Solving Problems in the Panel Regression Model for Truncated Dependent Variables: A Constrained Optimization Method

Min-Hua Huang

Department of Political Science

Texas A&M University, College Station, TX 77843

email: admin@mhhspace.org

Mar 5, 2012

Abstract

Political science studies commonly uses panel data. In particular, many of these studies involve the analysis of a truncated dependent variable, such as aggregate-level voter turnout or a party's vote share. Unfortunately, panel regression, which is the standard method of analyzing panel data, contains three methodological problems: boundary violations, parameter estimation, and model specification. These issues raise concerns about the panel regression method's validity. In this article, I explain the nature of these problems and propose three models to solve boundary violations by applying constrained optimization in the least squares and maximum likelihood paradigm. Major findings indicate that the current method is sensitive to different centering methods and tends to generate false significance results. Throughout a comparative study in the admissibility of parameter estimates, I demonstrate how the three revised models can conditionally or fully eliminate boundary violations. Methodological advice is also provided regarding when and how the new methods should be employed.

Keywords: Panel Regression, Truncated Normal Distribution, Constrained Optimization, Least Squares, Maximum Likelihood

1 Introduction

Panel (or time-series-cross-section, TSCS hereafter) data is commonly used in political science studies (Wawro, 2002; Beck and Katz, 1995; 2007; Beck, 2007; Adolph, Butler, and Wilson, 2005). In particular, many of these studies involve the analysis of a truncated dependent variable (Gomez, Hansford, and Krause, 2007; Knack, 1995; Baek, 2009; Boyne et al., 2009). In American politics, the study of political participation is associated with aggregate-level (state or county) data about voter turnout in multiple temporal units, such as Current Population Survey (CPS) (Sides, Schickler, and Citrin, 2008). In political economy studies, the sovereign bond rating uses the data from a limited-point scale across different countries in multiple years from S&P and Moody's (Biglaiser, 2007). In world politics, for the past two centuries, the cross-national data of military spending has been measured as a percentage of GDP in the Correlates of War Project (COW) (Fordham and Walker, 2005). In comparative politics, research on a party's vote share considers electoral datasets, such as the Democratic Electoral Systems Around the World dataset (DESAW) (Golder, 2005). All of the aforementioned studies apply data that simultaneously possess spatial and temporal characteristics. The target of investigation is always related to a dependent variable that has boundary restrictions.[1]

The standard method of analyzing panel data is panel data regression (Greene, 2008: 180-213) in which the within- and between-groups estimators are applied to the fixed-effects or random-effects model, such as the *xtreg* command in Stata (McCaffrey

[1]Most of the work does notify readers about the limited range of the dependent variable. However, few actually discuss the statistical property of the truncated random variable. In fact, many convenient properties for a normal dependent variable do not hold in the truncated normal random variable. For example, linear transformation of a truncated normal random variable does not generate a truncated normal random variable. It has been proven that the simple arithmetic operations, such as additivity, do not work for a truncated normal random variable. See Horrace (2005a; 2005b).

et al, 2010).[2] The basic idea is to purge between-groups variance by subtracting the group means from the pooled regression, and then the OLS method can be applied to the within-groups regression. In other words, all the constants of the spatial units, representing any omitted time-invariant variables, are canceled out after the de-meaning operation (Wooldridge, 2005). Hence, the OLS estimator is BLUE (Baltagi, 2011:308). This approach is equivalent to the least squares dummy variable estimation (LSDV), typically known as the fixed-effect model (FE) (Hsiao, 2003: 30-33). Its advantage is avoiding working with the large covariate matrix if the number of spatial units is plenty. This approach achieves its goal by utilizing the important property–the equivalence of differencing and dummying (Wicharaya, 1995:200).

However, when the dependent variable is distributed as truncated normal, differencing is not equivalent to dummying, since the time mean of the truncated dependent variable is a biased estimate of the district-level location parameter (Hsiao, 2003: 243). This means that using the time mean to characterize the contextual effect of the spatial units is not valid(Alan et al. 2011). Therefore, a methodological problem occurs if we apply the panel data regression to analysis of a truncated dependent variable.

An alternative approach is using the LSDV estimator instead. While LSDV does not suffer from the above problem, the boundary restrictions of the dependent variable still pose a challenge to the feasibility of the parameter estimates for the panel data regression. In fact, neither the OLS (equivalent to LSDV when dummy variables are specified) nor the truncated regression model (such as *truncreg*) (Cong, 2000) can solve the boundary violations problem. If we seek nonlinear programming techniques, such as constrained optimization (Bertsekas, 1996), to resolve this issue,

[2] We do not intend to discuss all estimators for panel regression. Rather, we want to employ the simplest method to illuminate the basic problems. For more discussions on the strategies for analysis of TSCS data, see Franzese (2005) and Franzese and Hays (2007).

the dimensionality of the parameter space then becomes a critical concern, given that k regressors will engender $(2k + 6)$ boundary constraints in a constrained optimization problem (hereafter COP).[3] If k is in the thousands, which is usually the case, solving a COP becomes nearly impossible given the limited capacity of any personal computer (Greene, 2008: 195).

The essential problem of the panel regression for a truncated dependent variable is an out-of-bounds violation of the predicted value– a common scenario in political science research that is seldom reported. For instance, in Benjamin Fordham and Thomas Walker's 2005 article published in *International Studies Quarterly*, "Kantian Liberalism, Regime Type, and Military Resource Allocation: Do Democracies SpendLess?", the authors analyze three dependent variables, "Military spending as a percentage of GDP," "Military personnel as a percentage of population," and "Regression-based index of military allocation" using 14 panel regressions in total. All three dependent variables have a lower bound value of 0. However, nine of the 14 models have negative predicted values and apparently suffer from boundary violations. While Fordham and Walker do identify this problem in a footnote (n.4, p.147), they do not evaluate the admissibility of the parameter estimates, nor do they discuss how we can meaningfully interpret those out-of-bounds results.[4]

The above discussion pinpoints the protracted problem regarding analysis of the TSCS data of a truncated dependent variable via panel regression. First, the most significant and pressing issue is boundary violations. Though the problem is widespread, the political science community pays it little attention.[5] Second, if we intend to solve

[3]For a truncated regression model with k covariates, we need to specify two constraints for the maximal and minimal predicted values, $(2k+2)$ constraints for the lower and upper limits for all beta coefficients, and two for boundary constraints of the scale parameter. See supplementary document A for detail (p.9).

[4]For similar cases, see the supplementary document A for detail (p.3).

[5]While scholars have been aware of this issue and some efforts were made in the field of econometrics, for example Honorè(1992), we have not seen the emergence of a standardized approach that

boundary violations with the within- and between-groups estimators by applying the constrained optimization technique, the demeaning operation generates an invalid estimate of the within-groups variations. Third, if we use the LSDV estimator, the large number of spatial units makes constrained optimization implementation impractical. Consequently, despite the fact that the panel regression model has the same problem as the linear probability model in terms of out-of-bounds violations (Aldrich and Nelson 1984:24), political scientists do not question the validity of the panel regression as they would to the linear probability model.[6]

In this article, I propose a constrained optimization method to revise the panel regression model. This method can be implemented under different settings in both the least squares and maximum likelihood paradigms. In section two, I provide a comprehensive discussion of the methodological problems of the panel regression model. Next, I explain how to revise the panel regression given different scenarios: (1) taking least squares or maximum likelihood assumptions (2) whether to adopt the demeaning-bias correction. Then, I carry out a comparative study in the admissibility of parameter estimates for the current panel regression and various revised models, which is followed by the discussion and conclusion sections.

is widely accepted and available in statistical packages such as SPSS, SAS, or Stata.

[6]The usage of the logit or probit has already become the norm in social science when a binary dependent variable is analyzed. However, almost no political science literature contains a truncated regression model. Since binary and truncated dependent variables share the same nature of the problem, truncated regression certainly receives much less attention than it should.

2 The Panel Regression's Methodological Problems

The methodological problems of the panel regression can be summarized and put into three categories: boundary violations, parameter estimation, and model specification. Throughout this paper, I use the example in the first column of Table 1 from Thomas Hansford and Brad Gomez's 2010 article "Estimating the Electoral Effects of Voter Turnout" from *American Political Science Review* to discuss these problems and present the results of the revised models. The dependent variable is county-level democratic vote share in presidential elections from 1948 to 2000. The independent variables include (1) *Partisan composition*, measured as the moving average of the Democratic vote share in the three most recent elections, (2) *Turnout*, meaning voter turnout, (3) the two interaction terms *Turnout*×*GOP Incumbent* and *Turnout*×*Partisan composition*, where *GOP Incumbent* is a dummy variable for a presidential election in which the incumbent is Republican, and (4) 13 time dummy variables indicating the temporal units from 1952 to 2000. The default category, represented by the constant, is the presidential election in 1948 in which the incumbent is Democrat. The dataset, comprised of typical TSCS data with a sample size of 27401 that covers 1964 counties and 14 presidential elections, was compiled by the two authors from various sources. We confine our discussion to the fixed-effect model.

2.1 Boundary Violations

The panel regression is specified as

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)\,\beta + (e_{it} - \bar{e}_i)\,. \qquad (1)$$

In Stata, the grandmeans $\bar{\bar{y}}$ and $\bar{\bar{x}}$ are added back to the regression model for estimation of the intercept (Gould, 2011).

There are two kinds of boundary violations– empirical and theoretical out-of-bounds predictions. Empirical boundary violation is defined as

$$(x_{it} - \bar{x}_i)\,\hat{\beta} < a \qquad or \qquad (x_{it} - \bar{x}_i)\,\hat{\beta} > b,$$

where a and b are the lower- and upper-bounds of the dependent variable y_{it}. In plain language, if the predicted value of any empirical observation falls outside the permissible boundary, it is a case of empirical boundary violation.[7]

Theoretical boundary violation refers to the case in which an out-of-bounds predicted value occurs for any possible observation, given the covariate space defined by the empirical data. For example, the range of the covariates *Partisan composition* and *Turnout* are [10.14%,88.98%] and [20.37%,100%]. Given the beta estimates as shown in Table 1, we can derive that the predicted value of democratic vote share is 128.96% for a case in which *Partisan composition* = 88.98% and *Turnout* = 100% in the 1964 presidential election if the incumbent is a Republican. In the same way, the predicted vote share is −20.7% for a case in which *Partisan composition* = 10.14% and *Turnout* = 20.37% in the 1972 presidential election if the incumbent is a Democrat. Both cases indicate theoretical boundary violations because the joint presence of covariate values is logically possible, despite their empirical nonappearance. The same problem can be demonstrated in less dramatic examples. For instance, in the 1964 presidential election, if the incumbent is a Republican and given

[7] The dependent variable has different meanings if different demeaning operations are applied. If y_{it} is used without additional operations, y_{it} denotes vote share and is bounded with 0 and 1. If y_{it} is demended by $\bar{y}_i$, $y_{it} - \bar{y}_i$ means within-groups deviations. If the grandmean is added back, $y_{it} - \bar{y}_i + \bar{\bar{y}}$ indicates within-groups variations plus a baseline vote share. That is how Stata reports panel regression with a constant measure.

that *Partisan composition* = 60% and *Turnout* = 60%, the predicted vote share is 108.22%. In the 1972 presidential election if the incumbent is a Democrat and given that *Partisan composition* = 40% and *Turnout* = 40%, the predicted vote share is -2.62%. Again, both cases are not outlier cases, and similar cases do exist in the existing temporal domain, except for the *GOP Incumbent* variable. For the former case, there are 37 cases in which both covariates deviate from 60% by a 1% margin in all temporal units except '52, '84, and '88. For the latter, there are seven cases in which both covariates deviate from 40% by a 1% margin in '56, '76, '88, '98, and '00.

Some scholars may argue that only empirical observations count, and we do not have to consider those theoretical out-of-bounds cases. However, such an argument contradicts the fundamental reasoning of statistical inference, that is, using a variable to decontexualize a concept for universal comparison (Kellstedt and Whitten, 2009: 7-14). If the *GOP Incumbent* variable in 1964 could only be 0 in order to reflect the historical truth, then all observations in our empirical data should be viewed as idiosyncratic.[8] Thus, statistical inference is not possible. For this reason, we should consider the theoretical boundary violation as a sign of invalid parameter estimates.

2.2 Parameter Estimation

Parameter estimation differences in the least squares and maximum likelihood paradigms are the formulation of the objective function. In the least squares paradigm, the objective function is simply the sum of squares of residuals. We can regard the boundary restriction of the dependent variable as the linear constraints. Parameter estimation

[8] We are not opposed to interpreting a time dummy as a composite estimate of the time effect in a given place. Thus, all the idiosyncratic effects have already been lumped into a single measure. However, this does not point to the uniqueness of time dummies; rather, the maximum or minimum of such measures reflect the greatest or least effects of the time factor that has ever occurred. Since this has happened previously, we have no reason to rule out the possibility that it will happen again.

can be specified as a quadratic programming problem (QP): (Vanderbei, 2008)

$$\begin{aligned} &\textit{Minimize} \qquad f(\beta) = \sum_{i=1}^{n}\sum_{t=1}^{T_i}\left[(y_{it}-\bar{y}_i)-(x_{it}-\bar{x}_i)\beta\right]^2 \\ &\textit{Subject to} \qquad (x_{it}-\bar{x}_i)\beta \le b \\ &\qquad\qquad\qquad -(x_{it}-\bar{x}_i)\beta \le -a, \end{aligned}$$

where the number of the spatial and temporal units is n and T_i, and the parameter space $\beta \in \Omega_\beta$. We have not exactly specified what Ω_β should be. This is critical to solve the above QP problem successfully. We will discuss this issue in section 3.

We can modify the objective function slightly by providing a distributional assumption to the demeaned dependent variable $(y_{it}-\bar{y}_i)$ in (1)

$$(y_{it}-\bar{y}_i) \sim TN\left[(x_{it}-\bar{x}_i)\beta, \sigma^2; p_1, q_1\right], \tag{2}$$

Here the setup of the lower and upper limits, p_1 and q_1, are vital to a successful maximum likelihood estimation. Theoretically, p_1 and q_1 should be set to their greatest and least possible values according to the available data information.

$$\begin{aligned} p_1 &= a - \left(\bar{\bar{y}} + t_{\sigma_b}^{\min}\cdot\sigma_b\right) \\ q_1 &= b - \left(\bar{\bar{y}} + t_{\sigma_b}^{\max}\cdot\sigma_b\right). \end{aligned}$$

where σ_b is the between-groups deviation estimated by the untruncated normal assumption, and $t_{\sigma_b}^{\min}$ and $t_{\sigma_b}^{\max}$ refer to the largest deviation of $\bar{y}_i$ in the negative and positive directions, respectively.

Given this distributional assumption, the objection function can be specified as

$$Maximize \qquad \log L \equiv -\sum_{i=1}^{n}\sum_{t=1}^{T_i} D_{it} - \frac{1}{2\sigma^2}\left[(y_{it}-\bar{y}_i)-(x_{it}-\bar{x}_i)\beta\right]^2,$$

where $D_{it} = \sqrt{2\pi}\sigma\left[\Phi\left(\frac{p_1-(x_{it}-\bar{x}_i)\beta}{\sigma}\right)-\Phi\left(\frac{q_1-(x_{it}-\bar{x}_i)\beta}{\sigma}\right)\right]$. We must be aware that one additional parameter σ is added in the above likelihood function, and the parameter space $\sigma \in \Omega_\sigma$ needs to be specified. The inequality constraints and the parameter space $\beta \in \Omega_\beta$ remain the same.

From the perspective of the likelihood paradigm, the two objective functions above are all problematic, since the panel regression is incorrectly specified in the first place.[9] To see why this is so, we first assume that the dependent variable y_{it} is distributed as truncated normal

$$y_{it} \sim TN\left(\mu_i, \sigma; a, b\right),$$

where μ_i is the district-level location parameter. The time mean of y_{it} is

$$E_t\left(y_{it}\right) = \mu_i - \frac{\sigma\left\{\exp\left[-\frac{(b-\mu_i)^2}{2\sigma^2}\right]-\exp\left[-\frac{(a-\mu_i)^2}{2\sigma^2}\right]\right\}}{\sqrt{2\pi}\left[\Phi\left(\frac{b-\mu_i}{\sigma}\right)-\Phi\left(\frac{a-\mu_i}{\sigma}\right)\right]}.$$

Unless $b-\mu_i = \mu_i - a$ or $(a,b) \to (-\infty,\infty)$, the time mean $E_t\left(y_{it}\right)$ (or noted as $\bar{y}_i$)

[9] In contemporary statistical science, the likelihood theory is a crucial paradigm of inference for data analysis (Royall, 1997:xiii). It provides a unifying approach of statistical modeling to both frequentists and Bayesians with the criterion of maximum likelihood (Azzalini, 1996). The rapid development of political methodology in the last two decades has also witnessed the establishment of the likelihood paradigm in the scientific study of politics (King, 1998). As a model of inference, the fundamental assumption of the likelihood theory is the likelihood principle, which states that "all evidence, which is obtained from an experiment, about an unknown quantity θ, is contained in the likelihood function of θ for the given function." (Berger and Wolpert,1984:vii) In other words, given the fact that the likelihood function is defined by the probability density (or mass) function, we must make a distributional assumption of the dependent variable to derive a likelihood function. The plausibility of such a distributional assumption is therefore vital to the validity of the statistical inference.

is a biased estimate of μ_i (Johnson and Kotz, 1970: 81).[10] If we intend to purge the between-groups variation, we should subtract μ_i from y_{it}. However, we mistakenly use the time mean $E_t(y_{it})$ to estimate μ_i, and the differencing operation $y_{it} - E_t(y_{it})$ fails to generate the valid within-groups variation. This illustrates the common problem for the previous panel regressions in (1) and (2).

If we want to specify a conceptually equivalent model as the panel regression, we can use the individual-level dependent variable to estimate district-level location parameters μ_i, and then perform the demeaning operation to derive the within-groups regression.[11] In this scenario, the dependent variable can be specified as

$$(y_{it} - \hat{\mu}_i) \sim TN\left(x_{it}^*\beta, \sigma^2; p_2, q_2\right),$$

where x_{it}^* represents the covariate matrix that is fixed at the minimum after being demeaned, and

$$\begin{aligned} p_2 &= a - \left(\hat{\mu} + t_{\sigma_{\hat{\mu}_i}}^{\min} \cdot \sigma_{\hat{\mu}_i}\right) \\ q_2 &= b - \left(\hat{\mu} + t_{\sigma_{\hat{\mu}_i}}^{\max} \cdot \sigma_{\hat{\mu}_i}\right). \end{aligned}$$

Applying a maximum likelihood estimation,we can derive the objective function as

$$Maximize \qquad \log L \equiv -\sum_{i=1}^{n}\sum_{t=1}^{T_i}\left\{D_{it} - \frac{1}{2\sigma^2}[(y_{it} - \hat{\mu}_i) - x_{it}\beta]^2\right\},$$

where $D_{it} = \sqrt{2\pi}\sigma\left[\Phi\left(\frac{p_2 - x_{it}^*\beta}{\sigma}\right) - \Phi\left(\frac{q_2 - x_{it}^*\beta}{\sigma}\right)\right]$. Since the demeaning operation is

[10] When $b - \mu_i = \mu_i - a$, the normal distribution is evenly truncated at both ends. When $(a, b) \to (-\infty, \infty)$, the variable is not truncated at all. Both situations rarely occur when the dependent variable is distributed as truncated normal.

[11] This involves a two-stage procedure. In the first stage, $\hat{\mu}_i$ is estimated by μ_{it} without covariates. In the next stage, we take $\hat{\mu}_i$ as the district-level property and subtract it to derive complete within-groups deviation.

achieved with the maximum likelihood estimates of $\hat{\mu}_i$, no demeaning on the covariate matrix is necessary. However, we will apply the same demeaned specification for the sake of comparability. In section 3, we will apply the technique of constrained optimization to solve the three optimization problems given different parameter constraints.

2.3 Model Specification

Proper model specification is critical for the successful application of constrained optimization to the panel regression. While the demeaning operation is easy to apply, it generates some questions regarding interpretability, as well as logical soundness. For the covariates that can vary independently at the individual level, the demeaning operation experiences no problem, since the time mean represents the aggregate-level measure, which reflects the characteristic of the spatial units across time. However, when the covariates cannot vary independently, such as when they function as a time dummy or have involvement with an interaction term, the demeaning operation is problematic and unnecessary.

In Hansford and Gomez's study, there are 13 time dummies, and their values after demeaning operation range from -0.26 to 1. This scenario contradicts our intuition because a time dummy is either 0 or 1. Mathematically, after being demeaned, a time dummy should remain as 0 or 1 since

$$\begin{aligned} x_{it}^d - \bar{x}_i^d + \bar{\bar{x}} &= x_{it}^d - \frac{1}{T} - \frac{n}{n \cdot T} \\ &= x_{it}^d, \end{aligned}$$

where the superscript d denotes a time dummy. And here, we assume the data consists of balanced panels. Therefore, if time dummies have values other than 0 or

1, it indicates a missing value problem. For instance, the time dummy *Yr2000* has a lowest value of -0.26 in the '48 observation #1471, since there are only three cases available in this county, including the case in 2000 (county id: #39075). On the other hand, the time dummy *Yr2000* has a positive value of 0.070 in the '00 observation #25095, since only the '00 case is missing in this county (county id: #41049). Given that the demeaning operation is unnecessary and the result is difficult to interpret, we use the original data for the time dummies.

Another common issue in political studies is associated with interaction terms. However, we often neglect the nonlinear nature of the interaction model, especially where different centering methods are involved.[12] If we center non-interaction variables before forming interaction terms, the regression result will be different from centering all variables after interaction terms were formed. Suppose $x_3 = x_1 \times x_2$ and x_1 and x_2 are normally distributed, then the above case can be illustrated by the following equations:

$$y = \beta_0^{(1)} + (x_1 - \bar{x}_1)\,\beta_1^{(1)} + (x_2 - \bar{x}_2)\,\beta_2^{(1)} + (x_1 - \bar{x}_1)\,(x_2 - \bar{x}_2)\,\beta_3^{(1)} \tag{3}$$

$$= \beta_0^{(2)} + (x_1 - \bar{x}_1)\,\beta_1^{(2)} + (x_2 - \bar{x}_2)\,\beta_2^{(2)} + (x_1 x_2 - \overline{x_1 x_2})\,\beta_3^{(2)}. \tag{4}$$

Apparently, $\beta_i^{(1)} \neq \beta_i^{(2)}$ for $i = 1, 2, 3$, regardless of whether y is distributed as un-truncated or truncated normal.

When applying constrained optimization to panel regression, we suggest applying the fixed-at-minimum model if the regression includes dummy variables or interaction terms. The reason is twofold: first, the value of time dummies in many observations will approach zero as the temporal units increase, which will occasionally cause numer-

[12]We refer the centering methods to any model specification that involves linear transformation of the regression model.

ical problems; second, if we formulate the interaction term after fixing the composing variables at the minimum, given that the range of all variables is in the positive domain, we can simplify the specification of boundary constraints by eliminating the cases in which two negative composing variables result in a positive interaction term. Centering interaction variables, on the other hand, would complicate the specification of the admissible parameter space for β, and thus cause difficulty in parameter estimation.

3 Solving Boundary Violations with Constrained Optimization

Constrained Optimization is an important quantitative method in (non)linear programming and numerical optimization. It has been widely applied in molecular biology, meteorology, physical oceanography, industrial management, and many engineering fields (Bonnans et al.,2006: 5-10). The fundamental issue in a constrained optimization problem (COP) is to achieve the optimality under a given set of objective functions and parameter constraints, including equality and inequality constraints. When the objective function is linear, we refer to it as a "linear programming" problem. When the objective function is quadratic or higher order, we call it a "nonlinear programming" problem.

Identification of the objective function is the first step in specifying a COP problem. We already completed this task in the previous section by specifying three objective functions under least squares and maximum likelihood paradigms. All three objective functions are nonlinear. Next, we need to specify proper constraints to achieve admissible parameter estimates that do not suffer from boundary violations.

The basic set of the constraints for the three objective functions is defined by the beta coefficient $\hat{\beta}_m$. We can generalize those constraints by

$$
\begin{aligned}
g_1 &= \beta_0 + \hat{y}_{\sim 0}^{\max} - b \leq 0 \\
g_2 &= a - \beta_0 - \hat{y}_{\sim 0}^{\min} \leq 0 \\
g_3 &= \beta_0 - b \leq 0 \\
g_4 &= -\beta_0 + a \leq 0 \\
&\vdots \\
g_{2m+3} &= \beta_m - \frac{b - \hat{y}_{\sim m}^{\max}}{x_m^{*\max}} \leq 0 \\
g_{2m+4} &= -\beta_m + \frac{a - \hat{y}_{\sim m}^{\min}}{x_m^{*\max}} \leq 0
\end{aligned}
$$

For maximum likelihood estimation, we need to add two more constraints on the scale parameter σ as

$$
\begin{aligned}
g_{2m+5} &= \sigma - b + a \leq 0 \\
g_{2m+6} &= -\sigma + \kappa \leq 0.
\end{aligned}
$$

The above specification is deduced from a regression model that fixes all the variables at the minimum level with the exception of the constant covariate. After being fixed, the covariate matrix is expressed with an asterisk sign x^*. The predicted value of the dependent variable is bounded within the lower and upper limits, a and b. The eligible parameter space of the beta coefficient β_m can be derived by (1) fixing the independent variables of interest at the minimum, while finding the greatest and least predicted values, $\hat{y}_{\sim m}^{\max}$ and $\hat{y}_{\sim m}^{\min}$, (2) varying the covariate value of interest from minimum to maximum to derive its largest or least possible contribution, b and a,

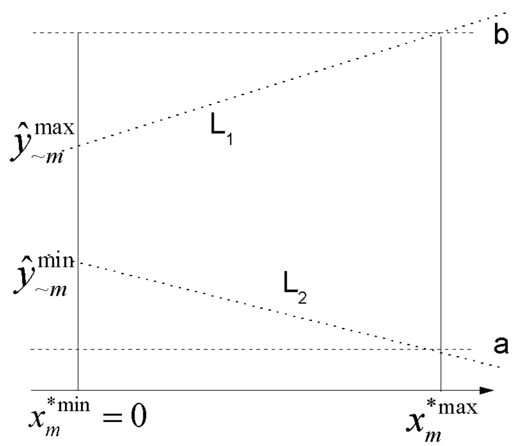

Figure 1: Boundary Constraints of β_m

and (3) dividing the largest and least possible contribution by the maximum range $x_1^{*\max} - x_1^{*\min}$, where $x_1^{*\min} = 0$.[13] Figure 1 illustrates this procedure as the above describes.

The first two constraints define the upper and lower bounds of the predicted value for the dependent variable. The third and four constraints specify the range of the constant, which refers to the level of the dependent variable when all the covariates are held at the minimum; and apparently, the upper and lower limits are the same as the dependent variable. For the rest of $2m$ equations, constraints are about the admissible parameter space of beta coefficients from x_1 to x_m. The two additional constraints specify the maximum and minimum value of the scale parameter. The maximum is set to the full range because the distribution will deviate from truncated normal and approach uniform distribution when $\sigma \gg |b - a|$. The minimum is set to an arbitrary small number κ to prevent negative variance from occurring.

We apply the sequential quadratic programming (SQP) algorithm to solve a COP

[13] If the interaction terms are fixed at the minimum at the same time as other variables, we will not be able to derive $x_1^* \times x_2^* = (x_1 \times x_2)^*$, as the previous discussion about (3) and (4) explains.

problem and derive parameter estimates. The idea of the SQP algorithm is to break down a complicated COP problem into a series of osculating quadratic problems (OQP) (Gilbert, 2009). Given that quadratic problems are much easier to solve, we can approach the optimal solution by solving the subproblems until we reach convergence throughout a sequence of iterations (Powell, 1978). To execute the SQP algorithm, we need to have information about the constraint vector, the first derivative matrix of the constraint vector, the gradient vector of the objective function, and the hessian matrix of the Lagrangian function. Beginning with giving the initial value,[14] we will derive a new set of solutions by updating those four matrices, and then we will evaluate whether the convergence is achieved by checking the Karush-Kuhn-Tucker conditions (Kuhn and Tucker, 1951), which are necessary terms for an optimal solution to exist in nonlinear programming. If the solution reaches convergence, we stop the iteration process and report the optimal solution.[15] If the maximum number of iterations runs out, we check the admissibility of the best available solution and report it.[16] If no admissible solution is available, then we report the estimation as failed.

Supplementary documents A and B explain the details of model specification, mathematical exposition, and simulation findings. According to previous research presented in these documents, applying constrained optimization to the modified truncated regression model exhibits unconditional superiority in terms of eliminating boundary violations and deriving the best admissible likelihood measure. Comparing the OLS method and the current truncated regression model, the regression model that incorporates constrained optimization can always find an admissible solution and

[14] The initial parameter values are set by the solution of the current panel regression.

[15] We adopt the SQP algorithm suggested by Bonnans et al. (2006: 257).

[16] Due to the large sample size, we reduce the maximum number of iterations to 31. The tolerance value applied to check the KKT conditions is set to $10^{-}4$. The step-size parameter is set to $\tau = 50$.

effectively solves the out-of-bound violations, regardless of how they are empirically or theoretically defined.

4 Applying Constrained Optimization to the Panel Regression

In this section, we apply constrained optimization to the panel regression by replicating Hansford and Gomez's (2010) model as mentioned at the beginning of section 2. Five different models that use various settings will be presented. Each model has different assumptions in model specification or parameter estimation, or both. Model I and Model II differ in the centering methods in model specification, while they share the least-squares objective function. Additionally, neither model applies constrained optimization in parameter estimation. Model II and III adopt the same model specification, but they differ in whether or not to apply constrained optimization. Model III and IV both apply constrained optimization, but they differ in the specification of least-squares or maximum likelihood objective functions. Model IV and V share all the assumptions in parameter estimation, but they differ in whether to correct the demeaning bias in model specification. Table 1 sums up the similarities and differences for the five models' assumptions.

Model I is the original model in which we report all beta coefficients, including the time dummies. We replicate this model using the *xtreg* command in the Stata environment. The result is presented in Table 2, and we regard this as the base model, since Stata's *xtreg* command is one of the most widely used methods for panel regression. Due to the nature of the large sample size, all of the beta coefficients are statistically significant at the 0.01 level. Most of the beta coefficients are nearly zero or negative, except *Partisan composition*, *Turnout*$\times$*GOP Incumbent*, *Yr64*, and *Yr96*, and the constant.

Model II adopts a different specification by fixing all the covariates at the minimum level. In comparing the results of Model I and II, while we expect differences related to the constant and the interaction terms because the model specification changes, the actual difference is far greater than expected. Not only do three of the four main explanatory variables have significant differences, 7 of 13 time dummies also have a different estimate by at least a 3% margin. Moreover, two beta coefficients, *Turnout*$\times$*GOP Incumbent* and *Yr60*, now become insignificant. We reach the same conclusion: different centering specification significantly alters the panel regression results if the model specification involves interaction terms.

Thus far, we have not applied constrained optimization to the panel regression. In the next three models, we incorporate this nonlinear programming technique under different statistical settings. The first is Model III (LSCO), as reported in Table 3, in which we specify a least squares objective function without any distributional assumption. The model specification is the same as Model II, fixing covariates at the minimum, except that the constant is not comparable given different demeaning operations. We intend to compare Model III with Model II to see how they differ if constrained optimization is applied. The results indicate that the four main beta coefficients have substantial differences, while the time dummies are relatively stable. Specifically, *Partisan composition* drops .212% from .799% to .587%, and *Turnout*, *Turnout*$\times$*GOP Incumbent*, and *Turnout*$\times$*Partisan composition* all have different results of significance.

Model IV (MLCO1) is distinct from Model III when taking the distributional assumption into consideration and using the maximum likelihood estimation to form the objective function. The demeaning operation assumes an untruncated normal distribution to the dependent variable. While both models apply constrained optimization, we expect some differences, since the objective function and the parameter set are different. As Table 3 makes evident, the estimated beta coefficients are significantly different. The major difference is not simply the magnitude, but also the level of significance, as well as the sign. For instance, the small change of the beta coefficient is in the *Turnout*$\times$*GOP Incumbent*, where the difference is 0.012 when its coefficient is increased from a significant -.031 to insignificant -.019. Considering that the magnitude of the interaction term varies from hundreds to thousands, the 0.012 change is actually quite significant.[17] As for the sign of the beta coefficients, 2 of 18 change the sign. This striking difference illustrates the fact that the panel regression under the least squares and maximum likelihood paradigm could generate quite different results with the same model specification.

We further correct the district-level location parameter bias due to the demeaning operation by assuming a truncated normal distribution. Model V (MLCO2) reports the panel regression results with constrained optimization. Comparing Model V and Model IV, their difference is very limited. None of the 18 estimated beta coefficients have a noticeable difference in sign or magnitude, and all of the differences are under the 1% margin. Apparently, the use of the biased estimate, the time mean, to evaluate the district-level properties only causes a marginal difference.

In terms of the boundary violations, none of the five models generate empirical violations, due to the relatively stable statistics of democratic vote share. Unfortunately, in a two-party system, such as in United States, this nice data property obscures the out-of-bound problem that is common in other contexts, particularly within the scope

[17] We do not mean to evaluate the magnitude of contribution by multiplying an interaction term with the margin of the beta coefficient's change. Since the variation of interaction terms is not independent of its composing variables, we need to incorporate information of other relevant covariates and coefficients to evaluate the actual contribution.

of cross-national research that covers a very fragmented to one-party system. However, considering theoretical boundary violations as Table 4 shows, the two panel regressions without applying constrained optimization both suffer this problem at different levels. For the original specification, there are 14 boundary violations, while Model II has 37 violations. With regard to Model III, we cannot find an admissible solution in the original setting; therefore, we introduce a rescaled parameter η that reduces the between-groups variability by multiplying a fraction as follows:

$$
\begin{aligned}
p_1^* &= a - \left(\bar{\bar{y}} + t_{\sigma_b}^{\min} \cdot \sigma_b \cdot \eta\right) \\
q_1^* &= b - \left(\bar{\bar{y}} + t_{\sigma_b}^{\max} \cdot \sigma_b \cdot \eta\right),
\end{aligned}
$$

where p_1^* and q_1^* are new lower and upper bounds for the demeaned dependent variable $(y_{it} - \bar{y}_i)$. We start from 100% and decrease by 1% each time, and finally reach an admissible solution at $\eta = 60\%$. In this way, we derive a solution without theoretical boundary violation, but on the condition that it only applies to those counties where the average democratic vote share ranges from 26.48% to 63.67%, which numbers 95.57% in all 1964 counties.

At last, the two models that use maximum likelihood estimation successfully eliminate theoretical boundary violations. Their estimates are relatively stable, regardless of whether the demeaning bias is corrected. Findings in both models are far more conservative than the original panel regression, and only one of the four explanatory variables (*Partisan composition*) has a significant relationship. Moreover, the standard error measures are also larger, and hence lead to smaller t values. These findings indicate that, the current panel regression tends to generate lenient results and likely to suffer from the problem of false significance.

5 Discussions

The previous analysis can be summarized with the following findings: (1) Results of panel regression are sensitive to different centering methods in model specification, (2) Without applying constrained optimization, panel regression is likely to suffer from theoretical boundary violations, (3) Application of constrained optimization in panel regression greatly alters the results in the least squares, as well as maximum likelihood paradigm, (4) Despite the fact that constrained optimization can solve the boundary violation problem with a least-squares or maximum-likelihood estimation, the former only gives a conditional solution, while the latter achieves a complete solution, (5) Within the maximum likelihood paradigm, the demeaning bias only causes a marginal difference in a panel regression, (6) In comparison to the panel regression that applies constrained optimization, the current model tends to generate false significance results.

The above findings raise some potential problems that could significantly compromise the validity of political science research if the current panel regression method is applied to an analysis of a truncated dependent variable. The first concern is empirical boundary violations. Under no circumstances is a regression result that generates out-of-bound predicted value acceptable. Next, theoretical boundary violations must be considered. While some people might insist that regression results only explain the empirical data and do not extend to possible cases that do not appear in our sample, this view greatly limits the scope to which our analysis could apply. Third, in order to solve boundary violations with constrained optimization, different estimation methods must be into considered. The choice of least squares and maximum likelihood could generate very different results in panel data analysis. Fourth, if we select the least squares method, a full solution might not always be available, although we can always find a conditional solution that applies to a varying range of cases. On the other hand, despite the fact that maximum likelihood estimation is more likely to achieve a full solution, its results tend to be less significant.

I do not argue that constrained optimization has to be imposed to panel regression, but this article provides three different resolutions when empirical or theoretical boundary violation occurs. Given the same model specification, my analysis suggests that the current panel regression (Model II) and the maximum likelihood with correction of demeaning bias (Model V) can serve as the optimistic and conservative results in hypothesis testing, respectively. If a significance finding appears in both models, then the robustness of this finding is corroborated. However, if the results in both models do not agree, we need to be careful in interpreting them, since different results are associated with different methodological assumptions. Nevertheless, I do recommend adopting constrained optimization whenever panel regression suffers from empirical boundary violations, because their occurrence is illogical.

6 Conclusions

This article presents three methods under different scenarios with least squares or maximum likelihood assumptions. All three methods can conditionally or fully solve the boundary violations problem. The analysis indicates that the results of current panel regression are not only subject to some methodological problems, but also sensitive to different centering methods. By applying the technique of constrained optimization, those methodological problems can be targeted and the best possible solution can successfully be reached. I suggest a mandatory use in one of the three methods when an empirical boundary violation occurs. The robustness of the result can be also checked by comparing the results from the current method and the maximum likelihood method that corrects demeaning bias.

While the application of constrained optimization can eliminate boundary violations, it costs more time and computing capacity to execute the revised methods, especially with the maximum likelihood assumption. Given the scope of this article, a systematic assessment is yet to be done with regard to different methods's performance under various conditions. We expect more future work in simulations, as well as empirical studies to illuminate to what extent the current panel regression suffers from boundary violations, and to what extent the revised models can successfully eliminate these violations. By possessing this knowledge, scholars in this field can provide more definite criteria to prevent political science research from reporting an illogical out-of-bound finding.

References

[1] Adolph, Christopher, Daniel M. Butler and Sven E. Wilson. 2005. "Which Time-Series-Cross-Section Estimator Should I Use Now? Guidance from Monte Carlo Experiments." Paper presented at the 2005 Annual Meeting of the American Political Science Association, Washington, D.C.

[2] Alan, Sule, Bo E. Honore, Luojia Hu, and Soren LethVPetersen. 2011. Estimation of Panel Data Regression Models with Two-Sided Censoring or Truncation. Federal Reserve Bank of Chicago, Working Paper Series, WP-11-08.

[3] Aldrich, John H., and Forrest D. Nelson. 1984. *Linear Probability, Logit, and Probit Models*. Thousand Oaks, C.A.: Sage Publications.

[4] Azzalini, Adelchi. 1996. *Statistical Inference Based on the Likelihood.* London: Chapman & Hall.

[5] Baek, Mijeong. 2009. "A Comparative Analysis of Political Communication Systems and Voter Turnout." *American Journal of Political Science* 53(2): 376-93.

[6] Baltagi, Badi H. 2011. *Econometrics,* 5th edition. NY: Springer.

[7] Beck, Nathaniel. 2007. "From Statistical Nuisances to Serious Modeling: Changing How We Think About the Analysis of Time-Series-Cross-Section Data." *Political Analysis* 15(2): 97-100.

[8] Beck, Nathaniel, and Jonathan N. Katz. 1995. "What To Do (and Not To Do) with Time-Series Cross-Section Data." *American Political Science Review* 89(3): 634-647.

[9] Beck, Nathaniel, and Jonathan N. Katz. 2007. "Random Coefficient Models for Time-SeriesCross-Section Data: Monte Carlo Experiments." *Political Analysis* 15(2): 182-95.

[10] Berger, James O., and Robert L. Wolpert. 1984. *The Likelihood Principle*. Hayward, C.A.; Institutte of Mathematical Statistics.

[11] Bertsekas, Dimitri P. 1996. *Constrained Optimization and Lagrange Multiplier Methods*. Belmont, M.A.: Athena Scientific.

[12] Biglaiser, Glen, and Karl DeRouen Jr. 2007. "Sovereign Bond Ratings and Neoliberalism in Latin America." *International Studies Quarterly* 51(1): 121-138.

[13] Bonnans, Joseph-Frederic, Jean Charles Gilbert, Claude Lemarechal, and Claudia A. Sagastizabal. 2006. *Numerical Optimization: Theoretical and Practical Aspects*. New York: Springer.

[14] Boyne, George A., Oliver Jamesa, Peter Johna, and Nicolai Petrovsky. 2009. "Democracy and Government Performance: Holding Incumbents Accountable in English Local Governments." *Journal of Politics* 71(4): 1273-84.

[15] Cong, Ronna. 1999. "Truncated regression." *Stata Technical Bulletin* 52:47-52.

[16] Fordham, Benjamin O., and Thomas C. Walker. 2005. "Kantian Liberalism, Regime Type, and Military Resource Allocation: Do Democracies Spend Less?" *International Studies Quarterly* 49(1): 141-57.

[17] Franzese, Robert J. Jr. 2005. "Empirical Strategies for Various Manifestations of Multilevel Data." *Political Analysis* 13(4): 430-46.

[18] Franzese, Robert J. Jr., and Jude C. Hays. 2007. "Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data." *Political Analysis* 15(2): 140-64.

[19] Golder, Matt. 2005. "Democratic Electoral Systems around the World, 1946-2000." *Electoral Studies* 24(1): 103-121.

[20] Gomez, Brad T., Thomas G. Hansford, George A. Krause. 2007. "The Republicans Should Pray for Rain: Weather, Turnout, and Voting in U.S. Presidential Elections." *Journal of Politics* 69(3): 649-63.

[21] Gilbert, Charles J. 2009. "SQPlab– A Matlab software for solving nonlinear optimization problems and optimal control problems." INRIA Technical Report, RT-378.

[22] Gould, William. 2011. "Interpreting the Intercept in the Fixed-Effects Model", Stata Corporation, online article <http://www.stata.com/support/faqs/stat/xtreg2.html>

[23] Greene, William. 2008. *Econometric Analysis*, 6th ed. Upper Saddle River, N.J.: Pearson Prentice Hall.

[24] Honoré, Bo E. 1992. "Trimmed Lad and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects." *Econometrica* 60(3): 533-65.

[25] Horrace, William C. 2005a. "On Ranking and Selection from Independent Truncated Normal Distributions." *Journal of Econometrics* 126(2): 335-54.

[26] Horrace, William C. 2005b. "Some Results on the Multivariate Truncated Normal Distribution." *Journal of Multivariate Analysis* 94(1): 209-21.

[27] Hsiao, Cheng. 2003. *Analysis of Panel Data.* Cambridge: Cambridge University Press.

[28] Kellstedt, Paul M., and Guy D. Whitten. 2009. *The Fundamentals of Political Science Research.* Cambridge: Cambridge University Press.

[29] King, Gary. 1998. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference.* Ann Arbor Michigan University Press.

[30] Knack, Stephen. 1995. "Does "Motor Voter" Work? Evidence from State-Level Data." *Journal of Politics* 57(3): 796-811.

[31] Kuhn, H.W.,and A.W. Tucker. 1951. "Nonlinear programming". *Proceedings of 2nd Berkeley Symposium.* Berkeley: University of California Press, pp. 481V492.

[32] Johnson, Norman L., and Samuel Kotz. 1970. *Distributions in Statistics: Continuous Univariate Distribution-2.* New York: John Wiley & Sons.

[33] McCaffrey, Daniel F., J. R. Lockwood, Kata Mihaly, and Tim R. Sass. 2010. "A Review of Stata Routines for Fixed Effects Estimation in Normal Linear Models." Unpublish manuscript.

[34] Powell, M. J. D. 1978. "Algorithms For Nonlinear Constraints That Use Lagrangian Function." *Mathematical Programming* 14(1):224-248.

[35] Royall, Richard. 1997. *Statistical Evidence: A Likelihood Paradigm.* London: Chapman & Hall.

[36] Sides, John, Eric Schickler, and Jack Citrin. 2008. "If Everyone Had Voted, Would Bubba and Dubya Have Won?" *Presidential Studies Quarterly* 38(3): 521-39.

[37] Wawro, Gregory. 2002. "Estimating Dynamic Panel Data Models in Political Science." *Political Analysis* 10(1): 25-48.

[38] Wicharaya, Tamasak. 1995. *Simple Theory, Hard Reality: The Impact of Sentencing Reforms on Courts, Prisons, and Crime.* Albany, NY: SUNY Press.

[39] Wooldridge, Jeffrey M. 2005. "Fixed-Effects and Related Estimators for Correlated Random-Coefficient and Treatment-Effect Panel Data Models." *Review of Economics and Statistics* 87(2): 385-90.

[40] Vanderbei, Robert J. 2008. *Linear Programming: Foundations and Extension,* 3rd edition. NY: Springer.

	Model Specification			
	Add grandmeans	*Centering methods*	*Centering before or after forming interaction terms*	*Correcting demeaning bias*
Model I	Yes	Centered by groupmeans	Before	No
Model II	No	Fixed at minimum	After	No
Model III	No	Fixed at minimum	After	No
Model IV	No	Fixed at minimum	After	No
Model V	No	Fixed at minimum	After	Yes

	Parameter Estimation	
	Objective functions	*Application of COP*
Model I	Least Squares	No
Model II	Least Squares	No
Model III	Least Squares	Yes
Model IV	Maximum Likelihood	Yes
Model V	Maximum Likelihood	Yes

Table 1: Model Assumptions

Democratic Pres Candidate's Vote	I. Original Coefficient	I. Original *SE*	II. Fixing at Minimum Coefficient	II. Fixing at Minimum *SE*
Partisan composition	.595**	.029	.799**	.035
Turnout	-.066**	.021	.139**	.022
Turnout×GOP Incumbent	.155**	.009	-.007	.012
Turnout×Partisan composition	.001**	.000	-.003**	.001
Yr52	-10.012**	.166	-10.645**	.166
Yr56	-16.658**	.590	-5.935**	.554
Yr60	-11.503**	.651	-.534	.615
Yr64	17.164**	.230	16.553**	.229
Yr68	-4.467**	.168	-4.665**	.167
Yr72	-23.021**	.612	-12.592**	.547
Yr76	-7.174**	.567	3.070**	.510
Yr80	-4.584**	.192	-4.689**	.193
Yr84	-16.013**	.609	-5.966**	.521
Yr88	-8.226**	.586	1.649**	.498
Yr92	-2.640**	.608	7.437**	.526
Yr96	4.741**	.220	5.176**	.217
Yr2000	-6.168**	.229	-5.967**	.224
Constant	19.446**	1.506	-19.582**	.950
σ	4.890		4.763	
$\hat{y}^{\max}$	83.701		92.746	
$\hat{y}^{\min}$	12.696		3.804	
Theoretical Boundary Violations	14		37	
N	27401		27401	

Level of significance: $^{*}p \leq 0.05$; $^{**}p \leq 0.01$.

Table 2: Replication of Table 1 (Column 1), Hansford and Gomez (2010: 277)

	III. LSCO		IV. MLCO1		V. MLCO2	
Democratic Pres Candidate's Vote	Coefficient	*SE*	Coefficient	*SE*	Coefficient	*SE*
Partisan composition	.587**	.036	.238**	.048	.237**	.050
Turnout	.002	.022	.002	.027	.000	.029
Turnout×GOP Incumbent	-.031**	.012	-.019	.013	-.017	.013
Turnout×Partisan composition	.001	.001	.001	.001	.001	.001
Yr52	-10.526**	.169	-11.509**	.194	-11.461**	.196
Yr56	-4.847**	.546	-7.897**	.570	-7.964**	.575
Yr60	.853	.610	-2.958**	.636	-3.056**	.640
Yr64	16.103**	.232	10.927**	.298	11.169**	.297
Yr68	-4.793**	.169	-6.199**	.189	-6.188**	.190
Yr72	-11.450**	.542	-11.447**	.584	-11.026**	.593
Yr76	3.787**	.501	1.291*	.540	1.197*	.545
Yr80	-5.131**	.198	-8.544**	.264	-8.515**	.264
Yr84	-5.682**	.519	-10.178**	.585	-10.190**	.589
Yr88	1.789**	.497	-2.210**	.566	-2.304**	.570
Yr92	7.714**	.525	3.166**	.597	3.065**	.601
Yr96	4.481**	.228	1.582**	.308	1.565**	.309
Yr2000	-6.286**	.233	-7.750**	.294	-7.736**	.295
Constant	-12.659**	.981	-2.079	1.307	-2.288	1.333
σ	4.778		5.395		5.437	
$\hat{y}^{\max}$	92.475		91.055		91.021	
$\hat{y}^{\min}$	4.275		6.839		6.574	
Theoretical Boundary Violations	0		0		0	
N	27401		27401		27401	

Level of significance: $^{*}p \leq 0.05$; $^{**}p \leq 0.01$.

Table 3: Replication of Table 1 (Column 1), Hansford and Gomez (2010: 277)

	Model I		Model II		Model III		Model IV		Model V	
Predicted Value by Varying	mix	max	min	max	min	max	min	max	min	max
All Variables	**-7.3**	**103.8**	**-17.1**	**117.3**	2.4	100.0	1.5	100.0	1.3	100.0
Constant	15.1	77.0	**-4.5**	57.5	13.8	51.0	13.0	74.9	12.7	74.8
Partisan composition	**-7.3**	**103.8**	**-17.1**	**117.3**	34.4	100.0	14.4	100.0	14.2	100.0
Turnout	**-7.3**	**103.8**	**-17.1**	**117.3**	2.5	100.0	1.6	100.0	1.3	100.0
Turnout×GOP Incumbent	**-25.7**	**122.2**	**-17.1**	**117.3**	0.0	100.0	0.0	100.0	0.0	100.0
Turnout×Partisan composition	**-15.4**	**111.9**	**-17.1**	**117.3**	5.5	100.0	3.9	100.0	3.7	100.0
Yr52	4.8	90.3	**-15.2**	**100.7**	3.3	73.4	1.5	77.6	1.3	77.4
Yr56	**-1.4**	92.0	**-10.5**	**100.7**	9.0	79.0	5.1	81.2	4.8	80.9
Yr60	3.4	91.0	**-5.0**	**100.7**	14.7	84.7	10.0	86.1	9.7	85.8
Yr64	11.2	**103.8**	**-4.5**	**117.3**	29.9	100.0	23.9	100.0	23.9	100.0
Yr68	9.9	88.6	**-9.2**	**100.7**	9.0	79.1	6.8	82.9	6.6	82.6
Yr72	**-7.3**	90.4	**-17.1**	**100.7**	2.4	72.4	1.5	77.6	1.7	77.8
Yr76	7.4	88.6	**-4.5**	**103.8**	17.6	87.7	14.3	90.3	13.9	90.0
Yr80	9.8	88.3	**-9.2**	**100.7**	8.7	78.8	4.4	80.5	4.2	80.3
Yr84	**-0.8**	91.0	**-10.5**	**100.7**	8.1	78.2	2.8	78.9	2.6	78.6
Yr88	6.4	89.5	**-4.5**	**102.4**	15.6	85.7	10.8	86.9	10.4	86.5
Yr92	11.6	88.3	**-4.5**	**108.2**	21.5	91.6	16.1	92.2	15.8	91.9
Yr96	12.5	92.2	**-4.5**	**105.9**	18.3	88.4	14.6	90.7	14.3	90.4
Yr2000	8.4	89.9	**-10.5**	**100.7**	7.5	77.6	5.2	81.3	5.0	81.1
Number of Boundary Violations	14		37		0		0		0	

Bold numbers indicate out-of-bound predicted values

Table 4: Translation of Boundary Violations into Predicted Values For Model I to Model V