DEPARTMENT OF ECONOMICS

COLLEGE OF BUSINESS AND ECONOMICS

UNIVERSITY OF CANTERBURY

CHRISTCHURCH, NEW ZEALAND

THE DETERMINANTS OF U. S. STATE ECONOMIC GROWTH: A LESS EXTREME BOUNDS ANALYSIS

by W. Robert Reed

WORKING PAPER

No. 05/2006

Department of Economics, College of Business and Economics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand

WORKING PAPER No. 05/2006

THE DETERMINANTS OF U. S. STATE ECONOMIC GROWTH: A LESS EXTREME BOUNDS ANALYSIS

by W. Robert Reed*

Revised, February 2, 2006

Abstract

This study investigates U.S. state economic growth from 1970-1999. I innovate on previous studies by developing a new approach for addressing "model uncertainty" issues associated with estimating growth equations. My approach borrows from the "extreme bounds analysis" (EBA) approach of Leamer (1985), while also addressing concerns raised by Granger and Uhlig (1990), Salai-Martin (1997) and others that not all specifications are equally likely to be true. I then apply this approach to identify "robust" determinants of state economic growth. My analysis confirms the importance of productivity characteristics of the labor force and industrial composition of a state's economy. I also find that policy variables such as (i) size and structure of government and (ii) taxation are "robust" and economically important determinants of state economic growth.

* Professor, Department of Economics, University of Canterbury, Private Bag 4800, Christchurch, New Zealand, Phone 64-3-3642846, email: bobreednz@yahoo.com.

JEL Categories: 040, 051, H10, H20, H30, H70, R11, R58, and C51

Keywords: Growth, U. S. State Economic Growth, State Fiscal Policy, Economic Development, Taxes, Model Selection, SIC, AIC, AICc, Extreme Bounds Analysis

Acknowledgements: I thank Hua Chen, Elisabeth Reed, and Ying Xiao for able research assistance. I am grateful to Chris Chatfield, David Merriman, Cindy Rogers, and Steven Yamarik for providing helpful comments on an earlier draft of this paper. In addition, special thanks go to Steven Yamarik for generously supplying state-level capital stock data from Garofalo and Yamarik (2002).

I. INTRODUCTION

It is now well-established that economic growth studies reach different conclusions depending on model specification. This has been documented repeatedly in the literature on cross-country growth regressions[1] and in studies of growth in U.S. states.[2] In response, attempts have been made to identify "robust" variables, the "best" model specification, or ways of combining alternative model specifications (e.g., Levine and Renelt, 1992; Crain and Lee, 1999; Granger and Uhlig, 1990; Sala-i-Martin, 1997; Fernandez, et al., 2001; Sala-i-Martin, et al., 2004; Hoover and Perez, 2004; Hendry and Krolzig, 2004).

This study follows in this line of research by attempting to identify "robust" determinants of U.S. economic growth from 1970-1999. I innovate on previous studies by developing a new approach for addressing "model uncertainty" issues associated with estimating growth equations. My approach borrows from the "extreme bounds analysis" (EBA) approach of Leamer (1985), while also addressing concerns raised by Granger and Uhlig (1990), Salai-Martin (1997) and others that not all specifications are equally likely to be true. I then apply this approach by sifting through a very large number of explanatory variables in order to find "robust" determinants of state economic growth. My analysis confirms the importance of productivity characteristics of the labor force and industrial composition of a state's economy. I also find that policy variables such as (i)

[1] Studies that have examined the robustness of coefficient estimates in the context of cross-country growth regressions include Levine and Renelt (1992), Sala-i-Martin (1997), Fernández et al. (2001), Hendry and Krolzig (2004), Sala-i-Martin et al. (2004), and Hoover and Perez (2004).

[2] The following studies have highlighted the phenomenon of wide-ranging coefficient estimates across empirical specifications: Bartik, 1992; McGuire, 1992, Phillips and Goss, 1995; Wasylenko, 1997; and Crain and Lee, 1999.

the size and structure of government and (ii) taxation are "robust" determinants of state economic growth.

The paper proceeds as follows: Section II develops a theoretical model of state economic growth that provides the framework for the subsequent empirical analysis. Section III describes the full set of variables used in this study. Section IV presents my approach for identifying "robust" determinants of economic growth. Section V describes my data and discusses details about the estimation procedure. Section VI presents the empirical results. Section VII concludes.

II. A THEORETICAL MODEL OF ECONOMIC GROWTH

I assume that state income (Y_t) is determined by the following generalized Cobb-Douglas production function,

(1) $$Y_t = A_t K_t^{\alpha} (L_t Q_t)^{\beta} = A_t Q_t^{\beta} K_t^{\alpha} L_t^{\beta},$$

where Q_t is the efficiency of labor and A_t is a time-varying, scaling variable that includes factor-neutral technology shocks.[3]

Dividing both sides by N_t gives

(2) $$\frac{Y_t}{N_t} = A_t Q_t^{\beta} \left(\frac{K_t}{N_t}\right)^{\alpha} \left(\frac{L_t}{N_t}\right)^{\beta} N_t^{(\alpha+\beta-1)}.$$

This can be expressed in log form as

(3) $$ln(y_t) = \alpha\, ln(k_t) + \beta\, ln(\ell_t) + (\alpha+\beta-1)\, ln(N_t) + ln(A_t) + \beta\, ln(Q_t)$$

where $y_t = \frac{Y_t}{N_t}$, $k_t = \frac{K_t}{N_t}$, and $\ell_t = \frac{L_t}{N_t}$.

Differentiating Equation (3) with respect to time yields

[3] This specification is very similar to that employed by Lee and Gordon (2005).

(4) $$\frac{\dot{y}_t}{y_t} = \alpha\frac{\dot{k}_t}{k_t} + \beta\frac{\dot{\ell}_t}{\ell_t} + (\alpha+\beta-1)\frac{\dot{N}_t}{N_t} + \left(\frac{\dot{A}_t}{A_t} + \beta\frac{\dot{Q}_t}{Q_t}\right)$$

It follows that

(5) $$\begin{aligned}\ln(y_t) - \ln(y_{t\text{-}L}) \cong \alpha\left[\ln(k_t) - \ln(k_{t\text{-}L})\right] + \beta\left[\ln(l_t) - \ln(l_{t\text{-}L})\right] + \\ (\alpha+\beta-1)\left[\ln(N_t) - \ln(N_{t\text{-}L})\right] + C_t\end{aligned},$$

where $C_t = \left[ln(A_t) - ln(A_{t\text{-}L})\right] + \beta\left[ln(Q_t) - ln(Q_{t\text{-}L})\right]$ and L = the length of the time period minus 1 (i.e., for a five-year period with t measuring calendar years, $L = 4$).[4][5]

This theoretical structure specifically identifies changes in capital, employment, and population as important determinants of economic growth. However, the last term, C_t, is sufficiently general that any number of variables could be argued for inclusion.

III. POTENTIAL DETERMINANTS OF STATE ECONOMIC GROWTH

TABLE 1 lists a number of variables that have been suggested in previous studies of economic growth, primarily U.S. state economic growth. The empirical task of this paper consists of identifying which of these should be included in a growth equation along with capital, employment, and population variables.

I group the variables into 4 major categories: (A) Population/Labor Force characteristic, (B) Economy characteristics, (C) Public Sector characteristics, and (D) Political Control characteristics. Variables included in the Population/Labor Force category include educational attainment, percent of the population that is working-aged (ages between 20 and 64), percent of the population that is nonwhite or female, and total

---

[4] In the subsequent empirical work, the difference in log values is multiplied by 100.

[5] An alternative specification solves for the steady state value of y as a function of state parameters, and then introduces convergence through the inclusion of a lagged value of the dependent variable. This both (i) imposes additional restrictions on the model and (ii) raises econometric issues of inconsistency from using both fixed effects and the lagged dependent variable as explanatory variables. Nevertheless, the approach of this paper is readily applied to selecting control variables for this, and other, specifications.

population. Economy characteristics include population density, degree of urbanization, the relative importance of various industries within the state, percent of the workforce that is unionized, and a measure of industrial diversity.

Public Sector characteristics are divided into three subcategories: (i) Size and Structure variables, (ii) Tax variables, and (iii) Expenditure variables. Each of these can be thought of representing a particular component of public policy. Size and Structure variables include the size of the (i) federal and (ii) state and local government sectors of the economy, measured both by share of total earnings and employment. Also included are the amount of federal government revenue received by state and local governments; the degree to which expenditures are made at the local, as opposed to the state, level; and the number of governments.

Tax variables include a measure of the overall importance of state and local taxes in the state's economy ("tax burden"), measured as a share of state personal income. Also included are specific types of taxes, such as property, sales, individual income and corporate income taxes. Ideally, I would have liked to have measured these latter tax variables as shares of total tax revenues. This would have been most appropriate for investigating the compositional effects of the tax burden. Unfortunately, data problems prevented me from doing this.[6] Instead, the specific types of taxes are also measured as shares of personal income.

Expenditure variables measure the compositional effects of state and local government spending. The specific expenditure categories are primary and secondary education, higher education, public health, and highways. Each of the respective

---

[6] A problem arises in that sales and income tax data are not separately reported for local governments, so that the shares of the respective tax subcategories do not sum to one.

expenditure variables is measured as a share of total state and local (direct general) spending.

Finally, Political Control variables measure the influence of political parties. These include how often the Democratic and Republican parties control the state legislature, and how often the governor is a Democrat.

These preceding variables attempt to capture the economic influences represented by $C_t = [\ln(A_t) - \ln(A_{t\text{-}L})] + \beta[\ln(Q_t) - \ln(Q_{t\text{-}L})]$ in Equation (5). One immediate issue is whether the 32 variables in TABLE 1 should be entered in (i) level or (ii) difference form. Because economic theory is not sufficiently specific to answer this question, this becomes an empirical issue. Restricting the Political Control variables to be entered in level form,[7] and recognizing that the change in population is already included in the core specification of Equation (5) (i.e., $[\ln(N_t) - \ln(N_{t\text{-}L})]$), leads to a total of 60 possible explanatory variables.

There are approximately 1.15×10^{18} ways to combine 60 variables. Each of these permutations, appending a core set of "free" variables, can be thought of as a single model. Thus, the empirical problem consists of choosing the best model, or set of models, from these 1.15×10^{18} possibilities. One might think that it was computationally unfeasible to estimate so many models. While this is true, there exist algorithms that allow me to circumvent this problem.

[7] Unlike the other variables in TABLE 1, the Political Control variables represent the average number of years in which a political party is in control during the respective 5-year period. There is no analog to 5-year differences that would correspond to the 5-year differences for the other variables in TABLE 1.

IV. A PROCEDURE FOR DETERMINING "ROBUST" VARIABLES

The SIC and the AICc. The first step in my approach consists of identifying a "best" specification: I employ two model selection criteria for this purpose: the Schwarz Information Criterion (SIC); and the corrected version of the Akaike Information Criterion, known as AICc. While I will give a brief description of these criteria, more detailed discussions can be found in McQuarrie and Tsai (1998), Burnham and Anderson (2002), and the references therein.

The SIC and the AICc respectively represent the two main, competing schools of thought regarding how to conceptualize the task of selecting the "best" model. If the researcher believes that the true model is included within the set of candidate models, then a desirable property of a model selection procedure is that it be "consistent." That is, that it selects the true model with probability converging to one as the sample size becomes infinitely large. The SIC is by far the most commonly used of the several model selection criteria that possess this property (other consistent criteria include the Hannan and Quinn [HQ] criterion, and the Geweke and Meese [GM] criterion).

Alternatively, if the researcher believes that the true model is not included within the set of candidate models, then a desirable property of a model selection procedure is that it be "efficient." That is, that it selects the model that is "closest" to the true model, where "closest" is defined by some distance or information criterion. A selection procedure is said to be "asymptotically efficient" if it selects the model closest to the true model with probability converging to one as the sample size becomes infinitely large.

A number of model selection procedures have been developed that have the property of asymptotic efficiency, including Akaike's Final Prediction Error (FPE), Mallow's Cp criterion, and the Akaike Information Criterion (AIC). Of these, the AIC is by far the most widely employed. However, many researchers have noted that the AIC suffers from over-fitting in finite samples, incorporating too many variables in its best models. As a result, a number of finite sample corrections have been developed for the AIC. Of these, the most preferred is a version known as AICc (Sugiura, 1978; Hurvich and Tsai, 1989).

Monte Carlo studies of finite sample performance have demonstrated that both the SIC and AICc perform well relative to alternative procedures (cf. McQuarrie and Tsai, 1998). While there are a number of equivalent formulations, this study uses the following formulae:

(6) $$SIC = T \cdot ln\left(\frac{SSE}{n}\right) + k \cdot ln(T), \text{ and}$$

(7) $$AICc = T \cdot ln\left(\frac{SSE}{n}\right) + T \cdot \left(\frac{T + k - 1}{T - k - 1}\right),$$

where T is the number of observations; k is the number of coefficients in the model, including the intercept; and SSE is the sum of squared residuals from the estimated model. Note that SSE and k are the only parameters that vary across models, since sample size and the dependent variable do not change. The SIC and AICc make different tradeoffs between these parameters. Generally, the SIC penalizes additional explanatory variables more severely than the AICc, producing "best" models with fewer variables.

Conceptually, I need a program that will sort through all 1.15×10^{18} possible linear combinations of the 60 variables (level plus difference forms) identified in TABLE 1 in order to select the best model specification according to each selection criterion. For this task, I use the SELECTION = RSQUARE option within the REG procedure available through SAS. While this procedure does not actually estimate all possible regression specifications, it can identify the best specifications within each set of all possible specifications having the same number of regressors. In this, it relies on the "leaps and bounds" algorithm developed by Furnival and Wilson (1974). It is straightforward to use the output generated by this SAS program to calculate a ranked ordering of the *M* overall best specifications -- for any predetermined value of *M* -- according to either the SIC or AICc criterion.[8] The corresponding SAS program is easy to write and remarkably efficient in computational requirements. It required about an hour to run using a standard desktop computer.[9]

<u>Extreme Bounds Analysis and Bayesian Model Averaging</u>. My approach uses insights from both "extreme bounds analysis" (EBA; Leamer, 1983) and "Bayesian model averaging" (BMA; Hoeting et al., 1999). Therefore, it is useful to consider these before proceeding.

EBA is designed to study the sensitivity of coefficient estimates across different regression specifications. For example, suppose a researcher wants to measure the effect of variable X_1 on variable Y. EBA proceeds by estimating a large number of specifications that include X_1, noting the upper and lower bounds of the resulting range of β_1 estimates. If the range of $\hat{\beta}_1$ values are all same-signed and more than two standard

[8] The general principle of the "leaps and bounds algorithm" can be illustrated in the context of a "regression tree:" Consider the case of 5 "doubtful' variables, X_1 through X_5. At the top of the regression tree are models with only 1 regressor. At the bottom of the tree are models with more variables. Suppose the R^2 from the model having only one regressor, X_1, is larger than the R^2 from a model with the four regressors X_2 through X_5. In this case, the model with the highest R^2 must lie on the "node" below X_1. This eliminates the necessity of estimating large portions of the "regression tree," which greatly reduces the computational burden. Further details are given in Furnival and Wilson (1974). SAS uses this algorithm and sorts the "best" R^2 models within subsets of specifications having the same number of regressors. I calculate SIC and AICc values within these subsets, where highest R^2 equates with lowest SIC/AICc values, and then globally rank the 'best" specifications across all subsets.

[9] A copy of the program used in this analysis is given in Appendix B.

deviations away from zero, then variable X_1 is said to be "robust" (cf. Levine and Renelt, 1992; Crain and Lee, 1999).

The main criticism of EBA is that it weights all model specifications equally, so that a divergent coefficient estimate from a poorly specified equation can be sufficient to disqualify a variable as "robust." In recognition of this shortcoming, Granger and Uhlig (1990) propose "reasonable extreme bounds analysis," where the range of coefficient values is restricted to the set of specifications that produce R^2 values within a given δ-value of the maximum achieved R^2 across all specifications. However, they do not provide guidance for the choice of δ and acknowledge that the use of R^2 has problems.

BMA directly addresses the "all specifications weighted equally" criticism by developing a system for weighting model specifications. BMA starts by positing a prior distribution for the population value for some parameter of the model specification (usually a regression coefficient). This prior distribution is updated with the results from regression estimates across (theoretically) all possible model specifications to form a posterior distribution of parameter values. The updating procedure weights the corresponding specifications by model probabilities that can be thought of as the conditional probability that a given specification is the "true model."[10]

While the BMA approach is useful for weighting specifications for forecasting purposes, it is problematic when used to weight coefficient estimates. Consider the following example: Suppose a researcher is interested in the relationship between dependent variable y and an explanatory variable, X_1. Let the true model be given by

[10] It is a conditional probability because the probabilities are calculated over the set of "included" model specifications.

$y_t = \beta_o + \sum_{k=1}^{K} \beta_k X_{k,t} + \varepsilon_t$, $t = 1,...T$, where some β_k may equal zero (but not β_1); and $Cov(X_j, X_k) = 0$ for all $j \neq k$. There are 2^K possible linear combinations of these variables, and we suppose the researcher considers each combination a potentially true model. Define $P(M_j)$ as the prior probability that Model *j* is the true model and let $P(M_j) > 0$ for all *j*.

The BMA approach calculates the posterior probability of each model as

(8) $$P(M_j \mid y) = \frac{P(M_j) T^{-k_j/2} SSE_j^{-T/2}}{\sum_{i=1}^{2^K} P(M_i) T^{-k_i/2} SSE_i^{-T/2}},$$

where k_j and SSE_j are the number of included regressors and the sum of squared residuals in Model *j*. The corresponding (posterior) expected value of β_1 is given by

(9) $$E(\beta_1 \mid y) = \sum_{j=1}^{2^K} P(M_j \mid y) \cdot \hat{\beta}_{1,j},$$

where $\hat{\beta}_{1,j}$ is the estimate of β_1 in Model *j*.

In each specification in which X_1 appears, the preceding assumptions insure that the least squares estimate is unbiased, so that $E(\hat{\beta}_{1,j}) = \beta_1$. However, X_1 appears in only half of all possible specifications. In the other 2^{K-1} models, X_1 is excluded, and the BMA approach sets $\hat{\beta}_{1,j} = 0$.[11] It follows that $E(\beta_1 \mid y) < \beta_1$ even if $\hat{\beta}_{1,j} = \beta_1$ in every specification in which it appears. In other words, the BMA-based expectation is biased

---

[11] Compare Equations (8) and (9) with Equations (7) and (8) in Sala-i-Martin et al. (2004, p. 817) and note that in that context they write, "…any variable excluded from a particular model has a slope coefficient with a degenerate posterior distribution at zero."

towards zero. This follows directly from the fact that BMA "estimates" the value of β_1 to be zero in all specifications in which X_1 is not included.[12]

A Less Extreme Bounds Analysis. My approach borrows elements from both EBA and BMA. Like EBA, I estimate a set of specifications and report the corresponding ranges of coefficient estimates and *t*-ratios for those specifications including the respective variables. However, like BMA, I do not give all specifications equal weight. I follow a procedure developed by Poskitt and Tremayne (PK; 1987) to identify two categories of models: (i) "reasonable" models, and (ii) others. Only "reasonable" models are considered for extreme bounds analysis.

PK take as their point of departure that informational criteria such as the SIC and the AICc are themselves sample statistics, so that the model with the lowest SIC or AICc value may not be the best model. They argue that all "close competitors" be included in a "portfolio" of "reasonable" models.

Let I^* be the value of the information criterion for the best model, and let I^A be the corresponding value for an alternative model. The posterior odds ratio is defined as

$$(10) \qquad \Re = exp\left[-\frac{T}{2}\cdot\left(I^* - I^A\right)\right].$$

[12] There are other problems with using the BMA approach. First, the results are sensitive to assumptions about the prior parameter distribution. For example, in order to implement their version of BMA known as Bayesian Averaging of Classical Estimates (BACE), Sala-i-Martin et al. (2004) must first specify an "expected model size." While they claim that their final results are "robust" across different assumptions about this parameter, they acknowledge that this is not true in all cases: Some of the variables that are "significant" under a given assumed "expected model size" become "insignificant" under a different assumed "expected model size" -- and vice versa. Second, there are important computational issues. BMA does not actually estimate all possible specifications. Instead, it uses sampling procedures (e.g., Markov chain Monte Carlo procedures, of which the Gibbs sampler is the best known) to estimate the "probability" that a given specification is the true one. There is no standard sampling algorithm, which raises the possibility that the results will be idiosyncratic to the program used by the individual researcher. Finally, the weighting probabilities are derived from Bayesian statistical foundations and are closely related to the SIC criterion defined above. As we shall see below, alternative criteria, such as the AICc, produce different results.

Following Jeffreys (1961, p. 143) and Zellner (1977), PK characterize any model with $\Re < \sqrt{10}$ as a "close competitor" to the best model:

> "…any…specification satisfying $\Re < \sqrt{10}$ may be thought of as a close competitor. This intimates that it may be advantageous to extend the usual model building process. It suggests not only that the model minimizing the criterion should be selected, but also that any additional specifications closely competing…should not be discarded, thereby advancing the general notion of a portfolio of models" (Poskitt and Tremayne, 1987, p. 127).

PK go on to present Monte Carlo evidence that model portfolios constructed in this manner behave well in finite samples.

To summarize, my approach constructs separate model portfolios using SIC and AICc selection methods. For each portfolio, I identify "robust" variables in a manner similar to conventional EBA. In this respect, my approach is similar to Granger and Uhlig's (1990) "reasonable extreme bounds analysis," except that I use information criteria, not R^2, to evaluate models, and the set of evaluated models is determined by PK's $\Re < \sqrt{10}$ standard, rather than an arbitary δ value.

V. DATA AND FURTHER ESTIMATION ISSUES

My data consist of observations on 46 U.S. states from 1970-1999.[13] I decided on this particular time period because a longer time frame would have required me to omit many variables of interest. The respective thirty years of data were grouped into six, 5-year

[13] Alaska and Hawaii were omitted, as usual in studies of U.S. state economic growth. Nebraska and Minnesota were also eliminated because the variables *Democratic Legislature* and *Republican Legislature* could not be constructed for these two states over the full-time period: In Nebraska, state representatives do not formally affiliate with political parties; whereas Minnesota had a unicameral state legislature through 1970.

periods. (1970-1974, 1975-1979, …, 1995-1999). Data for most of these variables were collected from original data sources.[14]

Using 5-year over annual data offers several advantages: It (i) averages out "business cycle effects" (Grier and Tullock, 1989); (ii) minimizes errors from misspecifying lag effects; and (iii) reduces time-specification issues. Time-specification issues arise because data can have different start and end periods within a given calendar year. For example, state income data is defined over calendar years; state fiscal data is defined over fiscal years (which are different for different states); and other variables (e.g. employment, population data) may be measured at different points within the year (beginning/middle/end). In addition, a number of variables (e.g., variables based on decennial Census data) require interpolation in order to get a balanced panel. For all of these reasons, the use of 5-year data should entail fewer estimation problems.

Following Equation (5), the general specification for the empirical models is[15,16]:

$$DLNY_t = \begin{bmatrix} \beta_0 + \beta_1 DLNK_t + \beta_2 DLNL_t + \beta_3 DLNN_t \\ + \textit{state fixed effects} + \textit{time fixed effects} \end{bmatrix} + \sum_l \lambda_l X_{l,t-4} + \sum_d \delta_d \left(X_{d,t} - X_{d,t-4} \right) + \sum_p \pi_p \overline{X}_{p,t} + \varepsilon_t , \quad (11)$$

where $t = 1974, 1979, 1984, 1989, 1994, 1999$; $DLNY_t$, $DLNK_t$, $DLNL_t$, and $DLNN_t$ are the respective difference quantities from Equation (5) multiplied by 100 (to give percent); $X_{l,t-4}$ is the value of the explanatory variable at the beginning of the five-year period

[14] Appendix A presents statistical descriptions of all the variables used in this study.

[15] Note that because (i) the dependent variable is expressed in logs and (ii) the annual price deflator is only available for the nation as a whole, and not for individual states, inflationary effects are captured by the time period dummies. Thus, there is no need to convert the dependent variable to real values.

[16] In the estimated specification of Equation (6), I do not impose the restriction that $\beta_3 = (\beta_1 + \beta_2 - 1)$ for two reasons. First, population growth could also be a factor included in C_t which, if true, would invalidate the restriction. Second, as a practical matter, this restriction is consistently rejected below the 1-percent significance level in all of the top model specifications.

(“level” form); $\left(X_{dt} - X_{d,t-4}\right)$ is the change in the explanatory variable over the five-year period (“difference” form); and $\overline{X}_{p,t} = \frac{\overline{X}_{p,t-1} + \overline{X}_{p,t-2} + \overline{X}_{p,t-3} + \overline{X}_{p,t-4} + \overline{X}_{p,t-5}}{5}$ is the 5-year average over the period (*t-5* to *t-1*) for the Political Control variables *Democratic Legislature*, *Republican Legislature*, and *Democratic Governor*.[17]

The 2^{60} possible model specifications each include the variables listed in brackets in Equation (11) but allow for alternative configurations of the last three sets of variables ($X_{l,t-4}$, $\left(X_{dt} - X_{d,t-4}\right)$, and $\overline{X}_{p,t}$), since the theory is non-specific about which variables belong in C_t (cf. Equation [5]).

VI. EMPIRICAL RESULTS

The empirical analysis proceeds in three steps. In the first step, I estimate a regression that uses the full set of regressors. In the second step, I use model selection criteria to identify the best SIC and best AICc specifications among the $2^{60} \cong 1.15 \times 10^{18}$ total possible models. These best models are then compared to the model with all variables. Finally, I identify separate SIC and AICc portfolios of “reasonable” models and perform EBA to identify which determinants of U.S. state economic growth are “robust.”[18]

Specification with All Variables. TABLE 2 reports the results of estimating Equation (11) using the full set of explanatory variables. Because the focus is on the TABLE 1 variables, I do not report estimates for *DLNK*, *DLNL*, and *DLNN*, nor state and time fixed effects (see Equation [11]). However, I note that the estimated coefficients for

[17] This last adjustment is made to account for the fact that it takes at least a year for political representation to get translated into legislation (cf. Poterba, 1994; Gilligan and Matsusaka, 1995; Reed, 2006), and it is the latter that is assumed to matter for economic growth.

[18] The fact that theory determines which variables are “doubtful,” and that all possible combinations of “doubtful” variables are considered, should address concerns about subjective classification (McAleer et al., 1985).

DLNL, *DLNK*, and *DLNN* are each positive in sign and highly significant (p-value < 0.0001). Likewise, the state and time fixed effects are each jointly highly significant (again, p-values < 0.0001).

Variables are identified by number to facilitate reference to TABLE 1. "D" and "L" indicate whether the specific variable is in "difference" or "level" form. Only seven of the sixty coefficient estimates in TABLE 2 have t-ratios larger than 2.0: three of these represent Population/Labor Force characteristics (*Education-L*, *Working Population-D*, and *Female-D*); two represent Economy characteristics (*Agriculture-D* and *Mining-D*); and two represent Public Sector (Policy) variables (*Federal Revenue-L* and *Sales Tax-L*). More will be said about specific coefficient estimates below, but for now I note that these results provide little evidence that public policy affects growth.

The most noteworthy result from TABLE 2, however, is the large number of insignificant coefficients. Most researchers would consider a regression specification with 53 insignificant coefficients to be poorly specified. But how many variables, and which ones, should be dropped? Towards that end, I use the SIC and AICc selection criteria to select "best" models.

<u>Best SIC and AICc Specifications</u>. TABLE 3 reports results from the best specifications as chosen by the SIC and the AICc model selection criteria. In this case, best means that these specifications have the lowest SIC and AICc values out of the full set of approximately 1.15×10^{18} possible specifications.

A comparison with TABLE 2 is illuminating. The coefficient estimates are generally similar for the variables that appear in both tables. Thus, at least in these cases, changing the set of regressors does not have a large impact on the respective point estimates. However, it does affect the significance of the coefficient estimates.

A total of thirteen variables are included in the Best SIC specification (where I count a variable as appearing once if it appears in both level and difference form); while fourteen are chosen using the AICc criterion. All of these variables have *t*-statistics greater than 2.0. Only seven variables from TABLE 2 had *t*-statistics greater than 2.0.

As a result, TABLE 3 presents a different picture about the determinants of U.S. state economic growth. Most importantly, it suggests a larger role for public policy. The variables *Federal Government*, *Federal Employees*, *Decentralization*, *Tax Burden* (in both level and difference forms), *Sales Tax*, and *Corporate Income Tax* each appear in both the Best SIC and Best AICc specifications. What is not clear is how robust this latter picture is to "reasonable," alternative model specifications.

Extreme Bounds Analysis. Following EBA convention, I identify as "robust" any variable whose coefficient estimates are all same-signed and lie more than two standard deviations away from zero. However, two features of my approach differ from standard EBA analysis: (i) I analyze two "portfolios of models" (one for SIC, one for AICc), and (ii) not every variable appears in every specification within a given portfolio. Accordingly, I also require "robust" variables to appear in at least 50% of the specifications in either portfolio.

The SIC portfolio consists of twenty-seven different models, the Best SIC specification and twenty-six "close competitors" as defined by the $\Re < \sqrt{10}$ criterion. The results from analyzing this portfolio of models are reported in TABLE 4A. Variables are ranked in descending order of number of appearances within the portfolio.

"Robust" variables are identified with an "R." A total of 18 different variables are analyzed in TABLE 4A. Some, like *Education*, appear in all twenty-seven models. Others, like *State & Local Employees* and *State & Local Government* appear in only a very few models (though both have high *t*-ratios when they do appear). Not surprisingly, there is a high overlap between (i) the set of variables that appears in at least 50% of the models in the SIC portfolio, and (ii) the set of variables having a range of *t*-ratios all same-signed and larger than 2.0.

TABLE 4B reports that fifty-seven models are included in the AICc portfolio. A total of 23 different variables appear in at least one of these models. However, many of these appear in only a few models and some, like *Individual Income Tax* and *Higher Education Spending*, appear only once.

TABLE 5 collects the "robust" variables from these extreme bounds analyses and reports them, along with a "mean estimated effect" calculated as the simple average of the respective means from TABLES 4A and 4B. To interpret the respective sizes of these effects, recall that the dependent variable is the 5-year growth rate in state Per Capita Personal Income. For my sample of 30 years (six, 5-year time periods) and 46 states (yielding 276 observations), the mean growth rate is 27.01 percent. Thus, a one percentage point increase in the 5-year growth rate equates approximately to a 3.7 percent increase in growth.

Given the underlying theoretical model of Equation (5), the variables of TABLE 5 should be related to the term, $C = \left[\ln(A_t) - \ln(A_{t-L})\right] + \beta\left[\ln(Q_t) - \ln(Q_{t-L})\right]$. Since A_t and Q_t represent production function parameters, theory suggests that these variables affect the rate of invention and adoption of new technologies that transform the production function over time. This includes effects on resource allocation.

Difference variables are indicated by “D” and represent changes in that variable during the 5-year period. Level variables are indicated by “L” and represent the value of that variable at the beginning of the 5-year period. A variable that appears in both difference and level form has both an immediate and a lagged effect. The difference form indicates the immediate effect, since changes during the 5-year period impact economic growth during that same period. The level form indicates a lagged effect, since changes that get reflected at the beginning of the period show up later, in the subsequent 5-year growth period.

TABLE 5 identifies three Population/Labor Force variables as “robust’ determinants of state economic growth: *Education*, *Working Population*, and *Female*. All have the expected signs. *Education* appears in level form. The mean estimated effect indicates that a one percentage-point increase in the percent of the population that is college-educated at the beginning of a five-year period is associated with a 0.97 percentage-point increase in that state’s subsequent 5-year growth rate. This effect is relatively small in economic terms given that the average 5-year growth rate is 27.01 percent.

The difference form of *Working Population* is also identified as a “robust” variable. The corresponding estimated positive effect indicates that a one-percentage point increase in the share of the population that is aged 20-64 during a given 5-year period is associated with an approximate 0.90 percentage-point, contemporaneous increase in economic growth during that period. Of course, one of the variables being held constant in the estimation is employment (specifically, *DLNL*). Thus, this variable likely reflects higher worker quality within the labor force. Increases in the female share of a state's population (*Female)* are also estimated to have a contemporaneous, albeit negative impact on economic growth. Again, since employment is being held constant, this may reflect productivity differences between men and women in the labor force.

TABLE 5 identifies two Economy characteristic variables: *Agriculture* and *Mining*. The coefficient for *Agriculture* is positive in both level and difference forms, indicating that states with larger and growing agricultural sectors (as measured by earnings share) grew faster than other states. The sources of increased agricultural productivity are debated, but lower input prices, public and private research, increased specialization and changes in farm size have all been identified as contributing factors (cf. Evenson and Huffman, 1997). In contrast, the coefficients for *Mining*, which also appear in both difference and level forms, are each negative. This is consistent with research that finds that the mining industry contributes negligibly, or even negatively, to aggregate TFP growth (cf. Jorgenson and Stiroh, 2000).

TABLE 5 includes seven Public Sector variables: *Federal Government, Federal Employees, Federal Revenue, Decentralization, Tax Burden, Sales Tax,* and *Corporate Income Tax.* The first two variables measure the size of the federal government's presence in a state, measured by earnings share and employment per capita, respectively. The corresponding coefficients for both variables indicate that a larger federal government sector is associated with lower economic growth, ceteris paribus. This may be due to the fact that, relative to the private sector, resources in the public sector are less likely to be allocated to where they will produce income growth (cf. Barro, 1990).

The mean estimated effect for the difference form of *Federal Government* indicates that a one percentage-point increase in this variable – corresponding to roughly a 15% increase in the size of the federal government sector over a 5-year period -- is associated with a contemporaneous 0.83 percentage-point decline in state economic growth. The corresponding estimate for the level form of *Federal Employees* implies that doubling the number of federal employees per capita would lower the subsequent 5-year growth rate of that state by 4.48 percent. While not "robust," it is interesting to note that I estimate similar-sized effects for both *State & Local Government* and *State & Local Employees* (cf. TABLES 4A and 4B).

The variable *Federal Revenue* measures the size of federal aid to states. The sample mean of *Federal Revenue* is 3.90. A one percentage-point increase in this variable would represent approximately a 25% increase in federal aid. The mean estimated effect reported in TABLE 5 for this variable indicates that an increase of this size would raise a state's subsequent 5-year growth rate by 1.16 percentage-points.

The variable *Decentralization* measures the share of total state and local public spending made at the local level. I estimate that a one-percentage point increase in the share of local control is associated with a contemporaneous decrease of 0.11 percentage points in a state's 5-year growth rate. Given that the sample mean of *Decentralization* is 55.0 percentage-points, this constitutes a very small effect. It is consistent with the fact that other studies have had difficulty finding significant effects for this variable (cf. Xie, Zou and Davoodi, 1999).

The remaining three variables are tax variables. The negative coefficients for *Tax Burden* indicate that an increase in state tax revenues as a share of state personal income (i.e., average tax rate) results in lower economic growth. The fact that both level and difference forms of the variable are identified as "robust" determinants indicates that the effect of taxes is both immediate and persistent. A one percentage-point increase in *Tax Burden* over a five-year period is associated with a contemporaneous decrease in state economic growth of 0.63 percentage points. In addition, it is estimated to lower growth by 0.73 percentage points over subsequent 5-year periods. As a gauge of size, a one-percentage point increase in *Tax Burden* equates approximately to a 10% increase in overall taxes.

While not huge, these effects are larger than estimated by previous studies (cf. Wasylenko, 1997). First, they imply both an immediate and long-lived effect of taxes. Second, the estimated effects represent the net effect of taxes and spending. Previous studies, following Helms (1987), commonly estimated "government budget constraint" specifications, so that public expenditures were held constant. The associated tax estimates did not incorporate the corresponding positive effects related to stimulative spending. In contrast, my specifications do not hold constant the level of public expenditures, and thus represent significantly larger negative tax effects.

In contrast, the estimated coefficients for both *Sales Tax* and *Corporate Income Tax* are each positive. Note that a one percentage-point increase in these variables represents approximate increases of 30% and 200%, respectively. The positive effects for these two variables indicate that sales and corporate income taxes are less distortionary than other taxes, such as individual income and property taxes. A further factor may be in play when it comes to business taxes in general, and corporate income taxes in particular: Corporate profits may be more likely than other sources of income to be exported outside the state. Taxing corporate profits may serve to channel economic activity within the state, thus contributing to economic growth.

VII. CONCLUSION

This study examines the determinants of U. S. state economic growth from 1970-1999. Using a generalized Cobb-Douglas production function framework, it considers a large number of potential explanatory variables, including Population/Labor Force characteristics, Economy characteristics, Public Sectory (Policy) variables, and Political Control variables. Counting both difference and level forms, a total of 60 possible explanatory variables are considered, in addition to the capital, employment, and population variables specified by the theory. This yields a total of $2^{60} \cong 1.15 \times 10^{18}$ possible linear combinations of variables, each representing a potentially true model.

I devise an approach for sorting through these different model specifications in order to identify "robust" determinants of state economic growth. My approach is related to the "reasonable extreme bounds analysis" of Granger and Uhlig (1990). Unlike their study, however, I use information criteria (the Schwarz Information Criterion and the corrected version of the Akaike Information Criterion) to choose "portfolios of reasonable models," as suggested by Poskitt and Tremayne (1987). I then perform conventional extreme bounds analysis within these portfolios. An advantage of my approach is that it uses standard SAS procedures and can be easily implemented by other researchers.

My analysis identifies twelve "robust" determinants of U.S. economic growth over the thirty-year period from 1970-1999. Among these are (i) college attainment within the population, (ii) share of the population that is "working age," (iii) population gender share, and the size of the (iv) agricultural and (v) mining sectors of the economy. I also find that a relatively large number of public sector variables are significantly correlated with growth. Among these are (vi and vii), the size of the federal sector within a state, (viii) federal aid, (ix) decentralization, and (x through xii) various categories of taxes. This latter finding highlights the importance of public policy as a determinant of economic growth. While one must be careful to draw causative inferences from these results, they provide further motivation to identify channels by which public policy directly impacts economic activity.

REFERENCES

Akai, Nobuo and Masayo Sakata. “Fiscal Decentralization Contributes to Economic Growth: Evidence From State-level Cross-section Data for the United States.” Journal of Urban Economics Vol. 52 (2002): 93-108.

Alm, James and Janet Rogers. “Do State Fiscal Policies Affect Economic Growth?” Mimeo, Department of Economics, Andrew Young School of Policy Studies, Georgia State University, 2005.

Aschauer, David Alan. “Is Public Expenditure Productive?” Journal of Monetary Economics Vol. 23 (1989): 177-200.

Barro, Robert J. “Government Spending in a Simple Model of Endogenous Growth.” Journal of Political Economy, Vol. 98 (1990): S103-S125.

Bartik, Timothy J. Who Benefits from State and Local Economic Development Policies? Kalamazoo, MI: W.E. Upjohn Insitute for Employment Research, 1991.

Besley, Timothy and Anne Case. “Does Electoral Accountability Affect Economic Policy Choices? Evidence From Gubernatorial Term Limits.” Quarterly Journal of Economics Vol. 110, No. 3 (1995): 769-798.

Besley, Timothy and Anne Case. “Political Institutions and Policy Choices: Evidence From the United States.’ Journal of Economic Literature Vol. 41 (2003): 7-73.

Burnham, Kenneth P. and David R. Anderson. Model Selection and Multimodel Inference, 2nd Edition. New York: Springer-Verlag New York, Inc., 2002.

Carroll, Robert and Michael Wasylenko. “Do State Business Climates Still Matter? – Evidence of a Structural Change.” National Tax Journal Vol. 47, No. 1 (1994): 19-37.

Caselli, Francesco and Wilbur John Coleman II. “The U.S. Structural Transformation and Regional Convergence: A Reinterpretation. Journal of Political Economy Vol. 109, No. 3 (2001): 584-616.

Chatfield, Chris. “Model Uncertainty, Data Mining, and Statistical Inference.” Journal of the Royal Statistical Society, Series A, Vol. 158, No. 3 (1995): 419-466.

Ciccone, Antonio and Robert E. Hall. “Productivity and the Density of Economic Activity.” American Economic Review Vol. 86, No. 1 (1996): 54-70.

Clark, David E. and Christopher A. Murphy. “Countywide Employment and Population Growth: An Analysis of the 1980s.” Journal of Regional Science Vol. 36, No. 2 (1996): 235-256.

Crain, W. Mark and Katherine J. Lee. "Economic Growth Regressions for the American States: A Sensitivity Analysis." Economic Inquiry Vol. 37, No. 2 (1999): 242-257.

Crown, William H. and Leonard F. Wheat. "State Per Capita Income Convergence Since 1950: Sharecropping's Demise and Other Influences." Journal of Regional Science Vol. 35, No. 4 (1995): 527-552.

Dalenberg, Douglas R. and Mark D. Partridge. "The Effects of Taxes, Expenditures, and Public Infrastructure on Metropolitan Area Employment." Journal of Regional Science Vol. 35, No. 4 (1995): 617-640.

Demekas, Dimitri G., and Zenon G. Kontolemis. "Government Employment and Wages and Labor Market Performance." Oxford Bulletin of Economics and Statistics Vol. 62, No. 3 (2000):

Evenson, Robert E. and Wallace B. Huffman. "Long-Run Structural and Productivity Change in U.S. Agriculture: Effects of Prices and Policies." Mimeo, Yale University, Economic Growth Center, Center Discussion Paper No. 773, 1997.

Fernández, Carmen, Eduardo Ley, and Mark F. J. Steel. "Model Uncertainty in Cross-Country Growth Regressions." Journal of Applied Econometrics Vol. 16 (2001): 563-576.

Fisher, Ronald C. "The Effects of State and Local Public Services on Economic Development." New England Economic Review (1997), March/April): 53-67.

Foster, Kathryn A. "Exploring the Links Between Political Structure and Metropolitan Growth." Political Geography Vol. 12, No. 6 (1993): 523-547.

Furnival, G. M. and Robert W. J. Wilson. "Regression By Leaps and Bounds." Technometrics Vol. 16 (1974): 499-511.

Garofalo, Gasper A. and Steven Yamarik. "Regional Convergence: Evidence From a New State-by-State Capital Stock Series." The Review of Economics and Statistics, Vol. 84, No. 2 (2002): 315-323.

Garcia-Milà, Teresa, and Therese J. McGuire. "The Contribution of Publicly Provided Inputs to States' Economies." Regional Science and Urban Economics Vol. 22 (1992): 229-241.

Gelb, A., J. B. Knight, and R. H. Sabot. "Public Sector Employment, Rent Seeking, and Economic Growth." The Economic Journal Vol. 101, No. 408 (1991): 1186-1199.

Gilligan, Thomas W. and Matsusaka, John G. "Deviations from Constituent Interests: the Role of Legislative Structure and Political Parties in the States." Economic Inquiry Vol. 33 (1995): 383-401.

Granger, Clive W. J. and Uhlig, Harald F. "Reasonable Extreme-Bounds Analysis." Journal of Econometrics Vol. 44, Nos. 1-2 (1990): 159-170.

Grier, Kevin B. and Gordon Tullock. "An Empirical Analysis of Cross-national Economic Growth, 1951-1980." Journal of Monetary Economics Vol. 24, No. 2 91989): 169-328.

Helms, Jay L. "The Effect of State and Local Taxes on Economic Growth: A Time Series-Cross Section Approach." The Review of Economics and Statistics Vol. 67, No. 4 (1985): 574-582.

Hendry, David F. and Hans-Martin Krolzig. "We Ran One Regression." Oxford Bulletin of Economics and Statistics Vol. 66, No. 5 (2004): 799-810.

Hirsch, Barry T., David A. MacPherson, and Wayne G. Vroman. "Estimates of Union Density by State." Monthly Labor Review Vol. 124, No. 7 (2001): 51-55.

Hoeting, Jennifer A., David Madigan, Adrian E. Raftery, and Chris T. Volinsky. "Bayesian Model Averaging: A Tutorial." Statistical Science Vol. 14, No. 4 (1999): 382-417.

Holtz-Eakin, Douglas. "Solow and the States: Capital Accumulation, Productivity, and Economic Growth." National Tax Journal Vol. 46, No. 4 (1993): 425-440

Hoover, Kevin D. and Stephen J. Perez, "Truth and Robustness in Cross-country Growth Regressions." Oxford Bulletin of Economics and Statistics Vol. 66, No. 5 (2004): 765-798.

Hurvich, C. M. and C. L. Tsai. "Regression and Time Series Model Selection in Small Samples." Biometrika, Vol. 76 (1989): 297-307.

Jeffreys, H. The Theory of Probability, 3rd Edition. Oxford: Clarendon Press, 1961.

Jorgenson, Dale W. and Kevin J. Stiroh. "Raising the Speed Limit: U.S. Economic Growth in the Information Age." Brookings Papers on Economic Activity, Vol. 1 (2000): 125-211.

Leamer, Edward E. "Sensitivity Analyses Would Help." American Economic Review Vol. 75 (1985): 308-13.

Lee, Young and Roger H. Gordon. "Tax Structure and Economic Growth." Journal of Public Economics Vol. 89 (2005): 1027-1043.

Levine, Ross and David Renelt. "A Sensitivity Analysis of Cross-Country Growth Regressions." American Economic Review Vol. 82, No. 4 (1992): 942-963.

Levitt, Steven D. and James M. Poterba. "Congressional Distributive Politics and State Economic Performance." Public Choice Vol. 99 (1999): 185-216.

McQuarrie, Allan D. R. and Chih-Ling Tsai. Regression and Time Series Model Selection. Singapore: World Scientific Publishing Co. Pte. Ltd., 1998.

McAleer, Michael, Adrian R. Pagan, and Paul A. Volker. "What Will Take the Con Out of Econometrics?" American Economic Review Vol. 75, No. 3 (1985): 293-307.

McGuire, Therese. "Review of 'Who Benefits From State and Local Economic Development Policies?'" National Tax Journal, Vol. XLV, No. 4 (1992): 457-459.

Miller, A. J. Subset Selection in Regression. London, Chapman and Hall, 1990.

Mitchener, Kris James and Ian W. McLean. "The Productivity of U.S. States Since 1880." Journal of Economic Growth, Vol. 8, (2003): 73-114.

Mofidi, Alaeddin and Joe A. Stone. "Do State and Local Taxes Affect Economic Growth?" Review of Economics and Statistics Vol. 72, No. 4 (1990): 686-691.

Mullen, John K. and Martin Williams. "Marginal Tax Rates and State Economic Growth." Regional Science and Urban Economics Vol. 24 (1994): 687-705.

Nelson, Arthur C. and Kathryn A. Foster. "Metropolitan Governance Structure and Income Growth." Journal of Urban Affairs Vol. 21, No. 3 (1999):309-324.

Partridge, Mark D. and Dan S. Rickman. "The Role of Industry Structure, Costs, and Economic Spillovers in Determining State Employment Growth Rates." Review of Regional Studies Vol. 26, No. 3 (1996): 235-64.

Phillips, Joseph and Ernest Goss. "The Effect of State and Local Taxes on Economic Development: A Meta-Analysis." Southern Economic Journal 62 (1995): 320-333.

Plaut, Thomas R. and Joseph E. Pluta. "Business Climate, Taxes and Expenditures, and State Industrial Growth in the United States." Southern Economic Journal Vol. 50 (1983): 99-109.

Poskitt, D. S., and A. R. Tremayne. "Determining a Portfolio of Time Series Models." Biometrika Vol. 74, No. 1 (1987): 125-137.

Poterba, James M. "States' Responses to Fiscal Crises: The Effects of Budgetary Institutions and Politics." Journal of Political Economy Vol. 102 (1994): 799-821.

Reed, W. Robert. "Democrats, Republicans, and Taxes: Evidence that Political Parties Matter." Journal of Public Economics (2006): Forthcoming.

Sala-i-Martin, Xavier X. "I Just Ran Two Million Regressions." American Economic Review Vol. 87, No. 2 (1997): 178-183.

Sala-i-Martin, Xavier, Gernot Doppelhofer, and Ronald I. Miller, "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach." American Economic Review Vol. 94, No. 4 (2004): 813-835.

SAS Institute Inc. SAS/STAT 9.1 User's Guide. Cary, NC: SAS Institute Inc., 2004.

Stansel, Dean. "Local Decentralization and Local Economic Growth: A Cross-sectional Examination of U.S. Metropolitan Areas." Journal of Urban Economics Vol. 57 (2005): 55-72.

Sugiura, N. "Further Analysis of the Data By Akaike's Information Criterion and the Finite Corrections." Communications in Statistics – Theory and Methods, Vol. 7 (1978): 13-26.

van der Ploeg, Frederick. "Rolling Back the Public Sector – Differential Effects on Employment, Investment and Growth." CESifo Working Paper Series No. 890. Working paper, March, 2003.

Wasylenko, Michael and Therese McGuire. "Jobs and Taxes: The Effect of Business Climate on State Employment Growth Rates." National Tax Journal Vol. 38, No. 4 (1985): 497-511.

Xie, Danyang, Heng-fu Zou, and Hamid Davoodi. "Fiscal Decentralization and Economic Growth in the United States." Journal of Urban Economics Vol. 45 (1999): 228-239.

Yamarik, Steven. "Can Tax Policy Help Explain State-level Macroeconomic Growth?" Economics Letters Vol. 68 (2000): 211-215.

Zellner, A. "Comments on Time Series Analysis and Causal Concepts in Business Cycle Research." In New Methods in Business Cycle Research, ed. C. A. Sims. Minneapolis: Federal Reserve Bank of Minneapolis, 1977.

TABLE 1: List of Potential Determinants of U.S. State Economic Growth[19]

NUMBER	NAME	DESCRIPTION (ALL VARIABLES CALCULATED AS 5-YEAR AVERAGES)	SELECTED STUDIES WHICH HAVE USED THIS OR A RELATED VARIABLE
A) Population/Labor Force Characteristics			
1	Education	Percent of population (aged 25 and up) who have completed college (SOURCE: Census)	Wasylenko and McGuire (1985); Garcia-Milà and McGuire (1992); Crown and Wheat (1995); Phillips and Goss (1995); Dalenberg and Partridge (1995); Partridge and Rickman (1996); Clark and Murphy (1996); Ciccone and Barro (1996); Crain and Lee (1999)
2	Working Population	Percent of population between 20 and 64 years of age (SOURCE: Census)	Wasylenko and McGuire (1985); Mofidi and Stone (1990); Dalenberg and Partridge (1995); Crain and Lee (1999)
3	Nonwhite	Percent of population that is nonwhite (SOURCE: Census)	Mofidi and Stone (1990); Partridge and Rickman (1996); Crain and Lee (1999)
4	Female	Percent of population that is female (SOURCE: Census)	Mofidi and Stone (1990); Partridge and Rickman (1996); Clark and Murphy (1996)
5	Population	Log of total population (SOURCE: Census)	Ciccone and Barro (1996); Alm and Rogers (2005)
B) Economy Characteristics			
6	Population Density	Population density (SOURCE: Census)	Wasylenko and McGuire (1985); Carroll and Wasyenko (1994); Clark and Murphy (1996); Ciccone and Hall (1996); Crain and Lee (1999)

[19] Descriptive statistics for all variables are reported in Appendix A.

NUMBER	NAME	DESCRIPTION (ALL VARIABLES CALCULATED AS 5-YEAR AVERAGES)	SELECTED STUDIES WHICH HAVE USED THIS OR A RELATED VARIABLE
7	Urban	Percent of population living in urban areas (SOURCE: Census)	Holtz-Eakin (1993); Partridge and Rickman (1996); Crain and Lee (1999)
8	Agriculture	Share of total earnings earned in "Farm" and "Other Agriculture" industries (SOURCE: BEA)	Crown and Wheat (1995); Caselli and Coleman (2001)
9	Manufacturing	Share of total earnings earned in "Manufacturing" industries (SOURCE: BEA)	Crown and Wheat (1995); Crain and Lee (1999); Caselli and Coleman (2001); Stansel (2005)
10	Service	Share of total earnings earned in "Service" industries (SOURCE: BEA)	Clark and Murphy (1996)
11	Mining	Share of total earnings earned in "Mining" industries (SOURCE: BEA)	Holtz-Eakin (1993); Crown and Wheat (1995); Clark and Murphy (1996); Mitchener and McLean (2003)
12	Union	Percent of nonagricultural wage and salary employees who are union members (SOURCE: Hirsch, McPherson, and Vroman, 2001)	Plaut and Pluta (1983); Mofidi and Stone (1990); Dalenberg and Partridge (1995); Phillips and Goss (1995); Partridge and Rickman (1996); Clark and Murphy (1996)
13	Diversity	A measure of industrial diversity, $Diversity = \sum_i \left(\frac{Earnings\ in\ Industry_i}{Total\ Earnings} \right)^2$ (SOURCE: BEA)	Mofidi and Stone (1990); Garcia-Milà and McGuire (1992); Partridge and Rickman (1996); Crain and Lee (1999)

C1) Public Sector: Size and Structure Variables

14	Federal Government	Share of total earnings earned in "Federal government" (SOURCE: BEA)	Crain and Lee (1999)

NUMBER	NAME	DESCRIPTION (ALL VARIABLES CALCULATED AS 5-YEAR AVERAGES)	SELECTED STUDIES WHICH HAVE USED THIS OR A RELATED VARIABLE
15	State & Local Government	Share of total earnings earned in "State and Local government" (SOURCE: BEA)	Crain and Lee (1999)
16	Federal Employees	Log of federal employees per capita (SOURCE: Census)	Gelb, Knight and Sabot (1991); Demekas and Kontolemis (2000); van der Ploeg (2003)
17	State & Local Employees	Log of state and local employees per capita (SOURCE: Census)	Gelb, Knight and Sabot (1991); Demekas and Kontolemis (2000); van der Ploeg (2003)
18	Federal Revenue	Intergovernmental revenue received by state and local governments from the federal government as a share of personal income (SOURCE: Census)	Carroll and Wasylenko (1994); Dalenberg and Partridge (1995); Xie, Zou, and Davoodi (1999); Akai and Sakata (2002)
19	Decentralization	Share of total state and local direct general expenditures made by local governments (SOURCE: Census)	Clark and Murphy (1996); Xie, Zou, and Davoodi (1999); Crain and Lee (1999); Akai and Sakata (2002)
20	Number of Governments	Number of state and local governments (SOURCE: Census)	Foster (1993); Nelson and Foster (1999); Stansel (2005)

C2) Public Sector: Tax Variables

NUMBER	NAME	DESCRIPTION	SELECTED STUDIES
21	Tax Burden	Total state and local tax revenues as a share of personal income (SOURCE: Census)	Plaut and Pluta (1983); Mullen and Williams (1994); Partridge and Rickman (1996); Xie, Zou, and Davoodi (1999); Yamarik (2000)
22	Property Tax	Total state and local property tax revenues as a share of personal income (SOURCE: Census)	Plaut and Pluta (1983); Helms (1985); Carroll and Wasylenko (1994); Yamarik (2000)
23	Sales Tax	Total state sales tax revenues as a share of personal income (SOURCE: Census)	Plaut and Pluta (1983); Wasylenko and McGuire (1985); Yamarik (2000)

NUMBER	NAME	DESCRIPTION (ALL VARIABLES CALCULATED AS 5-YEAR AVERAGES)	SELECTED STUDIES WHICH HAVE USED THIS OR A RELATED VARIABLE
24	Individual Income Tax	Total state individual income tax revenues as a share of personal income (SOURCE: Census)	Plaut and Pluta (1983); Wasylenko and McGuire (1985); Carroll and Wasylenko (1994); Yamarik (2000)
25	Corporate Income Tax	Total state corporate income tax revenues as a share of personal income (SOURCE: Census)	Plaut and Pluta (1983); Wasylenko and McGuire (1985); Carroll and Wasylenko (1994)

C3) Public Sector: Expenditure Variables

NUMBER	NAME	DESCRIPTION	SELECTED STUDIES
26	Local Education Spending	Total state and local spending on local schools as a share of total state and local expenditures (SOURCE: Census)	Plaut and Pluta (1983); Helms (1985); Wasylenko and McGuire (1985); Ashauer (1989); Mofidi and Stone (1990); Carroll and Wasylenko (1994); Fisher (1997)
27	Higher Education Spending	Total state and local spending on higher education as a share of total state and local expenditures (SOURCE: Census)	Plaut and Pluta (1983); Helms (1985); Wasylenko and McGuire (1985); Aschauer (1989); Mofidi and Stone (1990); Carroll and Wasylenko (1994); Dalenberg and Partridge (1995); Fisher (1997)
28	Health & Hospital Spending	Total state and local spending on health and hospitals as a share of total state and local expenditures (SOURCE: Census)	Helms (1985); Ashauer (1989); Mofidi and Stone (1990); Carroll and Wasylenko (1994)
29	Highway Spending	Total state and local direct spending on highways as a share of total state and local expenditures (SOURCE: Census)	Helms (1985); Ashauer (1989); Mofidi and Stone (1990); Carroll and Wasylenko (1994); Dalenberg and Partridge (1995); Fisher (1997)

D) Political Control Variables

NUMBER	NAME	DESCRIPTION	SELECTED STUDIES
30	Democratic Legislature	Percent of years that both houses of the state legislature were controlled by Democrats (SOURCE: National Conference of State Legislatures)	Levitt and Poterba (1997); Akai and Sakata (2002); Besley and Case (2003); Reed (2006)

NUMBER	NAME	DESCRIPTION (ALL VARIABLES CALCULATED AS 5-YEAR AVERAGES)	SELECTED STUDIES WHICH HAVE USED THIS OR A RELATED VARIABLE
31	Republican Legislature	Percent of years that both houses of the state legislature were controlled by Republicans (SOURCE: National Conference of State Legislatures)	Levitt and Poterba (1997); Akai and Sakata (2002); Besley and Case (2003); Reed (2006)
32	Democratic Governor	Percent of years that governor was a Democrat (SOURCE: National Conference of State Legislatures)	Levitt and Poterba (1997); Akai and Sakata (2002); Besley and Case (1995, 2003); Reed (2006)

TABLE 2
Regression Results Using All Variables

VARIABLE		COEFFICIENT (*t*-STAT)	VARIABLE		COEFFICIENT (*t*-STAT)
Education(1)	D	0.9883 (1.22)	State & Local Employees(17)	D	-0.4118 (0.08)
	L	1.1431 (4.38)		L	-3.9735 (0.77)
Working Population(2)	D	1.0171 (2.22)	Federal Revenue(18)	D	-0.1147 (0.29)
	L	0.1705 (0.58)		L	0.9760 (2.15)
Nonwhite(3)	D	0.5852 (0.98)	Decentralization(19)	D	-0.1199 (1.31)
	L	0.0510 (0.24)		L	0.0545 (0.52)
Female(4)	D	-4.3025 (2.37)	Number of Governments(20)	D	-2.7378 (0.91)
	L	1.5506 (1.58)		L	0.3560 (0.17)
Population(5)	L	5.5970 (1.51)	Tax Burden(21)	D	-0.3484 (0.72)
Population Density(6)	D	0.0214 (0.42)		L	-0.8502 (1.57)
	L	0.0026 (0.18)	Property Tax(22)	D	-0.1257 (0.19)
Urban(7)	D	-0.1832 (1.01)		L	0.6128 (0.89)
	L	-0.0916 (0.80)	Sales Tax(23)	D	0.3490 (0.90)
Agriculture(8)	D	0.4781 (2.95)		L	1.4607 (2.76)
	L	0.0886 (0.58)			

VARIABLE		COEFFICIENT (*t*-STAT)	VARIABLE		COEFFICIENT (*t*-STAT)
Manufacturing(9)	D	-0.3079 (1.56)	Individual Income Tax(24)	D	0.0514 (0.07)
	L	0.0137 (0.08)		L	0.8677 (1.24)
Service(10)	D	-0.2831 (1.08)	Corporate Income Tax(25)	D	-1.3218 (1.01)
	L	-0.2358 (0.92)		L	2.1011 (1.44)
Mining(11)	D	-1.4230 (4.36)	Local Education Spending(26)	D	0.0392 (0.30)
	L	-0.5063 (1.92)		L	-0.0154 (0.10)
Union(12)	D	0.0112 (0.13)	Higher Education Spending(27)	D	-0.0928 (0.48)
	L	-0.0917 (1.00)		L	-0.0684 (0.37)
Diversity(13)	D	0.4697 (1.81)	Health & Hospital Spending(28)	D	0.2375 (1.52)
	L	-0.1376 (0.49)		L	0.1968 (1.30)
Federal Government(14)	D	-0.6316 (1.92)	Highway Spending(29)	D	0.1173 (0.95)
	L	0.3208 (0.89)		L	0.1340 (0.92)
State & Local Government(15)	D	-0.5449 (1.36)	Democratic Legislature(30)		0.0073 (1.17)
	L	-0.2488 (0.61)	Republican Legislature(31)		-0.0033 (0.50)
Federal Employees(16)	D	-0.0719 (0.04)	Democratic Governor(32)		0.0033 (0.87)
	L	-4.5456 (1.95)			

Observations: 276 (= 46 states × 6 time periods)	$R^2 = 0.984$
SIC = 778.156	AICc = 804.288

NOTE: The regression equation is specified according to Equation (11) in the text and includes all possible variables. Variables are defined in TABLE 1. "D" and "L" stand for difference and level forms of the variables. In addition to the variables listed above, the model includes the variable DLNK, DLNL, DLNN, and state and time fixed effects. t-statistics are listed in parenthesis below each estimated coefficient.

TABLE 3
Regression Results from Best SIC and Best AICc Specifications

Variable Number	Variable Name[20]	D/L	Best SIC Specification	Best AICc Specification
1	Education	D	---	1.3994 (2.20)
		L	1.0353 (6.47)	0.9514 (6.84)
2	Working Population	D	0.9378 (2.88)	0.7838 (2.36)
4	Female	D	-5.8832 (4.54)	-5.8083 (4.28)
5	Population	L	---	4.2579 (2.47)
8	Agriculture	D	0.7486 (8.06)	0.7096 (7.82)
		L	0.3082 (4.60)	0.2808 (4.20)
10	Service	L	-0.3157 (2.31)	---
11	Mining	D	-1.1293 (4.65)	-1.1300 (4.66)
		L	-0.5044 (2.91)	-0.3595 (2.21)
14	Federal Government	D	-0.8684 (3.96)	-0.7497 (3.40)
16	Federal Employees	L	-5.8371 (4.26)	-3.8148 (2.45)
18	Federal revenue	L	1.1723 (4.17)	1.0943 (3.91)
19	Decentralization	D	-0.1087 (2.21)	-0.1111 (2.29)

[20] Variables are described in TABLE 1.

Variable Number	Variable Name[20]	D/L	Best SIC Specification	Best AICc Specification
21	Tax Burden	D	-0.6333 (3.32)	-0.7136 (3.84)
		L	-0.7405 (3.43)	-0.7643 (3.59)
23	Sales Tax	L	1.2511 (4.46)	1.1083 (3.94)
25	Corporate Income Tax	L	2.5037 (2.81)	2.4220 (2.76)
30	Democratic Legislature	--		0.0113 (2.23)
Number of observations			276	276
R^2			0.9806	0.9813
SIC			592.026	592.928
AICc			663.086	660.043

<u>NOTE</u>: The regression equation follows the general specification of Equation (11) in the text. Variables are defined in TABLE 1. "D" and "L" stand for difference and level forms of the variables. In addition to the variables listed above, the model includes the variables DLNK, DLNL, DLNN, and state and time fixed effects. t-statistics are listed in parenthesis below each estimated coefficient. The algorithm for determining the best specifications in described in Section IV.

TABLE 4A
Extreme Bounds Analysis for Portfolio of Top SIC Models

NUMBER (PERCENT)	ROBUST	VARIABLE	D/L	RANGE OF COEFFICIENT ESTIMATES			RANGE OF *t*-RATIOS		
				LOW	MEAN	HIGH	LOW	MEAN	HIGH
27(100%)	R	Education(1)	L	0.7100	0.9477	1.0932	5.16	6.38	6.85
27(100%)	R	Female(4)	D	-6.9309	-6.1086	-5.4376	4.10	4.61	5.30
27(100%)	R	Agriculture(8)	D	0.6360	0.7127	0.7693	6.21	7.57	8.25
27(100%)	R	Agriculture(8)	L	0.2440	0.3071	0.3582	3.48	4.57	5.37
27(100%)	R	Mining(11)	D	-1.2974	-1.1395	-0.9005	4.07	4.72	5.38
27(100%)	R	Federal Government(14)	D	-1.0244	-0.8805	-0.7497	3.39	4.00	4.74
27(100%)	R	Federal Employees(16)	L	-6.1410	-5.0072	-3.6200	2.31	3.48	4.46
27(100%)	R	Federal Revenue(18)	L	0.9085	1.1573	1.3387	3.24	4.02	4.52
27(100%)	R	Sales Tax(23)	L	0.9990	1.1636	1.2911	3.58	4.11	4.59
27(100%)	R	Corporate Income Tax(24)	L	2.2712	2.5939	3.3821	2.56	2.91	3.79
24(89%)	--	Working Population(2)	D	0.6515	0.9055	1.1352	1.97	2.78	3.37
24(89%)	R	Mining(11)	L	-0.7136	-0.4670	-0.3440	2.10	2.79	4.18
24(89%)	R	Tax Burden(21)	D	-0.8223	-0.6639	-0.5053	2.56	3.51	4.55

NUMBER (PERCENT)	ROBUST	VARIABLE	D/L	RANGE OF COEFFICIENT ESTIMATES			RANGE OF *t*-RATIOS		
				LOW	MEAN	HIGH	LOW	MEAN	HIGH
24(89%)	R	Tax Burden(21)	L	-0.9129	-0.7467	-0.6449	2.98	3.48	4.48
14(52%)	R	Decentralization(19)	D	-0.1321	-0.1112	-0.1012	2.07	2.27	2.68
11(41%)	--	Population(5)	L	3.2958	3.9428	4.5846	1.90	2.27	2.64
11(41%)	--	Service(10)	L	-0.4862	-0.3454	-0.2757	1.96	2.53	3.54
8(30%)	--	Education(1)	D	1.1661	1.4657	1.7840	1.84	2.31	2.84
7(26%)	--	State & Local Government(15)	D	-0.9016	-0.6049	-0.4388	1.75	2.40	3.39
7(26%)	--	Decentralization(19)	L	0.1020	0.1285	0.1422	1.90	2.44	2.71
6(22%)	--	Democratic Legislature(30)	-	0.0095	0.0110	0.0128	1.88	2.16	2.53
3(11%)	--	Health & Hospital Spending(28)	D	0.1894	0.2226	0.2588	1.63	1.91	2.23
2(7%)	--	State & Local Employees(17)	L	-7.2722	-7.1917	-7.1113	2.53	2.54	2.56
1(4%)	--	State & Local Government(15)	L	-0.6068	-0.6068	-0.6068	2.50	2.50	2.50

NOTE: *The algorithm for calculating "extreme bounds" is described in Section IV of the text.*

TABLE 4B
Extreme Bounds Analysis for Portfolio of Top AICc Models

NUMBER (PERCENT)	ROBUST	VARIABLE	D/L	RANGE OF COEFFICIENT ESTIMATES			RANGE OF *t*-RATIOS		
				LOW	MEAN	HIGH	LOW	MEAN	HIGH
57(100%)	R	Education(1)	L	0.8354	0.9912	1.1365	5.74	6.42	7.06
57(100%)	R	Working Population(2)	D	0.6760	0.8864	1.0386	2.04	2.70	3.23
57(100%)	R	Female(4)	D	-6.7496	-5.6811	-4.4992	3.26	4.21	5.13
57(100%)	R	Agriculture(8)	D	0.4968	0.6754	0.7602	4.02	7.01	8.21
57(100%)	R	Agriculture(8)	L	0.1783	0.2590	0.3071	2.40	3.80	4.58
57(100%)	R	Mining(11)	D	-1.3193	-1.1687	-1.0655	4.32	4.79	5.27
57(100%)	--	Mining(11)	L	-0.5221	-0.4173	-0.2859	1.74	2.46	3.03
57(100%)	R	Federal Government(14)	D	-0.9425	-0.7731	-0.6711	3.02	3.51	4.15
57(100%)	R	Federal Employees(16)	L	-5.7203	-3.9538	-3.4240	2.21	2.54	4.21
57(100%)	R	Federal Revenue(18)	L	0.9514	1.1672	1.3310	3.37	4.08	4.61
57(100%)	R	Tax Burden(21)	D	-0.7492	-0.6030	-0.4535	2.32	3.16	3.98
57(100%)	R	Tax Burden(21)	L	-0.8272	-0.7222	-0.6386	2.96	3.37	3.88
57(100%)	R	Sales Tax(23)	L	0.9668	1.0802	1.1669	3.36	3.81	4.13

NUMBER (PERCENT)	ROBUST	VARIABLE	D/L	RANGE OF COEFFICIENT ESTIMATES			RANGE OF *t*-RATIOS		
				LOW	MEAN	HIGH	LOW	MEAN	HIGH
56(98%)	--	Population(5)	L	3.1220	4.0412	5.0292	1.81	2.31	2.83
55(96%)	--	Decentralization(19)	D	-0.1284	-0.1093	-0.0959	1.96	2.25	2.64
52(91%)	R	Corporate Income Tax(25)	L	2.0656	2.4037	2.9069	2.31	2.72	3.26
37(65%)	--	Democratic Legislature(30)	-	0.0067	0.0103	0.0127	1.33	2.02	2.52
34(60%)	--	Education(1)	D	0.9103	1.2891	1.4924	1.43	2.02	2.33
34(60%)	--	Service(10)	L	-0.3761	-0.2798	-0.1997	1.44	2.03	2.62
26(46%)	--	State & Local Government(15)	D	-0.6745	-0.4750	-0.3118	1.23	1.87	2.56
17(30%)	--	Diversity(13)	D	0.1820	0.3384	0.4565	1.14	1.93	2.48
17(30%)	--	Health & Hospital Spending(28)	D	0.1544	0.1830	0.2297	1.34	1.58	1.87
13(23%)	--	Manufacturing(9)	D	-0.3544	-0.2642	-0.2211	1.72	1.96	2.36
13(23%)	--	Union(12)	L	-0.1085	-0.0887	-0.0571	1.05	1.60	1.94
8(14%)	--	Diversity(13)	L	-0.3253	-0.2776	-0.2003	1.25	1.74	2.05
5(9%)	--	Corporate Income Tax(25)	D	-2.6637	-2.4655	-2.2580	2.51	2.72	2.92
4(7%)	--	Union(12)	D	0.0667	0.0745	0.0853	1.17	1.31	1.51

NUMBER (PERCENT)	ROBUST	VARIABLE	D/L	RANGE OF COEFFICIENT ESTIMATES			RANGE OF *t*-RATIOS		
				LOW	MEAN	HIGH	LOW	MEAN	HIGH
3(5%)	--	Democratic Governor(32)	-	0.0039	0.0046	0.0049	1.19	1.42	1.54
2(4%)	--	Decentralization(19)	L	0.1065	0.1105	0.1146	1.99	2.06	2.14
1(2%)	--	Female(4)	L	0.8031	0.8031	0.8031	1.09	1.09	1.09
1(2%)	--	Service(10)	D	-0.2989	-0.2989	-0.2989	1.47	1.47	1.47
1(2%)	--	Individual Income Tax(24)	L	0.5450	0.5450	0.5450	1.39	1.39	1.39
1(2%)	--	Higher Education Spending(27)	D	-0.1359	-0.1359	-0.1359	1.16	1.16	1.16
1(2%)	--	Health & Hospital Spending(28)	L	0.1557	0.1557	0.1557	1.41	1.41	1.41

NOTE: The algorithm for calculating "extreme bounds" is described in Section IV of the text.

TABLE 5
"Robust" Variables and Mean Estimated Effects

CATEGORY	VARIABLE NUMBER	VARIABLE NAME	MEAN ESTIMATED EFFECT
Population/Labor Force characteristics	1	Education-L	0.97
	2	Working Population-D	0.90
	4	Female-D	-5.89
Economy characteristics	8	Agriculture-D	0.69
	8	Agriculture-L	0.28
	11	Mining-D	-1.15
	11	Mining-L	-0.44
Public Sector (Policy) variables	14	Federal Government-D	-0.83
	16	Federal Employees-L	-4.48
	18	Federal Revenue-L	1.16
	19	Decentralization-D	-0.11
	21	Tax Burden-D	-0.63
	21	Tax Burden-L	-0.73
	23	Sales Tax-L	1.12
	24	Corporate Income Tax-L	1.39

NOTE: *"Mean Estimated Effect" is the simple average of the "Mean" coefficient estimates in TABLES 4A and 4B.*

APPENDIX
Statistical Summary of Data

Number	*Name*[21]		*Mean*	*Std. Deviation*	*Minimum*	*Maximum*
Dep. Variable	5-Year Growth Rate in PCPI		27.01	10.30	6.49	64.40
1	Education	D	1.76	0.55	0.34	3.21
		L	16.39	4.95	6.66	30.21
2	Working Population	D	0.97	0.92	-1.22	2.93
		L	55.89	3.18	47.54	62.26
3	Nonwhite	D	0.55	0.52	-0.98	2.42
		L	12.08	8.79	0.36	37.35
4	Female	D	-0.02	0.15	-0.57	0.75
		L	51.24	0.79	48.77	52.76
5	Population	L	14.93	1.01	12.72	17.27
6	Population Density	D	5.09	6.77	-8.44	37.26
		L	167.72	234.21	3.44	1089.83
7	Urban	D	0.75	1.15	-1.97	3.96
		L	67.23	14.73	32.16	93.54
8	Agriculture	D	-0.04	2.42	-16.72	18.85
		L	3.12	3.88	-8.92	29.06
9	Manufacturing	D	-0.84	1.69	-6.09	3.37
		L	21.03	8.54	3.73	40.49
10	Service	D	1.47	1.27	-3.22	6.40
		L	19.56	5.71	10.93	41.55
11	Mining	D	-0.19	0.77	-3.29	4.27
		L	2.21	3.60	0.02	24.98
12	Union	D	-1.47	2.39	-10.6	5.0
		L	18.42	8.18	3.3	41.7
13	Diversity	D	-0.06	0.78	-5.42	4.66
		L	17.44	2.05	13.84	23.56
14	Federal Government	D	-0.57	0.82	-5.98	1.25
		L	7.02	3.63	2.05	23.45

[21] Variables are described in TABLE 1.

Number	*Name*[21]		*Mean*	*Std. Deviation*	*Minimum*	*Maximum*
15	State & Local Government	D	-0.10	0.88	-3.98	5.02
		L	11.98	1.66	8.47	18.40
16	Federal Employees	D	-0.04	0.09	-0.67	0.37
		L	4.70	0.38	3.99	5.93
17	State & Local Employees	D	0.03	0.06	-0.13	0.19
		L	6.20	0.13	5.86	6.66
18	Federal Revenue	D	0.10	0.77	-1.74	2.50
		L	3.90	1.22	1.67	8.31
19	Decentralization	D	-0.13	2.54	-10.37	6.35
		L	55.00	7.88	34.81	76.80
20	Number of Governments	D	-0.02	0.07	-0.36	0.40
		L	5.90	0.88	4.26	8.40
21	Tax Burden	D	0.12	0.88	-5.52	5.91
		L	10.84	1.37	7.92	19.27
22	Property Tax	D	-0.09	0.56	-2.97	3.21
		L	3.51	1.34	1.09	8.23
23	Sales Tax	D	-0.03	1.02	-3.55	2.92
		L	3.31	1.18	0.69	6.92
24	Individual Income Tax	D	0.19	0.29	-0.73	1.82
		L	1.65	1.09	0	4.23
25	Corporate Income Tax	D	0.01	0.14	-0.50	0.81
		L	0.46	0.25	0	1.18
26	Local Education Spending	D	-0.67	1.93	-7.51	4.38
		L	25.66	3.06	18.34	35.37
27	Higher Education Spending	D	-0.22	1.02	-4.65	2.89
		L	10.34	2.69	4.22	18.45
28	Health & Hospital Spending	D	0.25	1.13	-3.44	4.00
		L	8.14	2.88	2.46	18.37
29	Highway Spending	D	-1.21	1.74	-9.24	3.11
		L	10.96	4.04	4.27	25.59
30	Democratic Legislature	--	55.0	46.2	0	100

Number	***Name***[21]		***Mean***	***Std. Deviation***	***Minimum***	***Maximum***
31	Republican Legislature	--	24.9	39.0	0	100
32	Democratic Governor	--	55.94	40.91	0	100

APPENDIX B
SAS Program for determining *M* Best Specifications According to Both SIC and AICc Criteria

```
data new1 ;
set ying.data4;
if code ne 27; ***This eliminates Nebraska;
if code ne 31; ***This eliminates Minnesota;

proc reg outest = new1 noprint;
model dlny = dlnk dlnl dlnn s2-s20 s22-s24 s26-s48 t2-t6
   D_PCTCOLBR PCTCOLBR D_PCTMIDBR PCTMIDBR D_PCTNONBR PCTNONBR
   D_PCTFEMBR PCTFEMBR LNNBR D_DENSITY DENSBR
   D_PCTURBBR PCTURBBR D_PCTAGBR PCTAGBR D_PCTMANBR PCTMANBR
   D_PCTSERBR PCTSERBR D_PCTMIIBR PCTMIIBR D_PCTUNIBR PCTUNIBR
   D_DIVERSBR DIVERSBR D_PCTFGOBR PCTFGOBR D_PCTSLGBR PCTSLGBR
   D_FEMPPCBR FEMPPCBR D_SLEMPPBR SLEMPPBR D_PCTFREBR PCTFREBR
   D_DECENTBR DECENTBR D_GOVSPCBR GOVSPCBR D_TXBRDNBR TXBRDNBR
   D_PTXRATBR PTXRATBR D_STXRATBR STXRATBR D_IITAXRBR IITAXRBR
   D_CITAXRBR CITAXRBR D_LEDUCSBR LEDUCSBR D_HEDUCSBR HEDUCSBR
   D_HHOSPSBR HHOSPSBR D_HIWAYSBR HIWAYSBR
   LEGDEMBR LEGREPBR GOVDEMBR
    /
   include = 53
   selection = rsquare adjrsq aic sbc
   best = 1000;

**** This program sets the number of "best" specifications
**** (M) equal to 1000 ;                                           ;


data new2 ;
set new1;
SSE = ((_RMSE_)**2)*_EDF_ ;
AIC = 276*log(SSE/276) + 2*_P_ ;
SIC = 276*log(SSE/276) + (_P_ * log(276));
AICc= 276*(log(SSE/276) + ((276+ _P_ -1)/(276- _P_ -1)));

proc sort data=new2 out=new3;
   by SIC ;

data new3a;
set new3;
if _N_ le 1000;

data new3b;
set new3a;
if _N_ eq 1;
SICstar=SIC;
do i=1 to 1000;
output;
end;
keep SICstar;

data ying.bestsic46;
merge new3a new3b;
R=exp(-(1/2)*(SICstar-SIC));
if R < sqrt(10);
```

```
**** It is necessary that M (i.e., 1000) be larger than the number
**** of models that satisfy the R < sqrt(10) criterion;

proc sort data=new2 out=new4;
   by AICc ;

data new4a;
set new4;
if _N_ le 1000;

data new4b;
set new4a;
if _N_ eq 1;
AICcstar=AICc;
do i=1 to 1000;
output;
end;
keep AICcstar;

data ying.bestaic46;
merge new4a new4b;
R=exp(-(1/2)*(AICcstar-AICc));
if R < sqrt(10);

run;
```