

UCLA

UCLA Electronic Theses and Dissertations

Title

A Bayesian Multilevel Linear Regression Model and Its Application to Distributive Politics in Korea, 2005-2006

Permalink

<https://escholarship.org/uc/item/3x77k2z6>

Author

Koo, Bon Sang

Publication Date

2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

**A Bayesian Multilevel Linear Regression Model
and Its Application to Distributive Politics in
Korea, 2005–2006**

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

by

Bon Sang Koo

2013

© Copyright by
Bon Sang Koo
2013

ABSTRACT OF THE THESIS

**A Bayesian Multilevel Linear Regression Model
and Its Application to Distributive Politics in
Korea, 2005–2006**

by

Bon Sang Koo

Master of Science in Statistics

University of California, Los Angeles, 2013

Professor Mark Handcock, Chair

This thesis aims to properly test three hypotheses derived from existing political theories about distributive politics by employing Bayesian multilevel modeling. The specific case of intergovernmental grants, the *Special Local Allocation Grants*, in Korea at two nested levels (districts and provinces) from 2005 to 2006, verifies that unlike classical regression models, the Bayesian multilevel regression model can capture regional variations in the allocation and utilize substantive knowledge from previous literature. In particular, the model finds that a significant positive association between the amount of intergovernmental grants and being an electorally unstable province in a broad region affected by regional voting behavior (i.e., Electorally Unstable Provinces Hypothesis) even after controlling for the need-based criteria. It justifies the chief executive’s strategy to target an electorally unstable [swing] province even within a supporter region because people in the electorally stable province are strongly affiliated with a regional (or ethnic) identity so that they may be satisfied with the allocation of grants even if they are not the main beneficiaries. Thus, while the allocation is concentrated on core supporters that are well known quantities at the district level, the allocation at the higher level can be decided by the efficient targeting strategy. This find-

ing provides a strong implication for decentralized democratic governments under circumstances where significant regionally (or ethnically) affiliated-voting is observed. In Korea, the disproportional allocations of central government grants to electorally unstable province within its supporter region (Jeolla) from 2005 to 2008 helped the government party to increase its vote share in the unstable province (Jeonam) by 20% in 2008. The vote share in the electorally stable province (Jeonbuk) reduced only by 2.9% in comparison with that in the previous election. It was a remarkable outcome, considering the the government party was defeated by the wide margin 13% nationwide and experienced a swing against it by 20.3% across the country in the election.

The thesis of Bon Sang Koo is approved.

Ying Nian Wu

Juana Sanchez

Mark Handcock, Committee Chair

University of California, Los Angeles

2013

To my parents, and Ji Yean

TABLE OF CONTENTS

1	Introduction	1
2	Methods	3
2.1	Complete-Pooling and No-pooling Models	3
2.2	Partially Pooling Model	5
3	Application: Distributive Politics in Korea, <i>The Special Local Allocation Grants (2005–2006)</i>	9
3.1	Background Information about Korea	9
3.2	Existing Explanations of Distributive Politics in Korea	12
3.3	Testable Hypotheses	13
3.4	Data and Measures	17
3.5	Complete-Pooling and No-pooling Regression Models	19
3.6	Multilevel Linear Models in the Bayesian Framework	21
3.7	Computation with JAGS	24
3.8	Results	27
3.9	Sensitivity Analysis	31
3.10	Discussion	32
4	Conclusion	43
	References	51

LIST OF FIGURES

- 3.1 This figure shows the result of the Korean presidential election of 2002 by province. Blue areas represent provinces (or special or metropolitan cities) where Lee Hoi-chang, a candidate of the main conservative party (Grand National Party) earned the majority of the vote. Green areas indicate provinces (or special or metropolitan cities) where Roh Moo-hyun, a candidate of the ruling party (Unified New Democratic Party) won the majority of the vote. Roh Moo-hyun won the election (48.9% vs. 46.6%) nationwide. The percentages below provinces (or metropolitan cities) represent Roh's vote share in the provinces [ROK03]. 10

3.2 This figure shows an integrated theory about the relationship between intergovernmental grants and political players (represented as circles) who have different electoral incentives. We can either observe or proxy the solid lines. For instance, local governments directly request to favorably allocate the grants to their regions, or ask legislators elected in districts that belongs to their regions to influence the allocation process. Individual legislators may make a request to their party organizations for the allocation (path (i)). Senior members of the committee related to the allocation can directly influence the process (path (ii)). Legislators avowedly claim credit for successful allocation of the grants. Unlike legislators, the president has strong discretion over allocation of all kinds of government resources, but tends to evade being engaged in the allocation process. Therefore, it is not observed whether or not the president has influenced the allocation process (path (iii)). It is also not observable whether or not legislators who have a personal relationship with the president ask him to influence the process (dashed lines). 14

3.3 Posterior distributions of district-level predictors 37

3.4 This plot illustrates random effects represented as varying intercepts across provinces. Purple solid dots indicate the posterior medians of intercepts obtained by the informative priors. Dark green hollow dots represent the posterior medians of intercepts obtained by reference priors (i.e., uninformative priors). While thick lines represent 50% confidence intervals, thin lines indicate 95% confidence intervals. The dark red dashed-line indicates the median of intercepts that vary across provinces. The intercepts of three electorally unstable provinces (Chungnam, Jeonam, and Gyeongnam) are greater than their stable counterparts (Chungbuk, Jeonbuk, and Gyeongbuk, respectively). 38

3.5 This figure illustrates the results from the multilevel model that allows both intercepts and slopes to vary across provinces. Here intercepts and slopes used in this figure represent the mean values of the posterior distributions obtained from 1,000,000 iterations with the 10,000 burn-in. 39

3.6 This plot illustrates random effects represented as varying intercepts across provinces when we use two models shown in 3.2 and 3.3. Purple solid dots indicate the posterior medians of intercepts obtained by the multilevel linear regression with the group predictor. Dark green hollow dots represent the posterior medians of intercepts obtained by the multilevel linear regression without any group predictor. While thick lines represent 50% confidence intervals, thin lines indicate 95% confidence intervals. The dark red dashed-line indicates the median of intercepts that vary across provinces. The intercepts of three electorally unstable provinces (Chungnam, Jeonam, and Gyeongnam) are greater than their stable counterparts (Chungbuk, Jeonbuk, and Gyeongbuk, respectively). 40

3.7	This figure illustrates how sensitive regional intercepts (α_j 's) are to variance of coefficient for vote margin (β_4).	41
3.8	This figure illustrates how sensitive regional intercepts (α_j 's) are to variance of coefficient for vote margin (β_4).	42
4.1	This figure shows scatterplots between possible pairs of components of the special local allocation grants. <code>disaster05</code> and <code>disaster06</code> indicate the grants allocated for recovery from natural disasters in 2005 and 2006, respectively. <code>regproj05</code> and <code>regproj06</code> represent the grants allocated for the two components (regional needs, and pilot projects) in 2005 and 2006, respectively. Likewise, <code>well05</code> and <code>well06</code> indicate the grants allocated for financial incentives for well-performed local governments in 2005 and 2006, respectively. The black dotted-lines indicate linear best fit lines. Diagonals show densities of the components.	46
4.2	This figure shows posterior distributions of varying intercepts (α_j) when the MCMC runs 1,000,000 iterations (50,000 burn-in). Blue dashed-lines indicate 95% confidence intervals, and a red dashed-line presents the median of the posterior distribution in each province. The thick green line indicates the median of the intercepts that vary across provinces. We observe that despite the huge number of iterations the posterior distributions for some provinces such as Ulsan and Incheon are still not bell-shaped in t -distribution sampling model.	50

LIST OF TABLES

3.1	This table presents results of complete-pooling and no-pooling models.	20
3.2	This table presents the posterior distributions of coefficients for a province-level predictor by using the t -distribution model with a province-level predictor, the swing index by province. I used the informative priors, and 1,000,000 iterations (with the 1,000 burn-in).	30
3.3	This table summarizes three broad regions and six provinces where regional identities are found in Korea. The figures in parentheses indicate the rankings of posterior medians of random effects represented as varying intercepts across 16 provinces including Special City and Metropolitan cities, and the government party's vote shares in the party-list, respectively [ROK04]. As Figure 3.4 illustrates, the more <i>Special Local Allocation Grants</i> were allocated to electorally unstable [swing] provinces (Chungnam, Gyeongnam, and Jeonam) in comparison to the counterparts (Chungbuk, Gyeongbuk, and Jeonbuk). We also find the president's (or the government party's) targeting strategy that focuses on electorally unstable province (Jeonam) within his (or its) supporter region (the Jeolla region), considering Gwangju, a metropolitan city geographically located within the Jeonam province, is at the 1st in the Figure 3.4. By contrast, Daejeon metropolitan city placed within the Chungcheong region is at the bottom, and Daegu metropolitan city located within the Gyeongsang region, is ranked 15th out of 16 provinces.	33

4.1 This table presents different results of a complete-pooling model by dependent variable: four components of the special local allocation grants. (1), (2), (3), and (4) represent natural disasters, regional needs, pilot projects, and financial incentives to well-performed local units, respectively. (1)+(2)+(3)+(4) means the average total grants distributed to districts for the two years (2005-6). Likewise, (2)+(3) also means the average grants allocated for regional needs plus pilot projects, and (4) indicates the average grants allocated for financial incentives to well-performed administrative units. It is notable that the component of pilot projects was created in 2006. 47

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor Professor Mark Hancock for the continuous support of my M.S. study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the research and writing of this thesis. Besides my advisor, I would like to thank the rest of my thesis committee: Professor Yingnian Wu, and Professor Juana Sanchez. My sincere thanks also goes to Professor Jeffrey Lewis for his continuous support along the way. Without him I could not have imagined completing my M.S. study.

CHAPTER 1

Introduction

Distributive politics involves the link between electoral competition and allocations of public spending to geographical locations, and has been one of the main subjects in political science. In particular, comparative politics researchers who intend to explain differences as well as similarities among countries (or among regions within a country) have paid much attention to distributive politics [GM13]. Some theoretical models provide logics behind governments' (or government parties') tactical apportionment of targetable goods at an electoral district level [AP01, CM86, DL96, DL98, LW87]. For instance, emphasizing the machine-core constituent link that guarantees voters' response to rewards, [CM86] predict that political parties will distribute targetable goods to core supporters (or the electoral districts of their supporters).

These theoretical models have been empirically tested by limited methods, however. Traditional qualitative research emphasized either legislators' personal attributes such as seniority or voters' choices motivated by ethnic identities. Their arguments were supported by stylized episodes rather than by rigorous statistical methods. Even recent studies with quantitative methods ignore potentially valuable information from different nested-levels by focusing on either general patterns or distinctive features in some regions.

Considering the common feature that regional (or ethnic) identities are usually shared by voters beyond electoral district boundaries, a government's (or a government party's) targeting strategy at a broader regional level is not necessarily

identical to that at a district level. To identify targeting strategies which may differ by relevant level is more interesting to current comparative politics. These motivates us to pay attention to multilevel modeling as an alternative method.

This thesis aims to properly test hypotheses derived from existing political theories by utilizing Bayesian multilevel linear regression models. The Bayesian approach allows researchers to take advantage of our substantive information from previous qualitative as well as quantitative studies in the process of prior specification. The multilevel models will be applied to the Korean case which invites more careful statistical inferences as well as more suitable data in order not to be misled by spurious results from single-minded regressions. For this purpose, I will examine a specific case of intergovernmental grants, the *Special Local Allocation Grants*, in Korea at two nested levels (districts and provinces) from 2005 to 2006.

This thesis proceeds as follows. In chapter 2, I will review traditional regression models (complete-pooling and no-pooling regression models), and introduce multilevel linear regression models as an alternative method. In chapter 3 I will provide background information about distributive politics in Korea, and construct a theoretical framework to integrate fragmented explanations. Based on the framework, I will specify three testable hypotheses and underlying assumptions regarding the president (or parties) incentives and regionalism. Then I will present empirical models to test hypotheses on the allocation of intergovernmental grants by employing statistical techniques including Bayesian multilevel modeling. After presenting the results, I will discuss the substantial meaning of them, and then conduct a brief sensitivity analysis as well. In chapter 4, I will summarize my findings and discuss limitations of this research.

CHAPTER 2

Methods

2.1 Complete-Pooling and No-pooling Models

While complete-pooling regression does not include group indicators in a model, no-pooling fit separate models within each group [GH06]. The former regression completely ignore the group-level information such as variation among groups. The latter contains group-level indicators mainly written as categorical predictor, but does not include group-level predictors. These two extremes are traditional approaches to allocation of government resources in quantitative political science literature.

Classical complete-pooling regression models can be written as

$$y_i = \alpha + \mathbf{X}_i \boldsymbol{\beta} + \epsilon_i \quad (2.1)$$

where y_i is a column vector of observations, \mathbf{X} is the matrix with individual-level predictors, $\boldsymbol{\beta}$ is the column vector of the regression coefficients of the predictors. i is an individual-level indicator, and ϵ_i indicates residuals.

A generic form of no-pooling regression can be written as

$$y_i = \alpha_{j[i]} + \mathbf{X}_i \boldsymbol{\beta} + \epsilon_i \quad (2.2)$$

where $j[i]$ is the group corresponding to an individual unit i .

Suppose a simple scenario where there is no predictor \mathbf{X} , and observations y_i is normally distributed. Then the equation (2.2) can be expressed as

$$y_{ij}|\alpha_j, \sigma^2 \sim \mathcal{N}(\alpha_j, \sigma^2) \quad (2.3)$$

The complete-pooling model imposes the restriction that $\alpha_j = \alpha, \forall j$, and the maximum likelihood estimate of the grand mean can be written as

$$E(\alpha_j|\mathbf{y}) = \bar{y} = \frac{\sum_{j=1}^J \sum_{i=1}^{n_j} y_{ij}}{\sum_{j=1}^J n_j} \quad (2.4)$$

In a no-pooling model, the maximum likelihood estimate of the mean of y in group j can be obtained as follows:

$$E(\alpha_j|\mathbf{y}) = \bar{y}_j = n_j^{-1} \sum_{i=1}^{n_j} y_{ij} \quad (2.5)$$

These two estimators are traditional approaches to political phenomena in political science, and in particular the no-pooling models are called ‘fixed-effect’ regression models, which are preferred by comparative political scientists who tend to emphasize regional peculiarities. However, it is noted that we may give up information in the between-group distribution of the α_j when we use such no-pooling models.

2.2 Partially Pooling Model

The two traditional extremes which impose strong restrictions on either within-group variance or between-group variance are limited in distributive politics research which emphasizes regional variation in allocations of government resources as well as general distribution patterns. We can utilize all possible information by using a partially pooling model, and expect a better explanation of allocation of government resources. We can write partially pooling regression models as

$$y_i = \alpha_{j[i]} + \mathbf{X}\boldsymbol{\beta}_{j[i]} + \epsilon_i \quad (2.6)$$

The equation (2.6) can be differently expressed as

$$y_i \sim \mathcal{N}(\alpha_{j[i]} + \boldsymbol{\beta}_{j[i]}\mathbf{X}, \sigma_y^2) \quad (2.7)$$

Again, suppose a simple scenario where there is no predictor, and an observation y_i is normally distributed. Then the model can be expressed as

$$\begin{aligned} y_{ij} &\sim \mathcal{N}(\alpha_j, \sigma_j^2) \\ \alpha_j &\sim \mathcal{N}(\mu_\alpha, \sigma_\alpha) \end{aligned} \quad (2.8)$$

where $i = 1, \dots, n_j$ within group j , $j = 1, \dots, J$. Then we have a partially pooling model as follows:

$$\alpha_j | \mathbf{y}_j, \sigma_j^2, \mu_\alpha, \sigma_\alpha \sim \mathcal{N}(\mu_j^P, V_j) \quad (2.9)$$

where

$$\mu_j^P = \frac{\frac{1}{\sigma_\alpha^2}\mu_\alpha + \frac{n_j}{\sigma_j^2}\bar{y}_j}{\frac{1}{\sigma_\alpha^2} + \frac{n_j}{\sigma_j^2}}, \quad \text{and} \quad V_j = \left(\frac{1}{\sigma_\alpha^2} + \frac{n_j}{\sigma_j^2} \right)^{-1}, \quad (2.10)$$

and where $\bar{y}_j = n_j^{-1} \sum_{i=1}^{n_j} y_{ij}$ is the maximum likelihood estimate of α_j by equation (2.5).

Based on the logic above, we can think a multilevel model with conjugate priors in a Bayesian framework.

$$\begin{aligned}
y_{ij} | \alpha_j, \sigma^2 &\sim \mathcal{N}(\alpha_j, \sigma^2) \\
\alpha_j | \mu_\alpha, \sigma_\alpha^2 &\sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2) \\
\mu_\alpha &\sim \mathcal{N}(b_0, B_0) \\
\sigma^2 &\sim \text{Inverse-Gamma}(\nu_0/2, \nu_0 \sigma_0^2/2) \\
\sigma_\alpha^2 &\sim \text{Inverse-Gamma}(\kappa_0/2, \kappa_0 \sigma_{\alpha 0}^2/2)
\end{aligned} \tag{2.11}$$

where $\nu_0, \sigma_0^2, \kappa_0$, and $\sigma_{\alpha 0}^2$ are the hyperparameters.

To implement a Gibb sampler for the model, we need to build full conditional distributions included in Model (2.11). First, by Bayes rule, we obtain the full conditional density of α_j as follows:

$$p(\alpha_j | y_{ij}, \sigma^2, \mu_\alpha, \sigma_\alpha, b_0, B_0, \kappa_0, \sigma_{\alpha 0}^2) \propto p(\alpha_j | \mu_\alpha, \sigma_\alpha) \prod_{i=1}^{n_j} p(y_{ij} | \alpha_j, \sigma^2)$$

By (2.9), (2.10), and (2.11), we can easily derive the full conditional density of α_j which is normally distributed as follows:

$$\alpha_j | y_{ij}, \sigma^2, \mu_\alpha, \sigma_\alpha \sim \mathcal{N} \left(\frac{\frac{1}{\sigma_\alpha^2} \mu_\alpha + \frac{n_j}{\sigma^2} \bar{y}_j}{\frac{1}{\sigma_\alpha^2} + \frac{n_j}{\sigma^2}}, \left(\frac{1}{\sigma_\alpha^2} + \frac{n_j}{\sigma^2} \right)^{-1} \right) \tag{2.12}$$

where $\bar{y}_j = n_j^{-1} \sum_{i=1}^{n_j} y_{ij}$.

Second, we can obtain the conditional distribution of μ_α by using

$$p(\mu_\alpha | \alpha_j, \sigma_\alpha^2, b_0, B_0) \propto p(\mu_\alpha | b_0, B_0) \prod_{j=1}^J p(\alpha_j | \mu_\alpha, \sigma_\alpha^2)$$

As shown in (2.12), we can obtain the full conditional density of μ_α which is normally distributed as follows:

$$p(\mu_\alpha | \alpha_j, \sigma_\alpha^2, b_0, B_0) \sim \mathcal{N} \left(\frac{B_0^{-1} b_0 + \frac{J}{\sigma_\alpha^2} \bar{\alpha}}{B_0^{-1} + \frac{J}{\sigma_\alpha^2}}, \left(B_0^{-1} + \frac{J}{\sigma_\alpha^2} \right)^{-1} \right) \quad (2.13)$$

where $\bar{\alpha} = \sum_{j=1}^J \alpha_j / J$.

Third, the full conditional density of σ_α^2 which is assumed to be positive can be written as

$$p(\sigma_\alpha^2 | \alpha_j, \mu_\alpha, \kappa_0, \sigma_{\alpha 0}) \propto p(\sigma_\alpha^2 | \kappa_0, \sigma_{\alpha 0}) \prod_{j=1}^J p(\alpha_j | \mu_\alpha, \sigma_\alpha^2)$$

For mathematical convenience, we assume that σ_α^2 has an inverse-Gamma density as in (2.11). Thus, we can write the full conditional density of σ_α^2 as

$$\sigma_\alpha^2 | \alpha_j, \mu_\alpha, \kappa_0, \sigma_{\alpha 0} \sim \text{Inverse-Gamma} \left(\frac{\kappa_0 + J}{2}, \frac{\kappa_0 \sigma_{\alpha 0}^2 + \sum_{j=1}^J (\alpha_j - \mu_\alpha)^2}{2} \right) \quad (2.14)$$

Likewise the full conditional density of variance σ^2 can be expressed as

$$p(\sigma^2 | y_{ij}, \alpha_j, \nu_0, \sigma_0^2) \propto p(\sigma^2 | \nu_0, \sigma_0^2) \prod_{j=1}^J \prod_{i=1}^{n_j} p(y_{ij} | \alpha_j, \sigma^2)$$

The posterior density of σ^2 can be also an Inverse-Gamma density

$$\sigma^2 | y_{ij}, \alpha_j, \nu_0, \sigma_0^2 \sim \text{Inverse-Gamma} \left(\frac{\nu_0 + \sum_{j=1}^J n_j}{2}, \frac{\nu_0 \sigma_0^2 + \sum_{j=1}^J \sum_{i=1}^{n_j} (y_{ij} - \alpha_j)^2}{2} \right) \quad (2.15)$$

Employing the conditional distributions (2.12) - (2.15), we can implement the Gibb sampler for the conjugate normal multilevel model, and obtain marginal posterior distributions of all parameters in which we are interested in. Following the same steps we can build a Gibb sampler for multilevel linear models with predictors [GH06, Hof09, Jac09]. For convenience and efficiency I use **JAGS** to implement a Gibb sampler for several Bayesian multilevel linear models in the application chapter.

CHAPTER 3

Application: Distributive Politics in Korea, *The Special Local Allocation Grants (2005–2006)*

3.1 Background Information about Korea

Korea consists of six broad regions (Seoul-Metro, Chungcheong, Gangwon, Jeolla, Gyeongsang, and Jeju). It is worthwhile to note that two regions have long been political rivals. The Jeolla region is located in the southwest of Korea, and consists of two provinces: the northern Jeolla (Jeonbuk) and the southern Jeolla (Jeonam). The Gyeongsang region is situated in the southeast of Korea, and also comprised of two provinces: the northern Gyeongsang (Gyeongbuk) and the southern Gyeongsang (Gyeongnam).¹ The rivalry has clearly been revealed in presidential elections since the 1970s. For example, Kim Dae-jung, a native of the Jeolla region, won an 89% of the Jeolla vote but only 9% of the Gyeongsang vote in the presidential election of 1992. On the other hand, his main opponent, Kim Young-sam, who was born in the Gyeongsang region, won 69% of the Gyeongsang vote but only 5% of the Jeolla vote. Indeed, such a regional affinity to a particular political leader with the same regional background has been found to be the best predictor for Korean voters' behavior [Cho00, Lee98].

Under the political circumstance it has been widely believed that a president would allocate a disproportionately larger amount of intergovernmental transfers to regions with more loyal supporters [HL08]. From the 1960s to late 1990s, all

¹Historically, both the Gyeongsang and the Jeolla Regions have formed as administrative regions since 1018 when King Hyeonjong ruled during the Goryeo dynasty.

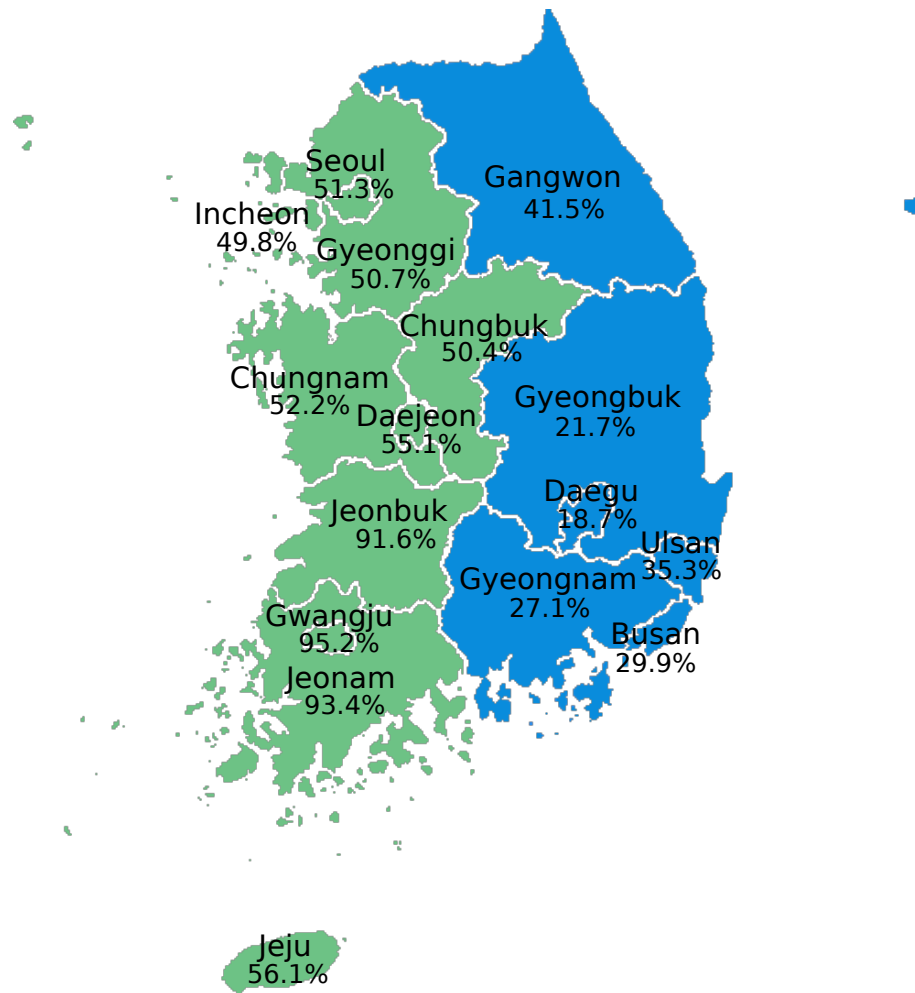


Figure 3.1: This figure shows the result of the Korean presidential election of 2002 by province. Blue areas represent provinces (or special or metropolitan cities) where Lee Hoi-chang, a candidate of the main conservative party (Grand National Party) earned the majority of the vote. Green areas indicate provinces (or special or metropolitan cities) where Roh Moo-hyun, a candidate of the ruling party (Unified New Democratic Party) won the majority of the vote. Roh Moo-hyun won the election (48.9% vs. 46.6%) nationwide. The percentages below provinces (or metropolitan cities) represent Roh's vote share in the provinces [ROK03].

of the presidents (both before and after the transition to democracy) came from one of the Gyeongsang provinces, which were the major beneficiaries of unfair economic policy. This belief was fortified even after the peaceful change of political power in 1997 by Kim Dae-jung, the long-term opposition leader. Due to his political indebtedness to his supporter region, the Jeolla region, which

had experienced unequal economic development during the presidencies of presidents from the Gyeongsang region, President Kim Dae-jung would compensate “his” provinces by disproportionately allocating central government expenditures for geographically concentrated public projects during his presidency, 1998–2003. Therefore, most of the literature mainly investigates a link between presidential election outcomes and the allocation of public expenditures [Kwo05, HL04].

The long-lasting belief has been steadily weakening since former human rights lawyer Roh Moo-hyun won the presidential election of 2002, however. First, since President Roh started his career as a politician after the transition to democracy, he was relatively freer of such political indebtedness to his supporter provinces than his predecessors. Second, although his hometown is in one of the Gyeongsang provinces, he was ironically supported by the rival provinces, the Jeolla provinces. In this sense, we may raise a general question about his preferences as to targeting regions: his hometown, the main supporter provinces, the electoral districts of his close circle, or swing provinces that could secure votes for his key policies. Third, his non-authoritative leadership style made it more difficult to observe his preferences regarding targeting regions. President Roh attempted to break from the convention, declaring that he would not be involved in the allocation process of *Special Local Allocation Grants*, which were regarded as a main resource to secure president’s influence in politics. It is true that the amount of the *Special Local Allocation Grants* in 2006 was gradually reduced to the half of what it was before. These new aspects of the Roh administration lead to recent approaches that focus on either individual legislators’ or major political parties’ roles in the allocation [HK09, CK08] rather than on the president’s role.

3.2 Existing Explanations of Distributive Politics in Korea

Many Korean scholars have focused on the president's role in the processes of intergovernmental transfers under the theoretical framework of so-called *imperial presidency*. For instance, Horiuchi and Lee [HL04] argue that government expenditures are disproportionately allocated to regions affiliated with the president. They implicitly set the president as a final decision-maker in the allocation process. Indeed, it was a Korean political tradition that presidents have implicitly or explicitly influenced the nomination of candidates in legislative elections, and placed their confidants on key positions in the party organization. As part of disclosing some political scandals, social activist organizations (e.g. the Citizens' Coalition for Economic Justice) pointed out that legislators' personal connection to the president has played an important role in the allocation processes in this institutional setting. Such fragmentary evidence, however, is not sufficient to prove the causal link between a personal connection with the president and the distribution of public expenditures itself.

Moreover, it is difficult to apply well-established theories in political science [CM86, DL98, LW87] to analyses that set a president as a key actor in distributive politics because such theories are constructed with one underlying assumption: a player who can decide who are targeted has an electoral incentive to be reelected. It is important to note that the president cannot seek reelection but still has the implicit discretionary power on the allocation process in Korea. Without any additional assumptions about the president incentive structure (i.e., incentive to be reelected), empirical tests as well as the theoretical logic at work might be flawed. That is, in order to consider the presidential role within the context of legislative politics and allocations of government expenditures, we need to specify an assumption about the president incentive structure when an incentive to be reelected is not involved.

Even after we theoretically specify the assumption about president's incen-

tives, we will still be faced with an empirical challenge: we cannot observe the actual role of the president in the allocation process. Instead, we can only infer the president's role in the unexplained variance in statistical models. One alternative way to statistically infer the president's explanations of the link between electoral incentives and the allocation of intergovernmental grants in Korea, I attempt to integrate hypotheses constructed at different levels, which is illustrated in Figure 3.2. I will test those hypotheses by employing a more suitable dataset on the *Special Local Allocation Grants* as well as appropriate statistical models.

3.3 Testable Hypotheses

This paper tests three hypotheses based on different levels. The hypothesis built at the lowest level is as follows:

***Legislators' capability hypothesis:** Legislators who are able to influence the allocation process are likely to deliver more grants to their districts.*

The logic behind this hypothesis is straightforward, and often raised by journalists and social activists who criticize Korean pork-barrel politics. Legislators who (i) are a member of the Public Administration and Security Committee which monitors the ministry of Public Administration and Security (MPAS), (ii) have senior status in the legislature, (iii) have a personal connection to the president, or (iv) a member of the government party, are able to allocate more grants to their districts.

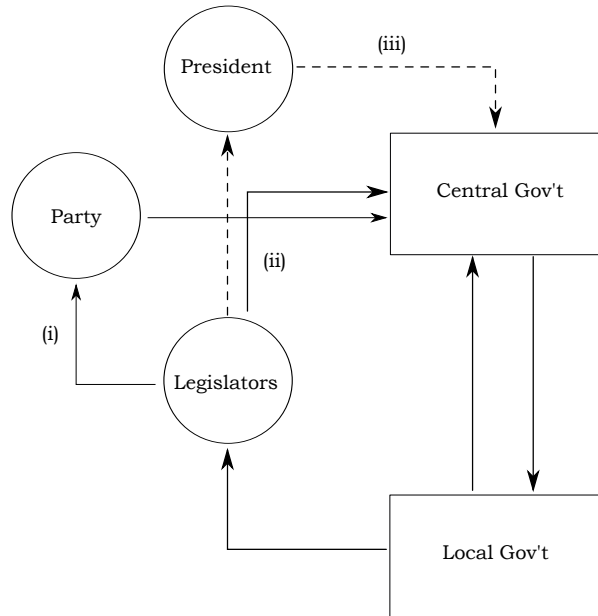


Figure 3.2: This figure shows an integrated theory about the relationship between intergovernmental grants and political players (represented as circles) who have different electoral incentives. We can either observe or proxy the solid lines. For instance, local governments directly request to favorably allocate the grants to their regions, or ask legislators elected in districts that belongs to their regions to influence the allocation process. Individual legislators may make a request to their party organizations for the allocation (path (i)). Senior members of the committee related to the allocation can directly influence the process (path (ii)). Legislators avowedly claim credit for successful allocation of the grants. Unlike legislators, the president has strong discretion over allocation of all kinds of government resources, but tends to evade being engaged in the allocation process. Therefore, it is not observed whether or not the president has influenced the allocation process (path (iii)). It is also not observable whether or not legislators who have a personal relationship with the president ask him to influence the process (dashed lines).

Unstable [Swing] electoral districts hypothesis: Parties are likely to deliver more grants to electorally unstable districts in order to maximize their seat shares in the legislature.

The second hypothesis is based on parties' electoral incentives. Electoral stability refers to the situation in which the number of supporters is much larger

than those of oppositions and swing voters. This can be empirically measured by vote margin (%) between the top two candidates under the single member district plurality (SMDP) rule. The logic behind this hypothesis is similar to the well-known *swing voter hypothesis* [DL98, LW87]. That is, parties tend to allocate resources to swing groups when these groups have a relatively large number of moderate voters who are ideologically indifferent between parties. However, it is notable that this hypothesis is distinguished from the swing voter hypothesis put forward by Lindbeck and Weibull. While the swing voter hypothesis is derived from theories to analyze redistribution *within districts*, the unstable electoral districts hypothesis focuses on geographical distribution *between districts*.

Unstable [Swing] provinces hypothesis: *Considering voting behavior based on regional identity in Korea, the president (or the government party) is likely to allocate more discretionary resources to an electorally unstable [potentially swing] province within a broader region than its stable counterpart. This is applicable to supporter, swing, and opposition regions that have experienced regional voting.*

Officially the minister of Public Administration and Security has the final say in the allocation of the *Special Local Allocation Grants*. Does this mean great ministerial discretion? It is worthwhile to note that the president has the authority to appoint the minister. Since the minister takes charge of the administration of elections on the national level, the president would choose one of his or her closest followers. In this sense, the minister can often be regarded as an agent of the president. In practice, since the *Special Local Allocation Grants* have been utilized by former presidents in order to construct policy coalitions or make legislators obedient to them, the grants are often called the 'disposable money in the president's pocket' (*Kukmin Ilbo*, September 20, 2007).

How can we theorize the president's role into the allocation process? Because the president cannot run for reelection in Korea, we need to specify what may induce the president to be involved in the allocation process. This paper assumes that the president wants his or her own policies to be secured (i.e., nonreversible by the legislative majority) even in the worst case scenario (i.e., his or her party losing its majority in the following legislative election). This assumption suggests that if the president needs some support from the opposition in order to get his or her policies secured, the president can allocate more resources to some of the opposition districts. In contrast, parties have an incentive to use their discretionary resources to allocate to their core supporters or swing groups rather than the opposition.

Considering regional voting in Korea, the rational president is likely to allocate more resources to electorally unstable provinces to potentially swing by election. The logic behind this hypothesis is that it is more efficient to allocate more resources to an electorally unstable province within a broader region where voters are strongly affiliated with the regional identity because people in the electorally stable province may be satisfied (or may not be angry) with the allocation of more grants to the relatively unstable counterpart sharing the same regional identity even if they are not a main beneficiary. This also implies that the president (or the government party) is likely to allocate relatively less resources to regions that do not show regional voting.

To test the third hypothesis, we need to carefully examine provincial variations in the allocation pattern. Methodologically, this motivates us to use the multi-level modeling techniques considering both fixed-effects and random effects. As Gelman and Hill [GH06] point out, the multilevel modeling is effective when it is close to complete-pooling assuming that each individual level unit (district level in this case) are homogenous regardless of regions, and ineffective when it is close to no-pooling. When the partially pooling estimate is close to complete-pooling

one, it can still capture variation between groups, which can be substantially more important to comparative politics research.²

3.4 Data and Measures

To test these hypotheses with a complete dataset, I gathered information from various reliable sources (Korean Statistical Information Service, and Korean Public Servants Union’s white paper on special local allocation grants from 2005 to 2006). The first thing to be considered is the discrepancy between electoral district units and local government units. To match units (245 electoral districts to 234 administrative units), I used Hur and Kwon’s criteria [HK09]. In cases where several administrative districts were embedded in an electoral district, I imputed the average of allocated grants in the administrative districts. In cases where several electoral districts were embedded in an administrative district, I selected the one electoral district with the smallest vote margin among them and seniority in the legislature. This is based on the underlying assumptions that (i) an incumbent legislator who won with the closest vote margin has a stronger incentive to deliver resources to his or her district than others and (ii) other things being equal, a senior legislator is more likely to be influential in the allocation process than others.

It is noted that in order to examine the government party’s targeting strategy Hur and Kwon excluded 28 administrative districts that have multiple electoral districts won by legislators affiliated to different parties. By deleting 28 cases ($\approx 12\%$) we may lose much information about electoral competition. Thus, I restored the 28 districts deleted in Hur and Kwon’s dataset.

²Besides, the partially pooling estimate can be close to complete-pooling for groups with small sample size and close to no-pooling for groups with large sample size, automatically performing well for both kinds of group (Gelman & Hill 2006, p. 270-1).

The outcome variable is the grants allocated to each local government (administrative unit). The special local allocation grants consist of four components: (1) special needs by natural disasters, (2) regional needs, (3) pilot projects, and (4) financial incentives to well-performed local governments. I checked for four possible outcome variables: the average of total grants ((1)+(2)+(3)+(4)), the average grants allocated for recovery from natural disasters ((1)), the average grants allocated for regional needs and pilot projects ((2)+(3)), and the average grants allocated for financial incentives to well-performed local governments ((4)). Figure 4.1 (see Appendix) shows the possible pairwise scatterplots of the components. The scatterplot between the grants allocated for natural disasters shows that (i) the allocation for this component in 2005 is not correlated with that in 2006, which reflects the uncontrollable characteristic of natural disasters, and (ii) the fitted line is strongly affected by some influential points. These imply that the component (1) can hardly be a suitable outcome variable.

As Choi and Kim [CK08] argue, the actual allocation for these two components are more likely to be affected by political, subjective evaluations than other components decided by relatively objective evaluations. The scatterplot between the grants allocated for regional needs and pilot projects ((2)+(3)) also shows a meaningful pattern between years (see Appendix). Hence, this paper focuses on these two components as an outcome variable.

Electoral stability is measured by vote margin in the list tier (PR rule) in the legislative election of 2004. In fact, Korea has adopted a mixed-member majoritarian electoral system [SW01]. The 2004 electoral reform adopted the double ballot system and lowered the threshold to earn a seat in the legislature from 5% to 3%. Hur and Kwon [HK09] use vote margin in the nominal tier (SMDP rule) as a measure of electoral stability. However, it is conventional wisdom that voters are less likely to vote strategically under PR than under SMDP [CGG09]. From the parties' perspective, vote margin in the list tier (PR rule) is a better measure

to capture electoral stability. The correlation between the two different vote margins is quite high (=0.79), but there exist some differences between them.

In order to measure legislators' personal connection with the president, I carefully examine the president's faction. As in Japan [PRS07], factions can be a good predictor for distribution of public expenditures in Korea. Indeed, when most presidents would form a policy coalition, they utilized the *Special Local Allocation Grants* as a tool to lure maverick legislators from the opposition parties. Before the electoral reform of 2004, it was convention for the president to seize control of the party organization including the authority to nominate legislative candidates. This dummy variable is measured by whether or not legislators belong to *Chin-Roh-Jikkye*, the political faction affiliated with the President Roh, who have personal access to him.³

Seniority is measured by counting how many times a legislator has been elected in legislative elections. This ranges from 1 to 5. The government party is measured by whether or not an incumbent legislator is affiliated with the government party (0–1).

3.5 Complete-Pooling and No-pooling Regression Models

The simplest regression model can be written as

$$\begin{aligned} \text{Average Grants} = & \gamma_0 + \gamma_1 \text{vote margin} + \gamma_2 \text{connection} + \gamma_3 \text{government party} \\ & + \gamma_4 \text{committee} + \gamma_5 \text{seniority} + \eta \text{Controls} + \epsilon \end{aligned} \tag{3.1}$$

where Controls include financial autonomy (0–100%), population (million),

³I classified legislators who belong to the *Cham-jeong-yeon* (Solidarity for True Participatory Politics) as *chin-Roh-Jikkye* (the closest circle). Most of them were former presidential secretaries.

Dependent Variable: Average amount of component (2) and (3)				
	OLS	Robust OLS	FE	FE
Predictor			(Regions)	(Provinces)
Financial Autonomy	-15.539** (3.672)	-17.661** (3.043)	-7.922 (4.062)	-5.788 (4.243)
Population	0.002** (0.001)	0.002** (0.000)	0.002** (0.000)	0.002** (0.000)
Population Density	-0.048** (0.008)	-0.044** (0.007)	-0.040** (0.008)	-0.025 (0.014)
Vote margin	5.897 (3.093)	4.463 (2.563)	0.581 (3.564)	0.959 (6.157)
Connection	81.453 (147.963)	87.799 (122.615)	156.119 (148.523)	207.553 (148.599)
Committee	333.437* (146.963)	390.585** (121.801)	410.891** (145.359)	528.756** (147.948)
Seniority	14.672 (37.516)	-1.751 (58.605)	27.113 (37.081)	21.636 (36.604)
Government Party	10.744 (89.596)	-12.023 (139.961)	24.541 (418.491)	47.998 (114.513)
Intercept	1295.071** (144.145)	1271.478** (119.451)	735.932** (213.159)	369.358 (345.258)
Observations	234	234	234	234
R^2	0.29	-	0.35	0.42

Note: ** and * indicate significance at $p < 0.01$, and $p < 0.05$, respectively. The numbers in parentheses represent standard errors.

Table 3.1: This table presents results of complete-pooling and no-pooling models.

and population density (million/ km^2). These control variables are intended to check that the allocation of special local allocation grants were carried out based on districts' need. We expect that while financial autonomy is negatively associated with the allocation, population is positively associated with it. As a proxy for urbanization, population density is expected to be negatively associated with the grant distribution. In case of no-pooling regression, so-called, fixed-effects model, additional sixteen dummies which represent provinces are included in the model.

However, such complete-pooling models cannot capture feasible omitted variables within regions as well as regional variations, which are substantively of

interest to researchers. Considering the importance of regionalism in Korean politics, these findings from complete-pooling models assuming that each region (or province) is homogenous may lead to misunderstanding. To test whether or not the hypotheses are still valid even when considering unobserved regional factors, we can utilize so-called “no-pooling models” or “fixed-effects” models [GH06]. The third column represents the fixed-effects model with six broad region-dummies (Seoul-Metro, Chungcheong, Gyeongsang, Jeolla, Gangwon, and Jeju). The result still supports most of the associations obtained from complete-pooling models. However, the positive association between vote margin and the allocation disappears. We find a strong association between the committee membership and the allocation, and this is robust to no-pooling models (see the third and fourth columns in Table 3.1).

What can we observe after controlling for provincial characteristics that are not observed? The fourth column in Table 3.1 shows that although R^2 improves to 0.42, all the political predictors except for committee membership lose their significance when including dummies for sixteen provinces. In particular, the negligible estimate of vote margin with inflated standard errors indicates that electoral stability is not generally associated with the allocation after controlling for province-dummies. This implies that the link between allocation of the grants and electoral stability in legislative elections may vary across provinces.

3.6 Multilevel Linear Models in the Bayesian Framework

Without checking over regional variations in the allocation of government resources, on the one hand, we cannot understand nuanced distributive politics in Korea. On the other hand, we may risk exaggerate the effect of regionalism if we do not control for variations within a broader region. Until now empirical studies

on distributive politics in Korea have failed to consider both a variation between regions and variations within a region at the same time [HK09, HL08]. This provides a sufficient justification for using multilevel models that consider *between group variation* as well as *within group variation* where the groups are provinces.

The sampling density of the outcome variable (i.e., the average grants allocated for regional needs and pilot projects) can be t -distribution with some outliers on the right side. This paper employs a simple multilevel model that allows intercepts to vary across groups (16 provinces) as follows.

$$\begin{aligned}
y_i &= t(\mu_i + \alpha_{j[i]}, \sigma_y^2, \text{df}_y), \quad \text{for } i = 1, \dots, 234. \\
\mu_i &\sim \mathcal{N}(\mathbf{Z}_i\boldsymbol{\beta}, v) \\
\sigma_y^2 &\sim \text{Inverse Gamma}(a, b) \\
1/\text{df}_y &\sim \text{Unif}(c, d) \\
\alpha_j &\sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2), \quad \text{for } j = 1, \dots, 16 \\
\sigma_\alpha^2 &\sim \text{Inverse Gamma}(e, f)
\end{aligned} \tag{3.2}$$

where \mathbf{Z} is the matrix with five district-level predictors (financial autonomy, population, population density, vote margin and membership of the committee), $\boldsymbol{\beta}$ is the matrix of the the regression coefficients of the predictors, and μ_α is a group-level average value of the intercept. v , a , b , c , d , e and f are specified by using information from Hur and Kwon’s [HK09], Horiuchi and Lee’s [HL08] and Choi and Kim’s [CK08] studies as well as [Hof09]’s way of specification of parameters in a inverse-gamma distribution (p.75).⁴

As Gelman and Hill [GH06] suggest, I also construct another multilevel model

⁴The prior for $\boldsymbol{\beta}$ is $(-16.5295, 0.002538, -0.06048, 5.87006, 207.57627)'$. The prior for the precision is $(0.0802386, 6968849, 1366.924, 0.0875, 0.0000519)'$. a and b are specified as 108 and 510731, respectively. The degrees of freedom df_y which must be positive as a normal distribution in this thesis. Considering that BUGS restricts a degrees of freedom greater than 2, Gelman and Hill [GH06] suggest a practical method to specify a degrees of freedom: $1/\text{df}_y \sim \text{Unif}(0, 0.5)$. e and f are specified as 2 and 200, respectively. It is note that the flat prior is used as the reference prior.

considering a group-level predictor. Hence, the second model changes the varying intercept $\alpha_j \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, for $j = 1, \dots, 16$ as follows:

$$\begin{aligned}
 \alpha_j &\sim \mathcal{N}(\gamma_0 + \gamma_1 u_j, \sigma_\alpha^2), & \text{for } j = 1, \dots, 16 \\
 \sigma_\alpha^2 &\sim \text{Inverse Gamma}(e, f) \\
 \gamma_0 &\sim \mathcal{N}(\mu_{\gamma_0}, \sigma_{\gamma_0}^2) \\
 \gamma_1 &\sim \mathcal{N}(\mu_{\gamma_1}, \sigma_{\gamma_1}^2)
 \end{aligned} \tag{3.3}$$

where u_j is a province-level predictor, the swing index by province. I constructed the swing index by summing up absolute values of differences in vote shares earned by two major parties between 1998 and 2002 legislative elections.⁵ It captures the swing of voters' choices in the two elections. γ_0 and γ_1 are the intercept and the slope of the regression line α_j , respectively, and μ_{γ_0} , $\sigma_{\gamma_0}^2$, μ_{γ_1} , and $\sigma_{\gamma_1}^2$ are the hyper-parameters.

As Bafumi and Gelman [BG06] point out, social scientists tend to avoid simple regression with varying intercepts when they suspect that predictors are correlated with units or groups (p.3). Following Bafumi and Gelman's suggestion, I also construct a multilevel model including the average of each district-level predictors in the province-level regression to avoid problems of bias and uncertainty caused by potential correlation between district-level predictors and group effects. Hence, the new multilevel model writes the varying intercept

⁵The reason why I employ absolute values is that both major parties can lose their vote shares in the multi-party system. For instance, the conservative party, Grand National Party (GNP) and the liberal party, New Millennium Democratic Party (NMDP) earned 43.27% and 45.06% in Seoul in 1998. GNP and Uri Party, a newly established major party, obtained 36.67% and 37.71% in Seoul in 2002. In this case the swing index of Seoul is calculated as $|-6.60| + |-7.35| = 13.95$.

$\alpha_j \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, for $j = 1, \dots, 16$ as follows:

$$\begin{aligned}
\alpha_j &\sim \mathcal{N}(\gamma_0 + \gamma_1 u_j + \bar{\mathbf{Z}}_j \gamma_2, \sigma_\alpha^2), \quad \text{for } j = 1, \dots, 16 \\
\sigma_\alpha^2 &\sim \text{Inverse Gamma}(e, f) \\
\gamma_0 &\sim \mathcal{N}(\mu_{\gamma_0}, \sigma_{\gamma_0}^2) \\
\gamma_1 &\sim \mathcal{N}(\mu_{\gamma_1}, \sigma_{\gamma_1}^2) \\
\gamma_2 &\sim \mathcal{N}(\mu_{\gamma_2}, \sigma_{\gamma_2}^2)
\end{aligned} \tag{3.4}$$

where u_j is a province-level predictor (the swing index by province), and $\bar{\mathbf{Z}}_j$ is a matrix of the average of each district-level predictor in j -th province.

Compared to normal distribution models, an implementation of t -distribution models with low degrees of freedom through the MCMC sampling may need heavy computations to obtain convergence. To evaluate whether the posterior distributions converge I use the Geweke’s convergence diagnostic as well as basic trace plots.⁶

3.7 Computation with JAGS

Suppose $\boldsymbol{\theta} = (\alpha_1, \dots, \alpha_j, \mu, \sigma_\alpha, \sigma)$. An iteration of the Gibbs sampler includes sampling from each full conditional distributions 2.12 to 2.15, and move from the vector $\boldsymbol{\theta}^s = (\alpha_1^{(s)}, \dots, \alpha_j^{(s)}, \mu^{(s)}, \sigma_\alpha^{(s)}, \sigma^{(s)})$ to $\boldsymbol{\theta}^{(s+1)}$ where s indicates each iteration as follows:

Step 1. sample $\alpha_j^{(s+1)}$ from 2.12 with $j = 1, \dots, J$.

Step 2. sample $\mu^{(s+1)}$ from 2.13. It is conditional on $\bar{\mu}$ and σ_α^2 set to $\bar{\mu}^{(s+1)}$ and $\sigma_\alpha^{2(s)}$, respectively.

⁶Besides JAGS, I also use STAN which is based on Hamiltonian Monte Carlo (HMC) sampling. HMC is expected to “accelerate both convergence to the stationary distribution and subsequent parameter exploration by using the gradient of the log probability function.” [Sta13] (p.3).

Step 3. sample $\sigma_\alpha^{2(s+1)}$ from 2.14. Here we use the $\alpha_j^{(s+1)}$ updated in Step 1 and $\mu^{(s+1)}$ updated in Step 2.

Step 4. sample $\sigma^{2(s+1)}$ from 2.15. Since it is conditioned on $\sum_{j=1}^J \sum_{i=1}^{n_j} (y_{ij} - \alpha_j)^2$, we use $\alpha_j^{(s+1)}$ updated in Step 1.

Taking **Step 1** to **Step 4** allow us to obtain a completely sampled $\theta^{(s+1)}$, and it will be the conditional arguments in the next iteration. These algorithm can be easily implemented in JAGS. The JAGS code for varying intercept model 3.2 can be written as

```

1 model{
2   for(i in 1:N){
3     y[i] ~ dt(y.hat[i],tau.y,df.y)
4     y.hat[i] <- a[X[i,1]] + inprod(b[],X[i,2:6])
5   }
6   tau.y ~ dgamma(aa,bb)
7   sigma.y <- pow(tau.y,-2)
8   df.y <- pow(inv.df.y,-1)
9   inv.df.y ~ dunif(0,0.5)
10
11  for(k in 1:K){
12    b[k] ~ dnorm(m[k],precb[k])
13  }
14
15  for(j in 1:J){
16    a[j] ~ dnorm(mu.a,tau.a)
17  }
18  mu.a ~ dnorm(ma0,preca0)
19  tau.a ~ dgamma(cc,dd)
20  sigma.a <- pow(tau.a,-2)
21 }

```


The model with a group predictor 3.3 can be written by modifying line 15 to 20 as follows:

```
1  for(j in 1:J){
2    a[j] ~ dnorm(a.hat[j],tau.a)
3    a.hat[j] <- g.0 + g.1*u[j]
4  }
5  g.0 ~ dnorm(500,0.0001)
6  g.1 ~ dnorm(0,0.001)
7  tau.a ~ dgamma(cc,dd)
8  sigma.a <- pow(tau.a,-2)
```

To implement a MCMC sampling for 3.4, the line 15 to 20 can be modified as follows:

```
1  for(j in 1:J){
2    a[j] ~ dnorm(a.hat[j],tau.a)
3    a.hat[j] <- g.0 + g.1*u[j] + inprod(g.2[],g.mean[j,])
4  }
5
6  g.0 ~ dnorm(500,0.0001)
7  g.1 ~ dnorm(0,001)
8
9  for(k in 1:K){
10   g.2[k] ~ dnorm(g.m[k],g.prec[k])
11 }
12
13 tau.a ~ dgamma(cc,dd)
14 sigma.a <- pow(tau.a,-2)
```

where `g.m` and `g.prec` indicate the mean and the precision hyperparameters for slope of the group predictor `g.2`, respectively. In line 3, `g.mean[j,]` represents a matrix of the average of each district-level predictor in j -th province

(= $\bar{\mathbf{Z}}_j$).

As mentioned above, we can also use a different MCMC sampling **STAN** for faster convergence to the stationary distribution.⁷ The **STAN** code for 3.3 can be added in Appendix.

3.8 Results

The main reason why I would carefully examine partially pooling regression models is that random effects may include political actors' (the president's or the government party's) influence at the higher level that cannot be included in no-pooling models.

The result is given not as point estimates but as posterior distributions. Not only do all the posterior distributions converge smoothly, but are also bell-shaped when it iterates greater than 500,000 times when employing the normal sampling model instead of t -distribution (see Appendix 4). By contrast, the posterior distributions given by t -distribution sampling model are not always bell-shaped, and we need to implement the MCMC sampling with iterations greater than 1,000,000.

Figure 3.3 presents the posterior distributions of coefficients for district-level predictors by two different sampling models (normal and t -distribution models without the province-level predictor, the swing index by province), the informative priors, and 1,000,000 iterations (with the 10,000 burn-in). All the posterior distributions pass the Geweke's convergence diagnostic for Markov chains. It is noted that the 95% confidence interval for the posterior distribution of the degrees of freedom (df_y) is [2.00, 2.05]. As the precision of the degrees of freedom df_y increases, the posterior means become closer to those obtained by normal distribution models. Figure 3.3 shows that control variables that represent the

⁷The results are similar to those from JAGS, but it converges faster. All the results are available at my Github site (<https://github.com/bsk245/sthesis>)

need-based criteria are associated with the distribution of the special local allocation grants, as expected. This means that the allocation has been carried out not so arbitrarily as journalists or social activists criticize.

The result also provides evidence against the *Unstable electoral districts hypothesis*. The posterior distribution of vote margin that represents electoral stability at the district level has a positive mean ($=9.511$), and the 95% confidence interval $[5.176, 13.840]$ is greater than zero. That is, we are 95% confident that district-level vote margin is positively associated with the allocation. Focusing on the mean value, we can say that an increase of 1% in the vote margin in the legislative election is associated with 9.5 million won on average. Considering the average of vote margin (17%), this can make substantial differences in the allocation.

In contrast, the *Legislators' Capability Hypothesis* is supported by the result. Based on the confidence interval, we are 95% confident that legislator who belong to the special committees (the Public Administration and Security Committee) can allocate the substantive amount of money to their districts.

Now let us compare this with the results from the complete-pooling regressions (see the first and second columns in Figure 3.1). Since the 95% confidence intervals for population, population density and committee membership include the point estimates obtained by the complete-pooling regressions, the multilevel model in the Bayesian framework is not distant from the complete-pooling models. However, the point estimate of vote margin is smaller than the posterior mean of vote margin. That is, the multilevel model provides a stronger evidence against the *Unstable Electoral Districts Hypothesis* than do the complete-pooling regression models.

Why were more grants allocated to electorally stable districts? One reasonable possibility is that risk-averse parties have accurate knowledge about the quantities of core supporters needed to guarantee reelections at the district level [CM86].

The 2004 electoral reform that adopted the double-ballot election system (SMDP in the nominal tier and PR system in the list tier) may help parties calculate accurate quantities of core supporters who are less likely to vote strategically under PR.

The most interesting result is found in random effects. The figure 3.4 displays random effects of sixteen provinces as varying intercepts. We see the provinces (Chungnam, Gyeongnam, and Jeonam) whose intercepts are obviously above the average intercept (the red dashed-line). These provinces are electorally more unstable than their counterparts (Chungbuk, Gyeongbuk, and Jeonbuk, respectively) within the swing (Chungcheong), the supporter (Jeolla), and the opposition (Gyeongsang) regions, respectively. This finding supports the *Unstable Provinces Hypothesis*: Considering voting behavior based on regional identity in Korea, the president (or the government party) is likely to allocate more discretionary resources to an electorally unstable province within a broader region than its stable counterpart. This is generally applicable to supporter, swing, and opposition regions that have experienced regional voting.

The coefficient plot also illustrates that less grants were allocated to provinces little affected by regional voting (e.g., Seoul, Gyeonggi, Incheon) even though the size and the number of units belong to the provinces are much larger than the three unstable provinces (Chungnam, Jeonam, and Gyeongnam).

I also tested two varying intercept, varying slope models with the same district level predictors as displayed in Figure 3.5. The left panel illustrates how intercepts and slopes vary when we believe no correlation between them. The right panel displays how intercepts and slopes vary when we believe that they are not completely independent of each other so that the prior for the correlation is specified as $\rho \sim U[0.1, 0.6]$. The 95% confidence interval for posterior distribution of ρ is [0.1023, 0.3187]. We see that both figures are substantially identical, which tells us that we do not have to add a more parameter ρ to the current model. Purple

Province-Level Predictor	Mean	Sd	2.5%	50.0%	97.5%
Intercept (g_0)	514.10	31.27	452.70	514.00	575.20
Slope (g_1)	20.54	6.25	8.32	20.57	32.72

Table 3.2: This table presents the posterior distributions of coefficients for a province-level predictor by using the t -distribution model with a province-level predictor, the swing index by province. I used the informative priors, and 1,000,000 iterations (with the 1,000 burn-in).

lines show how intercepts and slopes in electorally unstable provinces (Chungnam, Jeonam, and Gyeongnam) change across vote margins (%). Despite the different levels of slope the intercepts are higher than others (gray lines), which is consistent with the result from the simple varying intercept model (3.2).

The BIC (Bayesian Information Criterion) for the random intercept, random slope model is 3713 while that for the random intercept model is 3703. The log-likelihood is -1829 in both models, which implies the varying intercept, varying slope model does not improve the fit in spite of adding more variables.

The DIC (Deviance Information Criterion) for the varying intercept, varying slope model is 10037.800, which is much larger than that of the simple varying intercept model ($\text{DIC} = 3702$).⁸ Since the effective number of parameters (pD) estimated by subtracting $-2 \cdot \log$ -likelihood from the posterior mean of the deviance also is much larger than the number of observations in this case, the much lower value of DIC for the simpler model is more adequate to predict the distribution of the special local allocation grants in Korea.⁹

The result from the multilevel linear regression model that considers degree of province-level swing measured by the sum of absolute values of differences in vote shares earned by the two major parties in two legislative elections (1998 and

⁸This was implemented via `rjags` in JAGS. It is noted that the posterior mean of the deviance (\bar{D}) is not a measure of fit but a measure of adequacy (<http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/dicpage.shtml>).

⁹The DIC is valid only when the effective number of parameters in the model is much smaller than the number of observations [Plu08]. In order to use the DIC without any problem, all the posterior distributions obtained by Gibbs-sampler should be bell-shaped (i.e., close to normal).

2002).¹⁰ It is rational that president (or a government party) would allocate more grants to provinces which tend to swing across elections. The positive posterior mean of g_1 means that more grants are more likely to be allocated to electorally unstable provinces. Table 3.2 shows that the posterior distribution of g_1 are between 8.32 and 32.72.

Figure 3.6 shows varying intercepts across provinces when we use the multilevel linear regression models. The implementation of Gibbs samplers with informative priors confirms that despite some differences the intercepts of three electorally unstable provinces (Chungnam, Jeonam, and Gyeongnam) are greater than their relatively stable counterparts (Chungbuk, Jeonbuk, and Gyeongbuk, respectively), and they are main beneficiaries of the allocation, other things being equal, in both models.

I also tested 3.3 suggested by Bafumi and Gelman [BG06], but an inclusion of the average values of predictors in the model for varying intercept does not make any meaningful differences.¹¹

3.9 Sensitivity Analysis

How sensitive is the model's performance to changes in prior? This sensitivity analysis is expected to show how my model fits the data [GCS04]. I used an identical prior for the hierarchical part. In other words, priors are distinguishable only in parameters for the district-level predictors.¹² I found that the posterior distributions of control variables (β_1 , β_2 , β_3 and β_4) do not vary with the priors except for the mean of committee membership (β_5).

We see that 3.2 is sensitive to the variance (or the precision) of vote margin

¹⁰The two major parties are Grand National Party (GNP) and New Millennium Democratic Party (NMDP) in the election of 1998, and GNP and Uri Party in the election of 2002.

¹¹The figures are available at my website (<https://github.com/bsk245>).

¹²I did the analysis for all parameters, but I do not include the results in this paper.

as shown in Figure 3.7 and Figure 3.8. For example, Gyeongbuk (region 10) and Jeonbuk (region 13) show that variance of vote margin (β_4) at the district level can make huge differences in the posterior distribution of intercepts. From the president's (or the government party's) perspective, Gyeongbuk and Jeonbuk have been electorally stable over elections in the opposition region (Gyeongsang) and the supporter region (Jeolla), respectively. This sensitiveness can be explained in that a province-level vote margin is the accumulation of district-level vote margins. As uncertainty over the coefficient for vote margin at the district level reduces, province-level random effects expressed as varying intercepts are more likely to resemble the district-level coefficient. In particular, it is more appropriate in electorally stable provinces where election outcomes are less variable across districts. By the same logic, we do not find such sensitiveness in Chungbuk (region 2), the electorally stable province within the electorally unstable region (Chungcheong). In comparison with the opposition and supporter regions (Gyeongsang and Jeolla, respectively), a province-level vote margin is less likely to be the simple accumulation of district-level vote margins in the swing region filled with mixed election outcomes across districts.

3.10 Discussion

How consistent are these statistical findings on the three electorally unstable provinces with substantive knowledge? Let us approach the statistical finding substantively. The Gyeongsang region consists of two provinces: Gyeongbuk and Gyeongnam. While Gyeongbuk is northern part of the Gyeongsang region, Gyeongnam is southern part of the region. Both Gyeongsang provinces supported the strongest rival party (Grand National Party, GNP) against the president and his party (Uri Party) in the general election of 2004. However, the government

	Opposition Region (Gyeongsang)	Swing Region (Chungcheong)	Supporter Region (Jeolla)
Electoral Stable Province	Gyeongbuk (7th, 23.0%)	Chungbuk (10th, 44.7%)	Jeonbuk (11th, 67.4%)
Electoral Unstable Province	Gyeongnam (4th, 31.8%)	Chungnam (2st, 37.5%)	Jeonam (3rd, 46.8%)

Table 3.3: This table summarizes three broad regions and six provinces where regional identities are found in Korea. The figures in parentheses indicate the rankings of posterior medians of random effects represented as varying intercepts across 16 provinces including Special City and Metropolitan cities, and the government party’s vote shares in the party-list, respectively [ROK04]. As Figure 3.4 illustrates, the more *Special Local Allocation Grants* were allocated to electorally unstable [swing] provinces (Chungnam, Gyeongnam, and Jeonam) in comparison to the counterparts (Chungbuk, Gyeongbuk, and Jeonbuk). We also find the president’s (or the government party’s) targeting strategy that focuses on electorally unstable province (Jeonam) within his (or its) supporter region (the Jeolla region), considering Gwangju, a metropolitan city geographically located within the Jeonam province, is at the 1st in the Figure 3.4. By contrast, Daejeon metropolitan city placed within the Chungcheong region is at the bottom, and Daegu metropolitan city located within the Gyeongsang region, is ranked 15th out of 16 provinces.

party competed unexpectedly well with GNP in the Gyeongnam province (31.8% vs. 47.6% in the party-list tier) in comparison with its counterpart, the Gyeongbuk province (23.0% vs. 58.4%). In the nominal tiers, two out of the seventeen seats were earned in Gyeongnam while none out of fifteen seats in Gyeongbuk. This outcome may be caused by (i) public backlash against GNP’s irrational impeachment of President Roh, which can be evaluated as key variable to explain Uri Party’s legislative majority in 2004, and (ii) the two electoral districts (*Kimhae Kap* and *Kimhae Eul*) can be categorized as the president’s hometown. The government party earned two seats in the Gyeongnam province, and vote margins in these two districts (9.7%) are far smaller than the average vote margin in the province (18.46%). In sum, the Gyeongnam province is electorally more unstable than its counterpart, the Gyeongbuk province, and the result can be evidence for the *Unstable Provinces Hypothesis*: for the president (or the government party), it is more efficient to target an electorally unstable province *even within an opposi-*

tion region because people in the electorally stable province are strongly affiliated with a regional identity and they may be satisfied with the allocation of grants even if they are not the main beneficiaries.

Like the Gyeongsang region, the Jeolla region consists of two provinces: Jeonbuk and Jeonam. Jeonbuk is northern part of the Jeolla region, and Jeonam is the southern part of it. It is clear that both Jeolla provinces are the strongest supporter provinces of President Roh. However, the government party struggled with another opposition party (New Millennium Democratic Party, NMDP) in the Jeonam province (46.7% vs. 33.8% in the party-list tier) in compared with the Jeonbuk province (67.3% vs. 13.6%). In the nominal tiers, all the eleven seats were obtained in the Jeonbuk province while eight out of thirteen seats obtained in the Jeonam province. In terms of vote margin in the party-list tier, the difference between these two provinces is striking (52.45% in Jeonbuk vs. 13.76% in Jeonam). In sum, the Jeonam province is electorally more unstable than its counterpart, the Jeonbuk province, and thus the result can be also evidence for the *Unstable Provinces Hypothesis*, which suggests that it is efficient to allocate more resources to an electorally unstable province than to a stable province within a supporter region.

What happened in a swing but pivotal region, the Chungcheong region? Like the Gyeongsang and the Jeolla regions, the Chungcheong region consists of two provinces: Chungbuk and Chungnam. The Chungnam province, western part of the swing region, can be considered the unstable province within the swing region. The vote choices motivated by their regional identities in the Chungnam province have been evidently observed in presidential elections since 1987, when Korea revived the system of a direct presidential election. They supported Kim Jong-pil who is a native of the Chungnam province in the election of 1987. Whereas he earned only 8.1% nationwide, Kim obtained 45.0% in Chungnam. In 1992 no native presidential candidate ran for president, and three candidates shared the

vote (36.9% vs. 28.5% vs. 25.2%) in the province. With an alliance with Kim Jong-pil, Kim Dae-jung earned 48.3% in the region in the presidential election of 1997. However, many people in the region still supported two native candidates Lee Hoi-chang, the ruling party candidate and Lee In-je, a third party candidate (23.5% and 26.2%).

It was exceptional that Roh Moo-hyun earned 52.2% in the presidential election of 2002, which helped him elected in the election.¹³ The results of previous presidential election shows that Chungnam has tended to support Chungnam native candidates, but been an electorally unstable province that easily swings by election. It provides an explanation of President Roh's incentive to compensate voters in the Chungnam province. Figure 3.4 depicts significantly positive intercept in the Chungnam province. In this sense, previous studies based on traditional statistical models and presidential election outcomes cannot capture the significant association in the Chungnam province [HL04, HL08]. Hur and Kwon [HK09] and Choi and Kim [CK08] who pay attention to parties' electoral incentives and only two broad regions (Gyeongsang and Jeolla) do not fully consider regional variations so that they fail to find this association with rigorous statistical ways.

Assuming that President Roh desired to secure his reforms, we can also explain the president's incentive to allocate more intergovernmental transfers to the electorally unstable province. Only 47 legislators belonged to his party (Uri Party) before the general election of 2004, but the party won a narrow majority (152 out of 299 seats) due to victories in the metropolitan areas and the Chungnam province. It was the first time that a centrist (economically left-of-center but politically liberal) party won a majority since the 1963 general election under military dictatorship. It was obvious that support from the Chungnam province was

¹³Lee Myung-bak, the 17th president, earned 48.7% nationwide in the presidential election of 2007, but obtained 34.3% in the province. Although Lee Hoi-chang declared he would run for president on the 7th of November (42 days before the election day), he earned 33.2% in the Chungnam province. It demonstrates that regional identity still motivated vote choices in the region.

crucial to secure his reforms.

To win an legislative election in the swing region, the president employed a more efficient strategy to allocate government resources. The ruling party competed with two different opposition parties (GNP and United Liberal Democrats, ULD) in the Chungnam province in 2004. The races were closer in Chungnam, compared to those in Chungbuk. In the list tier, vote shares (%) of these three parties (Uri, GNP, and ULD) are 38.0, 21.1, and 23.8, respectively. In the nominal tiers, all the eight seats were obtained in the Chungbuk province while only six out of ten seats in the Chungnam province. In terms of vote margin in nominal tiers, the difference between these two Chungcheong provinces is striking (18.4% in Chungbuk vs. 5.2% in Chungnam). Consequently, since the Chungnam province is electorally more unstable than its counterpart, the Chungbuk province within the pivotal swing region, it was the efficient targeting strategy to allocate more intergovernmental transfers to the Chungnam province rather than the Chungbuk province. It consistently supports the *Unstable Provinces Hypothesis*.

The president's (or the government party's) strategy to target an electorally unstable province even within a supporter region can be partially justified. In practice, the disproportional allocations of central government grants to electorally unstable province within its supporter region (Jeolla) from 2005 to 2008 helped the government party to increase its vote share in the unstable province (Jeonam) by 20% in 2008. The vote share in the electorally stable province (Jeonbuk) reduced only by 2.9% in comparison with that in the previous election. It was a noteworthy outcome, considering the the government party was defeated by the wide margin 13% nationwide and experienced a swing against it by 20.3% across the country in the election.

Posterior Distributions of District-Level Predictors

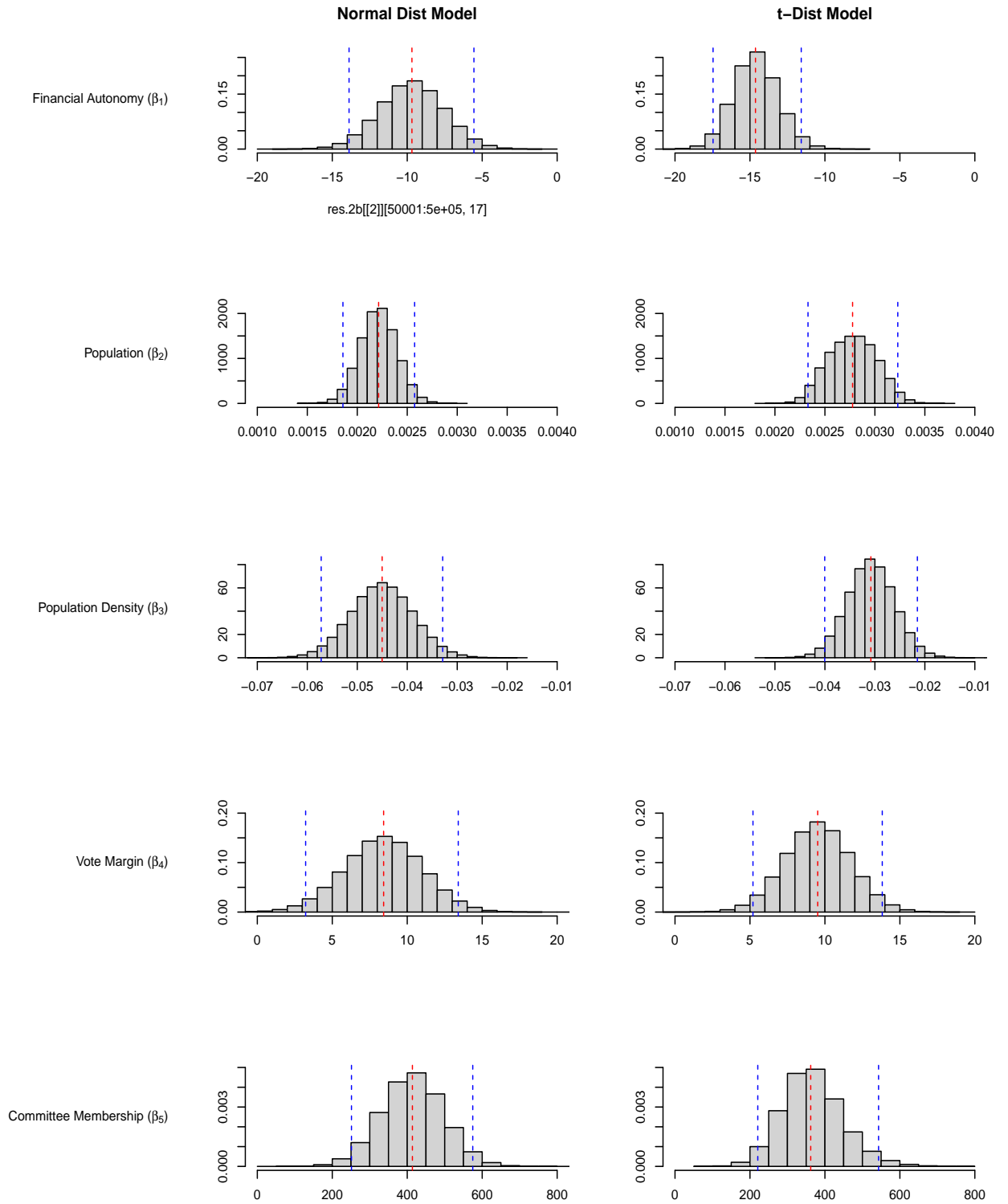


Figure 3.3: Posterior distributions of district-level predictors

Varying Intercepts (t-Distribution Model)

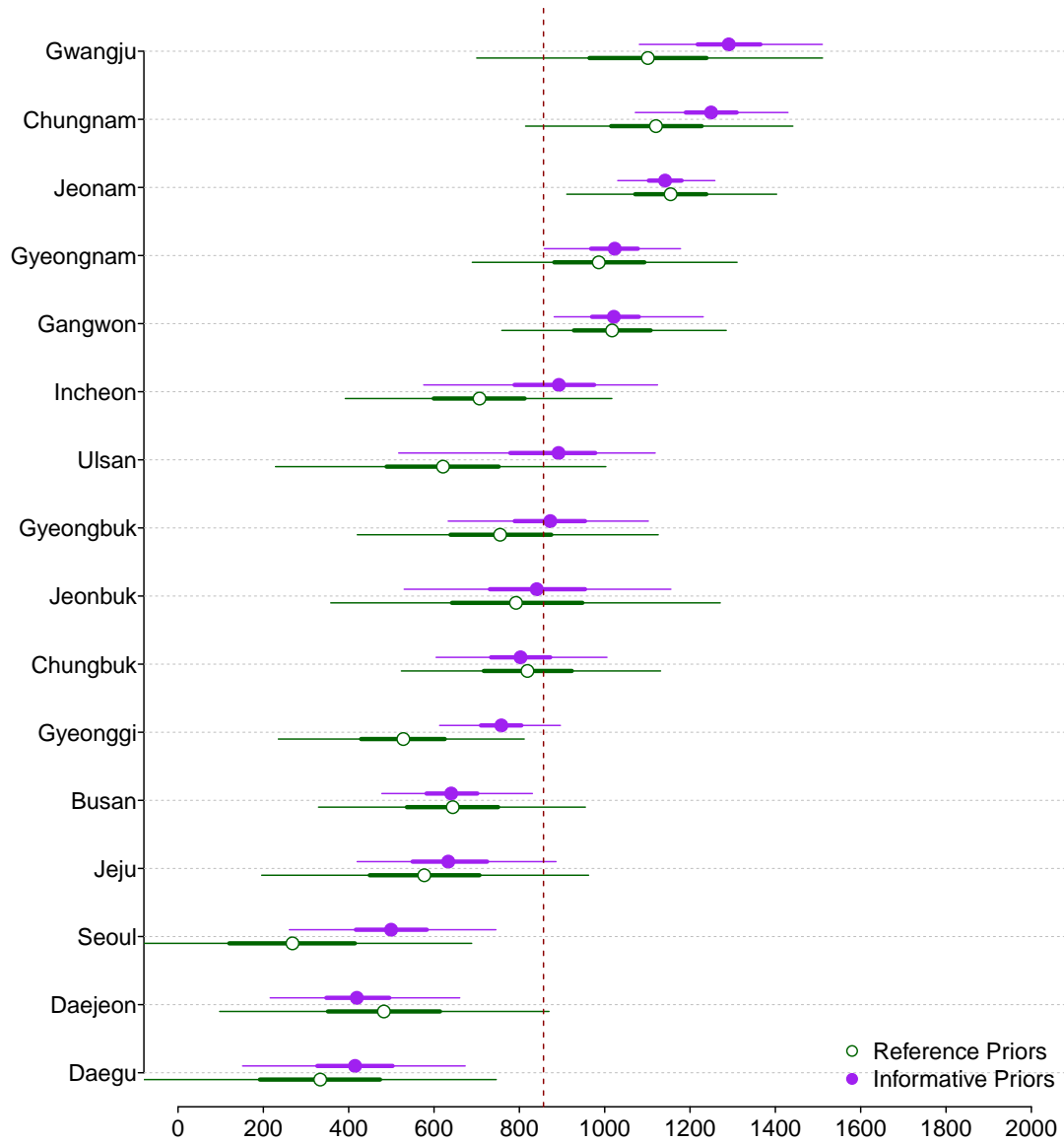


Figure 3.4: This plot illustrates random effects represented as varying intercepts across provinces. Purple solid dots indicate the posterior medians of intercepts obtained by the informative priors. Dark green hollow dots represent the posterior medians of intercepts obtained by reference priors (i.e., uninformative priors). While thick lines represent 50% confidence intervals, thin lines indicate 95% confidence intervals. The dark red dashed-line indicates the median of intercepts that vary across provinces. The intercepts of three electorally unstable provinces (Chungnam, Jeonam, and Gyeongnam) are greater than their stable counterparts (Chungbuk, Jeonbuk, and Gyeongbuk, respectively).

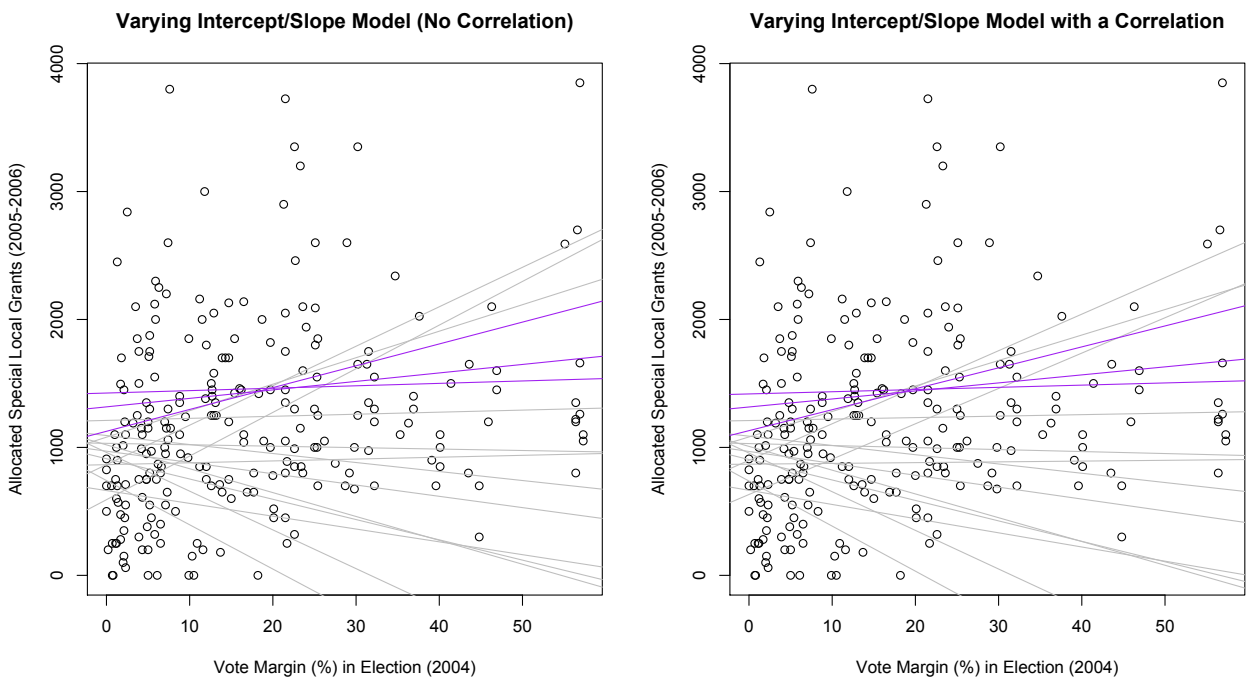


Figure 3.5: This figure illustrates the results from the multilevel model that allows both intercepts and slopes to vary across provinces. Here intercepts and slopes used in this figure represent the mean values of the posterior distributions obtained from 1,000,000 iterations with the 10,000 burn-in.

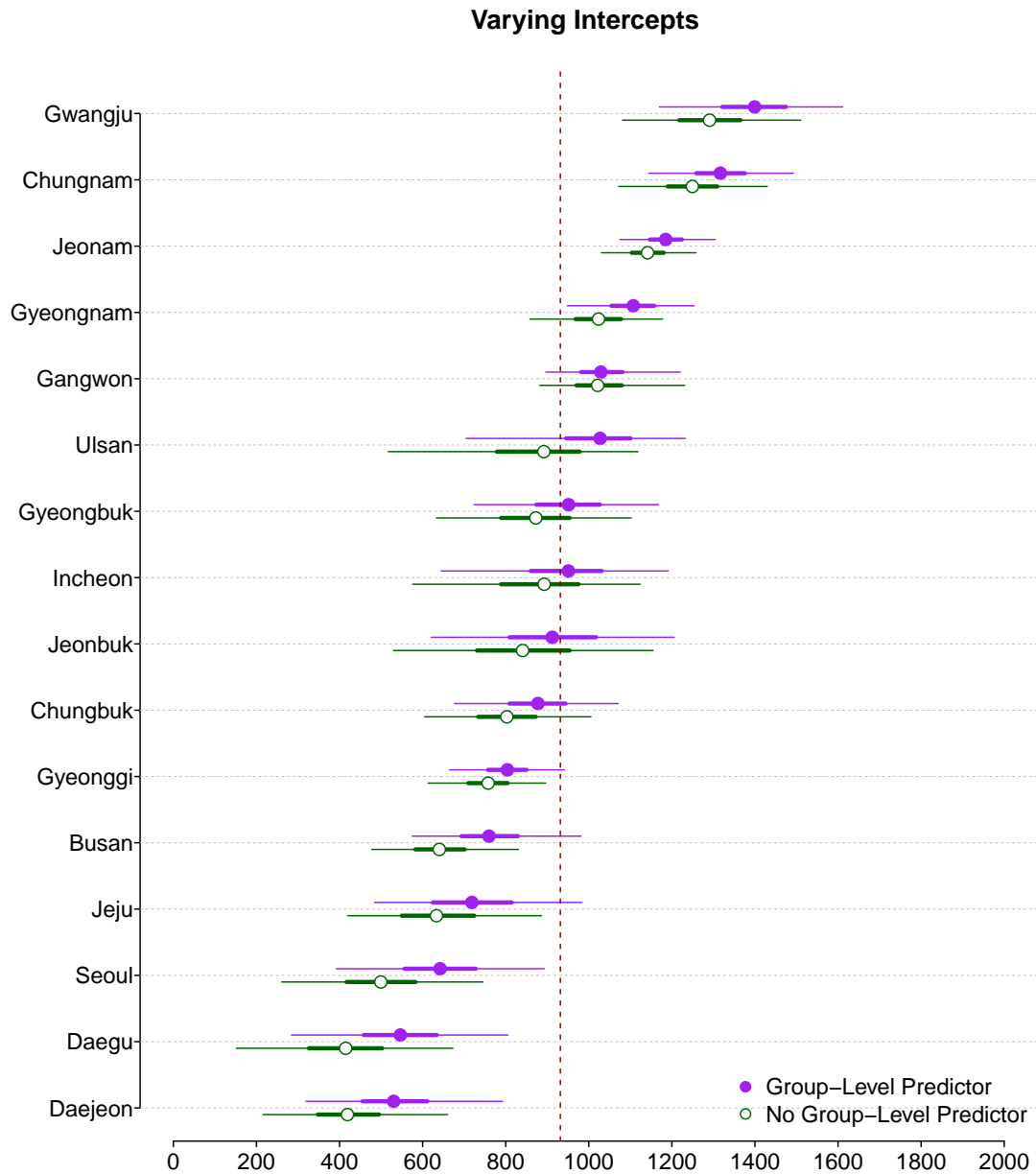


Figure 3.6: This plot illustrates random effects represented as varying intercepts across provinces when we use two models shown in 3.2 and 3.3. Purple solid dots indicate the posterior medians of intercepts obtained by the multilevel linear regression with the group predictor. Dark green hollow dots represent the posterior medians of intercepts obtained by the multilevel linear regression without any group predictor. While thick lines represent 50% confidence intervals, thin lines indicate 95% confidence intervals. The dark red dashed-line indicates the median of intercepts that vary across provinces. The intercepts of three electorally unstable provinces (Chungnam, Jeonam, and Gyeongnam) are greater than their stable counterparts (Chungbuk, Jeonbuk, and Gyeongbuk, respectively).

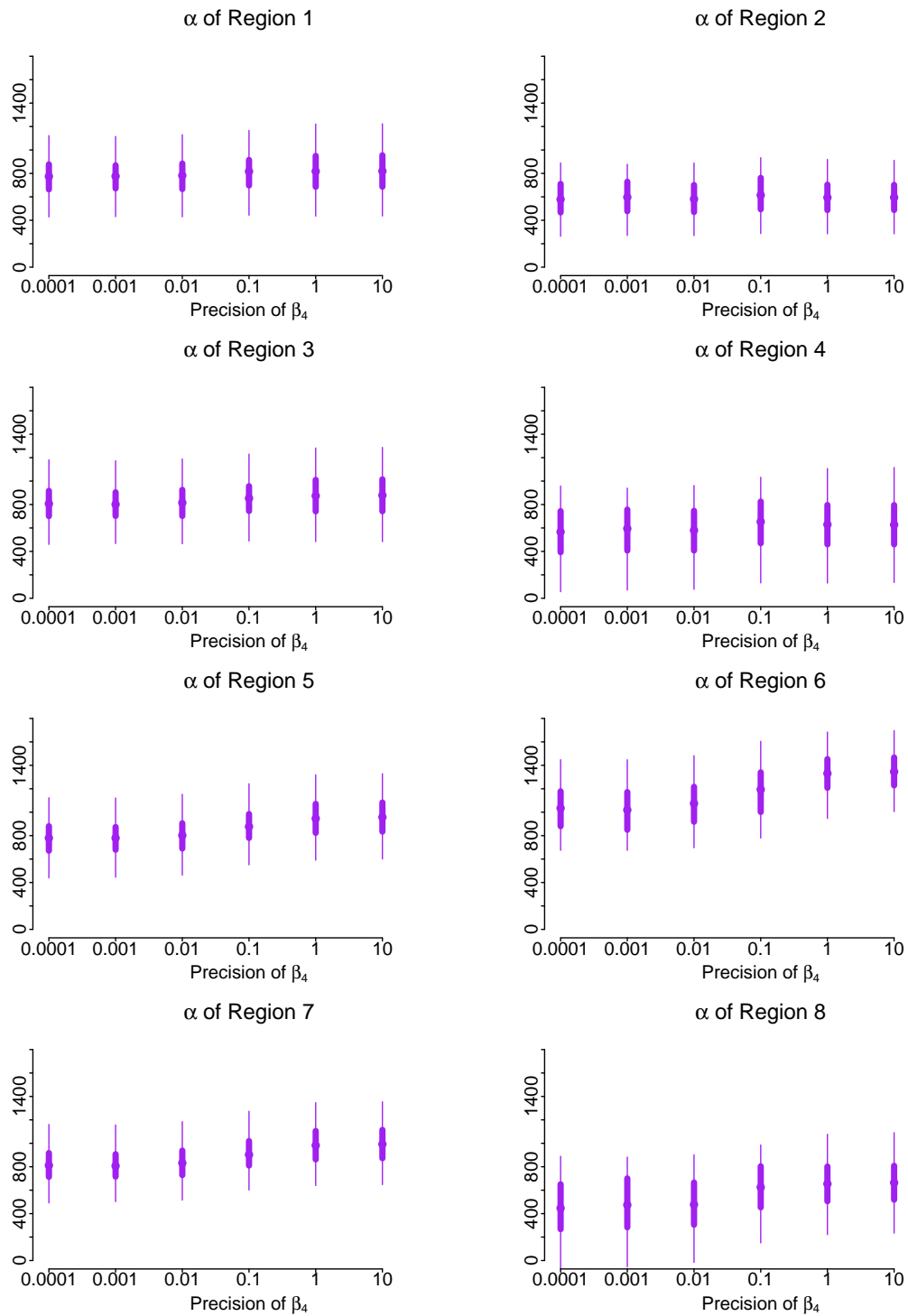


Figure 3.7: This figure illustrates how sensitive regional intercepts (α_j 's) are to variance of coefficient for vote margin (β_4).

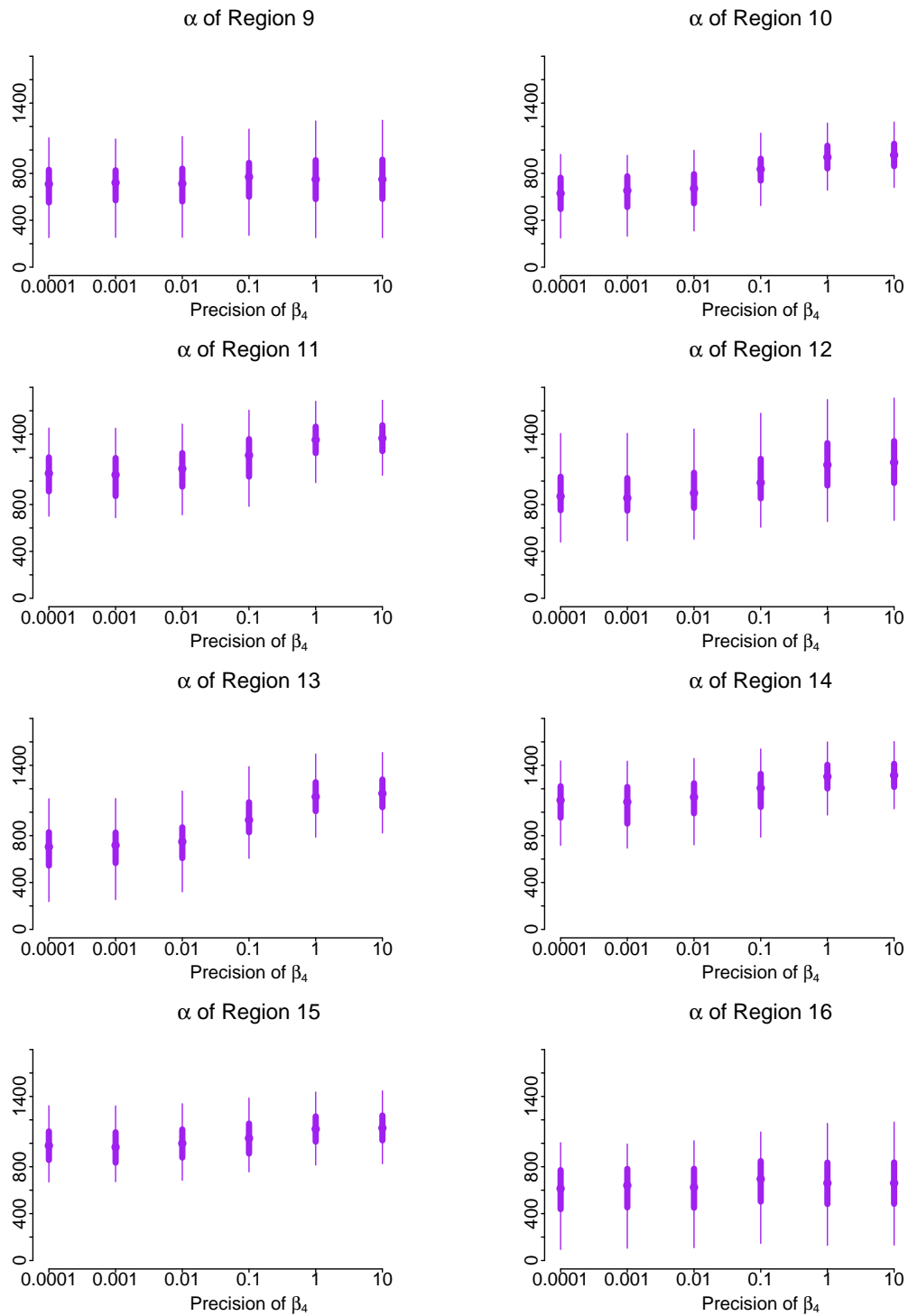


Figure 3.8: This figure illustrates how sensitive regional intercepts (α_j 's) are to variance of coefficient for vote margin (β_4).

CHAPTER 4

Conclusion

I have employed several regression models to test three hypotheses on the relationship between electoral incentives and the allocation of intergovernmental grants by focusing on the case of *special local allocation grants* in Korea from 2005 to 2006. The results from complete-pooling regression models and no-pooling linear regression models demonstrate that the *Unstable Electoral Districts Hypothesis* is not consistently supported and the tests are also very sensitive to selection on the outcome variable. The *Legislators' Capability Hypothesis* is partially supported. Most models show that being a member of the Public Administration and Security Committee is positively associated with the allocation. However, we can find a little substantive association between other capabilities (e.g., seniority, personal connection to the president) and the grant allocation.

To carefully test the *Unstable Provinces Hypothesis*, I employed the multilevel linear regression model in the Bayesian framework. This method substantively helped capture regional variations in the allocation and utilize substantive knowledge from literature on this issue. The result provides strong evidence against the *Unstable Electoral Districts Hypothesis*. The posterior distribution of vote margin that represents electoral stability at the district level has a substantively positive mean ($=9.511$), and the 95% confidence interval is also greater than zero. Considering the average of vote margin in the election, it could make substantial differences in the allocation. How can we explain this result against the *Unstable Electoral Districts Hypothesis*? There are some reasonable explanations. For ex-

ample, it is possible that risk-averse parties have accurate knowledge about the quantities of core supporters needed to guarantee reelections at the district level, which has been explained in classical qualitative literature.

As in the classical regression models, the posterior mean of legislators' membership of the Public Administration and Security is both substantively and statistically significant, which supports the *Legislators' Capability Hypothesis*.

My main finding, however, is that there is a significant positive association between the amount of intergovernmental grants and being an electorally unstable province (Chungnam, Jeonam, and Gyeongnam) in broader regions affected by regional voting behavior even after controlling for the need-based criteria (e.g., financial autonomy, population). That is, the result from Bayesian multilevel linear regression model supports the *Unstable Provinces Hypothesis*. In sum, while the intergovernmental grants are more likely to be allocated to stable electoral districts at the district level, the grants are more likely to be allocated to unstable provinces at the higher level. This apparently contradictory finding is new in quantitative works on this issue, but is consistent with findings by qualitative research that emphasizes the role of president in the allocation process. It also demonstrates that the Bayesian multilevel regression model which utilizes all information sources is an effective method to investigate distributive politics.

This finding also has a strong implication for decentralized democratic governments under circumstances where significant regionally (or ethnically) affiliated-voting is observed. That is, it can justify the president's (or the government party's) strategy to target an electorally unstable province even within a supporter region because people in the electorally stable province are strongly affiliated with a regional (or ethnic) identity so that they may be satisfied with the allocation of grants even if they are not the main beneficiaries. Thus, while the allocation is concentrated on core supporters that are well known quantities at the district level, the allocation at the higher level can be decided by the efficient targeting

strategy. In practice, the disproportional allocations of central government grants to electorally unstable province within its supporter region (Jeolla) from 2005 to 2008 helped the government party to increase its vote share in the unstable province (Jeonam) by 20% in 2008. It was a remarkable outcome, considering the the government party was defeated by the wide margin 13% nationwide in the election.

Nevertheless, this paper has some limitations. First, since my attempt to model distributive politics in Korea using the Bayesian approach is new, the prior for the hierarchical part unlike the other components is subjectively specified, which is the weakest part of this paper. Nonetheless it is meaningful that my results can be an informative prior for next research based on the Bayesian approach. Second, according to Dahlberg and Johansson [DJ02], the data used in this paper are not fully suitable for analyzing the relationship between electoral incentives and the allocation of intergovernmental grants in that the granting decisions are not made in close connection to an election. To fulfill Dahlberg and Johansson's criteria, we need the data on the special local allocation grants in 2007 (or 2008) in that the following legislative election was held in April, 2008.

APPENDIX

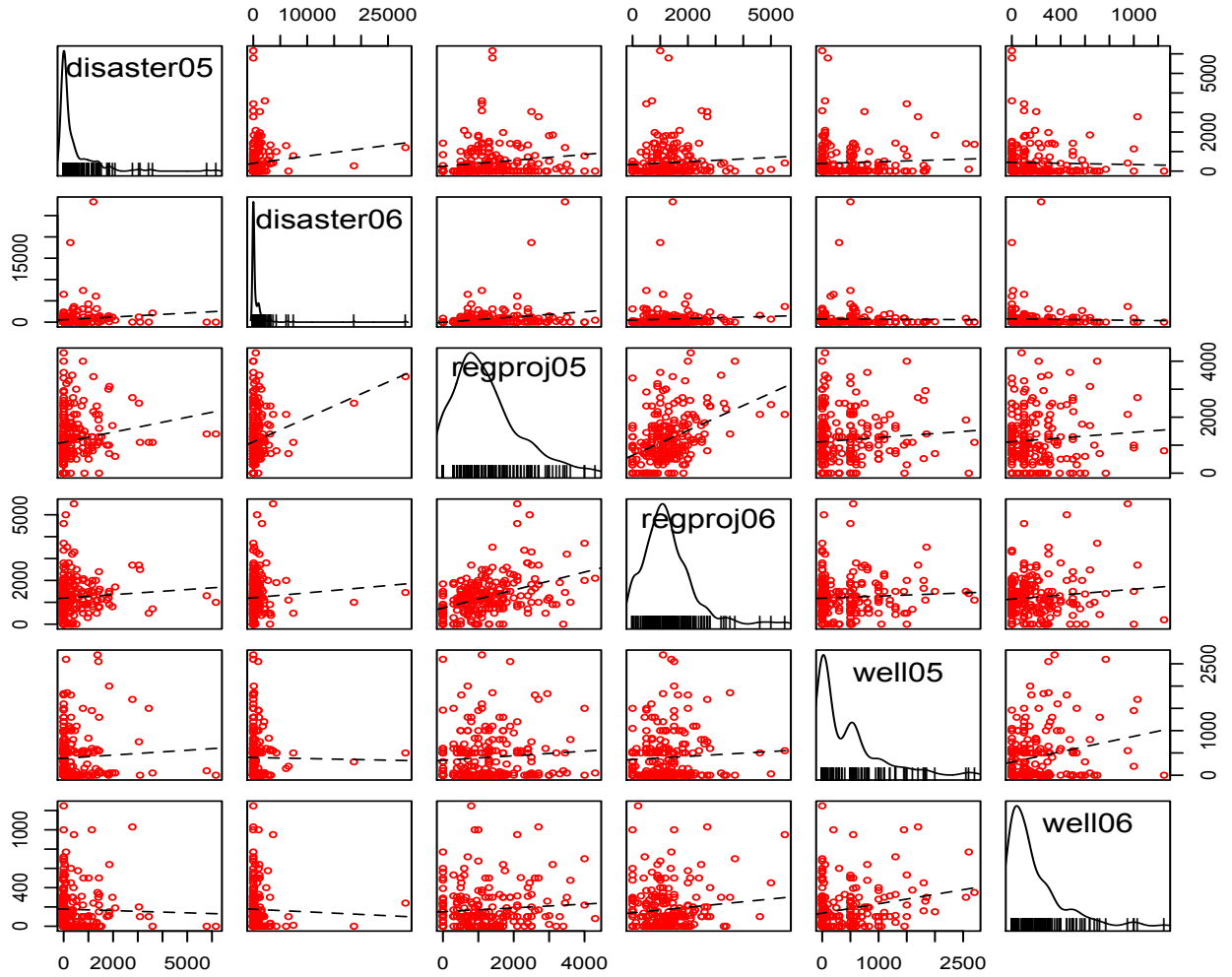


Figure 4.1: This figure shows scatterplots between possible pairs of components of the special local allocation grants. `disaster05` and `disaster06` indicate the grants allocated for recovery from natural disasters in 2005 and 2006, respectively. `regproj05` and `regproj06` represent the grants allocated for the two components (regional needs, and pilot projects) in 2005 and 2006, respectively. Likewise, `well05` and `well06` indicate the grants allocated for financial incentives for well-performed local governments in 2005 and 2006, respectively. The black dotted-lines indicate linear best fit lines. Diagonals show densities of the components.

Predictor	(1)+(2)+(3)+(4)	(1)	(2)+(3)	(4)
Financial Autonomy	-27.475* (8.888)	-10.840 (7.326)	-15.539** (3.672)	-1.096 (1.705)
Population	0.002* (0.001)	-0.001 (0.001)	0.002** (0.001)	0.001** (0.000)
Population Density	-0.098** (0.019)	-0.041* (0.016)	-0.048** (0.008)	-0.009* (0.004)
Vote margin	1.287 (7.487)	-5.912 (6.171)	5.897 (3.093)	1.302 (1.436)
Connection	137.919 (358.143)	162.472 (295.179)	81.453 (147.963)	-106.006 (68.692)
Committee	-59.705 (355.765)	-403.300 (293.219)	333.437* (146.981)	10.158 (68.236)
Seniority	-2.216 (90.807)	-19.924 (74.843)	14.672 (37.516)	3.037 (17.417)
Government Party	208.628 (168.119)	136.950 (178.739)	10.744 (89.596)	60.935 (41.595)
Intercept	2625.639** (348.902)	1150.410** (287.562)	1295.071** (144.145)	180.159** (66.919)
Observations	234	234	234	234
R^2	0.18	0.09	0.29	0.11

Note: ** and * indicate significance at $p < 0.01$, and $p < 0.05$, respectively.

The numbers in parentheses represent standard errors.

Table 4.1: This table presents different results of a complete-pooling model by dependent variable: four components of the special local allocation grants. (1), (2), (3), and (4) represent natural disasters, regional needs, pilot projects, and financial incentives to well-performed local units, respectively. (1)+(2)+(3)+(4) means the average total grants distributed to districts for the two years (2005-6). Likewise, (2)+(3) also means the average grants allocated for regional needs plus pilot projects, and (4) indicates the average grants allocated for financial incentives to well-performed administrative units. It is notable that the component of pilot projects was created in 2006.

The STAN code for 3.3 can be written as follows:

```
1 #####
2 # Varying Intercept with a group level predictor
3 # Note: (1) t-distribution sampling model
4 #       (2) informative prior for the group level predictor
5 #####
6 data {
7   int<lower=0> N;
8   int<lower=0> J;
9   int<lower=0> K;
10  vector[N] y;
11  real X[N,K];
12  int province[N];
13  vector[J] u;
14  real m[K];
15  real<lower=0> sigma_b[K];
16 }
17 parameters {
18   real a[J];
19   real b[K];
20   real g_0;
21   real g_1;
22   real<lower=0> sigmasq_y;
23   real<lower=0> sigmasq_a;
24   real<lower=0> inv_nu;
25 }
26 transformed parameters {
27   real<lower=0> sigma_y;
28   real<lower=0> sigma_a;
29   real<lower=0> nu;
```

```

30   sigma_y <- sqrt(sigmasq_y);
31   sigma_a <- sqrt(sigmasq_a);
32   nu <- 1/inv_nu
33 }
34 model {
35   g_0 ~ normal(500,10000);
36   g_1 ~ normal(0,100);
37   sigmasq_y ~ inv_gamma(108, 77182.75);
38   sigmasq_a ~ inv_gamma(2, 200);
39
40   inv_nu ~ uniform(0,0.5);
41
42   for (j in 1:J)
43     a[j] ~ normal(g_0 + g_1*u[j], sigma_a);
44
45   b ~ normal(m,sigma_b);
46
47   for (n in 1:N)
48     y[n] ~ student_t(nu,a[province[n]] + b[1]*X[n,1]
49                    + b[2]*X[n,2] + b[3]*X[n,3]
50                    + b[4]*X[n,4] + b[5]*X[n,5],
51                    sigma_y);
52 }

```

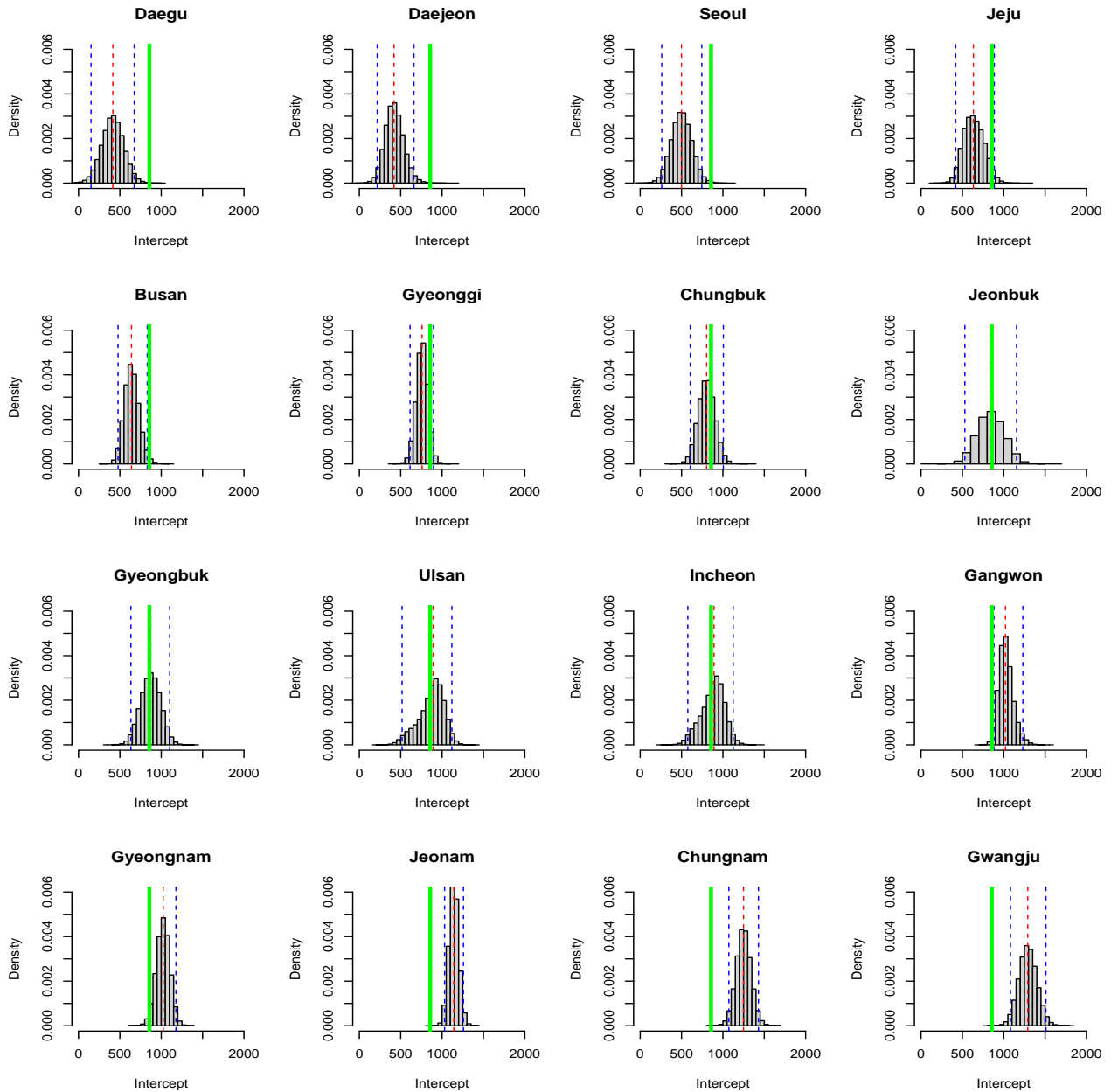



Figure 4.2: This figure shows posterior distributions of varying intercepts (α_j) when the MCMC runs 1,000,000 iterations (50,000 burn-in). Blue dashed-lines indicate 95% confidence intervals, and a red dashed-line presents the median of the posterior distribution in each province. The thick green line indicates the median of the intercepts that vary across provinces. We observe that despite the huge number of iterations the posterior distributions for some provinces such as Ulsan and Incheon are still not bell-shaped in t -distribution sampling model.

REFERENCES

- [AP01] Lizzeri Alessandro and Nicola Persico. “The Provision of Public Goods Under Alternative Electoral Incentives.” *American Economic Review*, **91**:225–39, 2001.
- [BG06] Joseph Bafumi and Andrew Gelman. “Fitting Multilevel Models When Predictors and Group Effects Correlate.” Unpublished Manuscript, March 2006.
- [CGG09] William Clark, Matt Golder, and Sona Golder. *Principles of Comparative Politics*. CQ Press, Washington D.C., 2009.
- [Cho00] Kisuk Cho. *Hangukui Jiyeokjuwi Seongeowa Haprijoek Yukwonja [Regional Voting and Rational Voters]*. Nanam, Seoul, 2000.
- [CK08] Yountae Choi and Sangheon Kim. “An Empirical Investigation of Distributive Politics in Special Local Allocation Tax.” *Korean Public Administration Review*, **42**:283–304, 2008.
- [CM86] Gary Cox and Mathew McCubbins. “Electoral Politics as a Redistributive Game.” *Journal of Politics*, **48**:370–89, 1986.
- [DJ02] Matz Dahlberg and Eva Johansson. “On the Vote-Purchasing Behavior of Incumbent Governments.” *American Political Science Review*, **96**:27–40, 2002.
- [DL96] Avinash Dixit and John Londregan. “The Determinants of Success of Special Interests in Redistributive Politics.” *Journal of Politics*, **58**:1132–55, 1996.
- [DL98] Avinash Dixit and John Londregan. “Ideology, Tactics, and Efficiency in Redistributive Politics.” *The Quarterly Journal of Economics*, **113**:497–529, 1998.
- [GCS04] Andrew Gelman, John Carlin, Hal Stern, and Donald Rubin. *Bayesian Data Analysis*. Chapman & Hall/CRC Press, New York, 2004.
- [GH06] Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Model*. Cambridge University Press, New York, 2006.
- [GM13] Miriam Golden and Brian Min. “Distributive Politics Around the World.” *The Annual Review of Political Science*, **16**:1–27, 2013.

- [HK09] Suk Jae Hur and Hyeok Yong Kwon. “Legislative Elections and Distributive Politics in Korea: The Case of Special Local Allocation Tax in the 17th National Assembly.” *Korean Political Science Review*, **43**:113–30, 2009.
- [HL04] Yusaku Horiuchi and Seungjoo Lee. “Regionalism and Redistribution in South Korea.” Policy and Governance Discussion Paper04-05, 2004.
- [HL08] Yusaku Horiuchi and Seungjoo Lee. “The Presidency, Regionalism, and Distributive Politics in South Korea.” *Comparative Political Studies*, **41**:861–82, 2008.
- [Hof09] Peter Hoff. *A First Course in Bayesian Statistical Methods*. Springer, New York, 2009.
- [Jac09] Simon Jackman. *Bayesian Analysis for the Social Science*. John Wiley & Sons, West Sussex, 2009.
- [Kwo05] Hyeok Yong Kwon. “Targeting Public Spending in a New Democracy: Evidence from South Korea.” *British Journal of Political Science*, **35**:321–41, 2005.
- [Lee98] Kap Yun Lee. *Hangukui Seongeowa Jiyeokjuui [Elections and Regionalism in Korea]*. Orum, Seoul, 1998.
- [LW87] Assar Lindbeck and Jörgen Weibull. “Balanced Budget Redistribution and the Outcome of Political Competition.” *Public Choice*, **52**:273–97, 1987.
- [Plu08] Martyn Plummer. “Penalized Loss Functions for Bayesian Model Comparison.” *Biostatistics*, **9**:523–39, 2008.
- [PRS07] Nicola Persico, Jose Carlos Rodriguez-Pueblita, and Dan Silverman. “Faction and Political Competition.” NBER Working Paper 13008, April 2007.
- [ROK03] ROKNEC. “A Comprehensive Bibliography on the 16th Korean Presidential Election (in Korean).” Government Publication, Republic of Korea National Election Commission, July 2003.
- [ROK04] ROKNEC. “A Comprehensive Bibliography on the 17th General Election (in Korean).” Government Publication, Republic of Korea National Election Commission, November 2004.
- [Sta13] Stan Development Team. “Stan Modeling Language: User’s Guide and Reference Manual.” Stan Version 1.3.0, April 2013.

- [SW01] Matthew Soberg Shugart and Martin Wattenberg. *Mixed-Member Electoral Systems: The Best of Both Worlds?* Oxford University Press, New York, 2001.