Rethinking Party Systems in the Third Wave of Democratization: The Case of Brazil*. Stanford: Stanford University Press.

Mainwaring, Scott and Matthew Søberg Shugart. 1997. *Presidentialism and Democracy in Latin America*. New York: Cambridge University Press.

March, James G. and Johan P. Olsen. 1984. "The New Institutionalism: Organizational Factors in Political Life." *American Political Science Review* 78(3):734–49.

March, James G. and Johan P. Olsen. 2006. Elaborating the "New Institutionalism". In *The Oxford Handbook of Political Institutions*. Oxford: Oxford University Press chapter 1, pp. 3–20.

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10(2):134–152.

McCarty, Nolan M., Keith T. Poole and Howard Rosenthal. 2001. "The Hunt for Party Discipline in Congress." *American Political Science Review* 95(3):673–688.

McCarty, Nolan M., Keith T. Poole and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge: MIT Press.

McDonald, Michael D., Silvia M. Mendes and Ian Budge. 2004. "What Are Elections For? Conferring the Median Mandate." *British Journal of Political Science* 34(1):1–26.

Meltzer, Allan H. and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *The Journal of Political Economy* 89(5):914–927.

Morgenstern, Scott. 2004. *Patterns of Legislative Politics: Roll-call Voting in Latin America and the United States*. New York: Cambridge University Press.

Morgenstern, Scott and Benito Nacif, eds. 2002. *Legislative Politics in Latin America*. New York: Cambridge University Press.

Müller, Wolfgang C. and Kaare Strøm, eds. 1999. *Policy, Office, or Votes?: How Political Parties in Western Europe Make Hard Decisions*. New York: Cambridge University Press.

Mundlak, Yair. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica: journal of the Econometric Society* 46(1):69–85.

Nickell, Stephen. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica: Journal of the Econometric Society* 49(6):1417–1426.

North, Douglass C. 1990. *Institutions, Institutional Change, and Economic Performance*. New York: Cambridge University Press.

Olson, Mancur. 1993. "Dictatorship, Democracy, and Development." *American Political Science Review* 87(3):567–76.

Pereira, Carlos and Bernardo Mueller. 2004*a*. "The Cost of Governing Strategic Behavior of the President and Legislators in Brazil's Budgetary Process." *Comparative Political Studies* 37(7):781–815.

Pereira, Carlos and Bernardo Mueller. 2004*b*. "A Theory of Executive Dominance of Congressional Politics: The Committee System in the Brazilian Chamber of Deputies." *The Journal of Legislative Studies* 10(1):9–49.

Persson, Torsten and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge, Mass.: MIT Press.

Persson, Torsten and Guido Tabellini. 2003. *The Economic Effects of Constitutions*. Cambridge, Mass.: MIT Press.

Plummer, Martyn. 2003. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*. pp. 20–22.

Plummer, Martyn. 2012. *JAGS Version 3.2.0 Manual (January 3, 2012)*.

Plümper, Thomas and Vera E. Troeger. 2007. "Efficient Estimation of Time-invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects." *Political Analysis* 15(2):124.

Plümper, Thomas, Vera E. Troeger and Philip Manow. 2005. "Panel Data Analysis in Comparative Politics: Linking Method to Theory." *European Journal of Political Research* 44(2):327–354.

Poole, Keith T. and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29(2):357–384.

Poole, Keith T. and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.

Powell Jr., G.Bingham. 2004. "The Chain of Responsiveness." *Journal of Democracy* 15:91–105.

Powell Jr., G.Bingham. 2007. Aggregating and Representing Political Preferences. In *The Oxford Handbook of Comparative Politics*, ed. Carles Boix and Susan C. Stokes. Oxford: Oxford University Press chapter 27, pp. 653–677.

Rasch, Georg. 1960. *Probabilistic Models for Some Intelligence and Attainment Tests*. The Danish Institute for Educational Research.

Raudenbush, Stephen W. and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage Publications, Inc.

Robert, Christian. 2001. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Springer.

Roodman, David. 2009. "A Note on the Theme of Too Many Instruments." *Oxford Bulletin of Economics and Statistics* 71(1):135–158.

Rosas, Guillermo. 2005. "The Ideological Organization of Latin American Legislative Parties An Empirical Analysis of Elite Policy Preferences." *Comparative Political Studies* 38(7):824–849.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41–55.

Ross, Michael L. 2006. "Is democracy Good for the Poor?" *American Journal of Political Science* 50(4):860–874.

Rubin, Donald B. 1973. "Matching to Remove Bias in Observational Studies." *Biometrics* 29(1):159–183.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66(5):688–701.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys.* John Wiley & Sons.

Rueschemeyer, Dietrich, Evelyn Huber Stephens and John D. Stephens. 1992. *Capitalist Development and Democracy.* Polity.

Saiegh, Sebastian M. 2009. "Recovering a Basic Space from Elite Surveys: Evidence from Latin America." *Legislative Studies Quarterly* 34(1):117–145.

Sartori, Giovanni. 1976. *Parties and Party Systems: A Framework for Analysis.* New York: Cambridge University Press.

Schofield, Norman and Itai Sened. 2006. *Multiparty Democracy.* New York: Cambridge University Press.

Sevestre, Patrick and Alain Trognon. 1985. "A Note on Autoregressive Error Components Models." *Journal of Econometrics* 28(2):231–245.

Shepsle, Kenneth A. 1979. "Institutional Arrangements and Equilibrium in Multidimensional Voting Models." *American Journal of Political Science* 23(1):27–59.

Shepsle, Kenneth A. 2006. Rational Choice Institutionalism. In *The Oxford Handbook of Political Institutions.* Oxford: Oxford Univ Press chapter 2, pp. 23–38.

Shepsle, Kenneth A. and Mark S. Bonchek. 1997. *Analyzing Politics: Rationality, Behavior, and Institutions.* New York: W.W. Norton.

Shor, Boris, Joseph Bafumi, Luke Keele and David Park. 2007. "A Bayesian Multilevel Modeling Approach to Time-Series Cross-Sectional Data." *Political Analysis* 15(2):165–181.

Shugart, Matthew Søberg. 2006. Comparative Executive-Legislative Relations. In *The Oxford Handbook of Political Institutions*. Oxford: Oxford University Press chapter 18, pp. 344–365.

Shugart, Matthew Søberg and John M. Carey. 1992. *Presidents and Assemblies: Constitutional Design and Electoral Dynamics*. New York: Cambridge University Press.

Shugart, Matthew Søberg, Melody Ellis Valdini and Kati Suominen. 2005. "Looking for Locals: Voter Information Demands and Personal Vote-earning Attributes of Legislators under Proportional Representation." *American Journal of Political Science* 49(2):437–449.

Skrondal, Anders and Sophia Rabe-Hesketh. 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Boca Raton: Chapman and Hall/CRC.

Smirnov, Oleg and James H. Fowler. 2007. "Policy-Motivated Parties in Dynamic Political Competition." *Journal of Theoretical Politics* 19(1):9–31.

Snyder, James M. and Michael M. Ting. 2002. "An Informational Rationale for Political Parties." *American Journal of Political Science* 46(1):90–110.

Somer-Topcu, Zeynep. 2009. "Timely Decisions: The Effects of Past National Elections on Party Policy Change." *The Journal of Politics* 71(1):238–248.

Song, Xin-Yuan, Zhao-Hua Lu, Yih-Ing Hser and Sik-Yum Lee. 2011. "A Bayesian Approach for Analyzing Longitudinal Structural Equation Models." *Structural Equation Modeling* 18(2):183–194.

Stimson, James A. 1985. "Regression in Space and Time: A Statistical Essay." *American Journal of Political Science* 29(4):914–947.

Stimson, James A., Michael B MacKuen and Robert S. Erikson. 1995. "Dynamic Representation." *American Political Science Review* 89(3):543–565.

Stokes, Donald E. 1963. "Spatial Models of Party Competition." *American Political Science Review* 57(2):368–377.

Strøm, Kaare. 1990. "A Behavioral Theory of Competitive Political Parties." *American Journal of Political Science* 34(2):565–598.

Su, Yu-Sung and Masanao Yajima. 2012. *R2jags ver. 0.03-08 (R package)*.
URL: *http://cran.r-project.org/web/packages/R2jags/*

Taagepera, Rein and Matthew Søberg Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems*. Yale University Press.

Tanner, Martin A. and Wing Hung Wong. 1987. "The Calculation of Posterior Distributions by Data Augmentation." *Journal of the American statistical Association* 82(398):528–540.

Tavits, Margit. 2007. "Principle vs. Pragmatism: Policy Shifts and Political Competition." *American Journal of Political Science* 51(1):151–165.

Tsai, Tsung-han and Jeff Gill. 2012. "superdiag: A Comprehensive Test Suite for Markov Chain Non-Convergence." *The Political Methodologist* 19(2):12–18.

Van Buuren, S. and K. Groothuis-Oudshoorn. 2011. "MICE: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45(3):1–66.

Walgrave, Stefaan and Michiel Nuytemans. 2009. "Friction and Party Manifesto Change in 25 Countries, 1945–98." *American Journal of Political Science* 53(1):190–206.

Weingast, Barry R. 1996. Political Institutions: Rational Choice Perspectives. In *A New Handbook of Political Science*, ed. Robert E. Goodin and Hans-Dieter Klingemann. Oxford University Press pp. 167–190.

Weingast, Barry R. 2002. Rational-Choice Institutionalism. In *Political Science: The State of the Discipline*, ed. Ira Katznelson and Helen Milner. Centennial ed. American Political Science Association pp. 660–92.

Wittman, Donald. 1977. "Candidates with Policy Preferences: A Dynamic Model." *Journal of Economic Theory* 14(1):180–189.

Wittman, Donald. 1983. "Candidate Motivation: A Synthesis of Alternative Theories." *The American political science review* 77(1):142–157.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT press.

Zellner, A. 1970. "Estimation of Regression Relationships Containing Unobservable Independent Variables." *International Economic Review* 11(3):441–454.

Zucco, Cesar. 2009. "Ideology or What? Legislative Behavior in Multiparty Presidential Settings." *The Journal of Politics* 71(3):1076–1092.

Zucco, Cesar. 2013. "Roll Call Votes from the Câmara de Deputados." *Cesar Zucco Dataverse* .

URL: *http://hdl.handle.net/1902.1/20270*

Zucco, Cesar and Benjamin E. Lauderdale. 2011. "Distinguishing between Influences on Brazilian Legislative Behavior." *Legislative Studies Quarterly* 36(3):363–396.