Modeling Economic Growth with Unpredictable Shocks: A State-Level Application for 1960-90

Stephan J. Goetz and Richard C. Ready*

Abstract

A Barro-type economic growth model is estimated for the 50 states in the U.S. using data for three decades beginning in 1960. Frontier estimation techniques are used to test for the presence of state-specific shocks to economic growth that are independent of the usual, normally-distributed random errors. We find that large, positive shocks to growth occurred during the period 1960–90. Our results indicate that the error term structure assumed under OLS may not be appropriate for modeling economic growth.

Key Words: economic growth, frontier estimation, shocks, U.S. states

The question of what causes economies to grow has occupied economists since the beginnings of the profession. In the early 1990s, interest in the subject escalated as researchers empirically verified testable hypotheses derived from dynamic optimization models of economic growth (Barro, Barro and Sala-i-Martin, Becker et al., Jorgenson and Yun, King and Rebelo, Mankiw et al., Romer, and Nissan). Applications have been carried out at the level of countries, regions and states of the U.S., using time series, cross-sectional and panel data.

These investigations typically involve estimation of a growth equation, in which a region's economic growth during a specified time period–usually measured as change in per capita income–is an estimated function of the conditions that exist in the region at the beginning of the period over which growth is calculated. Conditions or factors that affect economic growth typically include human capital (education of the residents of the region), public and private investment in physical infrastructure and other productivity-enhancing inputs, as well as political stability and the extent of government intervention in the economy. Several studies have shown that regions with lower initial income levels experience faster income growth than regions with higher initial income levels. A consequence of this tendency is that, over time, income levels tend to converge across regions (Barro).

In addition to the predictable influence that regional conditions exert on economic growth, growth is affected by stochastic processes (shocks) that are difficult or impossible to predict. For example, a large order for goods placed with an important employer in a region increases income growth over the period above what might otherwise be expected. Conversely, poor weather can depress income growth in an agriculturally-dependent region. The implicit assumption made in past investigations is that these positive and negative shocks to growth are numerous and that no one shock dominates, so that the total impact on growth

---

*S. J. Goetz is an associate professor and R. C. Ready an assistant professor of agricultural economics. Both are at the University of Kentucky. The authors gratefully acknowledge valuable comments and suggestions by anonymous *journal* reviewers.

J. Agr. and Applied Econ. 27 (2), December, 1995: 400-408

can be treated as a normally-distributed error term in the regression.

For a process such as economic growth, the assumption of small, independent shocks is questionable. Certainly, the negative influence of a natural disaster such as a drought or a flood on income growth would be large enough to dominate other shocks. Conversely, increased demand for a particular natural resource such as oil could have a dominating positive effect on income growth in regions that are rich in that resource. The econometric implication of such large shocks is that the error term of the growth equation regression may not be normally distributed, but instead could be skewed. The direction of the skew would depend on whether the economic growth process is dominated by positive shocks or by negative shocks. A practical set of questions here is, how large are these shocks, in which direction do they tend to operate, do they bias the coefficients in a model estimated by OLS, and is it therefore necessary to correct for them?

In this paper we distinguish between the kinds of shocks discussed above, which influence growth in only one region (such as a state), and macro-level shocks that affect economic growth in all regions. The influence of economy-wide recessions and recoveries can be controlled for when estimating a growth equation by including time period-specific shifters. We focus here on state-specific shocks, which vary not only from time period to time period, but also from state to state.

To allow for the possibility of large, unpredictable shocks to economic growth, we use a stochastic frontier model, which incorporates two types of errors: the usual normally-distributed error component and a second, asymmetric component that is either strictly positive (to model positive region-specific growth shocks) or strictly negative (to model negative region-specific growth shocks). We specify a Barro-type economic growth model for states in the U.S. using time series and cross-sectional data for all states over the three decades beginning in 1960. The model allows estimation of the net growth shock that each state experienced in each decade. Examination of these shocks may help us understand their relative importance in determining state-level growth.

The paper is organized as follows. Section I presents the stochastic model used here to allow for asymmetric errors in a growth equation regression. Section II describes the specific measure of economic growth and the explanatory variables used in the regression. In Section III, results of the estimated growth model are presented and compared with those of an OLS model. Estimates of the net growth shock experienced by individual states are also discussed.

I. Stochastic Framework

Consistent with earlier studies, we posit that the rate of economic growth in a geographic region (in our case, a state) depends on a vector of variables describing the conditions that exist in that region. Specifically, the rate of economic growth, y_{it}, in state $i \in \{50 \text{ states}\}$ over each decade beginning in year $t \in \{1960, 1970, 1980\}$ follows the relationship,

$$y_{it} = \theta(\mathbf{X}_{it};\beta)\exp(\varepsilon_{it}) \tag{1}$$

where $\mathbf{X}_{it}$ is a vector of k variables affecting economic growth in state i at the start of each decade t, β is the corresponding parameter vector, and ε_{it} is an error term unique to each state/decade combination. We further assume that eqn. (1) has a Cobb-Douglas form, so that

$$\ln(y_{it}) = \sum_{j=1}^{k} \beta_j \ln(X_{ijt}) + \varepsilon_{it} \tag{2}$$

The typical assumption about the error term, ε_{it}, is that it is normally distributed with zero mean. The parameter vector β can then be estimated from (2) using an OLS regression.

The assumption of normality is appropriate if ε_{it} is the sum of a large number of small, independent and unobservable stimuli and impediments (shocks) to economic growth. If, however, large individual shocks to growth occur that dominate the sum, then the error term in (1) will be non-normal. We use a frontier model consistent with that developed by Aigner et al. for the estimation. In this model, the error term consists of two components, and is constructed according to either

$$\varepsilon_{it} = v_{it} + u_{it} \tag{3a}$$

or

$$\varepsilon_{it} = v_{it} - u_{it}, \tag{3b}$$

where $v_{it} \sim iid\ N(0,\sigma_v^2)$ and u_{it} is a non-negative truncated random variable. We assume that u_{it} is distributed according to a half normal distribution with variance σ_u^2. The term v_{it} represents the usual, normally-distributed error term incorporating small, individual shocks to growth. The half-normal term u_{it} represents large shocks to growth. If these shocks are on balance positive, then (3*a*) will best model the observed error structure. If the shocks tend to be negative, then (3*b*) will best model the observed errors. The question of which structure is appropriate can be answered empirically.

This type of model is typically called a stochastic frontier model. For each state, the value of $\theta(\mathbf{X}_{it};\beta)$ is the deterministic frontier, and represents the rate of growth that would be expected absent unobservable shocks, conditional on the realization of vector X_{it} in each decade and state. Adding the normal error term yields the quantity $\theta(\mathbf{X}_{it};\beta)\exp(v_{it})$, which is the state's stochastic frontier. Growth above this level requires a positive shock u_{it}, yielding $\theta(\mathbf{X}_{it};\beta)\exp(v_{it}+u_{it})$. Growth below this level is caused by a negative shock $(-u_{it})$.

The term "frontier" comes from the production literature (Battese surveys applications of the frontier method in agricultural economics). When modeling a production function, the frontier represents the largest possible output that can be produced with a given amount of inputs. Firms may fall short of that frontier due to inefficiency. Such a model is consistent with the error structure (3*b*). When modeling a cost function, the frontier represents the minimum possible cost of producing a given amount of output. Here, inefficiency will cause costs to be above the frontier, consistent with (3*a*).

Other researchers have modeled growth by referring to a "growth production function," and interpreted right-hand-side variables as factors of production or inputs. This causes problems when one, for example, discusses voting preferences. The growth function modeled here has neither a production nor a cost function interpretation. The reason for this is that even though regressors such as private and public investment constitute *bona fide* "inputs" into the growth process, it is difficult to make the same argument for a regressor such as voting preferences.

A number of statistical packages are available to estimate the parameters β, σ_u^2, and σ_v^2 using maximum likelihood techniques, or the transformed parameters $\lambda = \sigma_u/\sigma_v$ and $\sigma = (\sigma_v^2+\sigma_u^2)^{½}$, which are easier to estimate. We used LIMDEP© for the estimation (Greene). A test of the null hypothesis λ=0 indicates whether the conventional OLS regression approach is appropriate.

After estimating the frontier model, it would be useful to identify those state/decade combinations that experienced very large shocks. Unfortunately, it is not possible to observe u_{it} directly. However, we do observe the combined error term, $\varepsilon_{it} = v_{it} \pm u_{it}$. Jondrow et al. show that the half normal component of the error term can be estimated from the conditional expectation

$$E[u_{it}|\varepsilon_{it}] = \frac{\sigma\lambda}{(1+\lambda^2)}\left[\frac{\phi(\varepsilon_{it}\lambda/\sigma)}{\Phi(\varepsilon_{it}\lambda/\sigma)} - \varepsilon_{it}\lambda/\sigma\right], \tag{4}$$

where ϕ is the *pdf* and Φ the *cdf* of the standard normal distribution. This conditional expectation is our best guess of the shock u_{it} given the observed value of $\varepsilon_{it} = v_{it} \pm u_{it}$. By examining which states had particularly strong or weak growth in each decade, we may be better able to understand the processes that affect regional economic growth.

II. Empirical Specification

We follow the conventional specification used in the economic growth literature, where economic growth is measured as the rate of real per capita income growth. Regressors in earlier studies have included starting income levels, school enrollment and other conditioning variables, which depend on the geographic units for which the model is estimated. In models estimated for individual countries, conditioning variables may include government investment or consumption, trade, inflation, political stability and other variables (e.g., Barro and Sala-i-Martin; Levine and Renelt). In our case the growth model is implemented using state-level data for the three decades starting in 1960, 1970 and 1980, which necessarily limits the variables available as regressors, relative to those used, for example, in cross-country studies.

The dependent variable is defined as $ln(Y_{it+10}/Y_{it})$, where Y_{it} is real (1982-84=100) average per capita income in state i at time t. Regressors ($\boldsymbol{X}_{it}$) include real starting income, $ln(Y_{it})$, and a vector of independent factors affecting economic growth ($\boldsymbol{Z}_{it}$), including real private and public investment spending per capita in each state; political preferences as reflected in the percent of elected representatives in the Upper and Lower houses of each state who are Democrats; and interaction terms for each of these variables.

We also include human capital stocks (percent of adults with 12 or more years of education), and the percent of earnings from different sectors, which are aggregated into government, construction, manufacturing, general services and natural resources. These variables are measured in percentage terms and entered in their natural form (without logs). Just as in a regular OLS regression, the sector-specific variables serve as controls, allowing the deterministic frontier to vary across states due to differences in the composition of the state's economy; each variable is also of interest in terms of its independent effect on the rate of economic growth. In addition, we include regional indicator variables for the Midwest, Northeast, South and West, and an indicator variable for each decade to allow for economy-wide, national recessions and growth periods. Summary statistics for the variables are reported in table 1.

This yields the empirical specification:

$$y_{it} = ln(Y_{it+10}/Y_{it}) = \beta_{k-1}Z_{it} + \alpha ln Y_{it} + (\nu_{it} \pm u_{it}) \quad (5)$$

Table 1. Summary Statistics for Variables Used in the Regression

Variable	1960s Mean	1960s StDev	1970s Mean	1970s StDev	1980s Mean	1980s StDev
Income growth (y_{t+10}/y_t)[a]	1.50	0.12	1.22	0.07	1.31	0.11
Personal income ($/capita)[a]	6,196	1,258	9,163	1,499	11,093	1,588
Private investment ($/capita)[b]	1,653	1,011	2,171	1,014	2,784	1,152
Public investment ($/capita)[c]	2,776	779	3,591	1,312	3,824	1,885
Voting record (% democrat)[d]	63.7	23.9	59.6	20.5	59.2	20.4
Human capital stock (%)[d]	41.7	7.3	53.1	8.1	67.5	7.6
Government earnings (%)[e]	17.6	6.6	19.9	6.3	17.3	4.3
Natural resources earnings (%)[e]	2.9	3.2	2.5	2.9	3.7	4.5
Construction earnings (%)[e]	7.6	2.1	7.5	1.4	7.7	2.0
Manufacturing earnings (%)[e]	25.8	11.3	23.9	10.2	23.0	9.5
Services earnings (%)[e]	46.2	6.2	46.1	5.7	48.4	5.8

Data sources (authors' calculations and additional explanations are presented in the appendix):

a. U.S. Department of Commerce, Bureau of Economic Analysis, *State Personal Income, 1929-87: Estimates and a Statement of Sources and Methods.* Washington, DC: U.S. Government Printing Office, July 1989.

b. U.S. Bureau of the Census, *Census of Manufactures.* Washington, DC: U.S. Government Printing Office, various years, and U.S. Department of Commerce, Economics and Statistics Administration, *Annual Survey of Manufacturers, Geographic Area Statistics.* Washington, D.C.: U.S. Government Printing Office, various years.

c. U.S. Bureau of the Census, *Census of Government.* Washington, DC: U.S. Government Printing Office, various years, and U.S. Department of Commerce, *Government(al) Finances.* Washington, D.C.: U.S. Government Printing Office, various years.

d. U.S. Bureau of the Census, *Statistical Abstract of the United States*, Washington, D.C., various years and editions.

e. U.S. Dept. of Commerce, Bureau of Economic Analysis, *State Personal Income, 1929-87: Estimates and a Statement of Sources and Methods.* Washington, DC: U.S. Government Printing Office, July 1989.

As discussed earlier, we depart from the conventional specification by using the frontier framework to model the error term so as to allow for large, asymmetric shocks to income growth, conditional on the starting conditions. An estimate of $\alpha < 0$ is consistent with the income convergence hypothesis, whereby richer states grow less rapidly in relative terms than poorer states. Based on findings in earlier studies, we expect education and investment to have positive effects on the growth rate, *cet. par.*

III. Results and Discussion

Maximum likelihood parameter estimates for the frontier growth model of eqn. (5) are reported in table 2 along with corresponding OLS estimates. The OLS residuals have a positive skew value of 0.358 (significant at the 10% level in a two-tailed test, with a critical value of 0.321), so that the normality assumption of the Gauss-Markov theorem is violated. For this particular data set, the systematic shocks to growth are on balance positive.

Table 2. Maximum Likelihood State-Level Frontier Economic Growth Model Estimates, 1960-90
Dependent variable = $ln[y_{it+10}/y_{it}]$

Variable	Parameter estimates (t-value) OLS	MLE
Constant	−5.176**	−4.656**
	(2.47)	(1.97)
ln[income/capita]	−0.212***	−0.222***
	(4.70)	(5.15)
ln[private investment/capita]	0.761***	0.700**
	(2.96)	(2.32)
ln[public investment/capita]	0.764***	0.692**
	(3.07)	(2.45)
ln[public invest./capita]×*ln*[private invest./capita]	−0.084***	−0.075**
	(2.72)	(2.01)
Voting record (% democrat)	0.022***	0.023***
	(2.91)	(2.89)
Voting×*ln*[public investment/capita]	−0.0019**	−0.0019**
	(2.15)	(1.98)
Voting×*ln*[private investment/capita]	−0.0009**	−0.0011***
	(2.17)	(2.87)
Human capital stock (%)	0.0015	0.0011
	(1.00)	(0.72)
Government earnings (%)	0.0099***	0.0096***
	(5.45)	(5.32)
Construction earnings (%)	−0.0060	−0.0055
	(1.30)	(1.29)
Manufacturing earnings (%)	0.0041**	0.0043**
	(2.38)	(2.44)
Services earnings (%)	0.0065***	0.0071***
	(3.70)	(3.53)
Northeast	0.074***	0.064***
	(3.48)	(2.71)
Midwest	0.029	0.030
	(1.56)	(1.19)
South	0.054**	0.045*
	(2.17)	(1.70)
d1970s	−0.189***	−0.183***
	(9.37)	(7.75)
d1980s	−0.093***	−0.077*
	(2.56)	(1.83)
$\lambda=\sigma_u/\sigma_v$		1.809***
		(3.99)
$\sigma^2=\sigma_v^2+\sigma_u^2$		0.075***
		(7.62)
Adjusted R^2	74.7	

*=significant at 10%, **=5%, ***=1% or lower in a two-tailed test of H_0: $\beta_j=0$.
Sample size = 150 (3 years × 50 states).

Consequently, the error term structure modeled as (3*a*) above is appropriate. In general, differences between OLS and MLE parameters in table 1 are small, but the parameters λ and σ^2 are individually statistically different from zero, indicating that OLS is not an appropriate estimator in this case. The parameter estimate for starting per capita income is significant and of the expected sign, confirming the convergence hypothesis. The coefficient estimate for human capital stocks has the expected sign, but is not significant.

The investment and voting variables are all statistically significant, as are their interactions. Although the parameter estimate for the voting record is positive, the effect of increasing the proportion of representatives who are Democrats is negative after taking into account interactions with private and public investment spending per capita, and using sample averages for the latter. A simulation shows that this result holds for values of public and private investment within one standard deviation of either or both means. Additional investments per capita of both private and public funds have a net positive effect on growth when the two other variables (including voting) are held constant at their means. For both types of investments, however, the marginal effect on growth eventually becomes negative if the other investment becomes large enough. Thus, if private investment is large, then public investment has a depressing effect on growth and *vice versa*. States with higher earnings in the government, manufacturing and services sectors, relative to natural resources, experienced faster income growth. States in the Northeast and South grew more rapidly than those in the West.

Table 3 shows estimates of the 15 largest systematic shocks affecting each state's economic growth, i.e., $E(u\,|\,\varepsilon_{it})$, calculated as shown in eqn. (4). In the 1960s, Hawaii, Arizona and Idaho had the largest shocks.[1] In the 1970s, Wyoming, Louisiana and Alaska each had larger-than average positive shocks to income growth, presumably because of oil price increases. Colorado and Texas appear to have similarly benefitted from the oil crisis. In the 1980s, the positive growth shock was largest in Connecticut, New Jersey, New Hampshire and Massachusetts. These states appear to have benefitted from the high technology revolution embodied in the so-called "Massachusetts miracle" of the 1980s in New England.

With the benefit of hindsight, it would be relatively straightforward to include an interaction term between natural resource earnings and the indicator variable for the decade of the 1970's. Such a term would capture the positive shock experienced by oil-rich states in that decade. Likewise, an interaction term between employment in defense and high-tech industries and the indicator variable for the decade of the 1980's would capture the "Massachusetts miracle." However, at the beginning of those decades, these positive shocks were generally unanticipated. If the goal is to forecast state economic growth, these types of shocks will not be known to the investigator, *a priori*. At best, a planner can predict the deterministic portion of the growth equation. It is not possible to predict the size of the shock that a state will experience.

In that context, it is useful to know how large these shocks can be, both in absolute terms and as a proportion of total growth (table 3). Planners can then bound their estimates of expected growth. For the three decades studied here, the proportionally largest shock occurred for Wyoming in the 1970s (17.8% of total growth), and the smallest for Hawaii in the 1970s (1.2%). The second and third largest proportional shocks occurred for Hawaii in the 1960s (15.1%) and Connecticut in the 1980s (13.5%). Consequently, these positive, unanticipated shocks contributed in a non-negligible manner to economic growth in each of the last three decades.

Further, the average size of these shocks was fairly stable across decades (0.049 in the 1960's, 0.050 in the 1970's and 0.048 in the 1980's). However, shocks experienced by individual states were not consistent across decades. The shock experienced by each state in decade t was uncorrelated to that experienced in decade t+10 (ρ=−0.055). Under the assumption that state economic growth continues to follow the pattern observed in the period 1960-1990, a planner predicting individual state growth would know that the state will experience a positive shock, would know the distribution of the shock, and thus would be able to construct a confidence interval for the

Table 3. State Rankings According to Size of Systematic Shock on Growth

State	Decade	Actual growth rate (y_{t+10}/y_t)	Growth rate without shock	Growth due to shock	Percent due to shock (%)
Hawaii	1960s	1.75	1.49	0.265	15.1
Wyoming	1970s	1.45	1.19	0.257	17.8
Connecticut	1980s	1.47	1.27	0.198	13.5
New Jersey	1980s	1.51	1.32	0.189	12.6
New Hampshire	1980s	1.49	1.30	0.186	12.5
Louisiana	1970s	1.37	1.20	0.172	12.6
Alaska	1970s	1.32	1.15	0.170	12.8
Massachusetts	1980s	1.49	1.33	0.158	10.6
Arizona	1960s	1.52	1.36	0.157	10.4
Idaho	1960s	1.58	1.43	0.150	9.5
Colorado	1970s	1.28	1.15	0.131	10.3
Michigan	1960s	1.44	1.31	0.129	9.0
Kentucky	1960s	1.65	1.52	0.128	7.8
Delaware	1980s	1.37	1.25	0.128	9.3
Texas	1970s	1.31	1.18	0.126	9.6

Source: Authors' calculations using results of eqn. (5).

shock, but would not be able to predict the actual size of the shock.

IV. Summary and Conclusion

Our results suggest the error term structure assumed under OLS is not appropriate for modeling the economic growth of states in the U.S. over the period 1960-90. More specifically, the OLS residuals in our original regression exhibited a skewed distribution, which was statistically significant at the ten percent level. Subsequent application of maximum likelihood techniques, which allowed for a normally-distributed error term component and an asymmetric component, yielded an overall positive value for the latter. This indicates that, in addition to random shocks, which are individually small and sum to zero over all states and time periods, states have experienced positive shocks to their economic growth rates that allowed them to grow more rapidly than they should have based on the values of the exogenous variables included to model growth.

This paper illustrates how shocks in an economic growth model can be specified and tested for. Further research using alternative specifications and time periods is required before more general conclusions about these systematic shocks and their directions can be derived. While we found the growth shock to be positive, there is no reason to expect that it will necessarily always be positive if, for example, different time periods or units of analysis are used (such as counties or countries), or if additional variables are included in the growth model.[2] To the extent that it will never be possible to control for all possible relevant variables, however, the error term decomposition used here should be considered in future studies of economic growth. As a minimum, a test for skewness in the OLS residuals would appear to be warranted.

The contribution of this study to the economic growth literature is that it begins to shed light on how *ex ante* unobservable and unmeasurable shocks to growth can be accounted for in applied work. More specifically, it is perhaps best not to model growth as proceeding smoothly from one equilibrium point to the next, as is conventionally assumed, but rather as a process which is constantly subjected to positive or negative shocks.

An appropriate analogy may be found in the field of ecology. There, steady-state models of ecological equilibrium have recently given way to new models which reflect growing awareness that ecosystems are repeatedly subjected to catastrophic and chaotic events. New insights into the process of economic growth may be gained if steady-state assumptions adopted in earlier studies are relaxed. A starting point would be to classify states in terms of the shocks they have experienced, find reasons for these shocks, and determine how the states have dealt with or adjusted to the shocks.

References

Aigner, D.J., C.A.K. Lovell and P. Schmidt. "Formulation and Estimation of Stochastic Frontier Production Functions Models," *J. Econometrics* 6(1977):21-37.

Barro, R.J., "Economic Growth in a Cross Section of Countries," *Q. J. Econ.* 106(1991):407-43.

Barro, R.J. and X. Sala-i-Martin, "Convergence Across States and Regions," *Brookings Pap. Econ. Act.* no. 1 (1991):107-82.

Battese, G.E. "Frontier Production Functions and Technical Efficiency: A Survey of Empirical Applications in Agricultural Economics," *Agr. Econ.* 7(1992):185-208.

Becker, G., K. Murphy and R. Tamura, "Human Capital, Fertility, and Economic Growth," *J. Polit. Econ.* 98(1990):S12-37.

Greene, B. *Econometric Analysis*, 2nd ed., McMillan Publishing Co., New York, 1993.

Jondrow, J., C.A.K. Lovell, I.S. Materov and P. Schmidt, "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model." *J. Econometrics*, 19(1982):233-8.

Jorgenson, D.W. and K.-Y. Yun, "Tax Reform and U.S. Economic Growth," *J. Polit. Econ.* 98(1990):S151-93.

King, R.G. and S. Rebelo, "Public Policy and Economic Growth: Developing Neoclassical Implications," *J. Polit. Econ.* 98(1990):S126-50.

Levine, R. and D. Renelt, "A Sensitivity Analysis of Cross-Country Growth Regression," *Amer. Econ. Rev.* 82(1992):942-63.

Mankiw, N.G., D. Romer and D.N. Weil, "A Contribution to the Empirics of Economic Growth," *Q. J. Econ.* 107(1992):407-37.

Nissan, E., "Convergence of Regional and State Rates of Growth of Income, Employment, and Population: 1969-2000," *Rev. Reg. Stud.* 22(1992):261-76.

Romer, P.M., "Endogenous Technological Change," *J. Polit. Econ.* 98(1990):S71-102.

U.S. Bureau of the Census, *Census of Population*, Washington, DC, various years.

U.S. Department of Commerce, Bureau of Economic Analysis, *State Personal Income by State: Estimates for 1929-89, and a Statement of Sources and Methods*, Washington, DC, 1989.

Appendix: Calculations

Real income growth rates are calculated over each of the three decades as the natural logarithm of the ratio of per capita income at the end of the decade to per capita income at the beginning of the decade.

Private investment is measured as real annual total new capital expenditures by manufacturers in each of the ten years prior to the decade for which the growth rate is calculated, divided by the total population at the beginning of the decade; this allows the investment expenditure to have a lagged effect on income growth. The first data source is for the census years 1952, 1957, 1962, 1967, 1972, and 1977. In 1979, 1980 and 1981, geographic area data were not tabulated due to budgetary constraints. For years in which data were missing, decade-average expenditures were substituted. National averages were substituted for Alaska and Hawaii for the 1960s, since data for these states were available only as of 1960.

Public investment is measured as real annual total capital outlay by state and local governments in each of the ten years prior to the decade for which the growth rate is calculated, divided by the total population at the beginning of the decade; this allows the investment expenditure to have a lagged effect on income growth. The first data source is for the census years 1957, 1962, 1967, 1972, and 1977. Data from the second are available only as of 1958, so that the average outlay in 1958, 1959 and 1960 was used for the missing years in the 1950s; also, the first Census of Government was administered in 1957. National averages were substituted for Alaska and Hawaii for the 1960s, since data for these states were available only as of 1960. In addition, data were not available for Minnesota in 1970 and 1971; Montana in 1951; Nevada in 1952; North Dakota in 1951 and 1952; South Dakota in 1955 and 1956; and Vermont in 1951. In these cases, national averages were substituted for each missing year.

The voting record is based on the percentage of elected representatives in the Upper and Lower Houses of each state that has a Democratic party affiliation. For most states, data are available at the start of each decade; in Hawaii and Maryland, elections were held in 1948 and 1958, rather than at the beginning of the decade. Party affiliation data are not reported for Nevada, which has a unicameral body of 49 legislators. Similarly, legislators in Minnesota were elected without party designation prior to 1980. In both latter cases national averages were substituted for the individual state.

Human capital stock is measured as the percent of the population 25 years or older having completed 12 years or more of formal education.

In the vector on earnings by sector, the excluded category (natural resources) consists of mining and agricultural services, forestry, fisheries and other related industry. General services include transportation and public utilities, wholesale and retail trade, finance, insurance and real estate, and services.

Endnotes

1. As suggested by a reviewer, statehood may explain the large shock for Hawaii.

2. A reviewer pointed out that unmeasured new technologies and agglomeration economies, which existed in New England in the 1980s, would positively affect growth. At the same time, dis-agglomeration, such as occurred when the steel industry moved out of the Midwest, would entail a negative shock. In general, it is not possible to predict *a priori* whether the positive or negative shock will dominate.