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

W. ROBERT REED*

[Correction added on June 27, 2008, after online publication: Equation (7) was incorrectly stated as $AICc = T \cdot ln\left(\frac{SSE}{T}\right) + T \cdot \left(\frac{T+k-1}{T-k-1}\right)$. The correct statement is $AICc = T \cdot ln\left(\frac{SSE}{T}\right) + T \cdot \left(\frac{T+k}{T-k-2}\right)$.]

This study investigates U.S. state economic growth from 1970 to 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" approach of Leamer while also addressing concerns raised by Granger and Uhlig, Sala-i-Martin, 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 (1) size and structure of government and (2) taxation are robust and economically important determinants of state economic growth. (*JEL* 040, 051, H10, H20, H30, H70, R11, R58, C51)

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., Crain and Lee 1999; Fernández et al. 2001; Granger and Uhlig 1990; Hendry and Krolzig 2004; Hoover and Perez 2004; Levine and Renelt 1992; Sala-i-Martin 1997; Sala-i-Martin, Doppelhofer, and Miller 2004). While not intended as a substitute for economic theory, these approaches can be useful when the theory is sufficiently broad such that a large number of variables are potential regressors.

This study follows in this line of research by attempting to identify robust determinants of U.S. economic growth from 1970 to 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), Sala-i-Martin (1997), and others that not all specifications are equally likely to be

ABBREVIATIONS

AIC: Akaike Information Criterion
AICc: Corrected Version of Akaike Information Criterion
BMA: Bayesian Model Averaging
EBA: Extreme Bounds Analysis
SIC: Schwarz Information Criterion

*I thank Hua Chen, Elisabeth Reed, and Ying Xiao for able research assistance. I am grateful to Chris Chatfield, David Merriman, Cindy Rogers, Steven Yamarik, and participants at the 2006 NZAE conference 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).

Reed: Professor, Department of Economics, University of Canterbury, Private Bag 4800, Christchurch, 8140 New Zealand. Phone +64 3 364 2846, Fax +64 3 364 2635, E-mail bobreednz@yahoo.com

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, Doppelhofer, and Miller (2004); and Hoover and Perez (2004).

2. The following studies have highlighted the phenomenon of wide-ranging coefficient estimates across empirical specifications: Bartik 1991, McGuire 1992, Phillips and Goss 1995, Wasylenko 1997, and Crain and Lee 1999.

Economic Inquiry
(ISSN 1465-7295)
Vol. 47, No. 4, October 2009, 685–700

doi:10.1111/j.1465-7295.2008.00127.x
Online Early publication April 8, 2008

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 (1) the size and structure of government and (2) taxation are robust determinants of state economic growth.

The paper proceeds as follows: Section II develops a framework for specification of the empirical growth models. 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 FRAMEWORK FOR SPECIFICATION OF THE EMPIRICAL GROWTH MODELS

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

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

where L_t and K_t are labor and capital, Q_t is the efficiency of labor, and A_t is a time-varying parameter that represents other variables that can influence state income (e.g., human capital variables). The textbook Solow model and the augmented human capital model of Mankiw, Romer, and Weil (1992) are both special cases of Equation (1).[3]

Dividing both sides by population, N_t, produces the following per capita expression:

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

This can be expressed in log form as:

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

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:

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

It follows that:

$$(5) \qquad \ln(y_t) - \ln(y_{t-L}) \cong \alpha[\ln(k_t) - \ln(k_{t-L})] + \beta[\ln(\ell_t) - \ln(\ell_{t-L})] + (\alpha+\beta-1)[\ln(N_t) - \ln(N_{t-L})] + C_t,$$

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

The preceding analysis identifies changes in capital, employment, and population as important determinants of economic growth. However, the last term, C_t, is sufficiently general that it allows for a large number of possible explanatory variables. It encompasses many of the models that have been used to estimate U.S. state economic growth (e.g., Garofalo and Yamarik 2002; Holtz-Eakin 1993; Lee and Gordon 2005).

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.[6] The empirical task of this paper consists of identifying which of these should be included

3. The textbook Solow model is $Y_t = K_t^{\alpha}(L_t Q_t)^{1-\alpha} = Q_t^{1-\alpha} K_t^{\alpha} L_t^{1-\alpha}$. Mankiw, Romer, and Weil's augmented version of the Solow model is $Y_t = K_t^{\alpha} H_t^{\beta}(L_t Q_t)^{1-\alpha-\beta} = H_t^{\beta} Q_t^{1-\alpha-\beta} K_t^{\alpha} L_t^{1-\alpha-\beta}$.

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 (1) imposes additional restrictions on the model and (2) 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.

6. For a detailed listing of studies that use these variables, see Table 1 in the expanded version of this paper available at http://ideas.repec.org/p/cbt/econwp/06-05.html.

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

Number	Name	Description
1	Education	Percentage of population (aged 25 and above) who have completed college (*Source:* Census)
2	Working Population	Percentage of population between 20 and 64 yr of age (*Source:* Census)
3	Nonwhite	Percentage of population that is nonwhite (*Source:* Census)
4	Female	Percentage of population that is female (*Source:* Census)
5	Population	Log of total population (*Source:* Census)
6	Population Density	Population density (*Source:* Census)
7	Urban	Percentage of population living in urban areas (*Source:* Census)
8	Agriculture	Share of total earnings earned in "Farm" and "Other Agriculture" industries (*Source:* BEA)
9	Manufacturing	Share of total earnings earned in "Manufacturing" industries (*Source:* BEA)
10	Service	Share of total earnings earned in "Service" industries (*Source:* BEA)
11	Mining	Share of total earnings earned in "Mining" industries (*Source:* BEA)
12	Union	Percentage of nonagricultural wage and salary employees who are union members (*Source:* Hirsch, MacPherson, and Vroman 2001)
13	Diversity	A measure of industrial diversity, Diversity = $\sum_i \left(\frac{\text{Earnings in Industry}_i}{\text{Total Earnings}}\right)^2$ (*Source:* BEA)
14	Federal Government	Share of total earnings earned in "Federal government" (*Source:* BEA)
15	State & Local Government	Share of total earnings earned in "State and Local government" (*Source:* BEA)
16	Federal Employees	Log of federal employees per capita (*Source:* Census)
17	State & Local Employees	Log of state and local employees per capita (*Source:* Census)
18	Federal Revenue	Intergovernmental revenue received by state and local governments from the federal government as a share of personal income (*Source:* Census)
19	Decentralization	Share of total state and local direct general expenditures made by local governments (*Source:* Census)
20	Number of Governments	Number of state and local governments (*Source:* Census)
21	Tax Burden	Total state and local tax revenues as a share of personal income (*Source:* Census)
22	Property Tax	Total state and local property tax revenues as a share of personal income (*Source:* Census)
23	Sales Tax	Total state sales tax revenues as a share of personal income (*Source:* Census)
24	Individual Income Tax	Total state individual income tax revenues as a share of personal income (*Source:* Census)
25	Corporate Income Tax	Total state corporate income tax revenues as a share of personal income (*Source:* Census)
26	Local Education Spending	Total state and local spending on local schools as a share of total state and local expenditures (*Source:* Census)
27	Higher Education Spending	Total state and local spending on higher education as a share of total state and local expenditures (*Source:* Census)
28	Health & Hospital Spending	Total state and local spending on health and hospitals as a share of total state and local expenditures (*Source:* Census)
29	Highway Spending	Total state and local direct spending on highways as a share of total state and local expenditures (*Source:* Census)
30	Democratic Legislature	Percentage of years that both houses of the state legislature were controlled by Democrats (*Source:* National Conference of State Legislatures)
31	Republican Legislature	Percentage of years that both houses of the state legislature were controlled by Republicans (*Source:* National Conference of State Legislatures)
32	Democratic Governor	Percentage of years that governor was a Democrat (*Source:* National Conference of State Legislatures)

BEA, U.S. Bureau of Economic Analysis.
[a]Descriptive statistics for all variables are reported in the Appendix.

in a growth equation along with capital, employment, and population variables.

I group the variables into four major categories: (1) Population/Labor Force characteristics, (2) Economy characteristics, (3) Public Sector characteristics, and (4) Political Control characteristics. Variables included in the Population/Labor Force category include educational attainment, percentage of the population that is working aged (ages between 20 and 64), percentage of the population that is nonwhite or female, and total population. Economy characteristics include population density, degree of urbanization, the relative importance of various industries within the state, percentage of the workforce that is unionized, and a measure of industrial diversity.

Public Sector characteristics are divided into three subcategories: (1) Size and Structure variables, (2) Tax variables, and (3) 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 (1) federal and (2) state and local government sectors of the economy, measured by both 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. These tax variables should be interpreted as measuring the *net growth effect* of increasing taxes to fund general spending.[7]

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 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 $C_t = [\ln(A_t) - \ln(A_{t-L})] + \beta[\ln(Q_t) - \ln(Q_{t-L})]$ in Equation (5). One immediate issue is whether the 32 variables in Table 1 should be entered in (1) level or (2) differenced form. Because economic theory is not sufficiently specific to answer this question, this becomes an empirical issue.[8] Restricting the Political Control variables to be entered in level form,[9] and recognizing that the change in population is already included in the core specification of Equation (5) (i.e., $[\ln(N_t) - \ln(N_{t-L})]$, 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.

IV. A PROCEDURE FOR DETERMINING ROBUST VARIABLES

A. Schwarz Information Criterion and the Corrected Version of the Akaike Information Criterion

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

7. Ideally, I would have liked to measure 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, 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 1.

8. The inclusion of level variables is consistent with an endogenous growth model with scale effects. Kocherlakota and Yi (1997) find support for such a model. Examples of other studies that have included variables in both differenced and level form are Mendoza, Milesi-Ferretti, and Asea (1997); Lee and Gordon (2005); and Miller and Russek (1997).

9. 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-yr period. There is no analog to 5-yr differences that would correspond to the 5-yr differences for the other variables in Table 1.

Criterion (AICc). While I 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 two 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 1 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 criterion and the Geweke and Meese 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 1 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, 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 overfitting 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 (Hurvich and Tsai 1989; Sugiura 1978).

Monte Carlo studies of finite sample performance have demonstrated that both the SIC and the 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:

$$(6) \qquad \mathrm{SIC} = T\ln\left(\frac{\mathrm{SSE}}{T}\right) + k\ln(T).$$

$$(7) \qquad AICc = T\cdot ln\left(\frac{SSE}{T}\right) + T\cdot\left(\frac{T+k}{T-k-2}\right),$$

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 differenced 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. This procedure does not actually estimate all possible regression specifications. Instead, it relies on the "leaps and bounds" algorithm developed by Furnival and Wilson (1974) to identify the specifications with the highest R^2 values among all possible specifications having the same number of regressors. It is straightforward to use the output generated by this SAS program to calculate a ranked ordering of the M best specifications across all possible variable combinations—for any predetermined value of M—according to either the SIC or the AICc.[10] The corresponding SAS program is easy to implement and remarkably efficient

10. The general principle of the "leaps and bounds algorithm" can be illustrated in the context of a "regression tree": Consider the case of five "doubtful" variables, X_1 through X_5. At the top of the regression tree are models with only one 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—noting that highest R^2 equates with lowest SIC/AICc values when the number of regressors is held constant—and then globally rank the best specifications across all subsets.

in computational requirements. It required about an hour to run using a standard desktop computer.[11]

B. EBA and Bayesian Model Averaging

My approach uses insights from both EBA (Leamer 1985) 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, calculating the confidence interval for each β_1 estimate. The highest and lowest values over all these confidence intervals define the "extreme" upper and lower bounds. If these bounds do not overlap zero (i.e., if they are same signed), then the variable X_1 is said to be robust (cf. Crain and Lee 1999; Levine and Renelt 1992).

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.[12] 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 based on information criteria. 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."[13]

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 $y_t = \beta_o + \sum_{k=1}^{K} \beta_k X_{k,t} + \varepsilon_t, \quad t = 1, \ldots, T$, where some β_k may equal 0 (but not β_1); and $\mathrm{Cov}(X_j, X_k) = 0$ 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:

$$(8) \qquad P(M_j|y) = \frac{P(M_j) T^{-k_j/2} \mathrm{SSE}_j^{-T/2}}{\sum_{i=1}^{2^K} P(M_i) T^{-k_i/2} \mathrm{SSE}_i^{-T/2}},$$

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:

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

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 ensure 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}

11. A copy of the SAS program used in this analysis is available from the author.

12. For example, suppose the true model is $Y_t = \alpha_0 + \alpha_1 Z_t + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$. Suppose further that X_{1t} and X_{2t} are both positively correlated and that $\beta_1 > 0$ and $\beta_2 < 0$. Last, suppose we now estimate the following three equations: (i) $Y_t = \alpha_0 + \alpha_1 Z_t + \beta_1 X_{1t} + \beta_2 X_{2t} + \varepsilon_t$, (ii) $Y_t = \alpha_0 + \alpha_1 Z_t + \beta_1 X_{1t} + \varepsilon_t$, and (iii) $Y_t = \alpha_0 + \alpha_1 Z_t + \beta_2 X_{2t} + \varepsilon_t$. It is possible that β_1 and β_2 could both be significant in Equation (i) but insignificant in Equations (ii) and (iii). EBA would classify these variables as not being robust, despite the fact that both variables are in the true model. The reason for this anomaly is that the latter two equations are "bad" specifications. EBA gives equal weight to "good" and "bad" specifications.

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

models, X_1 is excluded, and the BMA approach sets $\hat{\beta}_{1,j} = 0$.[14] It follows that $E(\beta_1|y) < \beta_1$ even if $\hat{\beta}_{1,j} = \beta_1$ in every specification in which it appears. In other words, the BMA-based expectation is biased toward zero. This follows directly from the fact that BMA "estimates" the value of β_1 to be 0 in all specifications in which X_1 is not included.[15]

C. 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 use information criteria to restrict the set of model specifications. I follow a procedure developed by Poskitt and Tremayne (PT; 1987) to identify two categories of models: (1) "reasonable" models and (2) others. Only reasonable models are considered for EBA.

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 I^A be the corresponding value for an alternative model. The posterior odds ratio is defined as:

$$(10) \qquad \Re = \exp\left[-\frac{1}{2}(I^* - I^A)\right].$$

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 reasonable EBA by Granger and Uhlig (1990), 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 to 1999.[16] I decided on this particular time period because a longer time frame would have required me to omit many variables of interest. The respective 30 yr of data were grouped into six 5-yr periods (1970–1974, 1975–1979,..., 1995–1999). Data

14. Compare Equations (8) and (9) with Equations (7) and (8) in Sala-i-Martin, Doppelhofer, and Miller (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."

15. 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, Sala-i-Martin, Doppelhofer, and Miller (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.

16. Alaska and Hawaii were omitted, as is usual for 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.

for most of these variables were collected from original data sources.[17]

Using over 5-yr rather than annual data offers several advantages: it (1) reduces the impact of "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 are defined over calendar years; state fiscal data are 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 these reasons, the use of 5-yr interval data should entail fewer estimation problems. Following Equation (5), the general specification for the empirical models is:[18,19]

$$
\begin{aligned}
(11)\quad \mathrm{DLNY}_t = [&\beta_0 + \beta_1 \mathrm{DLNK}_t + \beta_2 \mathrm{DLNL}_t \\
&+ \beta_3 \mathrm{DLNN}_t \\
&+ \text{State Fixed Effects} \\
&+ \text{Time Fixed Effects}] \\
&+ \sum_{l} \lambda_l X_{l,t-4} \\
&+ \sum_{d} \delta_d (X_{d,t} - X_{d,t-4}) \\
&+ \sum_{p} \pi_p \bar{X}_{p,t} + \varepsilon_t,
\end{aligned}
$$

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 5-yr period ("level" form); $(X_{d,t} - X_{d,t-4})$ is the change in the explanatory variable over the 5-yr period ("difference" form); and $\bar{X}_{p,t} = \frac{\bar{X}_{p,t-1} + \bar{X}_{p,t-2} + \bar{X}_{p,t-3} + \bar{X}_{p,t-4} + \bar{X}_{p,t-5}}{5}$ is the 5-yr average over the period ($t-5$ to $t-1$) for the Political Control variables Democratic Legislature, Republican Legislature, and Democratic Governor.[20]

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 ($X_{l,t-4}$, $(X_{dt} - X_{d,t-4})$, and $\bar{X}_{p,t}$), since the theory is nonspecific about which variables belong in C_t (cf. Equation (5)).

VI. EMPIRICAL RESULTS

A. Robust Determinants of State Economic Growth

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: (1) I analyze two "portfolios of models" (one for SIC and one for AICc) and (2) 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 27 different models, the Best SIC specification, and 26 "close competitors" as defined by the $\Re < \sqrt{10}$ criterion.[21] The results from analyzing this portfolio of models are reported in Table 2, Panel A. Variables are ranked in descending order of number of appearances

17. The Appendix presents statistical descriptions of all the variables used in this study.

18. Note that because (1) the dependent variable is expressed in logs and (2) 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.

19. 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% significance level in all of the top model specifications.

20. 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. Gilligan and Matsusaka 1995; Poterba 1994; Reed 2006), and it is the latter that is assumed to matter for economic growth.

21. The best variable specification according to the SIC is DLNK, DLNL, DLNN, State + Time Fixed Effects, Education-L, Working Population-D, Female-D, Agriculture-D, Agriculture-L, Service-L, Mining-D, Mining-L, Federal Government-D, Federal Employees-L, Federal Revenue-L, Decentralization-D, Tax Burden-D, Tax Burden-L, Sales Tax-L, Corporate Income Tax-L.

TABLE 2
EBA for Portfolio of Top SIC and AICc Models

				Range of Coefficient Estimates			Range of t Ratios		
Number (%)	Robust	Variable	D/L	Low	Mean	High	Low	Mean	High
A. SIC models									
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
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
B. AICc model									
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
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

continued

TABLE 2
Continued

Number (%)	Robust	Variable	D/L	Range of Coefficient Estimates Low	Mean	High	Range of t Ratios Low	Mean	High
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
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 criteria for determining which variables are robust are described in Section IV.

within the portfolio. Robust variables are identified with an "R." A total of 18 different variables are analyzed, as shown in Table 2, Panel A. Some, like Education, appear in all 27 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 (1) the set of variables that appears in at least 50% of the models in the SIC portfolio and (2) the set of variables having a range of t ratios all same signed and larger than 2.0.[22]

Table 2, Panel B, reports that 57 models are included in the AICc portfolio.[23] 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 3 collects the robust variables from these EBAs and reports them, along with a "mean estimated effect" calculated as the simple average of the respective means from Table 2, Panels A and B. To interpret the respective sizes of these effects, recall that the dependent variable is the 5-yr growth rate in state per capita personal income. For my sample of 30 yr (six 5-yr time periods) and 46 states (yielding 276 observations), the mean growth rate is 27.01%. Thus, a 1 percentage point increase in the 5-yr growth rate equates approximately to a 3.7% increase in growth.

Given the underlying theoretical model of Equation (5), the variables of Table 3 should be related to the term, $C_t = [\ln(A_t) - \ln(A_{t-L})] + \beta[\ln(Q_t) - \ln(Q_{t-L})]$. Since A_t and Q_t represent production function parameters, theory suggests that these variables affect the rate

22. t Statistics are calculated in the classic fashion.

23. The best variable specification according to the AICc is DLNK, DLNL, DLNN, State + Time Fixed Effects, Education-D, Education-L, Working Population-D, Female-D, Population-L, Agriculture-D, Agriculture-L, Mining-D, Mining-L, Federal Government-D, Federal Employees-L, Federal Revenue-L, Decentralization-D, Tax Burden-D, Tax Burden-L, Sales Tax-L, Corporate Income Tax-L, Democratic Legislature.

TABLE 3
Robust Variables and Mean Estimated Effects

	Variable		
Category	**Number**	**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 Table 2, Panels A and B.

of invention and adoption of new technologies that transform the production function over time. This includes effects on resource allocation.

Differenced variables are indicated by "D" and represent changes in that variable during the 5-yr period. Level variables are indicated by "L" and represent the value of that variable at the beginning of the 5-yr period. A variable that appears in both differenced and level form has both an immediate and a lagged effect. The differenced form indicates the immediate effect since changes during the 5-yr 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-yr growth period.

Table 3 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 1 percentage point increase in the percentage of the population that is college educated at the beginning of a 5-yr period is associated with a 0.97 percentage point increase in that state's subsequent 5-yr growth rate. This effect is relatively small in economic terms, given that the average 5-yr growth rate is 27.01%.

The differenced form of Working Population is also identified as a robust variable. The corresponding estimated positive effect indicates that a 1 percentage point increase in the share of the population that is aged 20–64 yr during a given 5-yr 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 3 identifies two economy characteristic variables: Agriculture and Mining. The coefficient for Agriculture is positive in both level and differenced 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 differenced and level forms, are each negative. This is consistent with research that finds that the mining industry contributes negligibly, or even negatively, to aggregate total factor productivity growth (cf. Jorgenson and Stiroh 2000).

Table 3 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 1 percentage point increase in this variable—corresponding to roughly a 15% increase in the size of the federal government sector over a 5-yr 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-yr growth rate of that state by 4.48%. While not robust, it is interesting to note that I estimate similar-sized effects for both State & Local Government and State & Local Employees (cf. Table 2, Panels A and B).

The variable Federal Revenue measures the size of federal aid to states. The sample mean of Federal Revenue is 3.90. A 1 percentage point increase in this variable would represent approximately a 25% increase in federal aid. The mean estimated effect reported in Table 3 for this variable indicates that an increase of this size would raise a state's subsequent 5-yr 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 1 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-yr 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 differenced forms of the variable are identified as robust determinants indicates that the effect of taxes is both immediate and persistent. A 1 percentage point increase in Tax Burden over a 5-yr 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-yr periods. As a gauge of size, a 1 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 a long-lived effect of taxes. Second, the estimated effects represent the net effect of taxes *and* spending. Previous studies, following Helms (1985), commonly estimated "government budget constraint" specifications, so that most categories of public expenditures were held constant.[24] 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 imply significantly larger negative tax effects.

In contrast, the estimated coefficients for both Sales Tax and Corporate Income Tax are each positive. Note that a 1 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

24. Following Helms (1985), most studies use welfare transfers as the omitted expenditure category, so that estimated tax effects measure the impact of tax-financed welfare expenditures.

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.

It should be noted that choosing a different interval length can produce different robust variables. When I repeated the analysis using 10-yr intervals, some of the robust variables from Table 3 continued to be chosen, but others were not.[25] This is not particularly surprising, given that previous research has demonstrated that estimates of economic growth equations for U.S. states can differ substantially when the interval length is changed (Reed 2008). Which interval length is most appropriate remains an unsettled research question.

B. *Comparison with Crain and Lee (1999)*

The only other study that searches for robust determinants of state economic growth is Crain and Lee (1999).[26] Crain and Lee (CL) implement the EBA approach of Levine and Renelt (1992) using annual data on U.S. states from 1977 to 1992.[27] A comparison of their robust variables and my robust variables reveals several differences.[28] Unlike CL, I find that Agriculture and Mining are robust determinants of economic growth and that Diversity and Service are not. Further, their "core variable" for Education is always insignificant, and sometimes negative, whereas I find that Education is robust and positively associated with economic growth.

There are, however, a number of similarities: while we measure it differently, we both find size of government variables to be robust and negatively associated with economic growth.[29] We both find that Decentralization is negatively associated with economic growth. Additionally, I find that Tax Burden is negatively associated with state economic growth, while CL obtain a similar finding using a measure that includes all state and local revenues, not just taxes.[30]

VII. CONCLUSIONS

This study examines the determinants of U.S. state economic growth from 1970 to 1999. It considers a large number of potential explanatory variables, including Population/Labor Force characteristics, Economy characteristics, Public Sector (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.

25. The robust variables selected from the 10-yr interval analysis, along with the sign of their mean estimated effects, are Decentralization-L(+), Democratic Governor(+), Diversity-D(−), Education-D(+), Education-L(+), Female-D(−), Female-L(+), Federal Employees-L(−), Federal Government-L(+), Federal Revenue-D(−), Highway Spending-D(+), Individual Income Tax-D(+), Manufacturing-L(−), Nonwhite-L(+), Number of Governments-D(+), Number of Governments-L(+), Population-L(+), State & Local Employees-D(+), Tax Burden-D(−), and Working Population-D(+). I thank an anonymous referee for suggesting that this analysis also be applied to 10-yr interval data.

26. Higgins, Young, and Levy (2008) also search for robust determinants of U.S. income growth, but they conduct their analysis at the county level.

27. Their sample employs data from 48 states, versus the 46 states used in my study. Further, while there is much overlap, our sets of variables differ both in kind (e.g., I include Political Control variables, they include Pressure Groups variables) and in form (e.g., their government expenditure variable is expressed as a share of state income, my government expenditure variables are broken down by categories and expressed as shares of total government expenditures).

28. This discussion is based on a comparison of their table II and my Table 3.

29. CL measure size of government by the combined earnings of local, state, and federal government. I have separate variables for the federal government (Federal Government, Federal Employees) and for state government (State & Local Government, State & Local Employees). My federal measures are both robust with estimated negative impact, while the state and local variables are consistently negative but not always statistically significant and hence not robust.

30. There are other differences in our studies, but these likely stem from the fact that we use different variable sets. For example, I find that Female is a robust determinant of economic growth, whereas CL do not include this variable. CL include a measure for political Pressure Groups ("Bus Assoc Revenue Share of Income") that I do not. CL include a measure of state and local expenditures as a share of state income ("Expenditure Share of Income"). In contrast, I do not include a separate variable for the size of state and local expenditures since I want Tax Burden to pick up the net effect of tax-financed expenditures. Last, I include separate components of state tax revenues (e.g., sales taxes, corporate income taxes) and a measure of federal aid to the states (Federal Revenue), while CL do not.

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 EBA of Granger and Uhlig (1990). Unlike their study, however, I use information criteria (the SIC and AICc) to choose "portfolios of reasonable models," as suggested by Poskitt and Tremayne (1987). I then perform conventional EBA within these portfolios. An advantage of my approach is that (1) it is a straightforward extension of a standard SAS program, (2) it requires relatively little computational time, and (3) its simplicity assures that different researchers using the same procedure will obtain identical results.

My analysis identifies 12 robust determinants of U.S. economic growth over the 30-yr period from 1970 to 1999. Among these are (1) college attainment within the population, (2) share of the population that is "working age," and (3) population gender share, and the size of the (4) agricultural and (5) 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 (6 and 7) the size of the federal sector within a state, (8) federal aid, (9) decentralization, and (10 through 12) 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.

APPENDIX

STATISTICAL SUMMARY OF DATA

Number	Name[a]	D/L	Mean	SD	Minimum	Maximum
Dependent variable	DLNY		27.01	10.30	6.49	64.40
—	DLNK		7.40	7.76	−26.92	55.43
—	DLNL		4.61	4.01	−7.22	14.98
—	DLNN		4.72	4.55	−8.63	21.45
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

continued

APPENDIX
Continued

Number	Name[a]	D/L	Mean	SD	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
31	Republican Legislature	—	24.9	39.0	0	100
32	Democratic Governor	—	55.94	40.91	0	100

[a]The numbered variables are described in Table 1. DLNY is derived from BEA Personal Income Data. DLNK measures the change in net private capital stock and comes from Steve Yamarik (Garofalo and Yamarik 2002). DLNL is computed from BEA data on the total number of full- and part-time jobs. DLNN comes from midyear population estimates provided by the Census.

REFERENCES

Barro, R. J. "Government Spending in a Simple Model of Endogenous Growth." *Journal of Political Economy*, 98, 1990, S103–25.

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

Burnham, K. P., and D. R. Anderson. *Model Selection and Multimodel Inference*. 2nd ed. New York: Springer-Verlag, Inc., 2002.

Crain, W. M., and K. J. Lee. "Economic Growth Regressions for the American States: A Sensitivity Analysis." *Economic Inquiry*, 37, 1999, 242–57.

Evenson, R. E., and W. 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, C., E. Ley, and M. F. J. Steel. "Model Uncertainty in Cross-Country Growth Regressions." *Journal of Applied Econometrics*, 16, 2001, 563–76.

Furnival, G. M., and R. W. J. Wilson. "Regression by Leaps and Bounds." *Technometrics*, 16, 1974, 499–511.

Garofalo, G. A., and S. Yamarik. "Regional Convergence: Evidence from a New State-by-State Capital Stock Series." *The Review of Economics and Statistics*, 84, 2002, 315–23.

Gilligan, T. W., and J. G. Matsusaka. "Deviations from Constituent Interests: The Role of Legislative Structure and Political Parties in the States." *Economic Inquiry*, 33, 1995, 383–401.

Granger, C. W. J., and H. F. Uhlig. "Reasonable Extreme-Bounds Analysis." *Journal of Econometrics*, 44, 1990, 159–70.

Grier, K. B., and G. Tullock. "An Empirical Analysis of Cross-national Economic Growth, 1951-1980." *Journal of Monetary Economics*, 24, 1989, 169–328.

Helms, J. L. "The Effect of State and Local Taxes on Economic Growth: A Time Series-Cross Section Approach." *The Review of Economics and Statistics*, 67, 1985, 574–82.

Hendry, D. F., and H.-M. Krolzig. "We Ran One Regression." *Oxford Bulletin of Economics and Statistics*, 66, 2004, 799–810.

Higgins, M. J., A. T. Young, and D. Levy. Forthcoming. "Robust Correlates of County-Level Growth in the U.S." *Applied Economics Letters*, 2008.

Hirsch, B. T., D. A. MacPherson, and W. G. Vroman. "Estimates of Union Density by State." *Monthly Labor Review*, 124, 2001, 51–55.

Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky. "Bayesian Model Averaging: A Tutorial." *Statistical Science*, 14, 1999, 382–417.

Holtz-Eakin, D. "Solow and the States: Capital Accumulation, Productivity, and Economic Growth." *National Tax Journal*, 46, 1993, 425–40.

Hoover, K. D., and S. J. Perez. "Truth and Robustness in Cross-Country Growth Regressions." *Oxford Bulletin of Economics and Statistics*, 66, 2004, 765–98.

Hurvich, C. M., and C. L. Tsai. "Regression and Time Series Model Selection in Small Samples." *Biometrika*, 76, 1989, 297–307.

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

Jorgenson, D. W., and K. J. Stiroh. "Raising the Speed Limit: U.S. Economic Growth in the Information Age." *Brookings Papers on Economic Activity*, 1, 2000, 125–211.

Kocherlakota, N. R., and K.-M. Yi. "Is There Endogenous Long-Run Growth? Evidence from the United States and the United Kingdom." *Journal of Money, Credit and Banking*, 29, 1997, 235–62.

Leamer, E. E. "Sensitivity Analyses Would Help." *American Economic Review*, 75, 1985, 308–13.

Lee, Y., and R. H. Gordon. "Tax Structure and Economic Growth." *Journal of Public Economics*, 89, 2005, 1027–43.

Levine, R., and D. Renelt. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review*, 82, 1992, 942–63.

Mankiw, N. G., D. Romer, and D. N. Weil. *The Quarterly Journal of Economics*, 107, 1992, 407–37.

McGuire, T. "Review of 'Who Benefits From State and Local Economic Development Policies.'" *National Tax Journal*, 45, 1992, 457–9.

McQuarrie, A. D. R., and C.-L. Tsai. *Regression and Time Series Model Selection*. Singapore: World Scientific Publishing Co. Pte. Ltd., 1998.

Mendoza, E. G., G. M. Milesi-Ferretti, and P. Asea. "On the Ineffectiveness of Tax Policy in Altering Long-Run Growth: Harberger's Superneutrality Conjecture." *Journal of Public Economics*, 66, 1997, 99–126.

Miller, S. M., and F. S. Russek. "Fiscal Structures and Economic Growth: International Evidence." *Economic Inquiry*, 35, 1997, 603–13.

Phillips, J., and E. Goss. "The Effect of State and Local Taxes on Economic Development: A Meta-Analysis." *Southern Economic Journal*, 62, 1995, 320–33.

Poskitt, D. S., and A. R. Tremayne. "Determining a Portfolio of Time Series Models." *Biometrika*, 74, 1987, 125–37.

Poterba, J. M. "State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Politics." *Journal of Political Economy*, 102, 1994, 799–821.

Reed, W. R. "Democrats, Republicans, and Taxes: Evidence that Political Parties Matter." *Journal of Public Economics*, 9, 2006, 725–50.

———. Forthcoming. "The Robust Relationship between Taxes and U.S. State Income Growth." *National Tax Journal*, 61, 2008, 57–80.

Sala-i-Martin, X. X. "I Just Ran Two Million Regressions." *American Economic Review*, 87, 1997, 178–83.

Sala-i-Martin, X., G. Doppelhofer, and R. I. Miller. "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach." *American Economic Review*, 94, 2004, 813–35.

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

Wasylenko, M. "Taxation and Economic Development: The State of the Economic Literature." *New England Economic Review*, March/April 1997, 37–52.

Xie, D., H.-fu Zou, and H. Davoodi. "Fiscal Decentralization and Economic Growth in the United States." *Journal of Urban Economics*, 45, 1999, 228–39.

Zellner, A. "Comments on Time Series Analysis and Causal Concepts in Business Cycle Research," in *New Methods in Business Cycle Research*, edited by C. A. Sims. Minneapolis: Federal Reserve Bank of Minneapolis, 1977, 167–73.