Bayesian Inference for Dynamics of Slowly Changing Variables in Time-Series Cross-Sectional Data Analyses*

Tsung-han Tsai†

August 25, 2011

Abstract

The time-invariant and/or rarely changing explanatory variables are of interest to political scientists, including both their short- and long-run effects. However, estimating these effects in the analysis of time-series cross-sectional (TSCS) data by the conventional estimators may be problematic when unit effects are included in the model. This paper discusses the advantages of using Bayesian multilevel modeling to estimate the dynamic effects of these slowly changing explanatory variables in the analysis of TSCS data and applies a Bayesian dynamic multilevel model to analyzing the effects of political regime on social spending in Latin America.

keywords: Bayesian inference, multilevel modeling, dynamic panel models, time-invariant variables, social spending, Latin America

*Paper prepared for presentation at the 2011 Annual Meeting of the American Political Science Association in Seattle, Washington, September 1-4.

†Graduate student, Department of Political Science at Washington University in St. Louis. Email address: t.tsai@wustl.edu

1 Introduction

The use of time-series cross-sectional (TSCS) data, which have both cross-sectional and intertemporal variations, allows researchers to analyze questions that cannot be addressed using pure cross-sectional or time-series data. However, the advantages of TSCS data structure come along with some problems of estimation. In specific, there exist serial correlations, heteroscedasticity, and contemporaneous correlations in TSCS data structure so that the standard assumptions underlying classical linear regression models are violated (Stimson, 1985; Beck and Katz, 1995).

Although a variety of estimators has been developed to analyze TSCS data, other problems arise. For example, when unit heterogeneity is taken into account in modeling TSCS data (e.g., the fixed-effects models and the dynamic panel models), it is problematic to estimate time-invariant and/or rarely changing variables with unit effects (Beck, 2001; Hsiao, 2003; Plümper, Troeger and Manow, 2005; Wilson and Butler, 2007). Most of the time, however, the rarely changing explanatory variables are of main interest to political scientists. For instance, political institutions such as political regime and electoral systems which are often-seen explanatory variables in the comparative political economy literature typically persist over time. In addition, because of the slowly changing characteristics, both short- and long-term effects of these variables are important.

This paper discusses a solution to estimating the dynamic effects of slowly changing variables with unit effects in TSCS data analyses. I use a dynamic panel model in the multilevel framework, which allows for subject-specific random coefficients, subject-specific effects, and time-specific effects. The dynamic multilevel model is estimated by adopting a Bayesian approach with Markov chain Monte Carlo (MCMC) techniques. Multilevel modeling is applied because TSCS data are in nature a multilevel structure and multilevel modeling appropriately accounts for both cross-sectional and intertemporal variations (Beck and Katz, 2007; Shor et al., 2003). A Bayesian approach to multilevel modeling offers flexibility for complex model specifications and resolves inferential problems that arise in non-Bayesian multilevel models (see, e.g., Gelman, 2006; Gelman and Hill, 2007; Gill, 2008; Raudenbush and Bryk, 2002). Furthermore, it has been shown that Bayesian multilevel models perform as well, or better than conventional estimators to deal with heterogeneity (Western, 1998) and heteroscedasticity and contemporaneous correlations (Shor et al., 2007) and in the estimation of the dynamic models (Hsiao, Pesaran and Tahmiscioglu, 1999; Zhang and Small, 2006) in Monte Carlo simulations and empirical examples.

To assess the performance of the Bayesian dynamic multilevel model presented in this paper, I employ a Monte Carlo study, in which I focus on the varying coefficients of the lagged dependent variable (LDV) and the slowly changing variables. The simulation results show that the Bayesian dynamic multilevel model performs well in the estimation of the dynamic models containing slowly changing variables and unit effects. Moreover, I employ the Bayesian dynamic multilevel model to analyzing the effects of political regime on social spending in Latin America, where some countries have experienced longer democracy than others. The results of analysis indicate heterogeneous effects of democratic regime, including both short- and long-run effects.

The contribution of this paper is two-fold. First, methodologically, this paper shows the advantages of Bayesian hierarchical models in the estimation of the dynamic models containing rarely changing variables in analyses of TSCS data. Furthermore, using the Bayesian methods allows researchers to easily estimate the quantities of interest with their uncertainty such as the dynamic multiplier and the long-run effect. Second, substantively, this paper contributes to our understanding of the effects of political regime on social spending in Latin American countries. On the one hand, the variation of democratic experience suggests heterogeneous effects of democracy on social spending among Latin American countries. On the other hand, this variation provides researchers opportunities to investigate the systematic effects on the spending on social programs in general, compared to OECD countries.

The remainder of this paper proceeds as follows. The next section discusses the short- and long-run effects of an almost time-invariant variable in the dynamic panel models and the Bayesian methods for these models in analyses of TSCS data, followed by a Monte Carlo study. The fourth section presents the application of the Bayesian dynamic model to social spending in Latin America. Finally, I summarize the findings in this paper and discuss the possibilities of future research.

2 The Model

Plümper and Troeger (2007) defined 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 for the period under analysis. In this paper, the time-invariant variables belong to the second category. In TSCS data, it is more often that these variables are time invariant in some countries and rarely changing in others. Hereafter, I use rarely changing, slowly changing, or almost time-invariant variables to refer to these variables.

2.1 The Dynamics of Slowly Changing Variables

For simplicity and the purpose of this paper, I start with a dynamic panel model, which is a model containing unit effects and a lagged dependent variable. Then I consider only one explanatory variable which is almost time invariant. A general model which includes more covariates will be discussed below. Moreover, I allow coefficients to vary across units and assume heteroscedastic, independent errors across units. Thus the model has the form:

$$y_{jt} = \phi_j y_{j(t-1)} + \beta_j z_{jt} + \delta_j + \varepsilon_{jt}, \tag{1}$$

where y_{jt} is the outcome variable subscripted for units (j) and time (t) and is assumed to be stationary for individual units; $y_{j(t-1)}$ is the lagged dependent variable and ϕ_j is the autoregressive coefficient; z_{jt} is the almost time-invariant explanatory variable with corresponding coefficient β_j; δ_j denotes the effects that are specific to individual units; the error term ε_{jt} is assumed to be independently, identically distributed over t with mean zero and variance $\sigma^2_{\varepsilon_j}$ and is independent across j. Furthermore, the coefficients $\boldsymbol{\theta}_j = (\phi_j, \delta_j, \beta_j)$ are assumed to be independently distributed across j with means $\bar{\boldsymbol{\theta}} = (\bar{\phi}, \bar{\delta}, \bar{\beta})$ and covariance matrix $\Sigma_{\boldsymbol{\theta}}$.

To explain how the inclusion of a lagged dependent variable can capture the dynamics of the explanatory variable, by some algebraic calculation, Equation (1) can be rewritten as:[1]

$$y_{jt} = \beta_j \sum_{q=0}^{\infty} \phi_j^q z_{j(t-q)} + \frac{\delta_j}{1-\phi_j} + \frac{\varepsilon_{jt}}{1-\phi_j}. \tag{2}$$

By recursive substitution and the calculation of the marginal effect of z_{jt} on $y_{j(t+p)}$, one can obtain the dynamic multiplier:

$$\frac{\partial y_{j(t+p)}}{\partial z_{jt}} = \phi_j^p \beta_j, \tag{3}$$

where p denotes the length of time between the input z_{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 but not on t. Therefore, $\frac{\partial y_{jt}}{\partial z_{j(t-1)}} = \frac{\partial y_{j(t+1)}}{\partial z_{jt}} = \beta_j$.

To derive the dynamic effects of the almost time-invariant explanatory variable on the outcome, one has to consider the values of the variable at different time periods. Usually, there are two ways to deal with the coding for the almost time-invariant explanatory variables. One way is to treat them as dummy variables. The other is to use ordinal scales by which the levels are meaningful (e.g., from -10 to 10).

Concerning the first case, suppose that, for a given unit j, z_{jt} takes 1 from time t to $t+p$ and 0 before t. Since $\frac{\partial y_{jt}}{\partial z_{j(t-1)}} = \frac{\partial y_{j(t+1)}}{\partial z_{jt}} = \beta_j$ and $z_{jt} = \cdots = z_{j(t+p)} = 1$, the impact on $y_{j(t+p)}$ of a permanent change in z_{jt} is given by

$$\frac{\partial y_{j(t+p)}}{\partial z_{jt}} = \beta_j(1 + \phi_j + \phi_j^2 + \cdots + \phi_j^p). \tag{4}$$

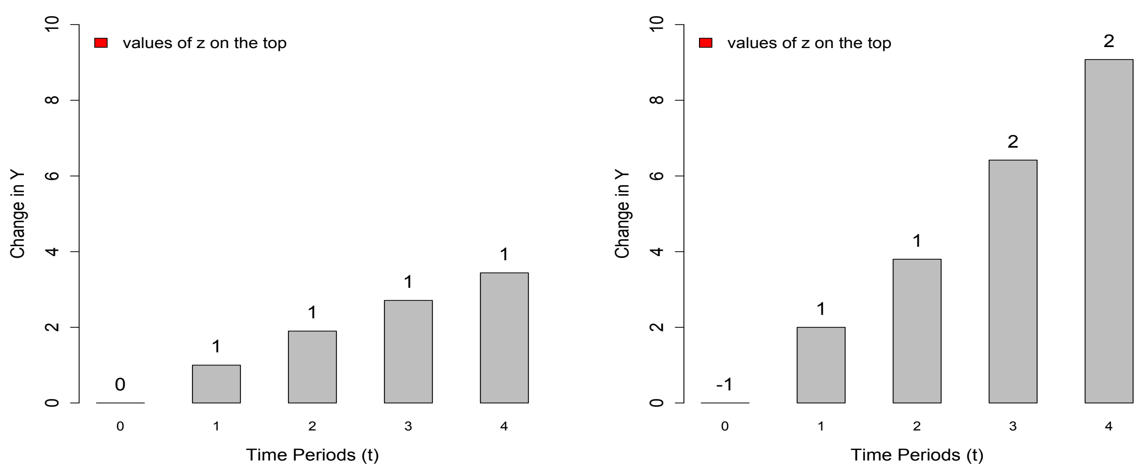
For a stationary process $\{y_{jt}\}$, that is, when $|\phi_j| < 1$, the limit of Equation (4) as p goes to infinity is the long-run effect of z_j on y_j given by the following:[2]

[1] The following results can be found in every textbook of time series analyses and are summarized from Enders (2004) and Hamilton (1994) in terms of panel models.

[2] Analogue to what Hamilton (1994) shows in univariate time series analysis, the cumulative effect on y of a transitory change in z in the limit is the same with the long-run effect, that is, $\sum_{p=0}^{\infty} \frac{\partial y_{j(t+p)}}{\partial z_{jt}} = \frac{\beta_j}{1-\phi_j}$.

$$\lim_{p\to\infty}\left(\frac{\partial y_{j(t+p)}}{\partial z_{jt}} + \cdots + \frac{\partial y_{j(t+p)}}{\partial z_{j(t+p)}}\right) = \frac{\beta_j}{1-\phi_j}. \tag{5}$$

An example is presented in the left panel in Figure 1, which shows a permanent change in the input variable (from 0 to 1) at time period $t = 1$ and the corresponding changes in the outcome, as presented in Equation (4), assuming $\phi_j = 0.9$ and $\beta_j = 1$. When the time period t goes to infinity, the change in the outcome approaches to the long-run effect of the input. The calculation of the dynamics of the outcome is the same if the input changes its value from 1 to 0 and 0 to 1 again.



(a) Binary Scales

(b) Ordinal Scales

Figure 1: Dynamics of the covariate and the outcome in two different ways of scaling. In this example, $\beta_j = 1$ and $\phi_j = 0.9$. The values of the input Z are on the top of these bars.

However, in the case that the measures are ordinal scales, the dynamic effects of the input variable are not obvious because the change in the levels of the variables should be taken into account. Suppose that the ordinal scale is from -10 to 10. As can be seen in the right panel in Figure 1, when the explanatory variable changes from -1 to 1 at time $t = 1$, the change in Y is $2\beta_j = 2$ for a given unit j. The dynamics of the outcome depending on the input can be obtained by multiplying Equation (4) by the change in the value of the input if the level of the input persists. Moreover, once the level of the input goes up, it will lead to an additional change in the outcome.[3] For example, the right panel in Figure 1 shows that $y_{j3} - y_{j2} > y_{j2} - y_{j1}$ while it is that $y_{j3} - y_{j2} < y_{j2} - y_{j1}$ in the left panel. This also changes the long-run effect of the input.[4]

2.2 The Dynamic Multilevel Models in TSCS Structure

Equation (1) can be extended and include more (both time-invariant and time-variant) covariates and the effects specific to each time period. The general dynamic multilevel model has the following form:

$$y_{jt} = \phi_j y_{j(t-1)} + \boldsymbol{Z}_{jt}\boldsymbol{\beta_j} + \boldsymbol{X}_{jt}\boldsymbol{\alpha}_j + \delta_j + \gamma_t + \varepsilon_{jt}, \tag{6}$$

where y_{jt} is the outcome variable and $y_{j(t-1)}$ is the lagged dependent variable with autoregressive coefficient $|\phi_j| \leq 1$; $\boldsymbol{Z}_{jt}$ is a $k \times 1$ vector of time-invariant explanatory variables (including the intercept) and $\boldsymbol{X}_{jt}$ is a $p \times 1$ vector of time-variant explanatory variables with unit-specific coefficients $\boldsymbol{\beta}_j$ and $\boldsymbol{\alpha}_j$, respectively; δ_j and γ_t are unit effects and time effects, respectively; the error term ε_{jt} is assumed to be iid over t with mean zero and variance $\sigma^2_{\varepsilon_j}$ and is independent across j. Moreover, the coefficients $\boldsymbol{\theta}_j = (\phi_j, \delta_j, \boldsymbol{\beta}_j, \boldsymbol{\alpha}_j)$ are assumed to be independently distributed across j with mean $\boldsymbol{\mu}_\theta$ and covariance matrix $\Sigma_{\boldsymbol{\theta}}$; γ_t is assumed to be iid over t with mean 0 and variance σ^2_γ.

The statistical model presented in Equation (6) has several advantages in the analysis of TSCS data. First, as presented in Equation (4), the inclusion of a lagged dependent variable in the model directly accounts for dynamics in equation specification rather than focusing on correction of the error terms (Wilson and Butler, 2007). It is argued that when including a lagged dependent variable appropriately captures dynamics of explanatory variables, autocorrelation can

[3]It is implicitly assumed that the marginal effects are constant no matter at what level the variable is when the ordinal-scale input is used. This assumption is problematic because, on the one hand, it ignores that the marginal effects may be larger at some level than at others. On the other hand, the marginal effects may be positive at some level and negative at others. One way to deal with these two problems is to use polynomial terms of explanatory variables.

[4]If the input changes from -1 to 1 and persists a long time, the long-run effect is $\frac{2\beta_j}{1-\phi_j}$. However, if the input changes from 1 to 2 at time $t = 3$, the long-run effect is $\frac{3\beta_j}{1-\phi_j}$.

be eliminated (Beck and Katz, 1996).[5] In particular, when researchers are interested in almost time-invariant explanatory variables and these variables are believed to have persistent effects which decay over time, the inclusion of a lagged dependent variable is appropriate (Keele and Kelly, 2006).

Second, unit effects and time effects specified in multilevel modeling accounts for unit heterogeneity and common shocks, respectively. Unit effects are often used to deal with omitted variable bias because they account for unit heterogeneity which is not explained by observed explanatory variables. Without unit effects, it basically implies that the unobserved heterogeneity does not exist. Moreover, the unit effects δ_j determine the levels of the outcome in the long run. For instance, as presented in Equation (2), the process $\{y_{jt}\}$ converges to $\frac{\beta_j+\delta_j}{1-\phi_j}$ in the long run, which means that the level of equilibrium might be different due to unit heterogeneity (Wilson and Butler, 2007). For the similar reason, common effects of particular time periods (e.g., years) across all units, which are a source of contemporaneous correlation, can be accounted for by including an indicator for time periods.

Third, multilevel modeling considers not only the mean effects but also the heterogeneous effects (Beck and Katz, 2007). The setup of varying coefficients basically assumes the existence of causal heterogeneity, including the short- and long-run effects, which is important in comparative politics (Western, 1998). From the discussion in section 2.1, we know that the varying slope β_j implies that the immediate effect of the explanatory variable differs across units and the varying autoregressive coefficient ϕ_j indicates that the ratio of the long-run effect to the short-run effect is different across units.[6] Moreover, these two coefficients together determine the dynamic effects of the explanatory variable for individual units, which is shown in Equation (4). Therefore, imposing any constraint on these coefficients (i.e., $\phi_j = \phi$ for all j to avoid the problem of identification),

[5] There is a debate over the inclusion of the lagged dependent variable (Achen, 2000; Beck and Katz, 1996, 2009; Keele and Kelly, 2006; Plümper, Troeger and Manow, 2005). Readers who are interested in the debate can find discussion in the literature.

[6] ϕ_j also indicates the decaying rate of the effect of the input. Although I allow the lagged dependent variable's coefficient (ϕ_j) to vary across units, the coefficient is identical across explanatory variables for each country, which implicitly assumes that decaying rates of all explanatory variables are not different. Moreover, the marginal effect β_j and the dynamic multiplier $\frac{\partial y_{j(t+p)}}{\partial z_{jt}} = \phi_j^p \beta_j$ do not depend on t, which implies that the dynamics of the explanatory variables are the same if the explanatory variables change from 0 to 1 at different time periods.

that is, assuming the lack of the heterogeneity, may lead to incorrect inference.

Fourth, the model presented in this section allows researchers to handle panel heteroscedasticity inherent in TSCS data. To account for heteroscedasticity, the error term is assumed to be iid over t and is heteroscedastic, independent across j. In a multilevel context, the varying variances can be estimated from the data and can be modeled hierarchically (Gelman and Hill, 2007).

Last but not least, the slowly changing explanatory variables and unit effects can be simultaneously estimated in multilevel models. It is a problem to estimate coefficients of the slowly changing variables with unit effects by estimators relying on asymptotic properties (e.g., the maximum likelihood (ML) method and the least squares dummy variable estimator (LSDV)).[7] However, this is not a problem in multilevel models since multilevel models allow both the within-unit and between-unit variances to be estimated conditional on the information at all levels (Gelman and Hill, 2007; Raudenbush and Bryk, 2002; Shor et al., 2007). With this "borrowing strength" from various levels, multilevel modeling attempts to find the best estimate in each unit, appropriately accounting for uncertainty (Gelman and Hill, 2007).

2.3 Bayesian Approach

The general dynamic multilevel model, which is presented here again,

$$y_{jt} = \phi_j y_{j(t-1)} + \boldsymbol{Z}_{jt}\boldsymbol{\beta_j} + \boldsymbol{X}_{jt}\boldsymbol{\alpha}_j + \delta_j + \gamma_t + \varepsilon_{jt},$$

is estimated within a Bayesian framework. It is problematic to estimate the general dynamic multilevel model by the classical estimators, and Bayesian multilevel modeling is preferred for several reasons. First, for a dynamic model containing varying autoregressive coefficients and

[7]In Anderson and Hsiao (1982), the authors discussed the estimation of panel models containing unit effects and time-invariant variables and focused on the asymptotic properties of the covariance method and the ML method conditional on the assumption about the initial observations in the process $\{y_{jt}\}$. Recently, Plümper and Troeger (2007) proposed a three-step procedure for the estimation of time-invariant and rarely changing variables in panel data models with unit effects, which is called the "fixed-effect vector decomposition" (FEVD) procedure. This FEVD procedure induced a controversy and readers can find a symposium on this procedure in the Spring 2011 edition of Political Analysis.

time-variant covariates (also called a dynamic random coefficients model), the classical estimators relying on asymptotic properties no longer provide consistent estimates of the mean of ϕ_j and $\boldsymbol{\alpha}_j$ even if $T \to \infty$ (Pesaran and Smith, 1995). It has been showed that a Bayesian hierarchical approach to the estimation of the dynamic random coefficients model performs fairly well even when T is small (Hsiao, 2003; Hsiao, Pesaran and Tahmiscioglu, 1999; Hsiao and Pesaran, 2004). Second, the estimates of the common coefficients (ϕ and α) in dynamic panel models are biased when T or N is small (Nerlove, 1971; Wilson and Butler, 2007). The biased estimates of ϕ (when T is finite) and β (when N is finite) can also be seen in Anderson and Hsiao (1981, 1982). Finally, the Bayesian methods do not rely on asymptotic properties and any quantity of interest (e.g., the long-run effects) with uncertainty can be estimated (Gill, 2008).

To estimate this model, I first assume that the error terms are iid over t and are independent across j with a normal distribution with mean zero and variance $\sigma^2_{\varepsilon_j}$, that is,

$$\varepsilon_{jt} \sim N(0, \sigma^2_{\varepsilon_j}).$$

Second, since the process $\{y_{jt}\}$ is assumed to be stationary, the autoregressive coefficient ϕ_j must be less than one in absolute value. Following Zhang and Small (2006), I assume that ϕ_j is drawn from a logit-normal distribution scaled to have support on (-1, 1) by the form of transformation $\phi_j = 2(\frac{\exp(\zeta_j)}{1+\exp(\zeta_j)} - 0.5)$, where ζ_j and other coefficients, $\boldsymbol{B}_j = (\zeta_j, \delta_j, \boldsymbol{\beta}_j, \boldsymbol{\alpha}_j)$, are assumed to have a multivariate normal distribution with mean $\boldsymbol{\mu}_B$ and covariance matrix $\boldsymbol{\Sigma}_B$.[8] Furthermore, I assume that γ_t is iid over t with mean 0 and variance σ^2_γ. Therefore,

$$\boldsymbol{B}_j \sim MVN(\boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B),$$
$$\gamma_t \sim N(0, \sigma^2_\gamma).$$

In the Bayesian context, I have to put prior distributions on parameters $\boldsymbol{\mu}_B$, $\boldsymbol{\Sigma}_B$, σ^2_γ, and $\sigma^2_{\varepsilon_j}$. I

[8]The values of the boundary never occur when a variable is logit-normal distribution. This property is appropriate for the autoregressive coefficient when the process is stationary.

assume priori independence of these parameters, that is, $f(\boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B, \sigma^2_\gamma, \sigma^2_{\varepsilon_j}) = f(\boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B) f(\sigma^2_\gamma) f(\sigma^2_{\varepsilon_j})$. Following the conventional method, I use the normal-inverse Wishart as the prior distribution for $\boldsymbol{\mu}_B$ and $\boldsymbol{\Sigma}_B$ with hyperparameters $(\mu_0, \boldsymbol{\Sigma}_B/\kappa_0; \nu_0, \Lambda_0)$, that is,

$$\begin{aligned}\boldsymbol{\Sigma}_B &\sim Inv-Wishart(\nu_0, \Lambda_0^{-1}),\\ \boldsymbol{\mu}_B|\boldsymbol{\Sigma}_B &\sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_B/\kappa_0),\end{aligned}$$

where ν_0 and Λ_0 are the degrees of freedom and the scale matrix for the inverse-Wishart distribution; μ_0 is the prior mean and κ_0 is the number of prior measurements on the $\boldsymbol{\Sigma}_B$ (Gelman et al., 2004). I assume inverse gamma priors for σ^2_γ and $\sigma^2_{\varepsilon_j}$ with parameters (a, b) and (c, d), respectively. Thus,

$$\begin{aligned}\sigma^2_\gamma &\sim IG(a, b),\\ \sigma^2_{\varepsilon_j} &\sim IG(c, d).\end{aligned}$$

In a Bayesian hierarchical framework, these parameters can be modeled hierarchically, i.e., using unit-specific covariates to model $\boldsymbol{B}_j$ or time-specific covariates to model γ_t. In addition, researchers can add their prior knowledge or impose constraints on these parameters through hyperparameters.

The joint posterior density of interest is as follows:

$$\begin{aligned}\pi(\boldsymbol{B}, \boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B, \boldsymbol{\sigma}^2_\varepsilon, \boldsymbol{\gamma}, \sigma^2_\gamma | \boldsymbol{y}, \boldsymbol{Z}, \boldsymbol{X}) &\propto f(\boldsymbol{y}|\boldsymbol{B}, \boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B, \boldsymbol{\sigma}^2_\varepsilon, \boldsymbol{\gamma}, \sigma^2_\gamma) f(\boldsymbol{B}, \boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B, \boldsymbol{\sigma}^2_\varepsilon, \boldsymbol{\gamma}, \sigma^2_\gamma)\\ &= f(\boldsymbol{y}|\boldsymbol{B}, \boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B, \boldsymbol{\sigma}^2_\varepsilon, \boldsymbol{\gamma}, \sigma^2_\gamma)\\ &\quad \times f(\boldsymbol{B}|\boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B) \times f(\boldsymbol{\mu}_B|\boldsymbol{\Sigma}_B) \times f(\boldsymbol{\Sigma}_B)\\ &\quad \times f(\boldsymbol{\gamma}|\sigma_\gamma) \times f(\sigma_\gamma)\\ &\quad \times f(\sigma^2_\varepsilon). \end{aligned} \tag{7}$$

With the priors assumed above, this posterior distribution is not a standard distribution from which samples can be easily drawn. Thus, Markov chain Monte Carlo (MCMC) methods are used (Casella and George, 1992; Chib and Greenberg, 1995) and can be implemented in a variety of programs (e.g., WinBUGS (Lunn et al., 2000) and JAGS (Plummer, 2003)).

3 Monte Carlo Experiments

In this section, I employ Monte Carlo simulations to assess the performance of a Bayesian approach to the estimation of a dynamic multilevel model containing slowly changing variables and unit effects. Since the performance of Bayesian inference for TSCS data structure (Shor et al., 2007) and for mean effects of random coefficients in dynamic models have been studied elsewhere (Hsiao, Pesaran and Tahmiscioglu, 1999; Zhang and Small, 2006), I focus on the varying coefficients of the lagged dependent variable and slowly changing variables.

In the simulations, I fix the number of units and the number of time periods as $J = 10$ and $T = 30$, respectively, and specify the following data generation process (DGP):

$$
\begin{aligned}
y_{jt} &= \mu + \phi_j y_{j(t-1)} + \beta_j z_{jt} + \delta_j + \varepsilon_{jt},\\
y_{j0} &\sim N(\frac{\mu+\delta_j}{1-\phi_j}, \frac{1}{1-\phi_j^2}),\\
\mu &= 3,\\
\phi_j &\sim U(-1,1),\\
\beta_j &\sim N(1, 0.5^2),\\
\delta_j &\sim N(0,1),\\
\varepsilon_{jt} &\sim N(0,1),
\end{aligned}
$$

where y_{j0} is assumed to be random and affect the equilibrium level (see Anderson and Hsiao, 1981, 1982). As can be seen in the DGP, I assume homoscedastic error terms and a lack of contemporaneous correlation, which makes OLS estimates consistent.

I design two basic experiments and the treatment is the number of changes in the explanatory

Table 1: RMSE of Coefficients ϕ and β

	Experiment 1 (Change to 1 once)				Experiment 2 (Change to 1 twice)			
	OLS		Bayesian		OLS		Bayesian	
Units	ϕ	β	ϕ	β	ϕ	β	ϕ	β
Unit 1	0.240	0.605	0.271	0.472	0.207	0.555	0.240	0.386
Unit 2	0.119	0.528	0.112	0.402	0.126	0.424	0.109	0.325
Unit 3	0.107	0.548	0.094	0.389	0.104	0.367	0.079	0.304
Unit 4	0.152	0.757	0.179	0.359	0.172	0.536	0.192	0.429
Unit 5	0.158	0.745	0.086	0.541	0.143	0.702	0.059	0.480
Unit 6	0.162	0.526	0.218	0.630	0.150	0.487	0.186	0.518
Unit 7	0.170	0.567	0.200	0.758	0.153	0.478	0.189	0.783
Unit 8	0.203	0.628	0.174	0.416	0.207	0.571	0.173	0.397
Unit 9	0.148	0.444	0.175	0.345	0.132	0.419	0.181	0.329
Unit 10	0.121	0.592	0.153	0.450	0.145	0.453	0.177	0.395
Average	0.158	0.594	0.166	0.487	0.154	0.498	0.158	0.435

variable. In the first experiment, I let the explanatory variable change from 0 to 1 at a certain time period and persist after that. A simple example is that the input changes from 0 to 1 at time period $t = 10$, which means that $z_{j10} = z_{j11} = \cdots = z_{j30} = 1$ and $z_{j1} = z_{j2} = \cdots = z_{j9} = 0$. In the second experiment, I let the explanatory variable change to 1 twice in some of the countries and change to 1 once in others. For example, the input z_{jt} changes from 0 to 1 at time period $t = 3$, changes from 1 to 0 at time period $t = 6$, changes from 0 to 1 at time period $t = 27$ again, and then persists. That is, $z_{jt} = 1$ for $t = 3, 4, 5, 27, 28, 29, 30$ and $z_{jt} = 0$ otherwise.[9] I performed 100 Monte Carlo trials for each experiment.

In the simulations, the quantities of interest are estimates of (ϕ_j, β_j) for each unit j. The dynamic model is estimated by two methods: OLS and Bayesian methods. Using OLS, these coefficients of each unit are estimated separately. The Bayesian model employed here is the one discussed in section 2.3 but without time-variant explanatory variables $\boldsymbol{X_{jt}}$ or year effects γ_t, which is estimated with MCMC techniques as implemented in JAGS 2.2.0 run under R (rjags). I

[9]The time at which the variable changes its value and the countries where the value of the variable changes twice are randomly assigned.

ran the Bayesian model with three chains of 40,000 iterations each. The first half of the iterations was discarded as a burn-in and 40 as thinning, which generated 1,500 samples in total.

I follow the literature in reporting root mean squared error (RMSE) of these two coefficients for each unit to compare these two methods. The RMSE captures both the bias and the efficiency of the estimators, which is calculated based on the formula $\sqrt{\frac{\sum_{s=1}^{100}(\beta_j^{(s)}-\beta_{\text{true}})^2}{100}}$, where s is the number of trials and β_{true} is the true value. The results, as presented in Table 1, show that the Bayesian approach performs as well, or better than the OLS estimator even though the data are generated based on a regular linear model.

4 Application: Social Spending in Latin America

In the comparative political economy literature, it is believed that government partisanship affects macroeconomic policies. Empirical studies on welfare states in advanced democratic countries have shown that the party composition of governments is related to government spending (e.g., Blais, Blake and Dion, 1993; Cameron, 1978; Hibbs Jr, 1977, 1992). More specifically, leftist governments are expected to spend more on social programs, including social security, education, and health, than rightist governments.

Although partisan theory is empirically supported in advanced democratic countries, it might not hold in other regions such as Eastern Europe (see Tavits and Letki, 2009) or Latin America (Huber, Mustillo and Stephens, 2008). One of the factors that differentiate between Latin American and advanced democratic countries in social expenditure is that, whereas most of the OECD countries have endurable democratic governments in the post-war era, many Latin American countries experienced regime change between democracy and authoritarianism (Avelino, Brown and Hunter, 2005; Huber et al., 2006; Huber, Mustillo and Stephens, 2008).

4.1 Previous Studies

Concerning the difference between Latin American and OECD countries, several studies have been conducted to investigate the determinants of social expenditure in Latin America. For example, Brown and Hunter (1999) found that democracies tend to spend more on overall social programs than authoritarian regimes.

Unlike previous studies which focused on the difference of social spending in two consecutive years, Huber, Mustillo and Stephens (2008) (hereafter HMS) investigated the determinants of social expenditure in the long term. They argued that the strength of democratic records is the significant determinant of social expenditure. In other words, a country that has a longer democratic regime tends to spend more. In their findings, the cumulative effect of democracy is positive on social spending. In contrast, repressive authoritarian regimes constrain social spending but only on education and health spending, compared to non-repressive authoritarian regimes.

Although HMS' argument is convincing, the method they employed might be problematic for several reasons. First, HMS investigated the long-term effects of democratic regime and government partisanship by cumulating the values of these variables from 1945 to the year of the observation. For repressive authoritarian regime, the yearly scores were cumulated over the five years prior to the year of the observation because the effects of authoritarianism would decrease over time. Cumulated yearly values are unable to capture dynamic effects of the covariates on the outcome.

Second, in order to obtain the properties of OLS estimation, HMS corrected for serial correlations by using feasible generalized least squares (FGLS) with a common autocorrelation coefficient, ϕ, for all countries (Kmenta, 1997; Parks, 1967). More specifically, they employed the Prais-Winsten transformation. However, ignoring the heterogeneity of the autocorrelation coefficient might lead to biased estimates (Pesaran and Smith, 1995). The generalized least squares (GLS) or FGLS corrections for serial correlations and contemporaneous correlations may also produce biased estimation of unit-specific serial correlation coefficient and biased estimation of standard errors if T is small (Beck and Katz, 1995, 1996). Moreover, the GLS-like estimators are problematic when the data have missingness or an unbalanced structure, which appears in the Latin American data set.

Third, HMS used PCSEs to correct for contemporaneous correlations and panel heteroskedasticity recommended by Beck and Katz (1995, 1996). The use of PCSEs corrects only for the standard errors of estimates but has nothing to do with model specifications.

Fourth, HMS did not include unit fixed effects in the statistical model because when the model is estimated by OLS these unit effects eliminate any variation in the outcome explained by time-invariant variables (Plümper, Troeger and Manow, 2005). As discussed before, excluding unit effects may be at the risk of having omitted variable bias.

Finally, since the democratic experience varies across Latin American countries, we would expect heterogeneous effects of political regime in Latin America, which are assumed constant in HMS.[10]

All of these problems can be solved by the Bayesian dynamic multilevel model. In the following, I apply the model presented in Equation (6) containing one lag of the explanatory variables to analyzing social spending in Latin America.[11]

4.2 Data and Measurements

The data set I used was collected by Huber et al. (2008).[12] Using this data set, HMS analyzed the long-term effects of political regime and government partisanship on social spending 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 political regime.

[10] In fact, Huber, Mustillo and Stephens (2008) emphasized the great variation in social policy and politics within Latin America. However, they ignored that the variation could result in varying effects of democracy on social spending.

[11] I started with a more general model and tested for different dynamic specification as recommended in previous studies (Beck and Katz, 2009; De Boef and Keele, 2008; Keele and Kelly, 2006; Wilson and Butler, 2007). Using Akaike information criterion (AIC) for MLE and deviance information criterion (DIC) for the Bayesian approach, the results showed that a model with more than one lag of explanatory variable is not supported.

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

Before I discuss the result, two points need to be illustrated. First, social security spending is measured as a percentage of GDP. 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 than the untransformed one. Moreover, although social security spending as a percentage of GDP is bounded between 0% and 100 %, a truncated normal distribution is not necessary because the mean and the standard deviation of the (untransformed) spending on social secutiry are very small (3.65 and 3.86, respectively) and the values of the outcome variable (log transformation of social security spending) lie between -3 and 3, which are not close to the boundary ($\log(0)=-\infty$ and $\log(100)=4.6$).[13]

The main explanatory variable is political regime, which are time-invariant or slowly changing, and coded as dummy variables: DEM, an indicator for democratic regime, and REPAUTH, an indicator for repressive authoritarianism, with non-repressive authoritarianism as the baseline category. The measure of political regime is based on the classification of regime types in HMS. I control for gross domestic production per capita (1,000 US dollars) adjusted for purchasing power parities (GDPPC), the percentage of the population that lives in the urban area (UBNPOP), the percentage of the aged population (POP65), export and import as a percentage of GDP (TRADE), foreign direct investment as a percentage of GDP (FDI), and IMF repurchase obligation (IMF=1 for each year if a country has repurchase obligations with the IMF and 0 otherwise).[14]

Following HMS, I include indicators for the debt crisis (1982-89) and for the recovery period

[13] A problem is that six observations (1981-1986 in Peru) have zero in the measurement of social security 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 and that the predictive values generated from the posterior distributions are very close to 0.

[14] In the data set collected by Huber et al. (2008), there are many missing values for the measurement of fiscal deficits in central governments. Therefore, I exclude this variable in my analysis. Moreover, I updated the data such as FDI and IMF repurchase obligation from World Development Indicators (WDI 2011) from the World Bank because there are many missing values in the data set. This updated data can be downloaded in `http://data.worldbank.org/data-catalog/world-development-indicators`. However, there were still some missing values in TRADE, and FDI. I impute these missing values by assigning these variables distributions in the Bayesian model (Gelman and Hill, 2007; Gill, 2008).

(1990-2000) to deal with the common shock for all countries. The baseline category is the period of 1970-81. These indicators are modeled as group-level covariates of time effects.

The Bayesian dynamic multilevel model presented above is estimated with MCMC techniques as implemented in JAGS 2.2.0 run under R (rjags). The model was run with three chains of 1,500,000 iterations each. The first half of the iterations were discarded as a burn-in period and 500 as thinning and thus 4,500 samples were generated. There is no evidence of non-convergence in these chains using standard diagnostic tools (e.g., Geweke, Gelman & Rubin, and graphical tools).

Table 2: Determinants of Social Security Spending

Variables	COEF	SE	95% CI	
Individual-level (country-year) predictors				
Intercept	−3.321	0.937	−4.807	−1.322
$\log(\text{LDV})_{t-1}$	0.700	0.120	0.448	0.865
DEM	0.092	0.100	−0.042	0.335
DEM_{t-1}	−0.067	0.050	−0.168	0.026
REPAU	0.032	0.087	−0.071	0.270
REPAU_{t-1}	−0.000	0.072	−0.145	0.152
GDPPC	0.000	0.000	−0.000	0.000
GDPPC_{t-1}	0.000	0.000	−0.000	0.000
UBNPOP	0.003	0.005	−0.001	0.014
UBNPOP_{t-1}	0.001	0.002	−0.001	0.006
POP65	0.013	0.085	−0.152	0.230
POP65_{t-1}	0.140	0.231	−0.220	0.561
TRADE	−0.001	0.001	−0.003	0.001
TRADE_{t-1}	0.000	0.001	−0.002	0.003
FDI	−0.015	0.010	−0.040	−0.001
FDI_{t-1}	−0.002	0.007	−0.018	0.010
IMF	−0.000	0.000	−0.000	0.000
IMF_{t-1}	0.000	0.000	−0.000	0.000
Group-level (year) predictors				
Debt crisis (1982-89)	−0.033	0.039	−0.108	0.048
Recovery (1990-2000)	0.061	0.056	−0.040	0.177

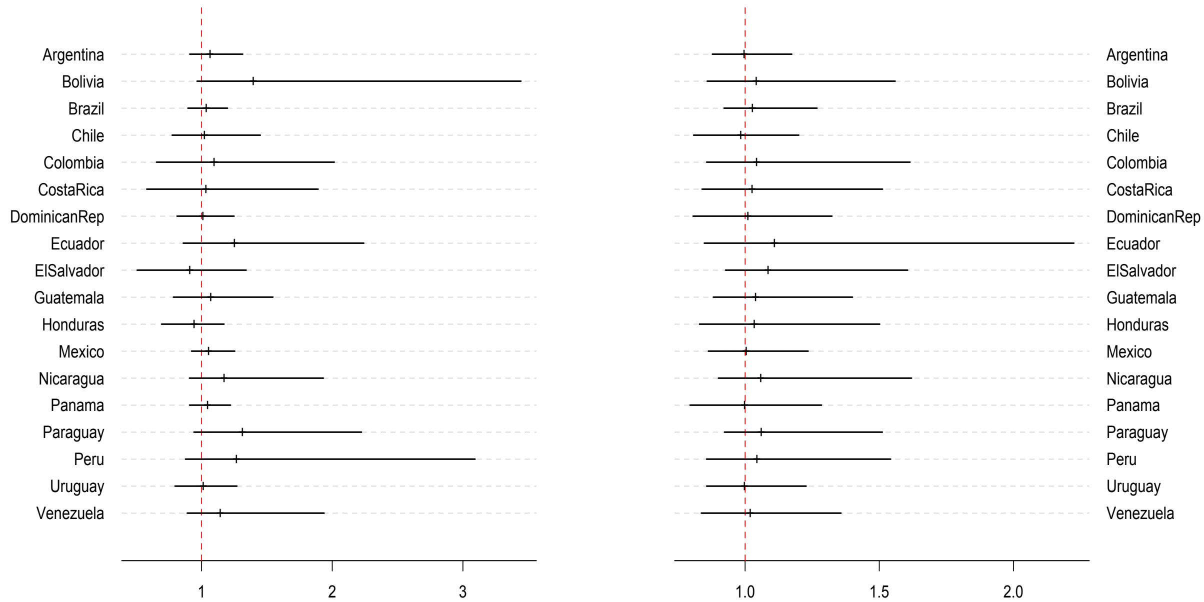
Note: estimates are not transformed by exponentiation.

4.3 Results of Analysis

The numerical results of the determinants of social spending are displayed in Table 2. First, concerning political regimes, in general, the result shows no evidence that democracy or repressive authoritarianism has immediate effects on social security spending, compared to non-repressive authoritarianism. Second, the mean immediate effect of FDI on social security and welfare spending is negative. The estimated effect of FDI is substantive ($e^{-0.015} = .0985$), which implies that a 1% increase of FDI leads to 1.5% decrease in security and welfare spending. The negative effect of FDI on social security spending is consistent with HMS's expectation although there is no evidence of the effect of FDI in their findings.

Next consider the heterogeneity of Latin American countries and I focus on the short- and long-run effects of political regime. First, with regard to the immediate effects of political regimes, the varying coefficients of political regimes are shown in Figure 2. As can be seen in the left panel, in most of these countries, democracy has an immediate and positive effect on social security spending although the effect is not significant at the conventional level. Contrary to the expectation, repressive authoritarianism does not immediately have a negative effect on social security spending in most of the countries, as presented in the right panel.

Second, to accurately capture the dynamics of political regimes, I consider countries that had experienced regime change between democracy and authoritarianism during the period of 1970 and 2000. According to the data set, 15 countries had experienced regime change from non-democracy to democracy. Based on Equation 4, Figure 3 shows the dynamic effects of democracy in these countries. As can be seen, the dynamics of democracy differs across countries and can be classified into four groups based on the trend towards the long term. In the first group (Bolivia, Ecuador, Nicaragua, Paraguay, and Peru), the short- and long-run effects of democracy on social security spending are positive. In the second group (Argentina, Chile, Dominican Republic, Mexico, Panama, and Uruguay), democracy has a positive short-run effect but a negative long-run effect. In the third group (El Salvator and Honduras), the short- and long-run effects of democracy are negative. In the fourth group (Brazil, and Guatemala), there is no large difference between short-



(a) Immediate Effect of Democracy (b) Immediate Effect of Repressive Authoritarianism

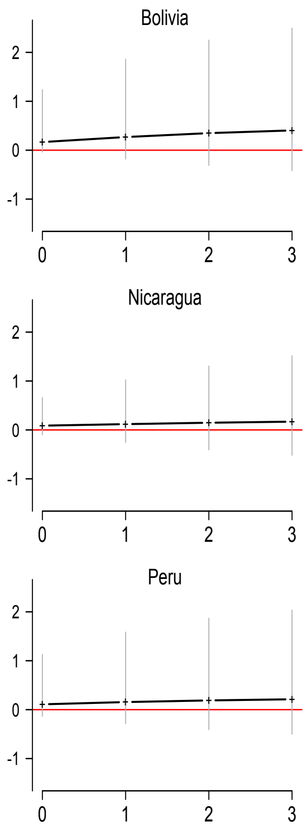
Figure 2: The 95% credible intervals of exponentiation of the immediate effects of political regime across countries. The left panel presents the immediate effect of democracy; the right panel shows the immediate effect of repressive authoritarianism. The red dotted lines represent the value of $e^0 = 1$, meaning no effect.

and long-run effects.

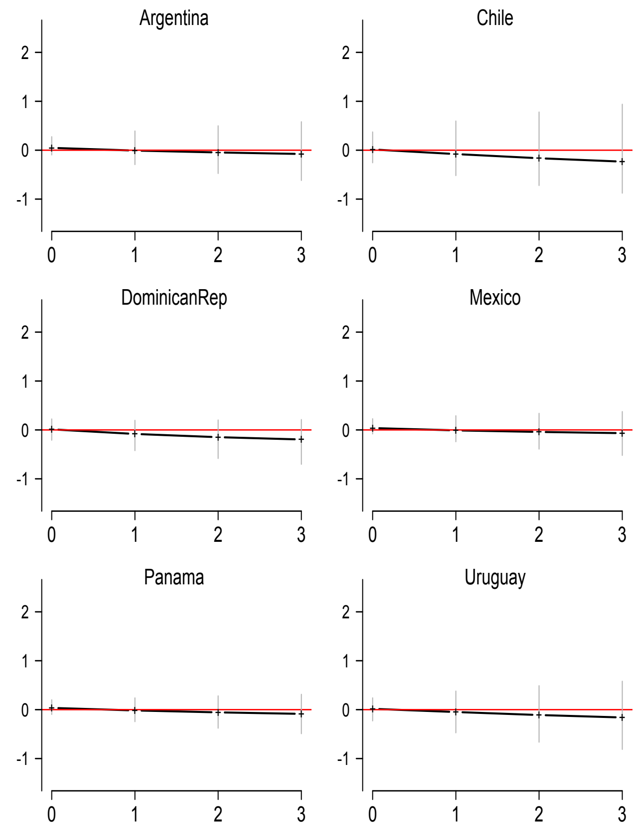
For countries that had experienced regime change from non-repressive authoritarianism to repressive authoritarianism, the dynamics of repressive authoritarianism are presented in Figure 4. It shows that repressive authoritarianism has a negative effect on social security spending in both the short and long terms although the effects are quite small and the uncertainty is large.

Three countries had been persistent democracies between 1970 and 2000: Colombia, Costa Rica, and Venezuela. Figure 5 presents the long-run effect of democracy in these three countries. As can be seen, there is no evidence of a positive effect of democracy on social spending in the long term in these three countries.

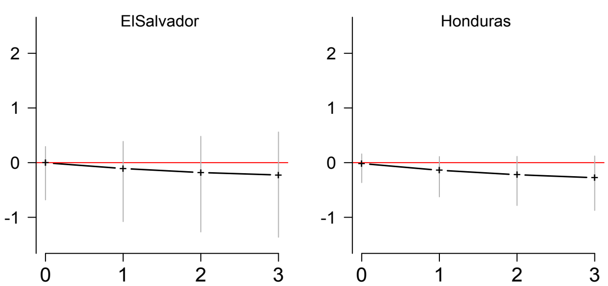
Finally, the left panel in Figure 6 presents the year effects on social security and welfare spending. It suggests that the spending on social security and welfare in the period 1984-1988 is substantially less than that in other time points. The result is consistent with the expectation of less spending on social security and welfare during the debt crisis period (1982-89) because gov-



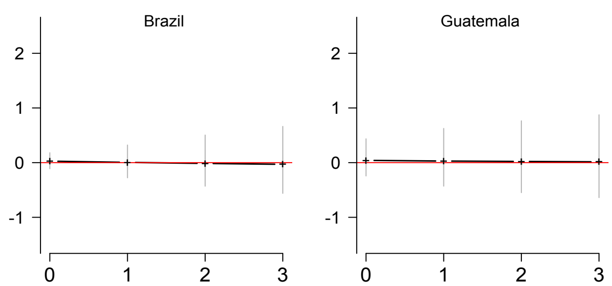
(a) Group 1



(b) Group 2



(c) Group 3



(d) Group 4

Figure 3: Dynamics of democracy in countries that had experienced regime change from non-democracy to democracy. The grey lines present 95% credible intervals and the red lines present 0 in the vertical axis. The horizontal axis presents the time periods. These fifteen countries are classified into four groups. In the first group, the short- and long-run effects of democracy on social security spending are positive; in the second group, democracy has a positive short-run effect but negative long-run effect; in the third group, the short- and long-run effect of democracy is negative; in the fourth group, there is no large difference between short- and long-run effects.

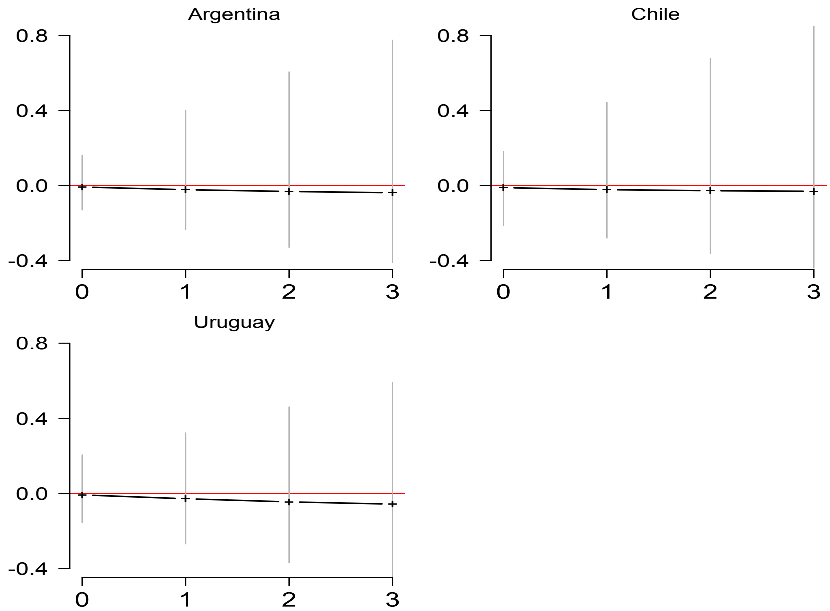


Figure 4: Dynamics of repressive authoritarianism in countries that had experienced regime change from non-repressive authoritarianism to repressive authoritarianism. The grey lines present 95% credible intervals and the red lines present 0 in the vertical axis. The horizontal axis presents the time periods.

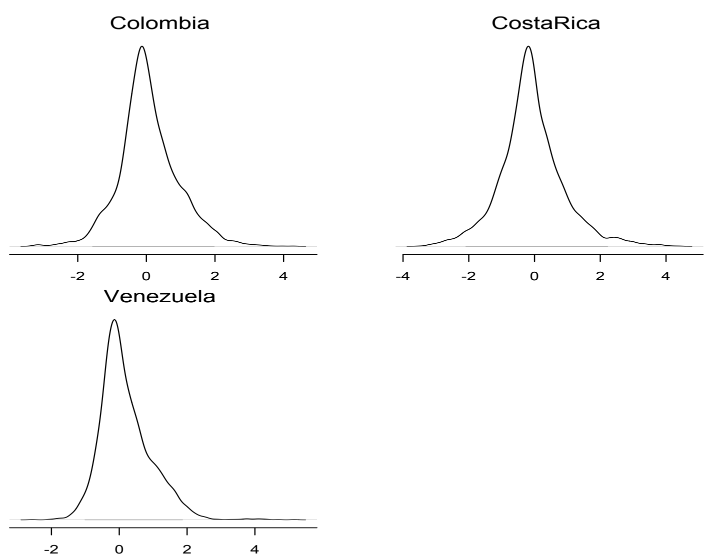
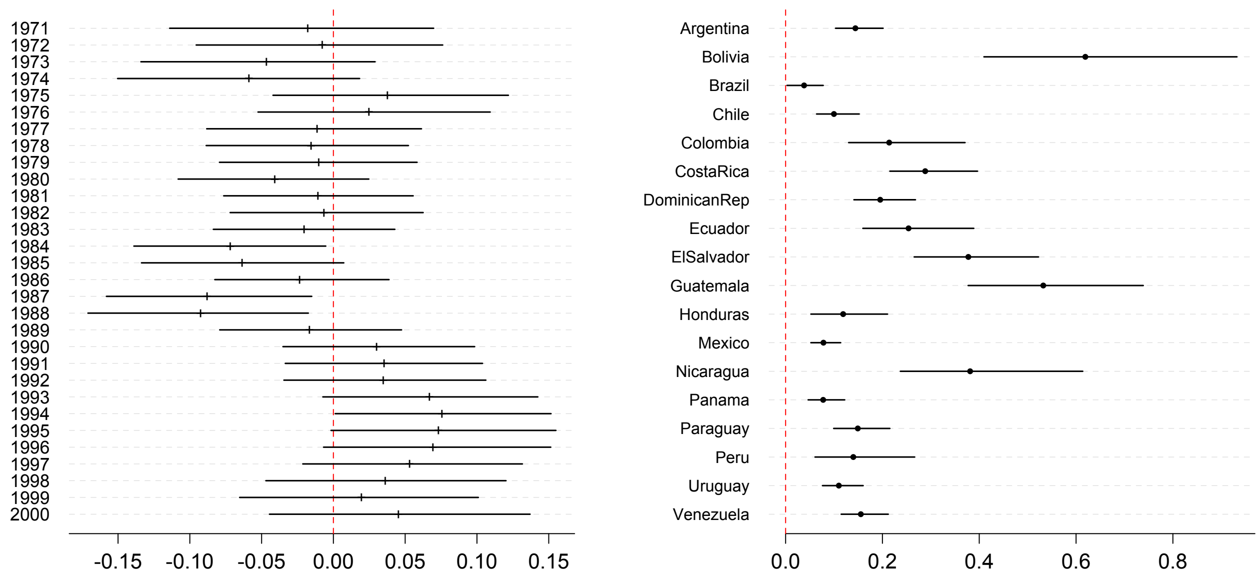


Figure 5: The long-run effect of democracy in countries that had experienced only democracy during the period of 1970 and 2000. The grey lines present 95% credible intervals.



(a) 95% Credible Intervals of Year Effects

(b) 95% Credible Intervals of SD of Security Spending

Figure 6: The 95% credible intervals of year effects and of the standard deviations of the outcome variable. The left panel presents the 30 year effects. The right panel presents the standard deviations of the outcome variable across countries.

ernments faced extreme constraints in finding resources for social security programs. Furthermore, it shows that spending on social security and welfare in the period 1993-1996 is more than that in other periods.

The right panel in Figure 6 provides the evidence of panel heteroscedasticity. It shows the standard deviations of security and welfare spending for each country and their uncertainty. As can be seen, Bolivia, Guatemala, El Salvador, and Nicaragua have larger variation than other countries. Most of the countries have standard deviations less than 0.3.

In sum, the immediate effect on social security spending of democracy is positive but not significant at the conventional level; the long-run effect on social security spending of democracy is not necessarily positive. Moreover, the immediate and the long-run effects of repressive authoritarianism on social security spending are not significant at the conventional level. However, these results may be driven by the fact that the Bayesian dynamic model reflects uncertainty more accurately (e.g., parameters and missing data). This uncertainty might be decreased when more information is included.

5 Discussion

The short- and long-run effects of slowly changing explanatory variables are often of interest in the studies of comparative political economy. However, the classical estimators are unable to estimate the effects of slowly changing explanatory variables in models containing unit effects. This paper argues that Bayesian multilevel models provide flexibility and advantages of the estimation of dynamic models in the analysis of TSCS data and shows that the Bayesian approach performs as well, or better than OLS estimators by conducting Monte Carlo simulations.

In the application of the Bayesian dynamic model to the social spending in Latin America, contrary to the findings in previous studies, this paper shows that, in general, political regime has no immediate or long-run effect on social security and welfare spending. This may be because the Bayesian dynamic model is more conservative than classical estimators in terms of estimating uncertainty. Moreover, we can observe the heterogeneous effects of democracy on social security and welfare spending, which is ignored in previous studies, although none of these reaches significant at the conventional level.

Considering the varying autoregressive coefficients opens up a potential avenue for future research in the study of dynamics of covariates. In this paper, I assume an identical autoregressive coefficient across all explanatory variables and the independence of the marginal effect of the rarely changing explanatory variables and the dates of the observations. Future studies could investigate the possibilities of relaxing these two assumptions.

A Appendix A: Model Diagnostics and Model Assessment

In this section, I do some diagnostics of the model specification. Generally, substantive theories provide little guide on which type of dynamic specification to employ when using TSCS. Thus, researchers have to consider different dynamic specifications and do some diagnostics (De Boef and Keele, 2008; Keele and Kelly, 2006; Wilson and Butler, 2007). I first test for autocorrelation of residuals. Second, I provide a residual plot to check for outliers. Third, Bayesian methods are criticized that the posterior distribution might be sensitive to the choice of prior distribution. To check the sensitivity to prior distribution, I change the prior distributions of variance terms. Finally, I assess the model quality by comparing the replicated data drawn from the posterior predictive distribution with the observed data.

Autocorrelation of Residuals

Figure 7 presents the plots of ACF of social security and welfare spending across countries. Some countries such as Argentina and Brazil have serial correlation while others such as Costa Rica and Ecuador do not. Therefore, assuming a common autocorrelation coefficient for all countries, as Huber, Mustillo and Stephens (2008) did, might not be appropriate.

Several approaches are available to test serial correlation of regression residuals, such as Durbin-Watson test or, more general, Breusch-Godfrey. However, there are missing values and the data are unbalanced, so the results of these tests might be misleading. To investigate the autocorrelation problem, I therefore present the autocorrelation function (ACF) plots of residuals, as shown in Figure 8.[15] As can be seen, Figure 8 shows that none of the residuals are serially correlated in these 18 Latin American countries. It suggests that the dynamic Bayesian multilevel model effectively provides an appropriate dynamic specification and, thus, eliminates the autocorrelation problem.

[15] Recall that some countries have missing values in the measure of social spending. For example, the autocorrelation coefficient at lag one in Chile is actually the autocorrelation between the current value and that of three lags.

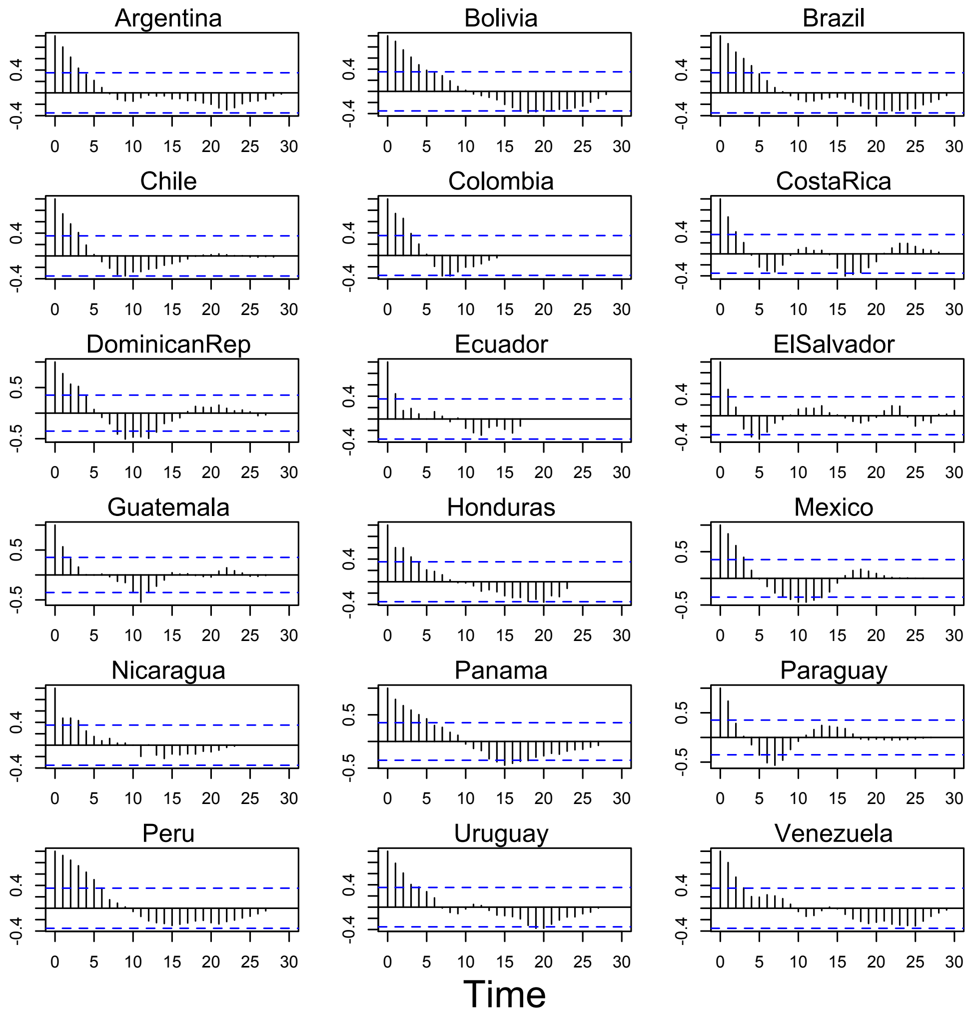


Figure 7: The ACF of social security/welfare spending across countries.

Standardized Residuals

Figure 9 presents the standardized residuals. The residuals are standardized because the variances within each country are not constant. The observations in 1983 Bolivia and 1988 Guatemala have large differences between the fitted values and the observed data. Generally, it shows that the standardized residuals are not far away from zero.

Sensitivity to Choice of Prior Distribution

Bayesian methods are criticized because the posterior distribution might be affected by the choice of prior distribution of parameters. To check the sensitivity to the choice of prior distribution, I

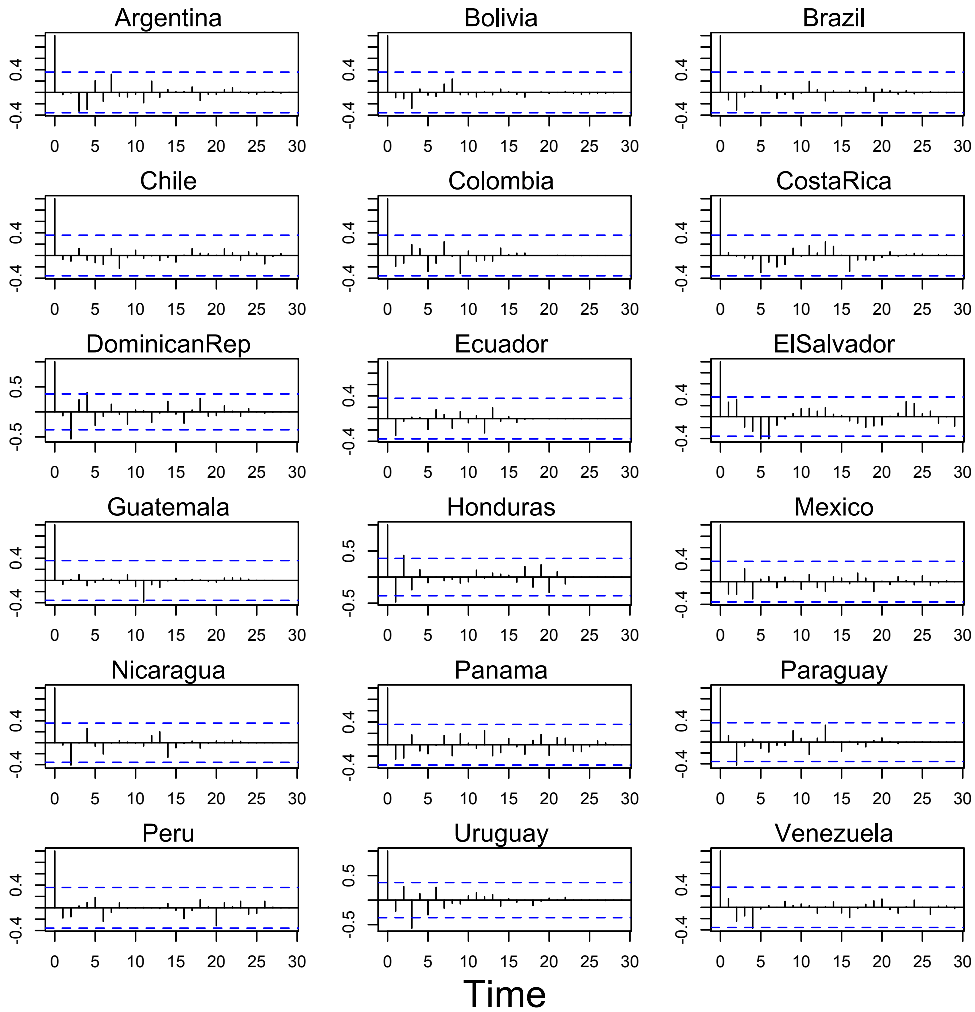


Figure 8: The ACF of residuals across countries.

change the distribution of variance from gamma distribution to uniform distribution. The uniform distribution is bounded between 0 and 100, which represents an uninformative prior. The results (which are not presented here) show that the posterior distribution of parameters are not affected by the choice of prior distribution.

Model Quality

To assess the model quality, I simulate replicated datasets from the fitted models and compare these replicated data to the observed data. These replicated datasets are drawn from the posterior predictive distribution. Figure 10 displays the observed data and the posterior distributions of

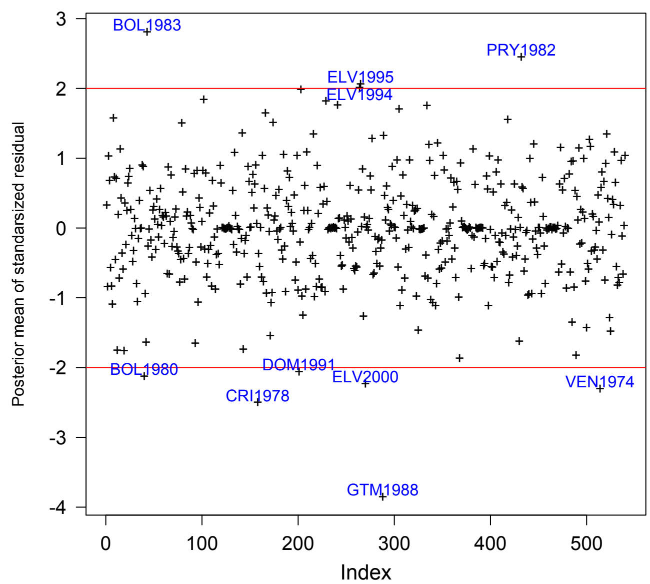


Figure 9: Scatterplot of standardized residuals.

the replicated data for each year within each country. In general, the model fits the data quite well. As can be seen in Figure 10, the replicated data are very similar to the observed data. In addition, these data capture the series of observed social security and welfare spending.

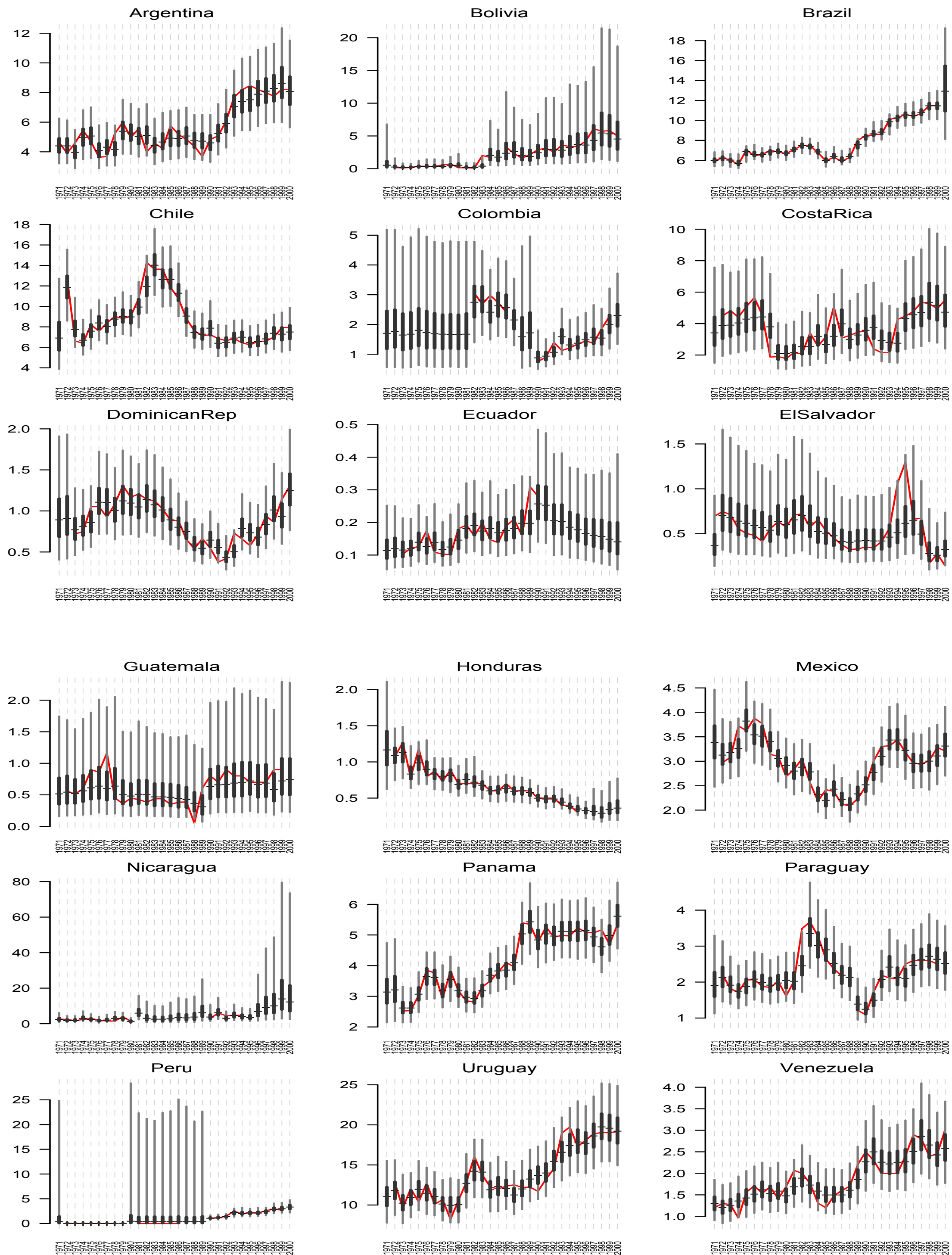


Figure 10: Replicated values and actual observations. The red lines represent the actual observations. The grey lines and black lines represent the 95% and 50% credible intervals, respectively.

B Appendix B: JAGS Code

```
model {
for (i in 1:N) {                                        # N OBSERVATIONS
y[i] ~ dnorm (yhat[i], tau.y[country[i]])
    yhat[i] <- beta0.raw + z.year[year[i]] + b[country[i]]*X.1[i,1] + inprod(B[country[i],2:M.1], X.1[i,2:M.1])

     # REPLICATED VALUES DRAWN FROM POSTERIOR PREDICTIVE DISTRIBUTION
        log.y.rep[i] ~ dnorm (yhat[i], tau.y[country[i]])
        y.rep[i] <- exp(log.y.rep[i])

        # RESIDUALS BY LOGARITHM TRANSFORMATION
        log.y.res[i] <- y[i] - yhat[i]
}

    # THE UNIT-SPECIFIC AUTOREGRESSIVE COEFFICIENT IS ASSUMED TO HAVE A LOGIT-NORMAL DISTRIBUTION SCALED TO HAVE SUPPORT ON (-1,1).
    for (j in 1:J) {
     b[j] <- 2*( exp(B[j,1] )/( 1+exp(B[j,1]) ) - .5)
    }

    # PRIORS FOR VARIANCES OF Y
    for (j in 1:J){
        tau.y[j] ~ dgamma (0.001, 0.001)
        sigma.y[j] <- 1/sqrt(tau.y[j])
     }

    # MULTIPLE VARYING COEFFICIENTS
    for (j in 1:J){                         # J UNITS
     for (k in 1:M.1){                   # M.1 PREDICTORS, INCLUDING INTERCEPT
     B[j,k] <- xi[k]*(B.raw[j,k] - mean(B.raw[,]))     # xi IS A SCALE PARAMETER.
     }
     B.raw[j,1:M.1] ~ dmnorm (mu.B.raw[], Tau.B.raw[,])    # THE UNSCALED COEFFICEINTS HAVE A MULTIVARIATE NORMAL DISTRIBUTION.
     }

    # UNSCALED MEAN PARAMETERS HAVE VAGUE HYPER PRIORS.
    for (k in 1:M.1){
     mu.B.raw[k] ~ dnorm(0, 0.0001)
     xi[k] ~ dunif (0, 100)
     mu.B[k] <- xi[k]*(mu.B.raw[k] - mean(mu.B.raw[]))     # SCALED MEAN PARAMETERS
     }

    Tau.B.raw[1:M.1,1:M.1] ~ dwish(W[,], df)   # UNSCALED PRECISION MATRIX HAS A WISHART DISTRIBUTION.
    df <- M.1+1
    Sigma.B.raw[1:M.1,1:M.1] <- inverse(Tau.B.raw[,])  # UNSCALED COVARIANCE MATRIX IS THE INVERSE OF THE PRECISION MATRIX

    # COMPUTE THE CORRELAITON BETWEEN COEFFICIENTS
    for (k in 1:M.1){
     for (k.prime in 1:M.1){
     rho.B[k,k.prime] <- Sigma.B.raw[k,k.prime]/sqrt(Sigma.B.raw[k,k]*Sigma.B.raw[k.prime,k.prime])
     }
     sigma.B[k] <- abs(xi[k])*sqrt(Sigma.B.raw[k,k])   # OBTAIN THE SCALED SD
     }

# REDUNDANT PARAMETER, GRAND MEAN
    beta0.raw ~ dnorm (0, 0.0001)

    # THE CONSTANT TERM
    constant <- beta0.raw + mean(B[,2]) + mean(z.year[])

    # THE GROUP-LEVEL FOR YEAR EFFECTS
    for (k in 1:K){
     beta.year[k] ~ dnorm (mu.year.raw[k], tau.year.raw)
     z.year[k] <- xi.year*(beta.year[k] - mean(beta.year[]))  # SCALED YEAR EFFECTS
     mu.year.raw[k] <- g0.raw + gamma.period.raw[2]*D.period2[k] + gamma.period.raw[3]*D.period3[k]
     }
    xi.year ~ dunif(0, 100)
```

```
#    tau.year.raw <- pow(sigma.year.raw, -2)
#    sigma.year.raw ~ dunif (0, 100)

    tau.year.raw ~ dgamma (0.001, 0.001)
    sigma.year.raw <- 1/sqrt(tau.year.raw)
    sigma.year <- xi.year*sigma.year.raw

    g0.raw ~ dnorm (0, 0.0001)

    # PRIORS FOR GROUP-LEVEL COEFFICIENTS
    gamma.period[1] <- 0
    gamma.period.raw[1] <- 0
    for (m in 2:3){
     gamma.period.raw[m] ~ dnorm (0, 0.0001)
     gamma.period[m] <- xi.year*gamma.period.raw[m]
     }

    # DUMMY VARIABLES FOR PERIODS
    for (k in 1:K){                 # K YEARS
     D.period2[k] <- equals(period[k], 2)  # CREATE A DUMMY VARIABLE FOR CRISIS PERIOD
     D.period3[k] <- equals(period[k], 3)  # CREATE A DUMMY VARIABLE FOR RECOVERY PERIOD
     }

    # THERE ARE MISSING VALUES IN EXPLANATORY VARIABLES (LDV, TRADE, AND FDI).
    for (i in 1:N) {
        X.1[i,1] ~ dnorm (mu.ldv[country[i]], tau.y[country[i]])
        X.1[i,13] ~ dnorm (mu.trade[country[i]], tau.trade[country[i]])
        X.1[i,14] ~ dnorm (mu.trade[country[i]], tau.trade[country[i]])
        X.1[i,15] ~ dnorm (mu.fdi[country[i]], tau.fdi[country[i]])
        X.1[i,16] ~ dnorm (mu.fdi[country[i]], tau.fdi[country[i]])
    }

    # PRIORS FOR MISSING VALUES IN EXPLANATORY VARIABLES
    for (j in 1:J){
        mu.ldv[j] ~ dnorm (0, 0.1)
        mu.trade[j] ~ dnorm (0, 0.1)
        mu.fdi[j] ~ dnorm (0, 0.1)

        tau.trade[j] ~ dgamma (0.01, 0.01)
        sigma.trade[j] <- 1/sqrt(tau.trade[j])
        tau.fdi[j] ~ dgamma (0.1, 0.1)
        sigma.fdi[j] <- 1/sqrt(tau.fdi[j])
     }
}
```

References

Achen, C.H. 2000. Why lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables. In *Annual Meeting of the Political Methodology Section of the American Political Science Association, UCLA*. pp. 20–22.

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

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

Avelino, G., D.S. Brown and W. 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.

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. 2007. "Random Coefficient Models for Time-Series-Cross-Section Data: Monte Carlo Experiments." *Political Analysis* 15(2):182.

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.

Blais, A., D. Blake and S. Dion. 1993. "Do Parties Make a Difference? Parties and the Size of Government in Liberal Democracies." *American Journal of Political Science* 37(1):40–62.

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

Cameron, D.R. 1978. "The Expansion of the Public Economy: A Comparative Analysis." *The American Political Science Review* 72(4):1243–1261.

Casella, G. and E.I. George. 1992. "Explaining the Gibbs Sampler." *American Statistician* 46(3):167–174.

Chib, S. and E. Greenberg. 1995. "Understanding the Metropolis-Hastings Algorithm." *American Statistician* 49(4):327–335.

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

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

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

Gelman, A., J. Carlin, H. Stern and D. Rubin. 2004. *Bayesian Data Analysis*. 2 ed. Chapman & Hall/CRC.

Gelman, A. and J. Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

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

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

Hibbs Jr, Douglas A. 1977. "Political Parties and Macroeconomic Policy." *The American Political Science Review* 71(4):1467–1487.

Hibbs Jr, Douglas A. 1992. "Partisan Theory after Fifteen Years." *European Journal of Political Economy* 8(3):361–373.

Hsiao, C. and M.H. Pesaran. 2004. "Random Coefficient Panel Data Models." *The Institute for the Study of Labor Discussion Paper Series* (1236).

Hsiao, C., M.H. Pesaran and A.K. Tahmiscioglu. 1999. Bayes Estimation of Short-Run Coefficients in Dynamic Panel Data Models. In *Analysis of Panels and Limited Dependent Variables*, ed. C. Hsiao, M.H. Pesaran, Kajal Lahiri and Lung Fei Lee. Cambridge U Press chapter 11, pp. 268–296.

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

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

Huber, E., J.D. Stephens, T. Mustillo and J. Pribble. 2008. *Social Policy in Latin America and the Caribbean Dataset, 1960-2006*. University of North Carolina.

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

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

Kmenta, Jan. 1997. *Elements of Econometrics*. University of Michigan Press.

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

Nerlove, Marc. 1971. "Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections." *Econometrica: Journal of the Econometric Society* 39(2):359–382.

Parks, R.W. 1967. "Efficient Estimation of A System of Regression Equations When Disturbances Are both Serially and Contemporaneously Correlated." *Journal of the American Statistical Association* 62(318):500–509.

Pesaran, M.H. and R. Smith. 1995. "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels." *Journal of Econometrics* 68(1):79–113.

Plummer, M. 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.

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

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

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

Shor, Boris, David Park, Joseph Bafumi and Andrew Gelman. 2003. Examining Time Series Cross Sectional Data with Bayesian Multilevel Models. 2003 APSA Annual National Conference (Philadelphia).

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

Tavits, Margit and Natalia Letki. 2009. "When Left Is Right: Party Ideology and Policy in Post-Communist Europe." *American Political Science Review* 103(4):555–569.

Western, Bruce. 1998. "Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach." *American Journal of Political Science* 42(4):1233–1259.

Wilson, S.E. and D.M. Butler. 2007. "A Lot More to Do: The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications." *Political Analysis* 15(2):101.

Zhang, Peng and Dylan Small. 2006. "Bayesian Inference for Random Coefficient Dynamic Panel Data Models." Mimeo.