Title
A Two-Part Bayesian Model with Elicited Priors to Analyze Longitudinal Government Expenditures in Latin America

Author
Santos, Felipe Nunes dos

Publication Date
2013

Peer reviewed|Thesis/dissertation

University of California
Los Angeles

A Two-Part Bayesian Model with Elicited Priors to Analyze Longitudinal Government Expenditures in Latin America

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Science in Statistics

by

Felipe Nunes dos Santos

2013

Abstract of the Thesis

A Two-Part Bayesian Model with Elicited Priors to Analyze Longitudinal Government Expenditures in Latin America

by

Felipe Nunes dos Santos

Master of Science in Statistics

University of California, Los Angeles, 2013

Professor Mark Handcock, Chair

Longitudinal government expenditure data are often used in political science to assess politicians' agenda, presidents' distributive strategies, and coalition bargaining. Typically, this type of data are semicontinuous; contains some temporal dependence; and present a positive correlation between the probability of expenditure disbursed and the expected level of expenditure by each unit. In this thesis I use a Bayesian two-part model to separately estimate the likelihood of municipalities receiving national investments over time and mean expenditure trajectories among those chosen to receive national investments. I use elicited priors from fieldwork interviews with national and local politicians to estimate more precise posterior distributions. This approach extends to government expenditure data desegregated by states or municipalities. Applications to data from Brazil and Mexico uncovers two temporal patterns in government expenditures: funds are targeted to municipalities governed by co-partisan mayors, regardless of the president's vote share. Regions lacking either supportive voters or local allies receive little or no distributive funds whatsoever.

The thesis of Felipe Nunes dos Santos is approved.

Hongquan Xu

Nicolas Christou

Mark Handcock, Committee Chair

University of California, Los Angeles

2013

To Jorge Alexandre Neves, my first statistics mentor.

Table of Contents

List of Figures

List of Tables

Acknowledgments

Thank you to my adviser Professor Mark Handcock for helping me reach this point in my graduate career. Mark has always had his door open to discuss my ideas and suggest improvements. More than a thesis advisor, Mark has been a great academic reference. I will never forget his pushes to make me think from first principles, and to imagine the data generating mechanism, even before thinking about any technique. The writing of this thesis was not a straightforward process, and he pointed me in the right direction when I started to feel lost, along with a healthy dose of perspective.

I would also like to thank the present members on my committee - Professors Nicolas Christou and Hongquan Xu - who provided me feedback and support, which helped me raise the quality of my work. Nicolas was my first stats professor at UCLA. His organization, motivation, and clearness serve as great incentives to pursue the master in statistics. Hongquan was my instructor in the experimental design class in the master program. He challenged me a lot to think harder about the social science designs when presenting the engineering models. He definitely inspired me to think harder about my own research.

Thesis are not only completed with the academic support. They are constructed in the everyday of the university. I thank Glenda Jones for her support and encouragement. She has always received me with a big smile and nice words in her office. Without her this M.S. would not be completed. I thank Professor Michael Thies, Kathleen Bawn, and Jeff Lewis from the political science Department for encouraging and supporting my plan of getting a M.S. in Statistics while doing a PhD in Political Science.

During this master I often ran into challenges, but I had friends and family to help me stay on course. I almost certainly would not have finished without them as a support system. Thank you to my classmates who were there to talk and to correct my grammar mistakes. Thank you to my colleague Antonio Pedro Ramos who was the first to motivate me to pursue this master. I'm glad I listened to him. Thank you to my perfect partner Paasha Mahdavi. Your friendship and your intellectual skills were essential components to inspire me to do better. Thank you to my parents who always instilled confidence and motivation and throughout my life, have encouraged me to pursue what makes me happy. Finally, thank you to Gabi, who is always there to listen and help, and for giving me more than I ever hoped for.

Finally, I want to express my gratitude to the full-term fellowship provided by the Brazilian government through Fulbright and CAPES. The financial support they provided allowed me to go further. I really hope that policies like that keep existing to proportionate opportunities otherwise missing in the life of many.

CHAPTER 1

Introduction

There is an important debate in the political economy literature about how politicians strategically allocate government goods and services to geographic localities to ensure electoral success (Cox, 2006; Dixit and Londregan, 1996). There are findings showing core voters as a strong predictor of municipal discretionary transfers (Gervasoni, 1998; Golden, 2003; Larcinese, Rizzo and Testa, 2006; Magaloni, Diaz-Cayeros and Estevez, 2007; Nichter, 2008; Stokes, 2005; Zucco, 2008); or that presidents have strong incentives to allocate public goods to obtain legislative support (Bonvecchi, 2009; Gibson and Calvo, 2000; Giraudy, 2007; Jones and Hwang, 2005; Sørensen, 1995). There are also research suggesting that transfers are used to benefit places with obvious infrastructure deficits, where the poverty rate is high (Litschig, 2008; Lodola, 2005; Luna, 2010; Luna and Mardones, 2010; Rovallion, 1998); or that transfers are politically manipulated to target president's local allies (Ames, 1994; Armesto, 2009; Arulampalam et al., 2009; Brollo and Nannicini, 2011; Fachelli and Ronconi, 2004; Levitsky, 2003; Nazareno and Stokes, 2006; Stokes, 2005).

Because most of the literature concentrates on the average effect of the aforementioned factors on the average amounts of pork, transfers, or public goods distributed by politicians, little is known about the set of strategies presidents pursue in allocating resources to localities to their greatest political advantage. The results in the literature are still not conclusive, and most of the time contradictory. To address this general important question it is necessary to take into account not only the average expenditures among those chosen to receive national investments, but also the likelihood of localities receiving national investments. Further, it is key to consider the over time trajectories of distribution and the political changes that it produces along the way. In this thesis I use a Bayesian model with elicited priors that allows the proper estimation of the likelihood that localities receive national government investments over time, and the mean national government expenditure trajectories among those chosen to receive national investments.

The model presented here is appropriate for analyzing how presidents distribute discretionary resources because it deals with three key features of longitudinal government expenditure data. First, the data are semicontinuous, with non-negative support and a spike at zero given the large number of localities that did not receive government investments whatsoever. Another important feature of the data concerns repeated measurements. In longitudinal government data, each municipality contributes an observation for each of the years for which we have data for, introducing within-subject correlation. Moreover, in each year, there are two outcomes per municipality that needed to be taken into account: municipality's access to national government investments, and, if so, the level of access measured by expenditures. Further, it may be reasonable to assume that the probability of access to investments is correlated with the expected level of expenditures.

One modeling strategy to deals with semicontinuous data that present correlation over time is to apply a longitudinal two-part model (Ghosh and Albert, 2009; Neelon, O'Malley and Normand, 2011; Olsen and Schafer, 2001; Tooze, Grunwald and Jones, 2002). Two-part models are mixtures of a point mass at zero followed by a right-skewed distribution for the nonzero values. The two mixture components are modeled in stages. First, the probability of disbursement of investments is modeled via mixed effects probit or logistic regression. Next, conditional on some investment being made, the expected expenditure level is modeled through a lognormal mixed effects model. The random effects for the two components are typically assumed to be correlated, to avoid potential biased inferences (Su, Tom and Farewell, 2009).

I build on these previous literature to suggest the application of a Bayesian two-part model in order to characterize the effect of electoral cycles and political alliances over the chances and the conditional expectation of presidential expenditures in two Latin American countries: Brazil and Mexico. Within each class of municipalities which receives the disbursements, I fit a probit-lognormal model with municipality random intercepts and year random slopes. An attractive feature of the model is that it permits the random effect covariance to vary across the two processes. Such specification allows, for example, a process of allocation that comprises municipalities with frequent high expenditures (positive correlation between the probability of receiving appropriations and the actual amount spent), and another class in which municipalities have frequent but modest investments (negative correlation between probability of receiving investments and amount spent).

In addition, I incorporate all information available to deal with my research question. The Bayesian approach employed here also gives me the opportunity to incorporate informative priors from in-depth interviews conducted with political elites in Brazil and Mexico into my quantitative analysis. As advocated by Gill and Walker (2005), the elicitation of priors has the potential to tie together 'the seemingly antithetical research approaches of qualitative area studies with data-oriented work based on statistical methods, perhaps then mending a recent rift in political science.' To combine fieldwork information collected through interviews (qualitative) with fiscal records (quantitative), as well as to ensure a well-identified model with proper posteriors, I pursued a systematic procedure to recover elicited priors for the quantities in which I am most interested. The Bayesian literature claims that the elicitation of prior beliefs, avoidance of asymptotic approximations, and practical estimation of parameter contrasts and multidimensional credible regions are, among others, main advantages for 'going Bayesian'. However claimed, elicitation of prior beliefs is not commonly practiced in political science. As I argue below, analyzing data revealing political strategies of allocation offer a critical opportunity to employ a Bayesian estimation incorporating politicians' beliefs about how they do things.

The remainder of this thesis is organized as follows: Section 2 reviews the literature on how to model longitudinal expenditure data, and details the reasons for why the solution employed here is appropriate for it; Section 3 presents the two-part Bayesian model that will be employed in this study; Section 4 discuss the results of a simulation study employed to evaluate the performance of the model proposed. Section 5 applies the aforementioned model to data in Brazil and in Mexico. The conclusion section discusses the uses and applications of this model and the substantive implications from the analysis pursued here.

CHAPTER 2

Estimation Challenges: How to Analyze Longitudinal Expenditure Data?

Government longitudinal expenditure data can be generally characterized as a semicontinuous random variable that combines a continuous distribution with point masses at one or more locations. Particularly, expenditures tend to be distributed as a mixture of 0's and continuously distributed positive values. In this thesis I demonstrate the usefulness of applying a Bayesian two-part model to estimate parameters from a regression when the researcher is motivated by the interest to recover the profile of strategies actors have when deciding on how to allocate such resources. Table 2.1 provides a description of the total spending data for a sample of municipalities in Brazil and Mexico from 2003 to 2010.

The case of Brazil is severe. Over 70% of municipalities had no annual discretionary federal investment whatsoever, while a small fraction had large expenditures, specially in 2008. The percentage of municipalities receiving national investments is roughly constant over time, while mean spending increased considerably in the last three years of Lula's government. In Mexico, roughly 20% of municipalities do not receive discretionary investments from the president, and the majority of localities have been receiving more investments per capita over time.

Table 2.1: Descriptive Statistics for Government Spending, Brazil and Mexico

	Brazil		Mexico	
Year	% Non-receivers	Mean (Real cap)	% Non-receivers	Mean (Pesos cap)
2003	0.75	3.86	0.10	683.79
2004	0.66	6.13	0.16	618.74
2005	0.68	9.59	0.20	612.66
2006	0.77	4.84	0.25	698.51
2007	0.74	10.13	0.29	820.98
2008	0.72	11.41	0.15	1159.23
2009	0.76	9.81	0.16	1227.41
2010	0.86	7.01	0.22	1174.97

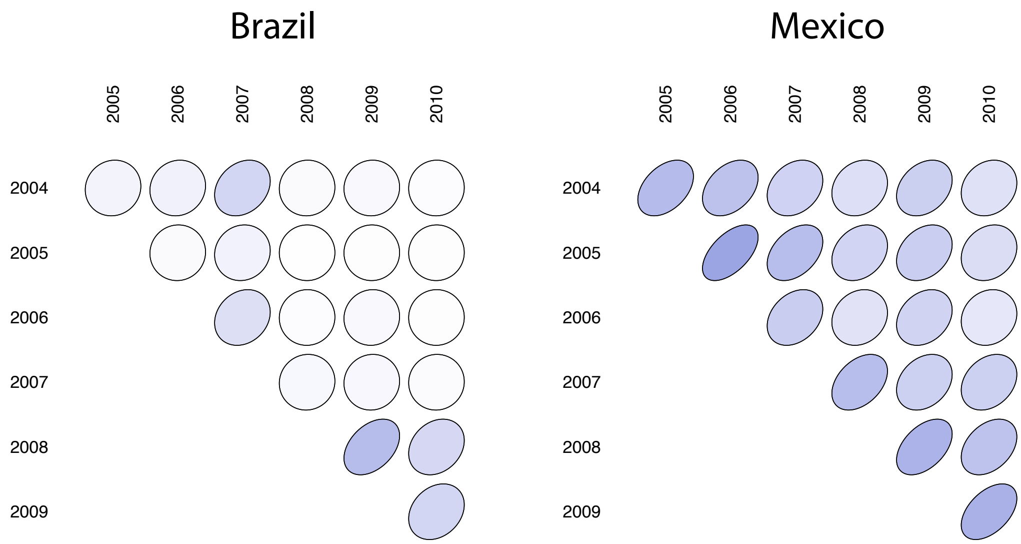
Source: Secretaria do Tesouro Nacional (STN), Brazil

Methods for estimating the moments of positive random variables with discrete probability mass at the origin were investigated by Aitchison (1955). Two-part regression models for variables of this type have appeared in econometric analyses for nearly three decades in Biostatistics and in Medicine, but they are still scarce in political science. Duan et al. (1983) described expenditures for medical care by a pair of regression equations, one for the logit probability of expenditure and one for the conditional mean of log expenditure (given that one exists). This application could have been transported to the social sciences very easily, because this type of question - how much and how frequent expenditures happen - have interested many researchers in political economy and in distributive politics. Another advantage of this type of model is that in cross-sectional applications, the two models may be fit separately with standard logit and linear regression. Although this is less efficient and does not allow the two components to vary depending on their conditional values, it is still better than not address this question at all.

Sample selection models, including the Tobit model (Amemiya, 1984; Tobin, 1958) and Heckman's selection model (Heckman, 1976), are often applied to limited or censored dependent variables. These posit an underlying normal random variable that is censored by a random mechanism; the mean of the underlying variable and the probability of censoring are jointly modeled as functions of covariates. As noted by Duan et al. (1983), two-part models are easier to interpret than selection models when 0's represent actual data, because the meaning of the underlying normal variable becomes dubious when 0 is a valid response rather than a proxy for a negative or missing value. In contrast, the second part of a two-part regression describes the conditional mean of the response given that it is non-0, a quantity that is highly meaningful. The choice of the two-part model in this thesis is highly influenced by this argument. If one quantity of interest is what strategy politicians have when allocating resources, then, models of the data should reflect that question completely. Allowing the two-parts to be estimated jointly, the model advocated here contributes to the estimation of the proper quantities of interest of many studies in the social sciences, but particularly in comparative politics.

In the typical longitudinal expenditure data sets, each unit who receives expenditures contributes an observation for each of the years in the data set, introducing within-subject correlation. Figure 2.1 presents the correlation matrix for government expenditures over time for the same samples of municipalities in Brazil and Mexico. Each ellipse represents the correlation of two years of the dependent variable. Narrower and darker the color of ellipses, bigger the one-year

Figure 2.1: Correlation Matrix for Government Expenditures Over Time



temporal dependency between the data. The ellipses also points the direction of the correlation, either positive or negative. Overall, the serial correlation is bigger in the Mexican than it is in the Brazilian case (note the shade of the circles). However, the correlation is strong among consecutive years in Brazil. Note how the ellipses get more clear in the years of 2006-2007, 2008-2009, and 2009-2010. The correlations vary from .01 to .40 in Brazil, whereas in Mexico they range from .14 to .56. This behavior of the data suggests we need to be sensitive to concerns about temporal dependency. An appropriate statistical model should address these multiple sources of correlation.

Ignoring that observations are correlated over time may lead to: exaggerated goodness of fit measures, estimated standard errors that tend to be smaller than the 'true' standard errors, and possible parameter inconsistency. The literature suggests many approaches for dealing with this. One approach is to assume that each unit has a fixed but unknown intercepts that can be estimated by adding a dummy variable for each unit — the 'fixed effects' (FE) model (Angrist and Pischke, 2009). However, the fixed effects remove any of the average unit-to-unit variation from the analysis, and simply ask whether intra-unit changes in y are associated with intra-unit changes in x (Beck and Katz, 2011). Moreover, using fixed effects also makes it difficult to discern the impact of variables that change only slowly (Blackwell, 2012; Greene, 2003).

Two other approaches are more flexible and have been presented good results in the literature. One can model the error structure of the regression model directly, instead of correcting errors. Doing this, the researcher is able to exclude the temporal correlation of the dependent variable from the error structure of the regression, making the model assumptions appropriate. Another approach is to assume that each unit has random effects that are part of a distribution family, and that these random variables could be further modeled properly. This second approach — also known as the 'mixed model' or 'random effects model' — adds a hierarchical structure to the random effects and models them accordingly. I use a mixed model with municipal random intercepts and year random slopes that not only deals with the problems mentioned above, but that also allows each municipality to have its own over time expenditure trajectory. A major complaint lodged against random effect models relates to the restrictive assumption that independent variables be orthogonal with the random effects term. Since a variable varies both within and between municipalities, many argue that this is an unrealistic assumption, because unobserved heterogeneity will almost always be correlated with the independent variables. I acknowledge this possible limitation in the approach used here, but since the results supporting such claim are still weak and inconclusive, I do not find reason to not pursue this analytical strategy (Weiss, 2005).

CHAPTER 3

Bayesian Two-Part Model with Elicited Priors

As explained before, spending data takes nonnegative values with a substantial proportion of zeroes. A standard way to deal with these kinds of data — often called censored — is to use a Tobit model (Tobin, 1958). In its simplest version, the Tobit model assumes there exist a latent variable that is observed whenever it is positive, but it is censored at zero when it is negative. The process that I am modeling here, by contrast, has a different nature that a Tobit model cannot capture. I want to model a two-part process of presidential spending that is composed of two correlated moments: (1) the decision of whether or not to disburse some portion of resources to a municipality, followed by (2) a decision of how much to allocate in case the first decision was positive.

This type of process is not censored in a typical fashion, but it is characterized as semicontinuous. That is, there may exist a gap between the zero values (reflecting a decision not to spend) and the positive values of spending (reflecting the level of discretionary spending). In contrast with what the Tobit model assumes, in my data a zero value is the result of a presidential choice not to allocate any resources to a certain municipality. My data are positive, by contrast, because after deciding on whether or not to allocate, the president decides *how much* to allocate. One solution for when the data contain zeroes and positive values coming from different decision processes is to use a two-part model. The first stage models whether the response is positive or zero. Conditional on the first stage being positive, the second stage models its level.

3.1 Two-part model

The two-part model for semicontinuous data is a mixture of a degenerate distribution at zero and a positive continuous distribution, such as a lognormal (LN), for the nonzero values. The model can be viewed as a two-part mixture model as the structure and the values of the error term of the random effects may be correlated across allocation processes. The model fitted here can be represented as follows:

$$f(y_{it}|\phi_{it}, \mu_{it}, \tau^2) = [(1-\phi_{it})^{1-d_{it}} \phi_{it}^{d_{it}}] \times [LN(y_{it}; \mu_{it}, \tau^2)]^{d_{it}} \tag{3.1}$$

$$\text{Probit}(\phi_{it}) = \mathbf{x}'_{it}\boldsymbol{\alpha} + \mathbf{z}'_{it}\boldsymbol{\gamma} + \mathbf{b}_{1i}$$

$$\mu_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{z}'_{it}\boldsymbol{\eta} + \mathbf{b}_{2i} + (\text{Year})\mathbf{b}_{3i}$$

$$\mathbf{b}_i \sim N_3(0, D)$$

$$D_{3x3} = \begin{bmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{bmatrix}$$

where

- y_{it} is the t^{th} spending allocation for municipality i (t $= \{1, ..., n_i\}$)
- d_{it} is an indicator that $y_{it} > 0$, and $\phi_{it} = Pr(Y_{it} > 0)$
- $\mathbf{x}'_{it}$ and $\mathbf{z}'_{it}$ are vectors of fixed effect covariates
- $LN(.)$ is lognormal density evaluated at y_{it}, with mean μ_{it} and precision τ^2

- $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\eta}$, and $\boldsymbol{\gamma}$ are fixed effect coefficients.

- $\mathbf{b}_i$ is a stacked vector of random effects for municipality i, with covariance D. I allow that $\mathbf{b}_{1i}$ and $\mathbf{b}_{2i}$ are correlated to capture possible dependence between the two parts of the decision process. D contains up to nine parameters to capture this dependency.

When $d_{it} = 0$, the only contribution to the likelihood is the first term in equation 3.1 — the probit probability that no allocation was made to municipality i at time t. When $d_{it} = 1$, the second term is activated and the contribution to the likelihood includes both the probit probability that spending occurs and also the value of the log-normal density at the level of observed spending. The model that I present here can be viewed as a three-level hierarchical model where the components or stages are modeled stochastically. It is a two-part model for semicontinuous data with random intercept for municipalities and random slopes for time in which the variance-covariance matrix of the random effects includes cross-covariances between the random effects of the two components. As inspection of the likelihood reveals, if the correlations were zero, the likelihood would be separable and the probit and the log-normal parts of the two-part model could be estimated separately. However, because the same unobserved factors are likely to affect each of the two parts of the allocative choice, I do not assume these correlations are zero and estimate the two-parts of the model jointly (Neelon, O'Malley and Normand, 2011).

3.2 Computation

A two-part mixed model with correlated random effects, like the one presented here, is an attractive approach to characterize the complex structure of longitudinal semicontinuous data. Since we have a full specified likelihood, it would be natural to compute maximum likelihood estimators for the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. It would also be natural to compute asymptotic approximations of standard errors exploring the Hessian matrix. However, there is no guarantee that the likelihood will be log-concave, and hence, finding the global maxima tends to be difficult (Geyer and Thompson, 1992).[1] Given the computational challenges to fit these complicated models to longitudinal data using maximum likelihood, and to avoid problems with algorithm convergence, I estimate the model using Bayesian Markov Chain Monte Carlo (MCMC). By sampling rather than optimizing, MCMC makes estimation and inference simpler (Jackman, 2000). First, the Bayesian methods are well suited to the large number of parameters to be estimated in each part of the model and the hierarchical structure that the data presents (Browne and Draper, 2006).

The posteriors are estimated using JAGS calling *rjags* 3-10 from *R* 2.15 (code available on the Appendix). I ran 5 MCMC chains for 200,000 iterations each, discarding the first 50,000 as a burn-in to ensure that a steady-state distribution had been reached, and retained every 50th draw to reduce autocorrelation. Convergence is monitored by running multiple chains from dispersed initial values and then applying standard Bayesian diagnostics, such as trace plots; autocorrelation statistics; Geweke (1991) Z-diagnostic, which evaluates the mean and variance of parameters at various points in the chain; and the Brooks-Gelman-Rubin scale-reduction statistic $\hat{R}$, which compares the within-chain variation to

[1]The use of maximum likelihood could still be reasonable. The validity and its justification, however, depend on assymptotic arguments whose strength depend upon the case. To evaluate the quality of the maximum likelihood I would have to code a bootstrap process to reproduce the data generating mechanism, which in this case will be too complex. I chose not to do it here, but to leave this for a future project.

the between-chain variation (Gelman et al., 2004). As a practical rule of thumb, a 0.975 quantile for $\hat{R} \leq 1.2$ is indicative of convergence. In the application below, convergence diagnostics were performed using the R package *boa* (Smith, 2007).

A well-known computational issue for Bayesian mixture models is 'label switching', which is caused by symmetry in the likelihood of the model parameters during the course of the MCMC run (Celeux, 2012; Jakobsson and Rosenberg, 2007). In some cases, label switching can be avoided by placing constraints on the class probabilities (Lenk and DeSarbo, 2000) or on the model parameters themselves (Congdon, 2005). As an alternative, Stephens (2000) proposed a relabeling algorithm that minimizes the posterior expected loss under a class of loss functions - which is the one used in the application below.

In addition, I would like to incorporate all information available to deal with my research question. The Bayesian approach employed here also gives me the opportunity to incorporate informative priors from in-depth interviews conducted with political elites in Argentina, Brazil, Mexico, and Venezuela into my quantitative analysis. As advocated by Gill and Walker (2005), the elicitation of priors has the potential to tie together 'the seemingly antithetical research approaches of qualitative area studies with data-oriented work based on statistical methods, perhaps then mending a recent rift in political science.' To combine fieldwork information collected through interviews (qualitative) with fiscal records (quantitative), as well as to ensure a well-identified model with proper posteriors, I pursued a systematic procedure to recover elicited priors for the quantities in which I am most interested.

3.3 Elicited Priors from Political Elites' Knowledge

The scholars who advocate in favor of the Bayesian approach to statistical inference usually claim that we ought to adopt this approach because it is simple and direct, it allows the correct interpretation of frequentist hypothesis tests, it avoids asymptotic approximations, and it permits practical estimation of parameter contrasts and multidimensional credible regions (Jackman, 2009). But the most common justification for a Bayesian framework is that the Bayesian inferential structure converts prior information to posterior evidence by conditioning on observed data. Therefore, prior information, as well as the correct model specification, plays a key role in the argument for why Bayesian estimation is superior to pure frequentist estimation.

Under a fully Bayesian approach, prior distributions are assumed for all model parameters. But priors could come in many forms. To make sure my work fully exploits Bayesian capabilities, instead of ensuring a well-identified model with proper posteriors determined almost entirely by the data, I decided to use elicited priors. Elicitation is the process of estimating a person's knowledge and beliefs about one or more uncertain quantities. The result of elicitation is a (joint) probability distribution over those quantities that characterize the person's beliefs. In the context of Bayesian statistical analysis, elicitation arises most often as a method for specifying the prior distribution for one or more unknown parameters of a statistical model. In this context, the prior distribution will be combined with a statistical model and data through Bayes' theorem to derive a posterior distribution. The literature on elicited priors is large and the justifications for using it abound (Garthwaite, Kadane and O'hagan, 2005; Gill and Walker, 2005; Kuhnert, Martin and Griffiths, 2010). One of the most frequently cited reasons to pursue it has to do with the fact that elicitation brings the analysis closer to the application by demanding attention to what is being modeled, and what is reasonable to believe about it. Further, elicited priors help researchers to read better conclusions when the quantitative data themselves are not necessarily trustworthy, but also when the data contain missing values that are not missing at random (Jackman, 2009; Western and Jackman, 1994).

Another advantage of elicitation of priors comes from the utility of using expert priors in order to understand decision-making processes (Garthwaite, Kadane and O'hagan, 2005). Often a reasonable goal for elicitation is to capture the big message in the expert's opinion. The details, for example the exact shape of the expert's opinion, may not matter for the decision to be reached. Even when the decision is sensitive to the exact shape of the elicited distribution, it is not the decision, but rather the expected utility that matters. A second reason why elicitation is worthwhile has to do with the use of elicitations to make inferences, and in particular for making possible the calculation of posterior distributions. In such a situation, elicitation encourages the expert and the researcher to consider the meaning of the parameters being elicited. This has two helpful consequences. Again, it brings the analysis closer to the application by demanding attention to what is being modeled, and it helps to make the posterior distributions, once calculated, into meaningful quantities.

Bayesian inference is relatively new in political science and there have been only a handful of instances of the use of prior elicitation (Gill and Walker, 2005; Jackman, 2004; Western and Jackman, 1994). These papers highlight the substantive value of using elicited priors because the 'expert' views are tempered by actual data through the mechanics of Bayesian inference and produce more substantively comprehensive estimates of our quantities of interest. Otherwise, the best that we can do is describe such opinions anecdotally or with imprecise summary impressions because the number of available experts typically insufficient for standard statistical analysis. As noted by Gill and Walker (2005), 'this is unfortunate because elicited priors can be a means of systematically integrating qualitative and quantitative empirical work in political science, thus reaching across a traditional divide in the discipline.' In comparative politics, it is essential and urgent to find a systematic way of augmenting the qualitative information we discover during fieldwork with the quantitative information we collect from various sources. From the quote above, it seems this method already exists, but has not being used fully in the social sciences at this point. Nearly all published work on elicited priors exists in the literatures on medical trials or engineering. As I argue below, analyzing data to reveal political allocation strategies offers a critical opportunity to employ a Bayesian estimation incorporating politicians' own beliefs about how they do things. In other words, the joint use of quantitative data and qualitative information gleaned from fieldwork can provide an way to evaluate whether what politicians say is compatible with what they indeed do.

Elicitation is properly conceived of as part of the familiar process of statistical modeling. I follow a standard procedure to produce elicited information from politicians: (1) I selected the expert(s) and identified the quantities of interest, (2) I elicited specific summaries of the experts' distributions over those quantities, (3) fitted a (joint) probability distribution to those summaries, and (4) assess the adequacy of the elicitation (see Appendix for more details).

Each of the experts I interviewed (see list in the Appendix), provided me with guesses on the propensity and the level of investments presidents allocate to municipalities of several types. I also asked the experts for a measure of the uncertainty of their estimates, an upper and lower bound, which I used as a reference in a range method to find the level of precision for the prior estimates.

As I conducted several interviews, I pooled the guesses using a logarithmic opinion pooling method (Garthwaite, Kadane and O'hagan, 2005). I recognize, however, that I am using a non-conservative measure for the uncertainty because experts tend to overestimate uncertainty values for very large intervals (Weiss, 2012).

Given the information (mean and range of distribution) gleaned from experts, I identified sensible and convenient parametric distributions for the experts' information. For the fixed effects, I assumed normal priors: $\alpha_\kappa \sim N(\mu_\alpha, \Sigma_\alpha)$ and $\beta_\kappa \sim N(\mu_\beta, \Sigma_\beta)$. Since α and β are assumed Normal, we can find Σ by $(\mu_{97.5} - \mu_{2.5})/4$ from $\mu \pm 2 \times \Sigma$. Each D is assumed to have a conjugate Inverse-Wishart $IW(\nu_0, D_0)$ distribution. Since the other parameters of the model do not have a scientific interest for this paper, I assigned weakly informative proper distributions to them. This will ensure a well-identified model with proper posteriors determined almost entirely by the data. For, the lognormal precisions, τ_κ^2, I assumed conjugate $Ga(\lambda, \delta)$ priors.

3.4 Assessment of the Final Model Fit

To assess the adequacy of the selected model, I use posterior predictive checking (Gelman, Meng and Stern, 1996), whereby the observed data are compared to data replicated from the posterior predictive distribution. If the model fits well, the replicated data, $\mathbf{y}^{rep}$, should resemble the observed data, $\mathbf{y}$. To quantify the similarity, we can choose a discrepancy measure, $T = T(\mathbf{y}, \boldsymbol{\theta})$, that takes an extreme value if the model is in conflict with the observed data. Popular choices for T include sample moments and quantiles, and residual-based measures.

The Bayesian predictive p-value (P_B) denotes the probability that the discrepancy measure based on the predictive sample, $T^{rep} = T(\mathbf{y}^{rep}, \boldsymbol{\theta})$, is more extreme than the observed measure T. A Monte Carlo estimate of P_B can be computed by evaluating the proportion of draws in which $T* > T$. A p-value close to 0.50 represents adequate model fit, while p-values near 0 or 1 indicate lack of fit. The cut-off for determining lack of fit is subjective, although by analogy to the classical p-value, a Bayesian p-value between 0.05 and 0.95 suggests adequate fit. To assess the fit of the binomial component, I use T_1 (the proportion of observations greater than zero). For the nonzero observations, I use a modification of the omnibus chi-square measure proposed by Gelman et al. (2004).

CHAPTER 4

Simulation Study

To verify the conditions of the model, I conducted a simulation study. First, I simulated 1000 datasets from a three-class model according to equation 3.1. The datasets contained 5000 subjects, each with 8 observations, for a total of 40,000 observations per dataset. Such dimensions were chosen as they resemble the real dimensions of the data sets I am using in the application section of this paper. Both the logit and lognormal components contained fixed-effect intercepts, linear fixed-effect time trends, linear fixed-effects for two covariates (i.e. political alliance and poverty) and random intercepts for each of the classes in the data. The data generating mechanism did not contemplated an intercept, but it does have a single covariate to the class membership probabilities (i.e. size of municipality). For the simulation study I assign weakly informative proper distributions to all class-specific parameters. The simulation took roughly 20 hours to be completed and some of the models were run more than once, given that adaptation was not accomplished in the first round of iterations.

Figure 4.1 shows the true and the distribution of the estimated values for the main parameters of interest across the 1000 simulations. The bias was extremely low for all the parameters, including the random effect variance components not displayed here. The simulated values for α_{23}, β_{11}, and β_{13} present not only unbiased but also very efficient estimates. The coverage rates ranged from 0.71 to 0.97, but for the most part, were close to the nominal value of 0.95. Variability in coverage rates was likely due to the size of the simulation. Estimated class percentages were similar to the true class percentages of 60%, 25%, and 15% for classes 1, 2, and 3, respectively. Next, I fitted one to four class models to each of the 1000 datasets and compiled the results. As the data generating mechanism assumed three classes, I observed, as expected that the average DIC across the 1000 simulations was lowest for the three-class fitted model.

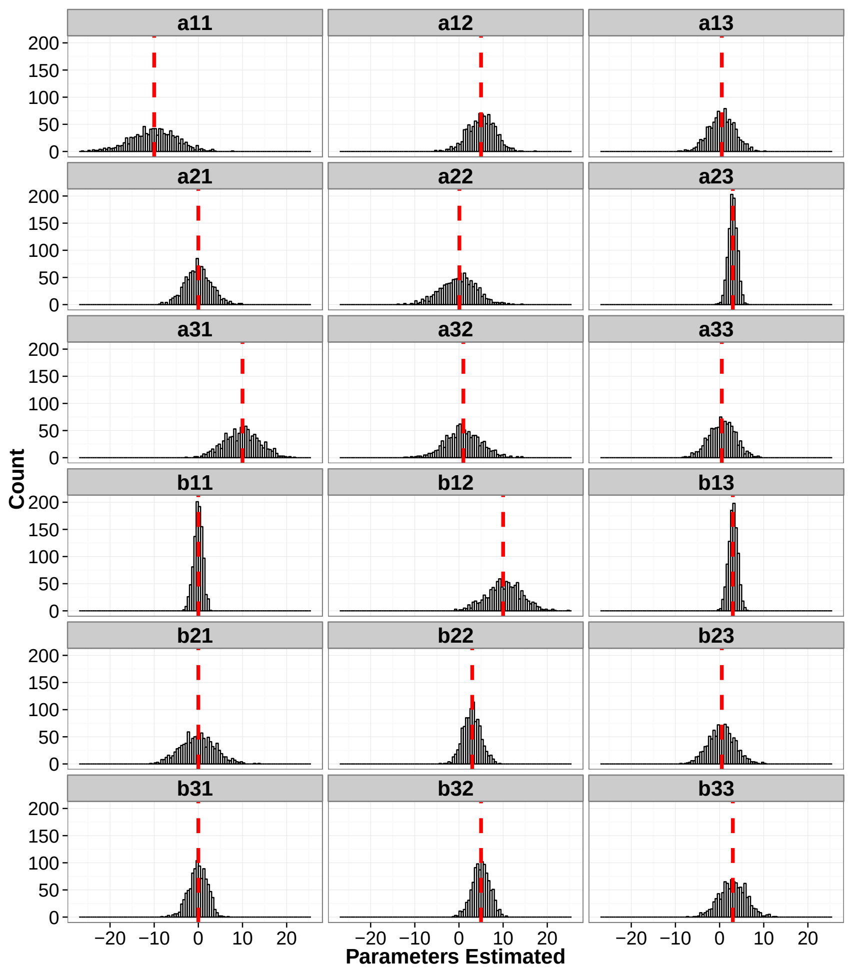


Figure 4.1: Distribution of the Estimated Values from Simulation Data of Parameters of Interest

CHAPTER 5

Government Expenditures in Brazil and Mexico

Besides the large amount of positive and negative powers they hold, Latin American presidents are famous for their power to decide on how to spend their citizens' taxes (Hallerberg, Scartascini and Stein, 2009; Mainwaring and Shugart, 1997; Mainwaring, 1993). The region is also well known for the weak ideological connection between voters and parties, and the importance of local governments in shaping national politics (Montero, Samuels and Helen Kellogg Institute for International Studies, 2004). One recent trend, however, has not been completely documented yet: traditional opposition parties have been re-elected for municipal governments, while those same constituents have voted to produce turnovers in the presidency. This paradoxical empirical trend raises, then, an interesting question: *How do newly inaugurated presidents allocate discretionary resources to constituencies that are voting inconsistently for his party at different government tiers?*

This is a particular puzzling question because it characterizes a dilemma for newly inaugurated presidents (Nunes, 2011). To the extent that voters respond to targeted spending, a president can use budgetary discretion to gain votes for himself or his designated successor. But federal transfer spending also improves re-election odds for incumbent mayors in targeted municipalities. The political benefits of transfer spending thus accrue not just to the president, but also to an array of local politicians who may or may not share the president's party and political goals. By targeting his own core voters and pursuing social policy goals, the president may be supporting the reelection goals of his political enemies at the local level. This dilemma looms larger when the president's party is weak at the municipal level, which was the case of Brazil and Mexico when, respectively, Lula (PT) and Fox (PAN) won their presidential elections. Should a newly inaugurated president allocates resources to localities governed by enemies, even when they might be able to steal credit from him? Or should he allocates resources to local allies, even though this would mean jeopardizing his electoral support in the majority of the country?

Empirically, this is an interesting empirical puzzle because, as discussed before, one of the main characteristics of expenditure data is the combination of a continuous distribution with mean different than zero, and a high frequency of zero values (Neelon, O'Malley and Normand, 2011; Olsen and Schafer, 2001). A semicontinuous variable is quite different from one that has been left-censored or truncated, because the 0's are valid self-representing data values, not proxies for negative or missing responses (Cooper et al., 2006; French, 2008; Xie et al., 2004). It is natural to view a semicontinuous response as the result of two processes, one determining whether the response is 0 and the other determining the actual level if it is non-zero. The two processes are qualitatively distinct and may be influenced by covariates in different ways. An added complication is that these two processes may be related, particularly if the semicontinuous response is observed at multiple time points; a high level of expenditure on one occasion may affect the probability of expenditures on another occasion (Neelon, O'Malley and Normand, 2011; Rizopoulos et al., 2008; Su, Tom and Farewell, 2009, 2011). Although government expenditure data is common in political science, we still lack a modeling strategy that deals properly with the characteristics of the data.

To demonstrate the applicability and usefulness of the model presented before, I analyze longitudinal government expenditure data from Brazil and Mexico during the presidential terms of Lula (BRA-PT), Fox (MEX-PAN) and Calderón (MEX-PAN). My aim here is to use this Bayesian approach to understand fundamentally what presidents do when, under scarce resources, they have to decide between rewarding his core voters or his political allies in the localities over the curse of their terms. Building upon the political economy literature on presidential pork barrel politics, this application reveals how national leaders, who often face the dilemma of the distribution, solve it finding an allocative strategy that works to their greatest political advantage. In the next section I describe the data used to assess this question and the estimation strategies employed here.

5.1 Data

To evaluate how presidents distribute resources to geographic areas in the country, I compare municipal and national executive election results to municipal-level observations of transfers allocation from 2003 to 2010 in Brazil and from 2000 to 2011 in Mexico. In Brazil, I look at the results of Lula's government. In Mexico, I analyze the distribution patterns of PAN under the leadership of Vicent Fox (2000-2006) and Felipe Calderón (2007-2011). The combination of municipalities mayors' party and president vote shares change every two years in Brazil and every four years in Mexico because of the non-concurrent schedule of elections. The term of mayors and presidents in Brazil is the same, four years, and they can be reelected for one extra term in elections that happen every other two years. National elections were held in 2002 and 2006, whereas municipal elections in 2000, 2004 and 2008. In Mexico the president has a single six year term, whereas mayors have single four year terms. Reelection is not allowed for any of the offices in Mexico. National elections were held in 2000 and 2006, whereas municipal elections in 2000, 2004 and 2008.

5.1.1 Dependent variable: Municipal Transfers

I construct annual indicators of transfer allocations per capita for all municipalities both in Brazil and in Mexico. The information for Brazil comes from the Budget Committee of the Brazilian Senate. They are responsible to audit the expenditures in the country, so the data seems to be reliable [1]. The transfers relate to the Brazilian fiscal and social security budget. The information about budget investments of national companies, (i.e. Petrobras) is not included. Given my puzzle, I am interested in transfers in which the president has discretionary power. Formally, I will be analyzing expenses called 'investments' according to the budget classification, which are investments in new buildings, new roads, hospitals, and so on.

The information for Mexico comes from the INEGI - the National Institute of Geography and Statistics of Mexico. The institute is internationally recognized as clean and non-corrupted by government officials. The investments relate to the Mexican fiscal and social security budget. The information related to oil and minerals is not included, but only the investments the president allocates to municipalities discretionally. Formally, I will be analyzing expenses called 'aportaciones' and 'inverson pública federal' according to the budget classification. These are discretionary amounts of resources invested by the national government direct to municipalities in order to create infra-structure in general.

[1] I also checked the data against the information provided by the Brazilian National Treasury at www.stn.gov.br

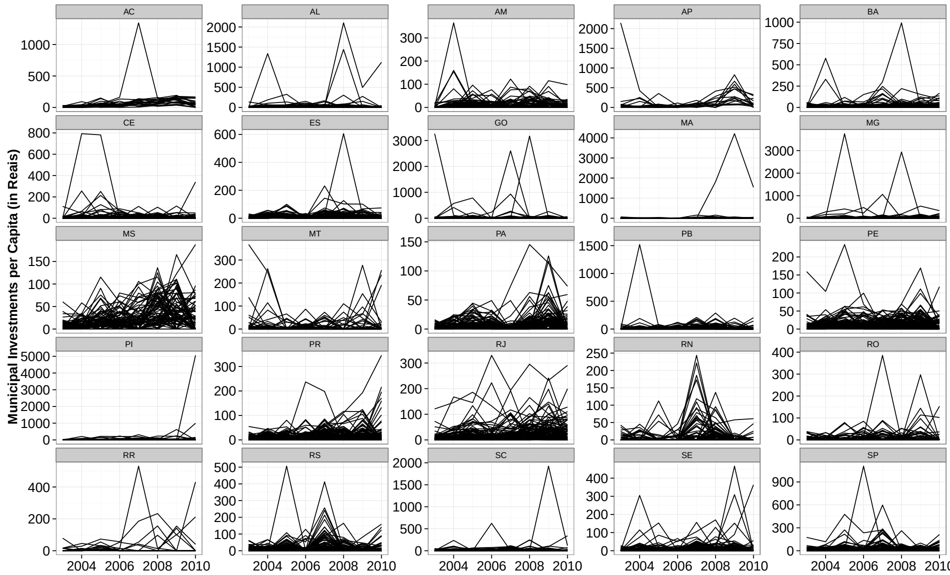


Figure 5.1: Municipalities that received federal investments by state, Brazil (2003-2010)

In Figure 5.1, I show the temporal trend of municipalities that received transfers by state in Brazil. Note how the y-axis change by state, indicating substantive differences in levels of investments per capita in each state. The trajectory of spending is not linear, nor follows a clear pattern that could be described by any of the well know error covariance models. This is a good indication that we need to use temporal fixed effects in order to capture the temporal trends properly (Weiss, 2005). The most clear pattern observed in this data set is the presence of investment picks followed by absence of spending in varies municipalities. Note the triangular shape of the lines connecting the dots. The Mexican case in Figure 5.2, although also presents the picks pattern observed in Brazil, has a group of municipalities receiving investments over time that seems to follow a linear trajectory. Note the block of lines varying monotonically together in the middle of the plots for Chiapas, Durango, Sinaloa, and Yucatan. A linear time trend could be sufficient to capture the temporal dependency of investments in Mexican municipalities, but given the presence of some unstructured tendencies of investments, I decided to use year fixed effects to model such temporal dependency as well.

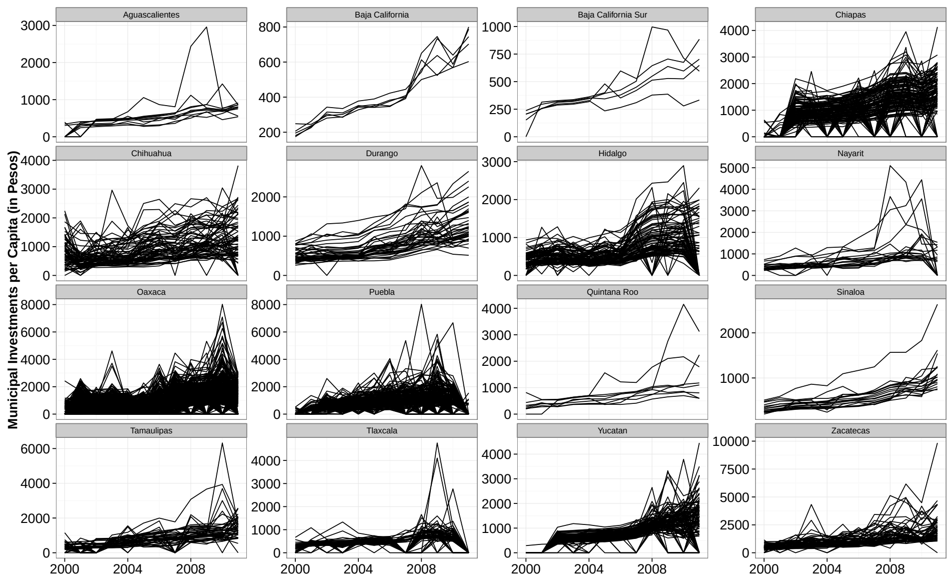


Figure 5.2: Municipalities that received federal investments by state, Mexico (2000-2011)

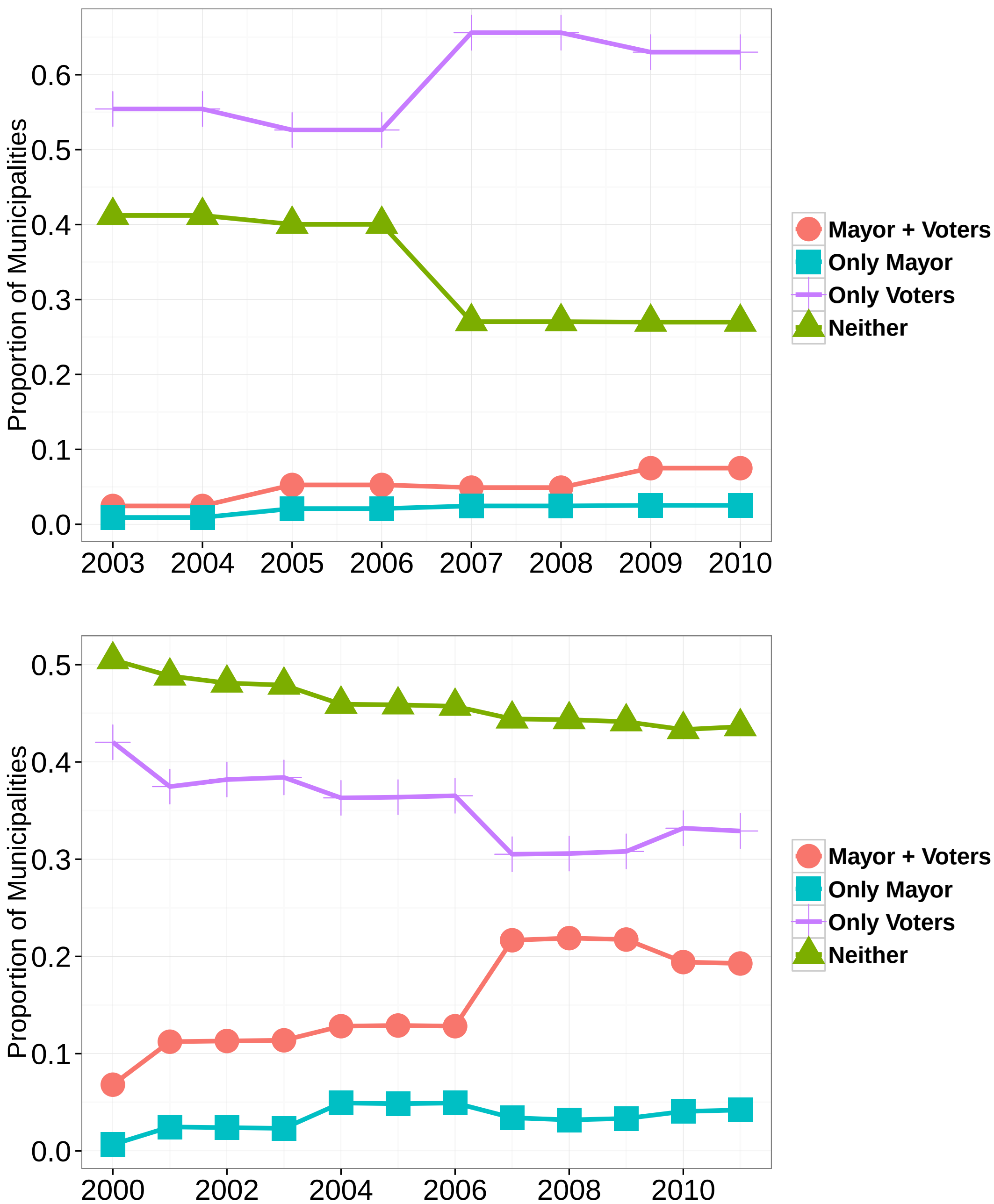
The most important feature of the investment data, however, is the presence of a large volume of zeros both across time and across municipalities. Note how close the lines are to zero in many periods of time, specially in Brazil. As presented before, 70% of municipalities in Brazil do not receive any investments over time. In Mexico, this number is much smaller although still significant: roughly 25%. Modeling the semicontinuity of the data over time helps understanding better how presidents decided on the distribution of investments for the municipalities over time. Given the large variation presented in the figures above, I expect to be able to identify precise estimates for the several trajectories that investments take in different municipalities and across the countries. The cost for estimating several parameters in both the probit and in the lognormal models is then compensated by the fact that we can clear identify the patterns differentiating the types of municipalities.

5.1.2 Independent Variable

My main variable of interest is the *district configuration*, which is a categorical variable that combines information about voters' preferences in the presidential and in the mayoral elections. The variable contains four mutually exclusive categories: municipalities in which most voters supported the president in the presidential election and also elected a mayor from the president's party, municipalities in which voters supported the president in the presidential election but did not elected a mayor from the president's party, municipalities in which most voters did not supported the president in the presidential election but did elected a mayor from the president's party, and municipalities in which most voters did not supported both the president in the presidential election and his mayoral candidate in the local election. The first dimension distinguishes voters choice in the national election. If the plurality or more of a municipality voted for the president, such district is coded as the president's core voters (=1); otherwise, as opposition supporters (=0). The second dimension identifies mayors' affiliation having as a reference the president's party (=1) or others (=0).

This variable changes by municipality and by year. The main purpose of creating it is to explore heterogeneous effects of mayoral party affiliation depending on the proportion of votes that presidents received in each municipality. As we can see in Figure 5.3 (top), Lula obtained the plurality of votes in the majority of

Figure 5.3: Municipal Political Configuration, Brazil and Mexico (2000-2011)



the municipalities, but his party was not able to accomplish even nearly the same thing. Lula obtained the plurality of votes in 61% of the municipalities in Brazil, whereas PT only governed 7% of them. However, the number of municipalities governed by PT more than doubled between 2002 and 2010 - it went from 200 to roughly 500 in eight years -, what can be considered a large accomplishment for the party. The type of municipality that decreased substantially, specially after 2006, were the ones governed by opposition mayors with the majority of voters against Lula. Overall, these numbers suggest that more voters supported Lula and his party over the years of his term. In Mexico (bottom of Figure 5.3), the situation was even more complicated when the inauguration of presidents' Fox and Calderón governments. Given the historical dominance of the PRI since the beginning of the 20th century, the newly inaugurated presidential party, the PAN, did not have the plurality of the votes in the majority of the municipalities. On the contrary, the PAN obtained a plurality in only 32% of the Mexican municipalities. Over time, PAN was able to govern 17% of the municipalities, far from what the PRI was able to accomplish during its term in government.

5.1.3 Controls

To reliably evaluate how president's distribution of resources comports with the main common findings in the literature on distributive politics, I had to consider variation in the following municipalities' features: (1) socio-economic conditions, (2) population size, (3) importance to any member of the presidents' legislative coalitions, and (4) the relative relevance of the district for the presidents' electoral successes. In other words, I controlled for factors that the literature says increases the likelihood that a municipality receives transfers.

To account for variations in the level of development across municipalities and states, I used measurements of *poverty rate* and *GDP per capita* for each municipality. The data for Mexico comes from the database SIMBAD of INEGI[2] and from the CONAPO[3]. Data for Brazil is from IPEADATA[4] and FIRJAN[5]. The municipal level of poverty for Brazil is only available for 2003. Therefore, I decided to use a better proxy that varies by year and by municipality: the *IFDM index*. The presence of infrastructure like a School or Medical Facility in the municipality might show a place has a lower need for transfers. Existing infrastructure may also reflect some latent factor associated with the ability of local residents and politicians to secure government projects in their municipalities. IFDM index controls for that. It is measured based on employment rate, income, education, and health levels of each municipality. It ranges from 0 to 1, with 1 being the best situation a municipality can have. In Brazil, I also control for the number of citizens that received *bolsa família*[6] each year after 2003. This is a way to control for the number of families that have received help from the government, and therefore would tend to support it. I also set in the model a control variable for municipality *population*, that identifies both the number of potential users of the benefits provided by the president and also the number of potential voters.

I also set a control variable to take into account the relative importance of each municipality for legislators that are members of the presidents' legislative coalition. Municipalities that deputies are accountable to are called the *Electoral Constituency*. Using legislative and electoral data from the TSE[7] and CIDAC's electoral database[8], I created a dummy variable that shows whether or not a

[2] http://sc.inegi.org.mx/sistemas/cobdem/

[3] See http://www.conapo.gob.mx/en/CONAPO/Indicadores

[4] See http://www.ipeadata.gov.br

[5] For more information access Firjan: http://www.firjan.org.br/data

[6] It is a Conditional Cash Transfer program created in 2004 to assist poor families in Brazil.

[7] See http://www.tse.gov.br

[8] `http://www.cidac.org/eng/Electoral_Database.php`

municipality can be considered the main electoral constituency of a legislator who composes presidents' legislative coalition. Municipalities that were identified in such way were coded as 1, the others receive 0 value. For Mexico, I build this variable based on the electoral outcomes on single-member districts. Every municipality that elected a deputy supporting the PAN in Congress is identified as electoral constituency of the government. For Brazil, I used two different strategies. First, from the entire set of municipalities from which a deputy in Lula's coalition received votes, I identified the one in which he/she received the bigger absolute number of votes, and coded it as 1.[9] Second, for each deputy in Lula's coalition, I identified the municipality in which he/she had the largest victory margin. It means, among all municipalities in which he or she received votes, I identify the one in which he got the biggest percentage of votes, as a share of the number of votes in each municipality.[10] Given that classification, I included a dummy variable in my data set (*Electoral Constituency*) that have values equal to 1 for the municipalities identified above, and 0 otherwise. Those were the municipalities coded as the areas in which governmental partners have influence. The motivation for such control was to isolate the alternative hypothesis that municipal transfers are used to secure governmental discipline in the Congress and in the Governor's cabinet, and then, build legislative coalitions.[11]

[9]For example, if deputy A was elected with 100 votes, being 30 from municipality A, 60 from municipality B, and 10 from C, municipality B was choosen to be coded = 1.

[10]For example, if deputy A was elected with 100 votes, being 30 from municipality A (population = 31), 60 from municipality B (population = 500), and 10 from C (population = 30), municipality A was choosen to be coded = 1.

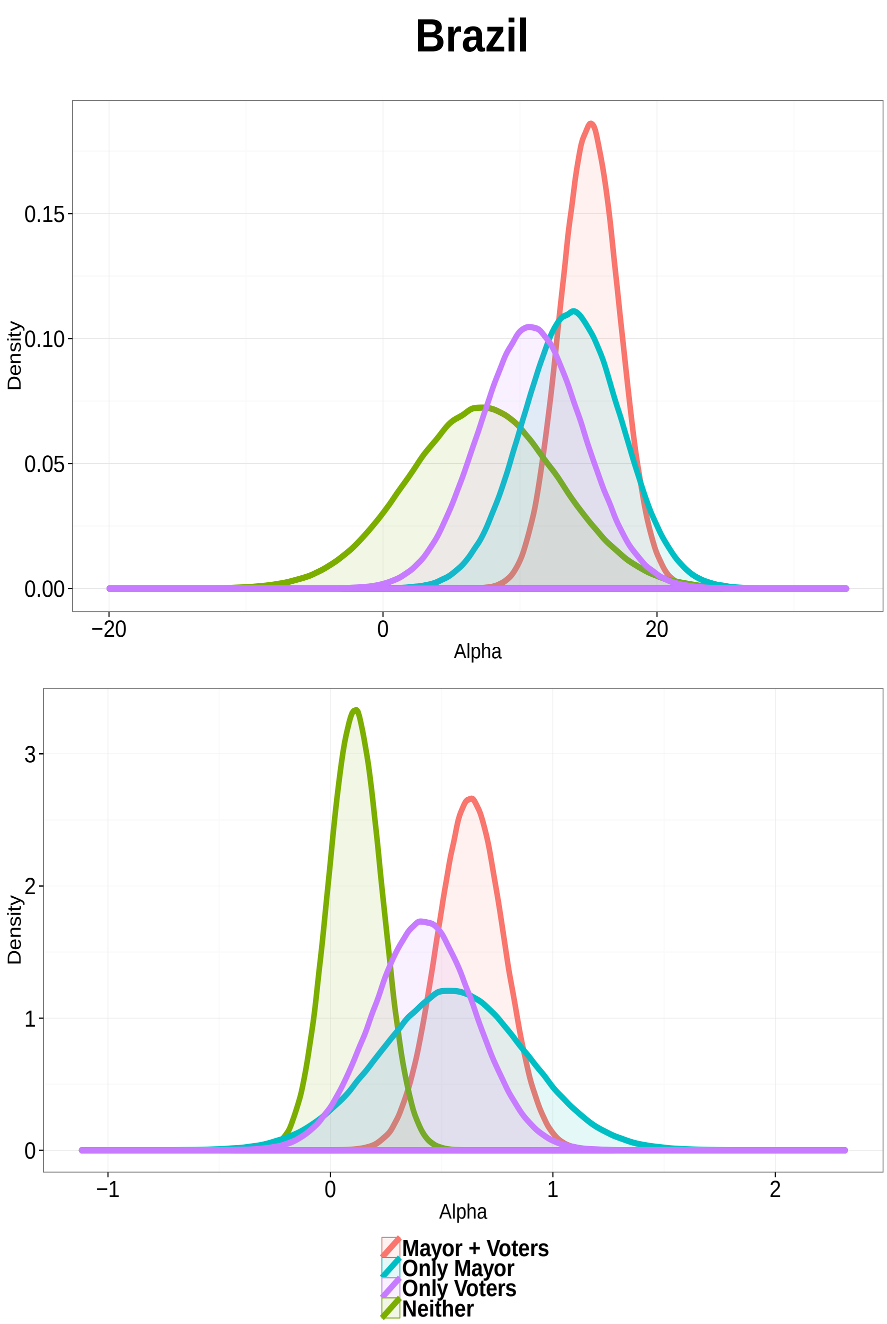
[11]The literature argues that presidents have incentives to distribute resources to municipalities in which its partners have votes, therefore, increasing their electoral chances. But if I show that PT mayors have been receiving more benefits than any other mayor, even when such control is set, I believe the argument claimed in this paper will look like more reasonable, than the alternative one.

To account for the municipal relative importance on the presidents' elections, I constructed a variable with the percentage of votes each municipality provided for each of the presidents, given the total number of votes they received. I am calling this the *President's Vote Weight*, as it takes into account how much electoral weight each municipality has in generating presidents' electoral victories. Two other controls are used to account for presidential term limits and mayor's power. First, I set a dummy variable for Lula's first (=0) and second (=1) terms, and for Fox (=0) and Calderón's terms (=1). Second, I calculated mayor's vote share in every municipal election that took place between 2000 and 2011. Such measurement gives me a good proxy for the municipal influence of mayors on their municipalities.

5.1.4 Priors

As described before, I interviewed several politicians and staff members of the government looking for elicited priors for my Bayesian analysis. Instead of assuming uninformative priors, as it is common to observe in the applied literature in social sciences (Jackman, 2009), I would like to contrast the 'experts' beliefs about discretionary allocation for political advantages with the actual data on political decisions about investments. I interviewed academics, politicians, and public officials from the national and subnational levels of government in Brazil and in Mexico, including former presidents Luis Inácio Lula da Silva and Fernando Henrique Cardoso (Brazil), and the former president of Mexico Vicente Fox. For the coefficients that I did not have expert priors, I use estimates from other papers that had estimated the effects of, for example, poverty, GDP, and the other control variables on presidential discretionary allocation in these four countries (Armesto, 2009; Brollo and Nannicini, 2012; O'Neill, 2006; Penfold-Becerra, 2004).

Figure 5.4: Prior Distributions for Two-part Bayesian Model, Brazil



I directly asked the subjects I interviewed about the prior means for the parameters α_κ and β_κ. The point estimates were then averaged to become what is reported in Table 5.1. I also ask the experts about their uncertainty in the estimates. To do so, I asked for an upper and or lower 'bound', which I used then as a reference in a range method to find the level of precision for the prior estimates. Again, the levels reported in Table 5.1 refer to the overall mean elicited from the experts I interviewed. I recognize I am using a non-conservative measures for the uncertainty since experts tend to overestimate such values for very large intervals.

To visualize the priors in their distributional forms, I plot in Figure 5.4 the priors for the different parameters for Brazil. In terms of magnitude, the municipalities who are assumed by experts in Brazil to receive the largest amounts of investments are the ones with both a plurality of voters supporting the president and a mayor from his party. Note, however, that its distribution mostly overlap with the prior distribution for municipalities governed by the PT, but without the plurality of voters supporting the president. The priors for the logit part of the model follow the same order of the priors for the lognormal effects. The most remarkable feature of these distributions, however, is that the prior for municipalities governed by the PT without electoral support for Lula has a very small precision, being the only non-informative prior in the case of Brazil.

In the case of Mexico (Figure 5.5), the experts' priors for the volume of resources invested in the municipalities follow the order: (1) municipalities with the plurality of voters supporting the PAN in the presidential election and with a PAN mayor, (2) municipalities governed by PAN but without majoritarian support for the president, (3) municipalities in which voters mostly voted for PAN in the presidential election, although the mayor is from an opposition party, and (4)

Figure 5.5: Prior Distributions for Two-part Bayesian Model, Mexico

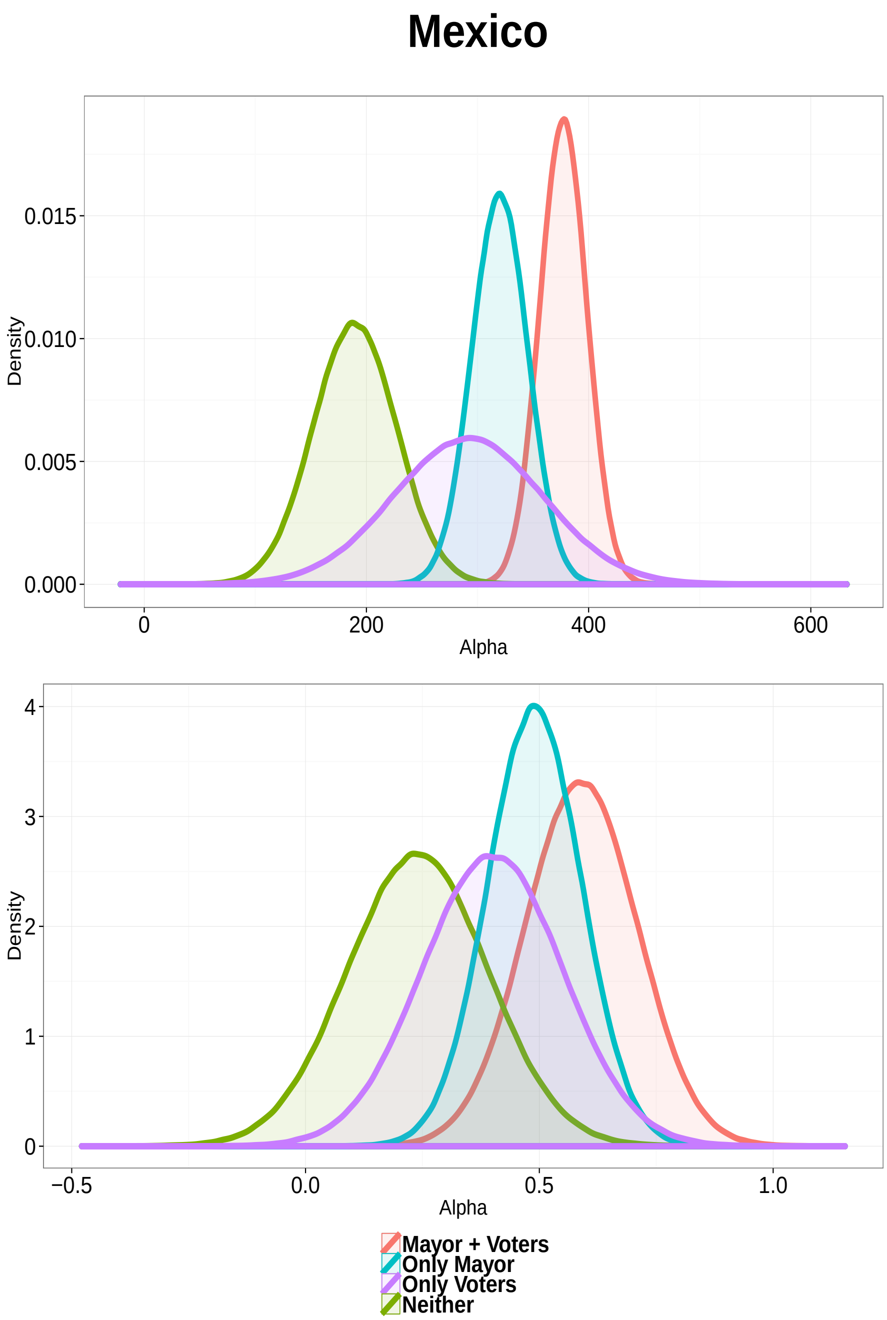


Table 5.1: Prior specification from fieldwork interviews and previous studies

		Brazil	Mexico
$\alpha_{\text{Mayor+Voters}}$	$\mu_{\alpha,\text{Mayor+Voters}}$	15.1	377.3
	$\Sigma_{\alpha,\text{Mayor+Voters}}$	2.15	21.1
$\alpha_{\text{Only Voters}}$	$\mu_{\alpha,\text{Only Voters}}$	10.8	291.7
	$\Sigma_{\alpha,\text{Only Voters}}$	3.8	67.1
$\alpha_{\text{Only Mayors}}$	$\mu_{\alpha,\text{Only Mayors}}$	13.8	320.1
	$\Sigma_{\alpha,\text{Only Mayors}}$	3.6	25.1
α_{Neither}	$\mu_{\alpha,\text{Neither}}$	7.3	189.5
	$\Sigma_{\alpha,\text{Neither}}$	5.5	37.5
$\beta_{\text{Mayor+Voters}}$	$\mu_{\beta,\text{Mayor+Voters}}$	0.63	0.59
	$\Sigma_{\beta,\text{Mayor+Voters}}$	0.15	0.12
$\beta_{\text{Only Voters}}$	$\mu_{\beta,\text{Only Voters}}$	0.42	0.40
	$\Sigma_{\beta,\text{Only Voters}}$	0.23	0.15
$\beta_{\text{Only Mayors}}$	$\mu_{\beta,\text{Only Mayors}}$	0.55	0.49
	$\Sigma_{\beta,\text{Only Mayors}}$	0.33	0.10
β_{Neither}	$\mu_{\beta,\text{Neither}}$	0.11	0.24
	$\Sigma_{\beta,\text{Neither}}$	0.12	0.15
τ^2	λ	2.26	3.12
	δ	4.51	5.13

municipalities not governed by PAN nor supporting the PAN president. Different than in Brazil, the experts' information in Mexico lead to more precise and non-overlaping prior distributions. The exception is the prior distribution for PAN/Opposition municipalities. The priors on the likelihood that a municipality would receive federal investments have the same order as presented before, but with highly overlapping tails.

Since the elicited priors are a form of information produced by previous knowledge from structured interviews with subjective area experts who have little or no concern for the statistical aspects of my study, it is possible to state that the densities described above are a good description about how politicians and bureaucrats understand the game of distributive politics in practice. My intent in eliciting these beliefs was to introduce qualitative and area-specific information into an empirical model in a systematic and organized manner in order to produce parsimonious yet realistic results. After presenting my findings I will come back to this discussion to compare the importance of the elicited priors to my conclusions.

5.2 Findings

In this section, I assess the main findings from my estimation exercise. The methodological choices are driven by a scientific question discussed and developed in another paper (see Nunes, 2013). In summary I propose that if strong local politicians (i.e., in more decentralized countries like Brazil and Mexico) are able to affect whether or not the president can fulfill her agenda, the president should use her discretionary power to allocate proportionally more resources to municipalities governed by co-partisan mayors, who tend to be her most loyal agents. By contrast, if weak local politicians cannot steal political credit or siphon resources

Table 5.2: Estimates for Presidential Discretionary Spending (Probit)

		Brazil	Mexico
Probit	Mayor + Voters	0.55	0.49
		(0.08)	(0.08)
	Only Voters	0.35	0.31
		(0.03)	(0.08)
	Only Mayor	0.56	0.48
		(0.03)	(0.10)
	Neither	0.23	0.33
		(0.03)	(0.04)
	GDP cap	-0.35	-0.52
		(1.00)	(0.44)
	Poverty Rate	-0.71	0.26
		(8.55)	(1.96)
	Electoral Constituency	-0.25	0.19
		(0.87)	(1.89)
	Benefitted by CCT	-0.31	-0.25
		(2.38)	(1.33)
	President's Constituency	0.05	0.05
		(0.91)	(3.16)
	Mandatory Resources	-0.35	-0.17
		(0.37)	(0.10)
	Mayor's Vote Share	Yes	Yes
	Governor's Political Party	Yes	Yes
	Year Fixed Effects	Yes	Yes
	Number of Cases	43,978	31,374
	MCMC Iterations	80,000	84,000

from the president and have less power to help or hurt her electorally, as in centralized systems, then we should observe the president allocating resources where her own core voters are settled, regardless of the local mayor's party affiliation.

To analyze the government expenditure data described before, I fitted a series of two-part models. For each part of the model, I fit fixed effects models with correlated random intercepts and a random slope for the lognormal component. As explained before, I assumed a probit-lognormal two-part model as in equation 3.1. For both components, the fixed-effect covariate vector x_{it} comprised the dummy variables with the four types of municipalities. These variables represent the combination of voters and mayors support to the president. The fixed-effect covariate vector z_{it} contains the control variables described before. Because my study included only eight measurements of time that are not clearly associated to any regular trend, I chose to model time categorically to allow for maximum flexibility in capturing the time trend. Alternative parameterizations of the time trend - such as polynomials or splines - may be appealing in other settings, particularly if there are a large number of time points.

Appendix Figure .2 presents post-burn-in trace plots for four representative parameters from the random intercepts model: α_{22} (change in log odds use at year 2 compared to year 1, class 2); β_{22} (increase in log-spending at year 2 for class 2); γ_{22} (log odds of class-two membership, large vs. small municipalities); and ρ_2 (class-2 random effect correlation). For clarity of presentation, I have graphed only one of the four MCMC chains. The overlapping trajectory lines suggest convergence and efficient mixing of the chains. The Geweke Z-diagostic p-values ranged from 0.11 (β_{22}) to 0.89 (α_{22}), indicating no significant difference in posterior means across regions of the chains; the 0.975 quantiles of the Brooks-Gelman-Rubin statistic were each less than 1.04, again indicating convergence of the chains. However, I did observe modest autocorrelation in the chains: the lag-10 autocorrelations ranged from 0.05 for α_{22} to 0.25 for ρ_2.

Table 5.2 presents the coefficients from the Bayesian two-part model that corresponds to the estimates from the probit model. Table 5.3 contains the results from the log-normal model. Each column presents outcomes for one of the countries analyzed here. I start analyzing the results from the first part of the model that estimate the probability that a municipality would receive any spending at all. Across all specifications, municipalities with presidents' co-partisans in decentralized systems are associated with a higher propensity to receive federal allocations, whereas municipalities with more of the presidents' core voters in centralized systems tend to have a higher propensity to receive national investments. These results are statistically significant and substantively meaningful when different municipal profiles are considered. When these political factors are taken into account, conventionally important variables such as municipality poverty rate, municipality GDP per capita, and number of municipal beneficiaries of conditional cash transfer programs present effects indistinguishable from zero.

Figure 5.6 reports the propensity effects from the estimated posterior distributions. For the propensity to receive any expenditures, the estimates show striking effects across the municipal categories. The estimates indicate that Brazilian and Mexican municipalities governed by presidents' co-partisans had as much as a 35 percent chance of receiving federal investments. These results are consistent with the elicited priors discussed before and corroborate the importance of decentralization in the politics of the region. In Mexico, the necessity to change the local dominance of the *Partido Revolucionario Institucional* (PRI), given its traditional power, seems also to be consistent with the higher propensity for PAN presidents to target only their allies. In Brazil, besides the relevance of decentralization, the fact that the PT was the most ideological political party in the country, and not in power for three decades, also helps explain the president's need to divert resources from the support bases of traditional ruling parties.

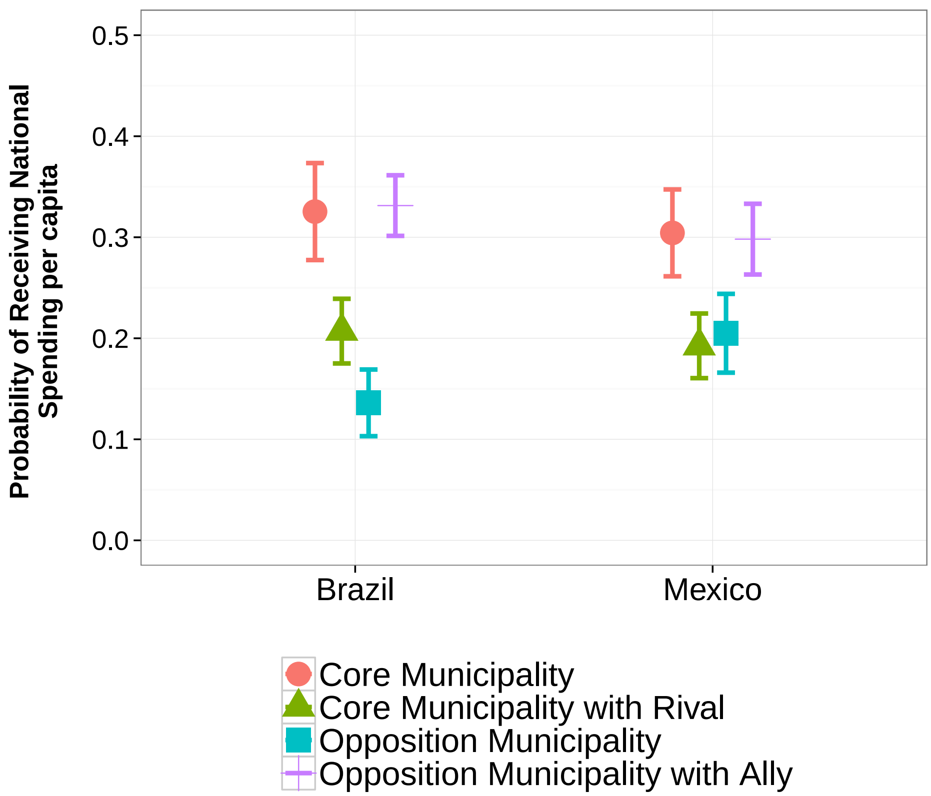


Figure 5.6: Estimated Propensity of Presidential Discretionary Investments per capita for Municipalities of Different Political Support, Brazil and Mexico

Controlling for poverty levels, legislators' electoral constituencies, and for many other important covariates, there are, then, important differences across municipal configurations for the probability of receiving investments. The effect of a co-partisan mayor in both Brazil and in Mexico is positive and substantially large. Moreover, there is no significant difference between municipalities without or with presidential voters in these systems, indicating mayoral partisanship trumps local presidential vote share for resource allocation decisions. In order to target core voters, the results suggest, presidents in Brazil and Mexico prefer to do so mostly where they have a co-partisan in the city hall.[12]

Next, I assess the average volume of investments each municipality receives, conditional on it having received something. The results are also consistent with what we have observed for the probit models (see Figure 5.7). The effects for each municipal profile are straightforward to interpret. I exponentiated the estimated coefficients to report meaningful values here. In a hypothetical district within Brazil or Mexico where voters support the president, being governed by a president's co-partisan increases by $15 per capita the amount a president invests in a municipality. When we compare two hypothetical districts governed by co-partisans of the president, one whose voters supported the president and the other whose voters did not, we observe a small difference in the amount of presidential expenditures. Municipalities with allies receive, on average, a much higher volume of transfers than municipalities with opposition mayors in both countries, regardless of the vote pattern observed there.

Different than with the probit model results, it is possible to distinguish the pattern of allocation between municipalities in the two countries using the log-normal posterior distributions. There is a clear centralization-based ordering of municipality types in terms of the level of investments each receives. The order is as follows: core districts, opposition districts with ally, core districts with ri-

[12]One interesting exception, however, is São Paulo. Even though it has been controlled by Lula opponents for 10 years, Lula has disbursed a lot of resources there in order to guarantee benefits to such voters. PT has been doing everything to persuade voters to elect a PT mayor again.

Table 5.3: Estimates for Presidential Discretionary Spending (Lognormal)

		Brazil	Mexico
Lognormal	Mayor + Voters	27.6	29.6
		(3.29)	(2.29)
	Only Voters	13.5	16.5
		(1.93)	(0.92)
	Only Mayor	24.5	22.9
		(1.83)	(1.50)
	Neither	2.6	3.8
		(1.59)	(1.73)
	GDP cap	-7.89	-1.98
		(1.22)	(5.08)
	Poverty Rate	1.23	6.63
		(0.10)	(0.87)
	Electoral Constituency	-3.78	-7.74
		(3.19)	(2.67)
	Benefitted by CCT	0.09	-2.01
		(0.01)	(2.87)
	President's Constituency	3.90	-0.01
		(1.75)	(0.99)
	Mandatory Resources	-6.75	-4.19
		(4.83)	(2.45)
	Mayor's Vote Share	Yes	Yes
	Governor's Political Party	Yes	Yes
	Year Fixed Effects	Yes	Yes
	Number of Cases	43,978	31,374
	MCMC Iterations	80,000	84,000

val, and opposition district. Core districts tend to receive $30 per capita and opposition districts with an ally, on average, $22. This difference is not large, suggesting the importance of allies in determining the volume of federal investments a municipality receives. Core districts governed by president's rival party receive, on average, $15 per capita, but opposition districts only $4. In Mexico, for example, opposition districts with an ally receive more spending than core districts governed by a rival party. Whereas the former receives $24 per capita, the latter only receives $17.

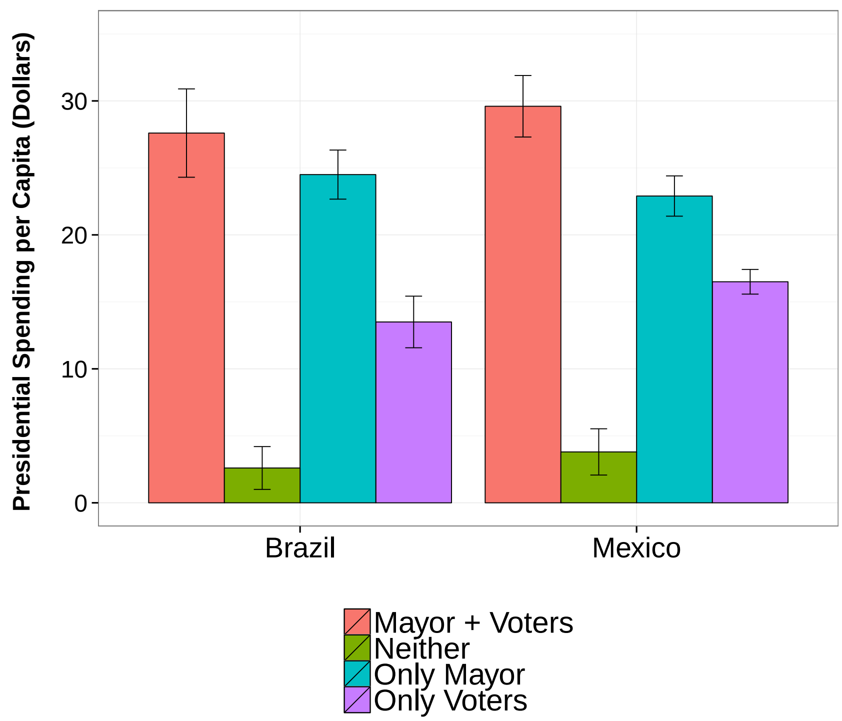


Figure 5.7: Estimated Amount of Presidential Discretionary Investments per capita for Municipalities of Different Political Support, Brazil and Mexico.

The results presented here are not only statistically significant, but are substantively important as well. For instance, in Brazil a change to a president's mayor is estimated to have the same effect on municipal spending as a 9 percentage-point increase in municipal poverty rate. To have a reference of the magnitude of this difference, it is equal to the poverty rate difference between Bento Gonçalves (Rio Grande do Sul) — a very rich city in the South of Brazil — and Macurure (Bahia) — one of the poorest cities in the Northeast of Brazil. These voters are being deprived because they are not governed by president's loyal allies. In Mexico, increasing the electoral support a president has in a municipality corresponds to an increase in spending that is equal to a 1 percentage-point increase in the level of poverty of a municipality — not a substantive difference.

These results are notable given the widespread idea that political decisions in Brazil and in Mexico are just a matter of how the government addresses the needs of the poor. When it comes to municipal transfers, the differences across mayors' affiliation are substantial, being the most positive effects in municipalities that elected local mayors of the presidents' parties. I demonstrate that when faced with the presidential dilemma, Lula, Fox and Calderón have pursued two different strategies: when the mayor in office is a political 'friend', presidents tend to reward him. When, on the other hand, presidents are facing opponents or coalition officers their strategies follows two steps. First, they appropriate transfers for opponents in which municipality voters supported them. Second, they avoid reinforcing opposition mayors' power. Therefore, Lula, Fox and Calderón seem to buy mayors opposition support when the municipality voted for them, and try to punish voters when allies govern the municipality.

Regardless of the appealing of such results, we cannot conclude that these patterns reflect a true causal effect of president's party power in the municipalities, since it could be biased by an omitted variable in which an unobserved factor is associated with both president's party electoral success in the municipalities, and higher transfer rates. While the inclusion of municipal fixed effects could absorb time-invariant factors that operate within regions, and the year dummies should account for broad temporal trends affecting the whole country, these statistical adjustments provide only a partially satisfying response to such concerns. Recognizing the causal limitations of my work is necessary to properly place the contribution of the model discussed here.

Elicited Priors vs. Data in the Bayesian Framework

These results highlight not only the relevance of decentralization to explain political allocations in Latin America, but also a substantive value of using elicited priors in political science. The prior beliefs that I elicited throughout interviews suggest that loyalty in the municipalities is a good indicator of discretionary distribution. This is possible to detect because the political elites' expressed views are tempered by actual data on discretionary allocation through the mechanics of Bayesian inference. On the one hand, without the actual spending data, we could only anecdotally report the politicians' and staff members' opinions because the number of available experts is insufficient for standard statistical analysis. On the other hand, without the elicited priors, we might rely too much on a particular draw of data and estimation technique. As always, qualitative and quantitative approaches complement each other, but the modeling technique used here brings the qualitative data to bear in a systematic and transparent way, instead of merely relegating it to the role of impressionistic reality check, as is so common in the field.

A posterior distribution summarizes our knowledge about the quantities of interest in Bayesian analysis. Analytically, the posterior density is the product of the prior density and the likelihood estimated from the quantitative data. Using elicited information requires careful decisions though (Gigerenzer, 1996). As highlighted in Burgman (2005), the perception of experts is inevitably subject to bias and depending on the nature of that bias, their opinions may influence the model estimates. It is therefore important to be aware of the impact that priors can have on models as this may influence our conclusions. There are several scenarios that can arise when combining the likelihood with priors generated from expert opinion. The amount of data, the mean value, the precision and the way in which the prior mean and precision are captured and incorporated into a model can influence the posterior estimate. In situations where data are limited, the expert's expressed opinion has the potential to drive model predictions. When data is abundant, by contrast, the priors tend to play a less critical role.

In most of the above analysis, the informative priors are just improving the precision of the estimates. Including them does not change the direction or the substantive interpretation at all. But there are also the instances in which the conclusions do change from null to positive or from positive to null when non-informative priors are used.

Although for the most part the incorporation of the elicited priors does not contradict the information in the data, there are some instances in which the expert knowledge influences the outcome distinguishing the levels of spending by municipal type. The choice to present the posterior results with elicited priors and not the noninformative ones relies on two arguments. First, the critique of the use of elicited priors in the Bayesian paradigm is reasonable when small samples are being analyzed. In the analysis reported here, the majority of the weight is automatically assigned to the data, given its sample size. Overall, I am analyzing more than 100,000 data points. Second, given the frequently expressed doubts about the reliability or completeness of official government data in some Latin American countries, either due to politically motivated massaging or to important, unofficial flows of resources; talking to people in a position to know and treating their answers as evidence can help us to evaluate the observational data with the proper circumspection.

Finally, a word of caution. Regardless of the appeal of these results, we cannot conclude that these patterns reflect a true causal effect of presidential support in the municipalities — they could be biased by some omitted variable that is associated with both president's party's electoral success in the municipalities and higher transfer rates. Although the inclusion of municipal fixed effects could absorb time-invariant factors that operate within regions, and the year dummies should account for broad temporal trends affecting the whole country, these statistical adjustments provide only a partially satisfying response to such concerns. Recognizing the causal limitations of my work is necessary to properly value the contribution of the model discussed here.

As robustness checks, I tried matching, fixed effects, and difference-in-difference models in the Brazilian data. With these non-parametric modeling approaches, I found similar results to those presented here. To mitigate the possibility of selection bias and reduce model dependence, I use matching to achieve balance between treatment (co-partisan mayors) and control groups across all observed covariates. Matching seeks to create a sample of treatment group observations that look as similar as possible to the control group. Balance between treatment and control groups allows analysis that is less sensitive to choices of functional form and model selection while also reducing bias and variance (Ho et al., 2007). Having identified a matched sample, I then ran analyses to estimate the treatment effect of PT power on municipal transfers. The coefficient on the PT treatment indicator was positive and statistically significant in every specification. The fact that PT mayors are effective at getting transfers for municipalities may reveal both their ability to persuade the president, and the presidential interests in maintaining and expanding PT influence in Brazilian municipalities. The evidence explored here is coherent, no matter what methodological strategy is pursued. But more important, these findings set forth a novel research agenda that will build on the empirical patterns observed here.

CHAPTER 6

Conclusion

This thesis uses a Bayesian model with elicited priors on a novel panel data set of local investments and transfer allocations to study variations in the provision of public goods benefits across municipalities in Brazil and in Mexico. I have described a Bayesian approach to fitting a two-part model for longitudinal government expenditure data. The model includes correlated random effects with, thus permitting the variance and correlation between the two model components to vary across allocation processes. Advantages of the approach include distinct modeling of zero and nonzero values; flexible modeling of time trends, and between-subject heterogeneity (both within and across processes); full posterior inference; and incorporation of prior information.

In my application, I was able to show that the Latin American presidents have to deal with a hard task: decide how to allocate resources when either voters or mayors do not support him. I portray such problem as the 'presidential dilemma' and present some evidence that shows a clear empirical pattern suggesting that co-partisan mayors are well more benefited than any other local officer. My results, therefore, join other empirical contributions that have demonstrated the importance of brokers in allocation choices. Using annual data on municipal transfers since the beginning of the century, I show that both the likelihood and the volume of municipal transfers is substantially and significantly higher in constituencies governed by president's party mayors. This is true regardless of the electoral support such localities provided to the president. However, when the Brazilian and the Mexican presidents had electoral support, but did not find co-partisans in local offices, they target oppositions in a weak situation, in what I interpret as an effort to transform opponents in potential partners.

I have focused primarily on the probit-lognormal two-part model commonly used to analyze expenditures in other areas of research (i.e. biostatistics, medicine, pharmacology). This model has the particular advantage of yielding closed-form full conditionals for many of the model parameters, resulting in efficient MCMC computation. Given the difficulty that common users have in writing MCMC algorithms, I implemented a JAGS code to use with this model that will allow many researchers to adapt to their studies. JAGS is a standard Bayesian software, available for any one interested in using Bayesian inference. The model can be easily adapted to allow for a logit link for the binomial component, as well as multivariate random effects (e.g., random slopes as well as random intercepts). The standard random-slope model implies a 4x4 covariance matrix that induces cross-correlations between the intercepts and slopes for the two components. The model can also be extended to accommodate multiple outcomes by assuming, for example, conditional independence between outcomes given class membership and subject-specific random effects. Alternatively, the factor-analytic approach recently proposed by Leiby et al. (2009) can be used to avoid the conditional independence assumption. Future work might also allow the covariance structures themselves to vary across classes, permitting, say, an AR1 structure in one class and a compound symmetric structure in another.

Substantively, I document the existence of tactical motivations in the allocation of federal transfers by the central government in four federal countries, aimed at targeting president's core voters. I demonstrate how presidential strategies about discretionary spending are conditioned by the level of governmental decentralization. Presidents under decentralized countries tend to target municipalities governed by co-partisan mayors, even when such municipalities are abundant in opposition voters. Presidents in centralized systems tend to distribute resources to municipalities with the plurality of core voters, even when such municipalities are governed by president's political rivals. In a decentralized system, the political benefits of federal spending accrue not just to the president, but also to local politicians. If powerful local politicians from parties different from the president's have access to resources, it is likely they will make voters believe their parties are responsible for the investments observed. This is a real threat to the president, as she is wasting resources and producing rivals for herself.

I believe that my results could extend to the politics of intergovernmental transfers in every presidential system where the president is strong, but has lost some budgetary authority. Particularly, my argument should apply where (a) the central government has some discretionary power in sharing central revenues among lower-level layers of government; (b) political credit spillovers exist between central and local governments in claiming credit for the transfers; and (c) political parties are not extremely weak and partisan affiliation shapes electoral competition at the local and national level. Perhaps, these results could also extend for other federal or unitary systems which levels of political, administrative and fiscal decentralization are high. The decentralized Bolivia, Colombia, and Peru should display discretionary spending being delivered to the presidents' local allies, whereas it should be possible to observe presidential voters targeted in the centralized Chile.

The evidence presented here shows that discretionary transfers in federal systems are allocated on political grounds even controlling for alternative explanatory variables for efficiency and equity. It is not that political factors explain all presidential choices. But the fact that lots os municipalities in Brazil and Mexico, as well as in Venezuela and in Argentina, do not receive these discretionary funds even having considerable levels of poverty, seems to suggest a special role that the political motivations play in determining presidential discretionary transfers. If one believes that redistribution should be a goal for newly inaugurated left presidents, then, the normative implications of my empirical findings call for a spending system that emphasizes constitutional rules based on necessity, rather than let the fiscal responsibility be with central governments, as they appear to be strongly influenced by political considerations. The results also suggests that presidents do not use public goods allocation to build legislative majorities in Congress, as voters and mayors political alignments to the president seem to matter more in explaining presidential allocative decisions.

From a more general perspective, my study emphasizes that (a) whoever allocates benefits may care about electoral outcomes at all levels, (b) capturing the top prize of the presidency in a decentralized context, although necessary, is not sufficient for a party interested in implementing a distributive agenda and enjoying the electoral benefits of it, and (c) conflictive political interests in multi-level systems may produce inefficient or slower changes. These emphases allow me to disclose a systematic intergovernmental pattern, namely, that presidents neglect core voters governed by opposition mayors in decentralized systems. But it also present a novel interpretation of federal politics, that is to say, the recognition that policy changes can be harder when politically motivated actors face decisions that not necessarily give them political credit. Hence my approach suggests that various theoretical models of distributive politics may benefit from taking a multilevel view.

The next steps of this research agenda involve mainly the investigation of the mechanisms to justify the actor's decisions laid out here. First, it is necessary to evaluate the claim that decentralization increases the political influence of local politicians in national politics. If local officers become more autonomous in deciding how resources are allocated in their districts, we should also expect them to have a bigger impact in the political life of such localities. Second, we need to investigate if voters respond to targeted spending rewarding the politician they believe is responsible for the provision of the benefits they care most about. In other words, how voters assign political credit for politicians when more than one officer has the legitimate authority to claim credit for outcomes observed locally. Both questions are part of a broader research agenda that I develop in other research papers.

Also, this paper projects the possibility of investigating whether or not these strategies pay off. The challenges for presidential spending presented by decentralization and vertically divided government characterizes president's party nationalization as a means to unified policy control. I expect that newly inaugurated presidents use central government authority to create a net of competent candidates to win subnational offices in Brazil and in Mexico, but induce opposition mayors to switch parties in Argentina and in Venezuela to jump at the presidential bandwagon. This tactical movement should explain in the long run why allocations vary by level of decentralization. Municipalities that swung to be governed by the president's party in future elections may have done so to receive more resources from the president.

Appendix

Eliciting Priors

There has been a recent surge in the use of expert knowledge in Bayesian models (Gill and Walker, 2005; Jackman, 2004; Western and Jackman, 1994). There are two reasons for this trend. First, the types of political science questions being proposed, particularly those pertinent to political decision-making, are characterized by uncertainty and paucity of empirical data. Even when data are available, they are invariably subject to error due to the size and complexity of political systems, resulting in parameter estimates with wide confidence intervals, leading to uninformative predictions. Second, decisions based on previous studies focussing on public policies and resource allocation are often required urgently. In situations such as these where hard data are lacking yet political decisions are required, the use of expert knowledge may provide a way forward. Yet for researchers wishing to use expert knowledge, questions remain regarding how to properly conduct an elicitation and use it in a model to address the political research questions.

Although frequentist techniques are evolving to accommodate expert knowledge (e.g. Lele and Allen (2006)), Bayesian methods are naturally suited to the incorporation of expert knowledge through priors; probability distributions representing what is known about the variable (Gelman et al. 2003). In brief, Bayesian modeling consists of four key elements: a prior probability distribution capturing prior knowledge about a parameter; data on the parameters captured through the likelihood; a model that describes the underlying process and incorporates both the likelihood and priors; and finally posterior estimates that result from combining the likelihood with the prior reflecting uncertainties captured by the model (McCarthy & Masters 2005; Cressie et al. 2009).

There has been considerable discussion about elicitation methods and how elicited information can be incorporated into a model as one or more priors to inform an analysis (Garthwaite & Dickey 1988; Steffey 1992; OHagan 1998). Here, I follow Kuhnert, Martin and Griffiths (2010) who provide a comprehensive guide to conducting an elicitation of priors through a modeling process that explicitly considers expert knowledge and its impact in a model. Between June and October of 2012 I conducted fieldwork research in Argentina, Brazil, Mexico, and Venezuela. During this period I collected fiscal, electoral, and socio-economic data for each of the municipalities of these four countries for the time period between 2000 and 2011. I also used this opportunity to carry out in-depth interviews with academics, politicians, and public officials from the national and subnational levels of government in the four countries (see list by country below). This information is used to produce elicited priors for the Bayesian model I am interested in fitting here.

Eliciting expert information needs careful structure, drawing on aspects of the social sciences (Gigerenzer 1996, 2002, 2007) to extract relevant information in an unbiased manner that is non-threatening to the expert. Furthermore, the process needs to align not only with the research question but with the model that will be used to incorporate the expert information. The main dangerous of elicitation is that experts are invariably subject to bias and depending on the nature of that bias, their opinion may influence conclusions. There are several scenarios that can arise when combining the likelihood with priors generated from expert opinion. Both the amount of data, mean, precision and the way in which the prior mean and precision is captured and incorporated into a model can influence the posterior estimate. In situations where data are limited, the expert's opinion has the potential to drive model predictions. The researcher therefore needs to be aware of the issues that can lead to bias and ensure that expert biases can be minimized.

As more data become available, the likelihood is moderated with the prior. However, in situations where the prior directly specifies the mean and precision, an informative prior can lead to a very informative posterior distribution, irrespective of the empirical data and how much data are collected (Lele & Allen 2006). If priors are incorporated into the model as an adjustment to an overall mean and precision, depending on their specification, the posterior estimates can be conservative. Here, the term adjustment refers to a shift in the mean or a rescaling of the precision, where the mean and precision are also considered random variables with appropriate priors attached. On my research I find value of using multiple experts in an elicitation exercise as the aggregation of multiple responses leads to an estimate of the uncertainty around the elicited quantity. It also represents a natural mechanism for feedback through the discussion and revision of opinion amongst experts.

Each of the experts I interviewed (see list below), provided me with guesses on the propensity and the level of investments presidents allocate to municipalities of the following types: municipalities in which the plurality of voters supported the current president in the presidential election and also elected a mayor from the president's party (core districts), municipalities in which the plurality of voters supported the president but did not elect a mayor from the president's party (core districts with rivals), municipalities in which the plurality of voters did not support the president but did elect a mayor from the president's party (opposition districts with allies), and municipalities in which the plurality of voters did not support either the president or her party's mayoral candidate in the most recent round of elections (opposition districts). These guesses were summarized in the Table 5.1.

Given the information (mean and range of distribution) gleaned from experts, I identified sensible and convenient parametric distributions for the experts' information. For the fixed effects, I assumed normal priors: $\alpha \sim N(\mu_\alpha, \Sigma_\alpha)$ and $\beta \sim N(\mu_\beta, \Sigma_\beta)$. I assumed that the prior hyper-parameters are identical across classes. Each Σ is assumed to have a conjugate Inverse-Wishart $IW(\nu_0, D_0)$ distribution. Since the other parameters of the model do not have a scientific interest for this paper, I assigned weakly informative proper distributions to them. This will ensure a well-identified model with proper posteriors determined almost entirely by the data. For, the lognormal precisions, τ^2, I assumed conjugate $Ga(\lambda, \delta)$ priors. Following Garrett and Zeger (2000) and Elliott et al. (2004) I assumed $\gamma \sim N_r[0, (9/4)I_r]$, which induces a prior for π_i centered at $1/K$ and bounded away from 0 and 1.[1]

[1]Prior precision specified through range method. Since α and β are assumed Normal, we can find Σ by $(\mu_{97.5} - \mu_{2.5})/4$ from $\mu \pm 2 \times \Sigma$.

Computation and Assessment of Model Fit

The posteriors are estimated using JAGS calling *rjags* 3-10 from *R* 2.15 (code available upon request). I ran 5 MCMC chains for 200,000 iterations each, discarding the first 50,000 as a burn-in to ensure that a steady-state distribution had been reached, and retained every 50th draw to reduce autocorrelation. Convergence is monitored by running multiple chains from dispersed initial values and then applying standard Bayesian diagnostics, such as trace plots; autocorrelation statistics; Geweke (1991) Z-diagnostic, which evaluates the mean and variance of parameters at various points in the chain; and the Brooks-Gelman-Rubin scale-reduction statistic $\hat{R}$, which compares the within-chain variation to the between-chain variation (Gelman et al., 2004). As a practical rule of thumb, a 0.975 quantile for $\hat{R} \leq 1.2$ is indicative of convergence. In the application below, convergence diagnostics were performed using the R package *boa* (Smith, 2007).

Figure .2 presents post-burn-in trace plots for twelve representative parameters from the random intercept and slope models: α_{11}, α_{12}, α_{21}, α_{22}, α_{31}, α_{32}, β_{11}, β_{12}, β_{21}, β_{22}, β_{31}, and β_{32}. For clarity of presentation, I have graphed only one of the four MCMC chains. The trajectory of lines suggest convergence and efficient mixing of the chains. The Geweke Z-diagostic p-values ranged from 0.11 (β_{22}) to 0.89 (α_{22}), indicating no significant difference in posterior means across regions of the chains; the 0.975 quantiles of the Brooks-Gelman-Rubin statistic were each less than 1.04, again indicating convergence of the chains. However, I did observe modest autocorrelation in the chains: the lag-10 autocorrelations ranged from 0.05 for α_{22} to 0.25 for ρ_2.

A well-known computational issue for Bayesian mixture models is 'label switching', which is caused by symmetry in the likelihood of the model parameters during the course of the MCMC run (Celeux, 2012; Jakobsson and Rosenberg, 2007). In some cases, label switching can be avoided by placing constraints on the class probabilities (Lenk and DeSarbo, 2000) or on the model parameters themselves (Congdon, 2005). As an alternative, Stephens (2000) proposed a relabeling algorithm that minimizes the posterior expected loss under a class of loss functions - which is the one used in the application below. Figure .1 presents good computational evidence for no auto-correlation in the several iterations of the MCMC, but it does show a moderate to low label switching trend. Even when I used Stephens' algorithm the problem persisted.

To assess the adequacy of the selected model, I use posterior predictive checking (Gelman, Meng and Stern, 1996), whereby the observed data are compared to data replicated from the posterior predictive distribution. If the model fits well, the replicated data, $\mathbf{y}^{rep}$, should resemble the observed data, $\mathbf{y}$. To quantify the similarity, I can choose a discrepancy measure, $T = T(\mathbf{y}, \boldsymbol{\theta})$, that takes an extreme value if the model is in conflict with the observed data. Popular choices for T include sample moments and quantiles, and residual-based measures.

The Bayesian predictive p-value (P_B) denotes the probability that the discrepancy measure based on the predictive sample, $T^{rep} = T(\mathbf{y}^{rep}, \boldsymbol{\theta})$, is more extreme than the observed measure T. A Monte Carlo estimate of P_B can be computed by evaluating the proportion of draws in which $T* > T$. A p-value close to 0.50 represents adequate model fit, while p-values near 0 or 1 indicate lack of fit. The cut-off for determining lack of fit is subjective, although by analogy to the classical p-value, a Bayesian p-value between 0.05 and 0.95 suggests adequate fit. In some cases, a stricter range, such as (0.20, 0.80), might be more appropriate. For the latent class two-part model, I follow the recommendations proposed by Neelon, O'Malley and Normand (2011). To assess the fit of the binomial component, I use T_1 (the proportion of observations greater than zero). For the nonzero observations, I use a modification of the omnibus chi-square measure proposed by Gelman et al. (2004).

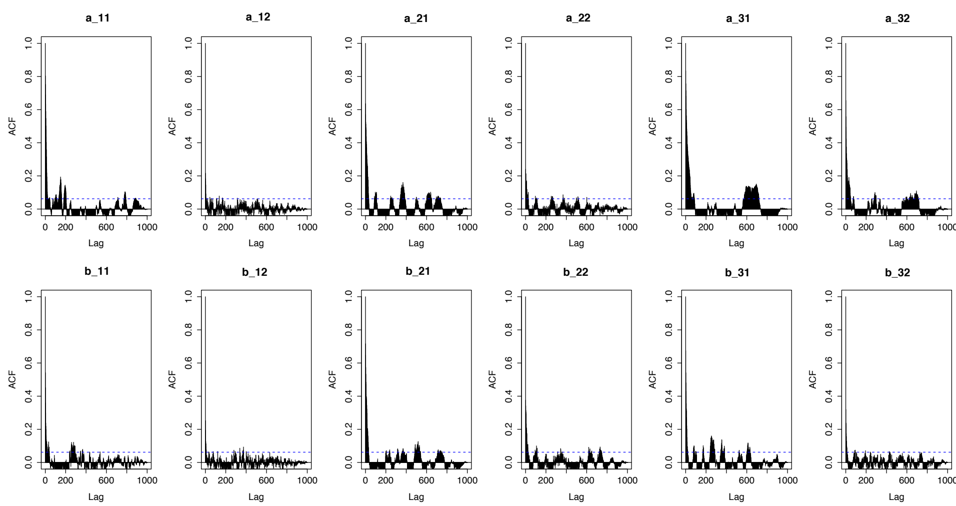


Figure .1: Auto-correlation plots for twelve representative parameters from the random intercept and slope correlated models, Brazil, and Mexico.

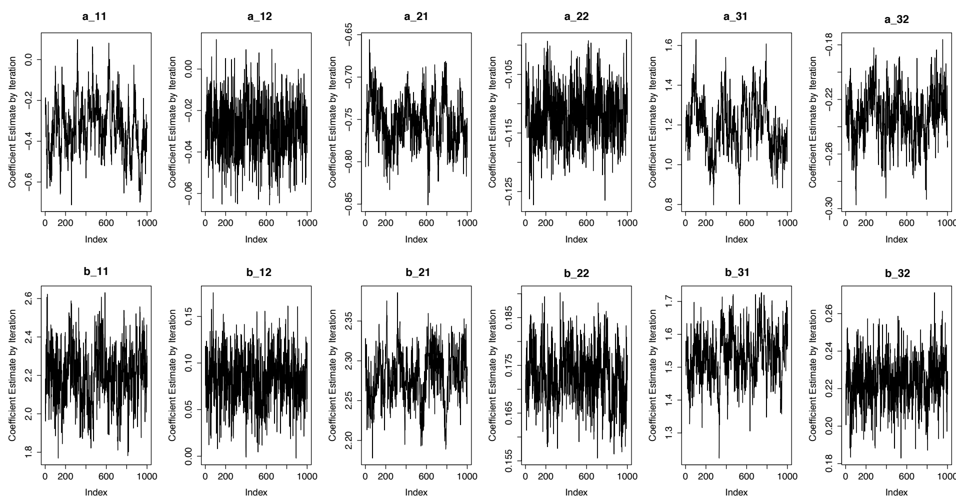


Figure .2: Trace plots based on one MCMC chain for twelve representative parameters from the random intercept and slope correlated models, Brazil, and Mexico.

Bibliography

Aitchison, John. 1955. "On the Distribution of a Positive Random Variable Having a Discrete Probability Mass at the Origin*." *Journal of the American Statistical Association* 50(271):901–908.

Amemiya, Takeshi. 1984. "Tobit Models: A Survey." *Journal of Econometrics* pp. 3–61.

Ames, Barry. 1994. "The Reverse Coattails Effect: Local Party Organization in the 1989 Brazilian Presidential Election." *The American Political Science Review* 88(1):95–111.

Angrist, J.D. and J.S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricistís Companion*. Princeton Univ Press.

Armesto, A. 2009. "Governors and mayors, particularism in the provision of local public goods in Mexico and Argentina". In *LASA 2009*.

Arulampalam, Wiji, Sugato Dasgupta, Amrita Dhillon and Bhaskar Dutta. 2009. "Electoral goals and center-state transfers: A theoretical model and empirical evidence from India." *Journal of Development Economics* 88(1):103–119.

Beck, Nathaniel and Jonathan N Katz. 2011. "Modeling Dynamics in Time-Series–Cross-Section Political Economy Data." *Annual Review of Political Science* 14(1):331–352.

Blackwell, Matthew. 2012. "A Framework for Dynamic Causal Inference in Political Science." *American Journal of Political Science* 57(2):504–520.

Bonvecchi, A. 2009. "The Dual Logic of Intergovernmental Transfers: Territorial Coalition Building in Argentina." *Latin American Studies Association* .

Brollo, Fernanda and Tommaso Nannicini. 2011. "Tying Your Enemy's Hands in Close Races: The Politics of Federal Transfers in Brazil." *IZA Discussion Paper* pp. 1–49.

Brollo, Fernanda and Tommaso Nannicini. 2012. "Tying Your Enemy's Hands in Close Races: The Politics of Federal Transfers in Brazil." *American Political Science Review* 106(04):742–761.

Browne, William J and David Draper. 2006. "A comparison of Bayesian and likelihood-based methods for fitting multilevel models." *Bayesian Analysis* 1(3):473–514.

Burgman, M. 2005. *Risks and Decisions for Conservation and Environmental Management*. Ecology, Biodiversity and Conservation Cambridge University Press.

Celeux, Gilles. 2012. "Bayesian Inference for Mixture: The Label Switching Problem." pp. 1–6.

Congdon, P P. 2005. *Bayesian Models for Categorical Data*. Wiley Series in Probability and Statistics Wiley.

Cooper, Nicola J, Paul C Lambert, Keith R Abrams and Alexander J Sutton. 2006. "Predicting costs over time using Bayesian Markov chain Monte Carlo methods: an application to early inflammatory polyarthritis." *Health Economics* 16(1):37–56.

Cox, G. 2006. "Swing voters, core voters and distributive politics." *Conference on Representation and Popular Rule, Yale University* .

Dixit, A and J Londregan. 1996. "The determinants of success of special interests in redistributive politics." *Journal of Politics* 58(4):1132–1155.

Duan, Naihua, Willard G Manning, Carl N Morris and Joseph P Newhouse. 1983. "A comparison of alternative models for the demand for medical care." *Journal of Business & Economic Statistics* 1(2):115–126.

Elliott, M R, J J Gallo, T R Ten Have, H R Bogner and I R Katz. 2004. "Using a Bayesian latent growth curve model to identify trajectories of positive affect and negative events following myocardial infarction." *Biostatistics* 6(1):119–143.

Fachelli, S and L Ronconi. 2004. "Poverty and employability effects of workfare programs in Argentina." *Poverty and Economic Policy Vol 34, n 2* .

French, B. 2008. "Analysis of longitudinal data to evaluate a policy change - French - 2008 - Statistics in Medicine - Wiley Online Library." *Statistics in Medicine* .

Garrett, Elizabeth S and Scott L Zeger. 2000. "Latent class model diagnosis." *Biometrics* 56(4):1055–1067.

Garthwaite, Paul H, Joseph B Kadane and Anthony O'hagan. 2005. "Statistical Methods for Eliciting Probability Distributions." *Journal of the American Statistical Association* 100(470):680–701.

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

Gelman, Andrew, Xiao-Li Meng and Hal Stern. 1996. "Posterior predictive assessment of model fitness via realized discrepancies." *Statistica Sinica* 6(4):733–760.

Gervasoni, Carlos. 1998. "El impacto de las reformas económicas en la coalición electoral justicialista (1989–1995)." *Boletín de la SAAP* 4(6).

Geweke, John. 1991. Evaluating the Accuracy of Sampling-based Approaches to the Calculation of Posterior Moments. In *Bayesian Statistics 4*, ed. J M Bernardo, J O Berger, A P Dawid and A F M Smith. Oxford University Press p. 42.

Geyer, Charles J and Elizabeth A Thompson. 1992. "Constrained Monte Carlo maximum likelihood for dependent data." *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 657–699.

Ghosh, P and P S Albert. 2009. "A Bayesian analysis for longitudinal semicontinuous data with an application to an acupuncture clinical trial." *Computational statistics & data analysis* .

Gibson, Edward L and Ernesto Calvo. 2000. "Federalism and low-maintenance constituencies: Territorial dimensions of economic reform in Argentina." *Studies in Comparative International Development (. . .* 35(3):32–55.

Gigerenzer, G. 1996. "The Psychology of Good Judgment: Frequency Formats and Simple Algorithms." *Medical Decision Making* 16(3):273–280.

Gill, Jeff and Lee D Walker. 2005. "Elicited priors for bayesian model specifications in political science research." *The Journal of Politics* 67(3):841–872.

Giraudy, Agustina. 2007. "The Distributive Politics of Emergency Employment Programs in Argentina (1993-2002)." *Latin American Research Review* 42(2):33–55.

Golden, Miriam. 2003. "Electoral connections: The effects of the personal vote on political patronage, bureaucracy and legislation in postwar Italy." *British Journal of Political Science* 33:189–212.

Greene, William H. 2003. *Econometric analysis*. Hydro Recordings.

Hallerberg, M., C.G. Scartascini and E. Stein. 2009. *Who Decides the Budget?: A Political Economy Analysis of the Budget Process in Latin America*. David Rockefeller/Inter-American Development Bank Inter-American Development Bank.

Heckman, James J. 1976. "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." *Annals of Economic and Social Measurement* 5(4):475–492.

Ho, Daniel, Kosuke Imai, Gary King and Elizabeth Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(3):199–236.

Jackman, Simon. 2000. "Estimation and inference via Bayesian simulation: An introduction to Markov chain Monte Carlo." *American Journal of Political Science* pp. 375–404.

Jackman, Simon. 2004. "Bayesian Analysis for Political Research." *Annual Review of Political Science* 7(1):483–505.

Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. John Wiley & Sons.

Jakobsson, M and N A Rosenberg. 2007. "CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure." *Bioinformatics* 23(14):1801–1806.

Jones, Mark and Wonjae Hwang. 2005. Provincial Party Bosses: Keystone of the Argentine Congress. In *Argentine Democracy. The Politics of Institutional Weakness*, ed. Steven Levitsky and María Victoria Murillo. University Park, Pennsylvannia: The Pennsylvania State University Press.

Kuhnert, Petra M, Tara G Martin and Shane P Griffiths. 2010. "A guide to eliciting and using expert knowledge in Bayesian ecological models." *Ecology Letters* 13(7):900–914.

Larcinese, Valentino, Leonzio Rizzo and Cecilia Testa. 2006. "Allocating the U.S. Federal Budget to the States: The Impact of the President." *The Journal of Politics* 68(2):447–56.

Lele, Subhash R and Kristie L Allen. 2006. "On using expert opinion in ecological analyses: a frequentist approach." *Environmetrics* 17(7):683–704.

Lenk, Peter J and Wayne S DeSarbo. 2000. "Bayesian inference for finite mixtures of generalized linear models with random effects." *Psychometrika* 65(1):93–119.

Levitsky, Steven. 2003. "From Labor Politics to Machine Politics: The Transformation of Party-Union Linkages in Argentine Peronism, 1983-1999." *Latin American Research Review* 38(3):3–36.

Litschig, S. 2008. "Financing Local Development: Quasi-Experimental Evidence from Municipalities in Brazil, 1980-1991." *Economics Working Papers* .

Lodola, German. 2005. "Protesta popular y redes clientelares en la Argentina: El reparto federal del Plan Trabajar (1996–2001)." *Desarrollo Económico* 44(176):515–35.

Luna, JP. 2010. "Segmented Party–Voter Linkages in Latin America: The Case of the UDI." *Journal of Latin American Studies* 42.

Luna, Juan Pablo and Rodrigo Mardones. 2010. Distributive Politics in a Non-Machine Party System: The Allocation of Targeted Social Funds and Subsidies in Chile (20002008). In *APSA 2009*.

Magaloni, Beatriz, Alberto Diaz-Cayeros and Federico Estevez. 2007. "Clientelism and portfolio diversification: a model of electoral investment with applications to Mexico". In *Patrons, clients, and policies: patterns of democratic accountability and political competition*, ed. Herbert Kitschelt and Steven Wilkinson. Cambridge University Press.

Mainwaring, S and Matthew Shugart. 1997. *Presidentialism and Democracy in Latin America*. Cambridge Studies in Comparative Politics Cambridge University Press.

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

Montero, A.P., D Samuels and Helen Kellogg Institute for International Studies. 2004. *Decentralization and democracy in Latin America*. Helen Kellogg Institute for International Studies University of Notre Dame Press.

Nazareno, M and S Stokes. 2006. *Interest and electoral dangers of public expenditure in Argentina*. Globalization and Inflation Desarrollo Economico-Revista De Ciencias . . . .

Neelon, Brian, A. James O'Malley and Sharon-Lise T. Normand. 2011. "A Bayesian Two-Part Latent Class Model for Longitudinal Medical Expenditure Data: Assessing the Impact of Mental Health and Substance Abuse Parity." *Biometrics* 67(1):280–289.

Nichter, Simeon. 2008. "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot." *American Political Science Association* 102(01).

Nunes, Felipe. 2011. "The Presidential Dilemma: How do Federal Governments Allocate Public Goods to Municipalities in Brazil?" Working Paper.

Olsen, MK and JL Schafer. 2001. "A two-part random-effects model for semi-continuous longitudinal data." *Journal of the American Statistical Association* 96(454):730–745.

O'Neill, Katheleen. 2006. *Who's Tugging at the Purse Strings? The Political Economy of Intergovernmental Transfers in Argentina, Mexico and Peru*. Edward Elgar pp. 257–277.

Penfold-Becerra, M. 2004. "Federalism and Institutional Change in Venezuela." *Federalism and democracy in Latin America* pp. 197–225.

Rizopoulos, Dimitris, Geert Verbeke, Emmanuel Lesaffre and Yves Vanrenterghem. 2008. "A Two-Part Joint Model for the Analysis of Survival and Longitudinal Binary Data with Excess Zeros." *Biometrics* 64(2):611–619.

Rovallion, Martin. 1998. "Reaching Poor Areas in a Federal System." *The World Bank* .

Smith, Brian J. 2007. "boa: an R package for MCMC output convergence assessment and posterior inference." *Journal of Statistical Software* 21(11):1–37.

Sørensen, RJ. 1995. "The demand for local government goods." *European Journal of Political Research* .

Stephens, Matthew. 2000. "Dealing with Label Switching in Mixture Models." *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 62(4):795–809.

Stokes, Susan. 2005. "Perverse accountability: A formal model of machine politics with evidence from Argentina." *American Political Science Review* 99(03):315–325.

Su, Li, Brian D. M. Tom and Vernon T. Farewell. 2009. "Bias in 2-part mixed models for longitudinal semicontinuous data." *Biostatistics* 10(2):374–389.

Su, Li, Brian Dm Tom and Vernon T. Farewell. 2011. "A likelihood-based two-part marginal model for longitudinal semi-continuous data." *Statistical methods in medical research* .

Tobin, James. 1958. "Estimation of relationships for limited dependent variables." *Econometrica: journal of the Econometric Society* pp. 24–36.

Tooze, Janet A, Gary K Grunwald and Richard H Jones. 2002. "Analysis of repeated measures data with clumping at zero." *Statistical methods in medical research* 11(4):341–355.

Weiss, Robert E. 2005. *Modeling longitudinal data.* Springer Verlag.

Weiss, Robert E. 2012. *Bayesian Methods for Modeling Data.* Unpublished manuscript.

Western, Bruce and Simon Jackman. 1994. "Bayesian inference for comparative research." *The American Political Science Review* pp. 412–423.

Xie, Haiyi, Gregory McHugo, Anjana Sengupta, Robin Clark and Robert Drake. 2004. "A Method for Analyzing Longitudinal Outcomes with Many Zeros." *Mental Health Services Research* 6(4):239–246.

Zucco, Jr, Cesar. 2008. "The President's 'New'Constituency: Lula and the Pragmatic Vote in Brazil's 2006 Presidential Elections." *Journal of Latin American Studies* 40(01):29–49.