10-2003

Convergence among the U.S. states: Absolute, conditional, or club?

Paul A. Johnson
Vassar College, pajohnson@vassar.edu

Lisa N. Takeyama
Amherst College

Convergence Among the U.S. States: Absolute, Conditional, or Club?

Paul A. Johnson*
Department of Economics
Vassar College
Poughkeepsie, NY 12604

Lisa N. Takeyama
Department of Economics
Amherst College
Amherst, MA 01002

October 2003

*Corresponding author. Email: pajohnson@vassar.edu. Tel: 845-437-7395. Fax: 845-437-7576.

Vassar College Economics Working Paper # 50

This paper attempts to ascertain which of the convergence hypotheses – absolute, conditional, or club – best describes the economic development of the U.S. states since 1950. We use regression tree analysis to identify convergence clubs among the states and argue that the club characterization of the data dominates the other two. We find three convergence clubs with a state's age and it's initial densities of post offices and telephone cable determining club membership. Abstracting from catch-up effects, those states with higher densities tend to grow faster.

JEL Classifications: O40, O51

1. Introduction

In Johnson and Takeyama [2001] we argue that data on per capita personal incomes in the US states over the period 1950 – 1993 are consistent with the club convergence hypothesis. We find evidence of three such clubs, each with a different long-run growth rate – one at the US average, one about ¼ percentage point above that average and one about ¼ percentage point below – implying divergence of state incomes. In that paper we assumed that the US states had a common set of economic fundamentals and differed only in their initial (1950) conditions. Thus, we tested the unconditional or absolute β-convergence hypothesis against the club convergence hypothesis. Our results contrast with those of Barro and Sala-i-Martin [1995] who cannot reject the (absolute) β-convergence hypothesis for the U.S. states.[1] This hypothesis is, however, also rejected in favor of the conditional β-convergence by Holtz-Eakin [1993] who finds a role for physical and human capital variables in explaining cross-state differences in growth rates. Similar rejections are reported by Carlino and Mills [1996] and Evans and Karras [1996a, 1996b]

Both types of rejections of the absolute β-convergence hypothesis – that in favor of the conditional β-convergence hypothesis and that in favor of the club convergence hypothesis – imply the existence of permanent cross-state differences in per capita income levels although for very different reasons. In the first case, the differences reflect cross-state heterogeneity in variables such as rates of capital accumulation and the appropriate econometric response is the introduction of "control variables" into the growth regression.[2] In the second case, the differences reflect states lying in different basins of attraction defined by initial conditions and the appropriate econometric response is dividing the states into groups using variables measuring

[1]Barro and Sala-i-Martin's tests for σ-convergence reach a similar conclusion as do the nonparametric approaches in Quah [1996] and Johnson [2000] who uses the method pioneered by Quah [1997]. Bernard and Jones [1996] test for β and σ-convergence in total and industry-specific labor productivity across the states and find affirmative results for the total as well as for mining and manufacturing. Using a finite mixture approach, Tsionas [2000] finds little mobility in the cross-state distribution of Gross State Product. Durlauf and Quah [1999] provide a review of the different approaches to the convergence hypothesis.
[2]In the cross-country growth regression literature, this is the approach taken by Barro [1991] and Mankiw, Romer and Weil [1992] among others.

initial conditions referred to as “split variables”.[3] Distinguishing between these two possibilities remains one of the challenges in the convergence literature.[4] Theory does not always provide a complete guide as to which variables ought to be included in cross-country growth regressions and the *ad hoc* inclusion of variables opens the possibility of incorrect inferences. The inclusion of a variable as a control for microeconomic heterogeneity when it's proper role is to determine the basin of attraction into which a country falls could lead to the incorrect conclusion that conditional convergence is occurring.[5] The converse is also true and theory is a far from complete guide in the selection of variables appropriate for splitting the sample into locally converging groups.

In this paper we relax the maintained hypothesis in Johnson and Takeyama [2001] – the assumption of a common set of economic fundamentals – and so allow for the possibility of conditional β-convergence among state incomes. We ask whether or not the states are converging absolutely and, after concluding that they are not, we seek to determine whether the implied permanent differences between the states are due to conditional convergence or convergence clubs. One advantage of using the U.S. states as the group of “countries” is that a wide variety of economic and social indicators suitable as candidate split variables are collected on a consistent basis for the states. One of the disadvantages is that there are markedly fewer states in the U.S. than there are countries in the world. In addition the states are relatively less heterogeneous than the countries and so there may not be sufficient variation in state characteristics to allow determination of which, if any, are pertinent for growth either as control or split variables.

---

[3]This is the approach taken by Durlauf and Johnson [1995] who, using the same cross-country data as Mankiw, Romer and Weil [1992], reject the single regime model in favor of the “convergence club” alternative.
[4]See Durlauf [1995] for a discussion of this issue.
[5]To see this, suppose that the growth process is such that economies with an initial value of some variable less than some threshold converge to one steady state while those with a value greater than the threshold converge to another steady-state with a higher value of income per capita than the first. Provided that countries from both sides of the threshold are included in the sample, the variable in question would enter a growth regression with a positive coefficient and appear to be a control variable in the same way that the saving rate is a control variable in the Solow model.

The next section of the paper outlines the empirical framework that we employ while section 3 details the data used. Section 4 presents our results and section 5 offers some concluding remarks.

2. Empirical Framework

Let y_{it} denote real per capita income in state i at time t and y_{ut} denote U.S. real per capita income at time t. As is well known, linearization of the neoclassical growth model implies

$$\gamma_{it} = g + \beta_t(\mathit{log}\,\widehat{y}_{i0} - \mathit{log}\,y_{i0}) \tag{1}$$

and

$$\gamma_{ut} = g + \beta_t(\mathit{log}\,\widehat{y}_{u0} - \mathit{log}\,y_{u0}) \tag{2}$$

where $\gamma_{it} = \frac{\mathit{log}\,y_{it} - \mathit{log}\,y_{i0}}{t}$, $\gamma_{ut} = \frac{\mathit{log}\,y_{ut} - \mathit{log}\,y_{u0}}{t}$, $\beta_t = \frac{1-e^{-\lambda t}}{t} > 0$, $\lambda > 0$ is the rate of convergence to the steady state, g is the (common and exogenous) rate of technological progress and $\widehat{y}_{i0}$ ($\widehat{y}_{u0}$) is the level of real per capita income that would have prevailed at time 0 were state i (the U.S.) in a steady state. Assuming a common steady state across states so that $\widehat{y}_{i0} = \widehat{y}_{u0}$ for all i (absolute convergence), subtracting (2) from (1) and adding an error term gives

$$\Gamma_{it} = \; -\beta_t(\mathit{log}\,y_{i0} - \mathit{log}\,y_{u0}) + \epsilon_i \tag{3}$$

where $\Gamma_{it} = \gamma_{it} - \gamma_{ut}$ is the rate of growth of per capita income in state i relative to that in the U.S.

All of the regressions estimated in this paper are of the form

$$\Gamma_{it} = \alpha - \beta_t(\mathit{log}\,y_{i0} - \mathit{log}\,y_{u0}) + \Pi X_i + \epsilon_i \tag{4}$$

where α is a constant, X_i is a vector of control variables, and Π is a parameter vector conformable with X_i.[6] In addition to $\beta_t > 0$, the absolute convergence hypothesis constrains α and β_t to be constant across states in addition to constraining Π to be zero. The conditional convergence hypothesis relaxes the latter constraint while the club convergence hypothesis relaxes the former as well as allowing cross-state variation in Π.[7]

3. Data

Nominal incomes for the 48 contiguous states, and the entire U.S., in 1950 and 1993 are measured using the per capita personal income data published by the Bureau of Economic Analysis [1994].[8] Consumer Price Index (CPI) data (all items) for 17 major cities, and the U.S. city average, published by the Bureau of Labor Statistics (http://stats.bls.gov) are used to convert the income data to real terms. While the published CPI data allow cross-regional comparisons of changes in consumer prices, in the absence of information about the relative levels of prices in some year, they do not permit cross-regional comparisons of levels of consumer prices. Roberts [1979] concludes that the dispersion of prices across 7 of the 9 census regions of the U.S. in 1950 was very small and so we assume that prices were equal across the cities in 1950.[9] This

[6]The constant is possibly nonzero because the U.S. income figure is not the average of the state income numbers. See footnote 9 below.

[7]See Durlauf and Johnson [1995] for a complete discussion of how the club convergence hypothesis implies "cross-country" variation of the parameters in equation (4).

[8]Note that the U.S. per capita personal income figure is that for the U.S. as a whole and not simply the average of the figures for the 48 contiguous states. The former includes Alaska and Hawaii, which did not become states until 1959, as well as the District of Columbia. "State personal income is defined as the income received by, or on behalf of, all the residents of the State. It consists of the income received by persons from all sources—that is, from participation in production, from both government and business transfer payments, and from government interest (which is treated like a transfer payment). … Personal income is defined as the sum of wage and salary disbursements, other labor income, proprietors' income with inventory valuation and capital consumption adjustments, rental income of persons with capital consumption adjustment, personal dividend income, personal interest income, and transfer payments to persons, less personal contributions for social insurance." [U.S. Department of Commerce, 1995, p. M-5]. Personal income is a concept much more akin to GNP than to GDP which is the preferred measure of output. This issue matters most for those states with large numbers of workers who earn income in another state most noticeably New Jersey and Connecticut, both of which have many workers who commute to New York. Many studies have used personal income as their output measure and an arguably more suitable measure of state incomes, gross state product, is available for a considerably shorter time span.

[9]Only 7 of 9 regions are used because the Mountain and Pacific regions are excluded due to a lack of suitable data. In 1950 the spread between the highest and lowest price regions of the remaining seven was less than 3 percentage points of the average. We show in appendix 1 that the asymptotic bias in the estimator of β in equation (3) depends on the dispersion of prices across the states in the base year. This result and Roberts' data suggest that the errors in

assumption and the published CPI data allow us to compute the relative price levels in the cities in 1993. These data are used to deflate the 1993 personal income data for the states and for the U.S., with the city data being assigned to states as described in Table 1.[10] Finally, these data are used to compute real per capita income levels in each of the 48 states relative to the U.S. average in 1950 and 1993 and the average growth rates over that period relative to that in the U.S.

Using this data, our baseline (that is, with Π constrained to zero) estimate of equation (4) is $\widehat{\alpha} = -.0900$ and $\widehat{\beta}_t = 1.3228$ with estimated standard errors of .0316 and .1045 respectively and $R^2 = 0.723$.[11] As $\widehat{\beta}_t$ is positive and significantly different from zero, these results imply that, over the sample period, the typical state that started below the U.S. average tended to grow faster than the U.S. average – a finding consistent with previous cross-sectional studies of convergence across the states.

The economic and social indicators that we use measure a variety of characteristics of each state in 1950.[12] They are based on data in the *Statistical Abstract of the United States* in 1952, 1953 and 1960. Table 2 gives the mnemonics, specific descriptions of the variables and the exact references to the various editions of the abstract employed. It also gives the U.S. average and sample coefficient of variation of each variable. While the choice of which variables to include is admittedly *ad hoc* our motivation was to include variables measuring a range of characteristics that could potentially be relevant for post-1950 growth in each state. Those variables designed to measure aspects of the physical capital stock are AIRAREA, AIRPOP,

---

our estimates from this source are small. The regional CPI data used to compute the real state per capita income numbers used here suggest a 1993 range of about 12 percentage points of the U.S. city average CPI.

[10]While far from perfect, this approach seems preferable to the use of nominal data or, equivalently in this context, deflating the nominal data with a U.S.-wide price index.

[11]The test statistic for the White test of the hypothesis of a homoskedastic error has a marginal significance value of 0.086 while that for the Bera-Jarque test of the hypothesis of a normally distributed error is 0.996. The test statistic for the Hausman specification test has a marginal significance value of 0.763. This test was performed using the logs of HOSPOP, KWPOP, PHPOP, VECHPOP, as instruments for *log* (y_{i0}/y_{u0o}) giving a first-stage R^2 of 0.825. These were the only physical or human capital stock variables with positive and significant coefficients in the first stage regression. Use of these physical capital stock variables as instruments means that this test can be considered a test against unobserved, systematic cross-state initial productivity differences as well as against measurement errors, which are not limited to those due to the deflation method. The latter are known to tend bias the cross-section convergence test toward rejection of the no-convergence null.

[12]In a small number of instances the characteristic is measured pre-1950 if a 1950 measure is not available.

BANKAREA, BANKPOP, BOOKS, CABAREA, CABPOP, HOSAREA, HOSPOP, KWPOP, PHAREA, PHPOP, POAREA, POPOP, RAILAREA, RAILPOP, and VECHPOP while HIWAY and IONVA proxy the rate at which that stock is accumulated. The variables BORROW, COLLEGE, and SCHOYRS and are designed to proxy the human capital stock, while HIED, INSCHO, PUPTCH, SCHEXP, SCHOUP, and VOC proxy the rate at which that stock is accumulated. Those designed to capture potential cross-state differences in production functions and factor productivity are AGEMP, FARM, FARMVAL, MANEMP, MFG, PROFEMP, and PUBEMP, while those designed to capture differences in the rate of technological progress are NEWBUS and PATENTS. Aspects of state government finance are measured by CAPFRAC, NOTAXES, and TAXES while AFDC, AGE, BILO, BIRTH, BOLI, DEATH, DEMOCRAT, FAMSIZE, INFANT, LUNCH, MALE, MARRIAGE, MARRIED, MURDER, POPDEN, POPINFAM, SAMHSE, UNION, and URBAN measure in one way or another a variety of state characteristics including population density, mobility and income distribution.

4. Results

While our baseline estimate is consistent with the absolute convergence hypothesis, we have yet to examine the data for evidence of conditional convergence or club convergence. The former task requires looking for statistically significant control variables among the state characteristic variables list above. The later requires using those variables to look for statistically significant splits in the sample.

A. Is there evidence of conditional convergence?

We begin our results with estimates of equation (4) with each of the state characteristic variables used separately as the control variable in both level and log form. As discussed earlier, some researchers have found evidence of permanent cross-state differences in income levels and allowing Π to differ from zero is consistent with the conditional convergence hypothesis.[13]

[13]Our approach differs somewhat from that in other studies such as Mankiw, Romer and Weil [1992] using cross-country data and Holtz-Eakin [1993] using cross-state data in that our control variables are beginning-of-sample-period values while theirs are averages over the sample period. Our approach is arguably econometrically

These results are presented in the left-hand columns of Table 3. Of the 59 variables used, 39 enter either the estimated level or log regression or both with a coefficient significantly different from zero at the 5% level. Some of the variables offer substantial increases in explanatory power over the baseline regression with several having R^2 values of 0.80 or higher.

Many of the significant coefficients have the expected signs but there are a few apparent anomalies. For example, the estimated coefficients on INSCHO, SCHEXP, and SCHOYRS, all human capital variables, have negative signs as do those on AIRPOP, HIWAY, POPOP, RAILAREA, and VECHPOP, all physical capital variables, while that on MURDER has a positive sign.[14] Most of the other physical capital variables have the expected positive coefficients. Among the technology variables, the positive coefficient on MANEMP is as expected as are the negative coefficients on AGEMP and PUBEMP. The negative coefficients on FARM and MFG may reflect sector-specific catch-up effects.[15] The significant estimated coefficient on TAXES is negative as expected.[16] Both POPDEN and URBAN have significantly positive estimated coefficients.[17] The estimated coefficient on UNION is positive and significant indicating that older states tend to have higher growth rates once catch-up effects are taken into account.

These results are consistent with the view that there is a role for additional control variables in U.S. state growth regressions as implied by the conditional convergence hypothesis.[18]

---

preferable (because there is likely to be less correlation between the error term and the explanatory variable) and makes for easier comparisons when the variables are used to split the sample below.

[14]The human capital variable results are partially consistent with those in Rappaport [1999] who, using county-level data, finds a negative relationship between growth and "medium" levels of initial human capital – the percentage of adults who are high school, but not college, graduates. He also a positive relationship between growth and "high" initial levels of human capital – the percentage of adults who are college graduates – which we do not find. Our result with the MURDER variable seems inconsistent with Rappaport's "unsurprising" find that violent crime is negatively correlated with growth.

[15]Bernard and Jones [1996] find convergence in manufacturing labor productivity across the states.

[16]Rappaport [1999] finds "robust" negative relationships between growth and measures of local government size.

[17]If we interpret URBAN as a measure of the within-state inequality in population density, these results are consistent with the model in Ciccone and Hall [1996] which predicts that productivity will be positively related to both population density and its inequality within a state if agglomeration effects are sufficiently strong.

[18]In each case the (unreported) estimate of β_t was positive and significantly different from zero.

B. Is there evidence of convergence clubs?

Our next set of results looks for instability across the states in the parameters of equation (4), with Π constrained to zero, as such instability is evidence that there are multiple basins of attraction in the growth process. For each of the state characteristic variables, we sort the sample into ascending order based on that variable and divide the sorted sample into thirds – bottom, middle and top. We allow the intercept and slope of the growth regression to vary across the thirds and estimate the model

$$\Gamma_{it} = \alpha + \alpha^B D^B_{ji} + \alpha^T D^T_{ji} - (\beta_t + \beta^B_t D^B_{ji} + \beta^T_t D^T_{ji})(log\, y_{i0} - log\, y_{u0}) + \epsilon_i \tag{5}$$

where $D^B_{ji} = 1$ if state i is in the bottom third of the sample when it is sorted on variable j and zero otherwise and $D^T_{ji} = 1$ if state i is in the top third of the sample when it is sorted on variable j and zero otherwise. The middle column of Table 3 presents Wald test statistics for H_0: $\alpha^B = \alpha^T = \beta^B_t = \beta^T_t = 0$, the hypothesis that the parameters of the growth regression do not vary across the thirds of the sample.[19] In 29 of the 60 cases this hypothesis is rejected.[20] We interpret these rejections as evidence consistent with the existence of multiple basins of attraction.

Notably, many of the successful split variables are proxies for the physical capital stock or its rate of accumulation while few are proxies for the human capital stock. This could reflect a more important role for physical capital or the lesser sample dispersion in the human capital variables which tend to have markedly lower coefficients of variation.[21] None of the state government finance variables produce significant splits but about half of the technology/productivity and general characteristic variables are successful in this regard. These

[19]The test statistics have (asymptotic) $\chi 2$ distributions as they are heteroskasticity consistent.
[20]There are 60 candidate split variables because we use the log of initial relative income, *log* (y_{i0}/y_{u0}), in addition to the 59 economic and social indicators.
[21]This information is given in Table 2.

results are consistent with the view that there is parameter instability across the states as implied by the club convergence hypothesis.

C. Regression tree estimates of convergence clubs

The results of the previous subsection show that there is instability in the estimated parameters across groups of states when several of the state characteristic variables are used to split the sample but they leave at least two issues unresolved. First, the splits rely on a mechanical splitting of the sample into thirds and there is no reason to believe that the true splits in the sample correspond to these boundaries. Second, we would like to be able to differentiate among the split variables and say something about which are the more important in determining long-run outcomes. In an attempt to resolve both issues we use regression tree analysis which allows us to search for an unknown number of splits of the data at unknown locations based on orderings given by one or more variables measuring initial conditions.

Here we give a brief outline of the method. A detailed discussion can be found in Breiman, *et al.* [1984] while Durlauf and Johnson [1995] present a treatment specifically tailored to the issue at hand. There are two stages to the regression tree procedure – growing and pruning. In the growing stage, equation (4), with Π constrained to zero, is estimated for each subsample defined by all possible binary splits of the sample based on each of the variables measuring initial conditions (subject to the restriction that within each subsample there must be at least 3 observations) and the total sum of squared residuals over all observations computed. The split with the minimum sum of squared residuals is taken as the first split of the data. This procedure is repeated separately on each of these subsamples further splitting the sample into four subsamples. Growing continues in this way until all degrees of freedom are exhausted. The result is a representation of sample as a tree with each node containing a subsample – a representation that is certain to severely overestimate the number of regimes in the data.

In the pruning stage, a cost of making a split is imposed. As this cost is increased from 0 to ∞ the tree from the growing stage is pruned until only the original sample remains. The optimal tree is one of the resultant sequence of pruned subtrees of the tree from the growing stage. To choose the optimal tree "cross–validated" estimates of the error variance of each model represented by a tree in the sequence are required. These are found by calculating the residual for each observation using estimates of the model parameters computed from the sample net of that observation. These residuals are used to estimate the error variance using the usual formula. There are two rules for selecting the optimal tree from this sequence. One is to choose that tree having the minimum cross–validated error variance. The other, which, of course, chooses a more parsimonious model, is to choose that tree having the largest cross–validated error variance that is less than the minimum plus its estimated standard error. No known asymptotic theory exists for testing for the number of regimes in the data but the procedure is consistent in the sense that, if there are finitely many splits, as the sample size goes to infinity, both rules will identify the correct model.

Our candidate split variables are those 29 variables in Table 3 which enable us to reject the hypothesis that the parameters of the growth regression are constant across the sample. We use the second of the two rules above as we choose to err on the side of parsimony given the number of observations available. This rule chooses a tree that divides the 48 states into four groups as illustrated in Figures 1 and 2. The first split in the tree in Figure 1 is on the variable UNION and divides the states into those older 28 states that joined the UNION before 1846 and those 20 newer states who joined in or after 1846. The second split divides the older states according to the value of POAREA. Those 17 with fewer than 35 post offices per 1000 square miles fall in the terminal node labeled "1" while the terminal node labeled "2" contains those 11 states with at least 35 post offices per 1000 square miles. As Figure 2 shows, the node 1 states are located in the south and mid-west while the node 2 states are located predominately in the northeastern U.S.[22] The U.S. average post office density is 11.5 per 1000 square miles so the

[22]Figure 2 reveals that there is a somewhat geographical nature to the state groupings that we find. To test against the possibility that the results are driven simply by spatial correlation we reran the regression tree procedure with the Bureau of Economic Analysis regions, the Bureau of the Census regions, a dummy for the original 13 colonies, and a dummy for the 11 confederate states as candidate split variables with no change in our results. Using, *inter alia*, data on the 48 contiguous U.S. states, Attfield et al. [2000, p. 112] find that "…geographic proximity per se does not have a key role to play in explaining the tendency of economic growth rates to cluster."

states in node 2 are very dense with post offices by U.S. standards. The third, and final, split in the optimal tree splits the newer states on the variable CABPOP. The terminal node labeled "3" contains those 10 states with fewer than 708 miles of telephone cable per 1000 persons. As the U.S. average is 970.6 miles per 1000 persons these states, all located west of the Mississippi River, are quite sparse with telephone cable by U.S. standards. The remaining 10 states, those with at least 708 miles of telephone cable per 1000 persons, fall in the terminal node labeled "4".[23]

In searching for splits, the regression tree procedure allows both the intercept and slope of the basic growth regression to vary across nodes. In order to discover which parameter differences are responsible for the splits that we find, we estimate variants of the dummy variable model implied by the splits in the regression tree shown in Figure 1. That is, we estimate variants of the model

$$\begin{aligned}\Gamma_{it} &= \alpha + \alpha^2 D_i^2 + \alpha^3 D_i^3 + \alpha^4 D_i^4 \\ &\quad - (\beta_t + \beta_t^2 D_i^2 + \beta_t^3 D_i^3 + \beta_t^4 D_i^4)(log\, y_{i0} - log\, y_{u0}) + \epsilon_i\end{aligned} \tag{6}$$

where $D_i^n = 1$ if state i is in node n, for $n = 2, 3, 4$. Table 4 presents estimates of this regression that allow intercepts, slopes and both to vary across nodes. Comparison of the R^2 values with that in the baseline regression shows that the differences in the intercept terms across nodes are largely responsible for the data's affinity for the splits discovered by the regression tree procedure. Allowing just the intercepts to differ increases R^2 from 0.72 to 0.89 while allowing both the intercepts and the slopes to differ gives an R^2 of 0.91.

Estimated standard errors are shown below each parameter estimate in Table 4 but, with the exception of the first column, it is not possible to use these to conduct inference in the usual

[23]The tree minimizing cross–validated error variance divides the sample into 5 groups and has a cross–validated error variance that is 9.1% less than that in Figure 1. The additional group arises because of a split of those states in node 1 into those 8 with PHPOP less than 183 and those 9 with PHPOP no less than this value. Given that PHPOP is similar in nature to POAREA and CABPOP, our use of the rule giving the more parsimonious tree in finite samples seems to have little qualitative implication for our findings.

way because we have mined the data (albeit in a statistically respectable fashion) to locate the splits.[24] However, the true marginal significance values are certainly higher than those implied by the standard errors shown so we can be sure that any estimate that is insignificant by conventional standards would also be insignificant were we able to conduct the proper test. In particular, the estimated values of both α^4 and β_t^4 are insignificant in the regression with both slope and intercept dummies.[25] The implication is that neither the intercepts nor slopes in nodes 1 and 4 are different – newer states with a relatively high amounts of phone cable per person behave similarly to older states with relatively low densities of post offices. We impose the restriction $\alpha^4 = \beta_t^4 = 0$ and estimate the regression with nodes 1 and 4 combined. The estimate of β_t^3 is not significant in this regression, which is shown in the next-to-last column of Table 4, implying that the slope in node 3 is the same as that in nodes 1 and 4.[26] We conclude that the important parameter differences across the nodes are, at most, the different intercepts between nodes 2, 3, and the combined nodes 1 and 4 as well as the different slope in node 2. Accordingly, the last column of Table 4 shows estimates of the model with α^4, β_t^4, and β_t^3 constrained to zero. Note that imposition of these constraints results in practically no reduction in the R^2 of the regression.

The regression tree procedure effectively divides the states into three groups with different growth behavior – those in node 2, those in node 3 and those in the combined nodes 1 and 4. The regression lines for each of these groups implied by the estimates in the final column of Table 4 are plotted in Figure 3 along with the data on Γ_{it} and *log* (y_{i0}/y_{u0}). The estimated differences in the intercepts show that, over the 1950 to 1993 period, abstracting from catch-up effects, the states in node 2 had about a quarter percentage point higher average annual growth

---

[24]These issues in no way undermine the validity of the use of regression trees here because we use them simply to locate splits rather than to establish the existence of splits. The latter is done by the tests in the "Split Tests" column of Table 3.
[25]The test statistic for the usual Wald test of the hypothesis that both parameters are zero is 0.48 and has a marginal significance value of 0.63.
[26]The test statistic for the usual Wald test of the hypothesis that $\beta_t^3 = \alpha^4 = \beta_t^4 = 0$ is 0.38 and has a marginal significance value of .77.

rate than those in nodes 1 and 4 while those in node 3 had about a quarter percentage point lower average growth rate. In addition, the lower slope of the regression line in node 2 when compared to the other nodes implies a rate of convergence to that node's steady state value that is about half that in the other nodes.[27]

D. Convergence clubs versus conditional convergence

The previous subsections have presented evidence consistent with both the conditional convergence and convergence club hypotheses. So, while we are able to reject the absolute convergence hypothesis, we have yet to determine whether the implied permanent differences between the states is due to conditional or club convergence. To do this we rerun the original conditional convergence tests while including dummy variables defining the convergence clubs in the regressions. Three dummies are included – two to allow for the different intercepts for the states in node 2 and those in node 3 and one to allow for the different slope in node 2. The results are presented in the right-hand columns of Table 3. Only one, MURDER, of the 39 variables that were found to be significant in the left-hand columns of Table 3 remain significant when the dummy variables are included although the increment to R^2 is small.[28] We interpret this dominance of split variables over control variables as evidence that the permanent differences in state growth behavior found here, and in other studies, reflect club rather than conditional convergence. In other words, initial conditions do matter for the states and the significant coefficients found in the left-hand columns of Table 3 occur because the variables are proxying for the roles of those conditions.[29]

5. Concluding Remarks

It is tempting to regard the successful split variables, UNION, POAREA, and CABPOP, as proxying for the amount of capital or infrastructure in each state – POAREA and CABPOP

---

[27]The implied convergence rate for nodes 1, 3 and 4 is 2.41% per year while that for node 2 is 1.24% per year.

[28]Note that, as before, the coefficient on MURDER is surprisingly positive. We have no good explanation for this or MURDER's continued significance but it should be pointed out that in 39 independent 5% tests of a true null hypothesis we'd expect almost two rejections so chance cannot be ruled out here.

[29]Of course, the caveats raised in Durlauf and Johnson [1995] apply here as well.

because they represent different ways of measuring the density of capital and infrastructure and UNION because older states have been accumulating for longer periods of time. With this interpretation we can think of the regression tree as dividing the states into a capital rich group (node 2), a relatively capital poor group (node 3), and a group in between (nodes 1 and 4). The conclusion of this paper would then be that initial conditions matter because initial capital or infrastructure stocks determine which states belong to which convergence clubs.[30] The mechanism here could be, for example, a threshold externality in capital or infrastructure per worker as suggested by Azariadis and Drazen [1990].

There are, however, at least two other conclusions consistent with our results. The thresholds that we find could reflect the influences of agglomeration effects or communication effects.[31] In the first case, the salient feature differentiating the states is the initial density of economic activity. POAREA is particularly interesting in this case as it is the product of POPDEN and POPOP so that states with high values of POAREA can be interpreted as states with both a high population density and a high capital-labor ratio. With this interpretation, an implication of our results is that different initial densities can have permanent effects. This interpretation is consistent with the conclusions of Ciccone and Hall [1996] who find that density is an important determinant of the level of labor productivity and that it may also be important for growth. Importantly, there is some correspondence between our classifications and the density index which they construct using data for 1988.[32]

In the second case, communication of ideas could be facilitated by the high density of communications networks – mail and telephone – which allows a more rapid flow of new ideas and hence increases the effective rate of technological progress. Canning [1999], for example stresses the role of communications in the “diffusion of technology” when describing research

[30]Positive effects of infrastructure investment on productivity have been found by Nadiri and Mamuneas [1994], Morrison and Schwartz [1996], and others.

[31]In either case, as before, UNION most likely captures the effects of the age of the state. Older states tend to be more dense because they were settled earlier. They also tend to have more established communications networks for the same reason.

[32]All of our Node 2 states except Kentucky, New Hampshire and Vermont have high values of the Ciccone and Hall [1996] density index while many of our Node 3 states have low values of the index.

suggesting "that telecommunications can have a large impact on subsequent economic growth" (p. 1). Using state and county level data from the US, Cronin, *et. al.* [1995] find feedback in both directions between telecommunications investment and economic activity, while Temple and Johnson [1998] find an important role for communications variables in explaining differences in cross-country growth behavior. More recently, Röller and Waverman [2001] find evidence of a positive causal relationship from telecommunications infrastructure to economic growth in the OECD countries. Consistent with the results here, that relationship is found to be stronger after a threshold in the amount of infrastructure is exceeded.

Of course, these explanations are not mutually exclusive as it is entirely likely that improvements in communication and all they involve enable the reaping of many of the benefits of agglomeration without incurring all of the congestion costs. For example, the U.S. Postal Service is regarded by some as being instrumental in the development of much of the U.S. transportation infrastructure.[33] States with higher densities of post offices would then tend to have lower transportation costs. In any case, it is clear that further work is necessary to determine the precise economic mechanisms underlying our statistical finding that initial conditions have played an important role in the economic development of the US states since 1950 and its implication that the observed permanent differences in state growth behavior reflects the existence of convergence clubs.

[33]The Postal Service claims that "… [it] has helped develop and subsidize every new mode of transportation in the United States. The postal role was a natural one; apart from postal employees themselves, transportation was the single most important element in mail delivery, literally, the legs of communication. … As mail delivery evolved from foot to horseback, stagecoach, steamboat, railroad, automobile, and airplane, with intermediate and overlapping use of balloons, helicopters, and pneumatic tubes, mail contracts ensured the income necessary to build the great highways, rail lines, and airways that eventually spanned the continent." United States Postal Service, [Date Unknown].

References

Attfield, C.L.F., Cannon, Edmund S., Demery, D., and Duck, Nigel W., [2000], "Economic Growth and Geographic Proximity," *Economics Letters*, 68:109-112.

Azariadis, Costas, and Drazen, Allan, [1990], "Threshold Externalities in Economic Development", *Quarterly Journal of Economics*, 105:501-526

Barro, Robert J., [1991], "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106:407-43

Barro, Robert J., and Sala-i-Martin, Xavier, [1995], *Economic Growth*, McGraw-Hill, New York.

Bernard, Andrew B., and Jones, Charles I., [1996], "Productivity and Convergence across U.S. States and Industries," *Empirical Economics*, 21:113-35.

Bernard, Andrew B., and Durlauf, Steven N., [1996], "Interpreting Tests of the Convergence Hypothesis," Journal of Econometrics, 71:161-73.

Breiman, Leo, Friedman, Jerome H., Olshen, Richard A., and Stone, Charles J., [1984], *Classification and Regression Trees*, Wadsworth, Pacific Grove, California.

Bureau of Economic Analysis, [1994], State Annual Summary Tables 1929-93 for the States and Regions of the Nation, Department of Commerce, Washington DC.

Canning, David, [1999], "Telecommunications, Information Technology, and Economic Development," *CAER II Discussion Paper 53*, Harvard Institute for International Development, Cambridge.

Carlino, Gerald, and Mills, Leonard, [1996], "Convergence and the U.S. States: A Time-Series Analysis," *Journal of Regional Science*, 36:597-616.

Ciccone, Antonio, and Hall, Robert E., [1996], "Productivity and the Density of Economic Activity," *American economic Review*, 86:54-70.

Cronin, Francis J., McGovern, Patricia M., Miller, Michael R., and Parker, Edwin B., [1995], The Rural Economic Development Implications of Telecommunications: Evidence from Pennsylvania," *Telecomunications Policy*, 19:545-59.

Durlauf, Steven N., [1993], "Nonergodic Economic Growth," *Review of Economic Studies*, 60:349-366.

Durlauf, Steven N., [1995], "On Growth and Indeterminacy: Some Theory and Evidence: A Comment," *Carnegie-Rochester Conference Series on Public Policy*, 43:213-23.

Durlauf, Steven N., and Johnson, Paul A., [1995], "Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics*, 10:365-84.

Durlauf, Steven N., and Quah, Danny, [1999], "The New Empirics of Economic Growth", in *Handbook of Macroeconomics*, vol. 1A, John B. Taylor and Michael Woodford (editors), North Holland, Amsterdam, p231-304.

Evans, Paul, and Karras, Georgios, [1996a], "Do Economies Converge? Evidence from a Panel of U.S. States," *Review of Economics and Statistics*, 78:384-388.

Evans, Paul, and Karras, Georgios, [1996b], "Convergence Revisited," *Journal of Monetary Economics*, 37:249-65.

Galor, Oded, [1996], "Convergence? Inferences from Theoretical Models," *Economic Journal*, 106:1056-69.

Holtz-Eakin, Douglas, [1993], "Solow and the States: Capital Accumulation, Productivity, and Economic Growth," *National Tax Journal*, 46: 425-39.

Izraeli, Oded, and Murphy, Kevin, [1997], "Convergence in State Nominal and Real Per Capita Income," *Public Finance Review*, 25:555-76.

Johnson, Paul A. [2000], "A Nonparametric Analysis of Income Convergence across the US States," *Economics Letters*, 69:219-23.

Johnson, Paul A. and Takeyama, Lisa N., [2001], "Initial Conditions and Economic Growth in the US States," *European Economic Review*, 45:919-27.

Mankiw, N. Gregory, Romer, David, and, Weil, David N., [1992], "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107:407-37.

Morrison, Catherine J., and Schwartz, Amy Ellen, [1996], "State Infrastructure and Productive Performance," *The American Economic Review*, 86:1095-111.

Nadiri, M. Ishaq, and Mamuneas, Theofanis P., [1994], "The Effects of Public Infrastructure and R & D Capital on the Cost Structure and Performance of U.S. Manufacturing Industries," *The Review of Economics and Statistics*, 76:22-37.

Quah, Danny, [1996], "Empirics for Economic Growth and Convergence," *European Economic Review*, 40:1353-75.

Quah, Danny, [1997], "Empirics for Growth and Distribution: Polarization, Stratification, and Convergence Clubs," *Journal of Economic Growth*, 2:27-59.

Rappaport, Jordan, [1999], “Local Growth Empirics,” *CID Working Paper No. 23*, Harvard University.

Roberts, Charles A., [1979], “Interregional Per Capita Income Differentials and Convergence: 1880-1950,” *Journal of Economic History*, XXXIX:101-12.

Röller, Lars-Hendrick, and Waverman, Leonard, [2001], “Telecommunications Infrastructure and Economic Development: A simultaneous Approach,” *American Economic Review*, 91:909-23.

Temple, Jonathan, and Johnson, Paul A., [1998], “Social Capability and Economic Growth,” *Quarterly Journal of Economics*, 113:965-90.

U.S. Bureau of the Census, [1952, 1960], *Statistical Abstract of the United States*, Washington, D.C.

U.S. Department of Commerce, Bureau of Economic Analysis, [1995], *State Personal Income, 1929-93*, U.S. Government Printing Office, Washington, DC. Available at http://www.bea.doc.gov/bea/ARTICLES/REGIONAL/PERSINC/Meth/spi2993.pdf.

U.S. Postal Service, [Date Unknown], *History of the U.S. Postal Service: 1775-1993*, http://www.usps.gov/history/his1.htm

Tsionas, Efthymios G, [2000], “Regional Growth and Convergence: Evidence from the United States,” *Regional Studies*, 34:231-38.

Appendix: The Effect of Using a Price Index to Deflate Nominal Incomes.

As in the text, let y_{it} denote real per capita income in state i at time t and y_{ut} denote U.S. real per capita income at time t. Let $Y_{it} =$ nominal per capita income in state i at time t, $Y_{ut} =$ U.S. nominal per capita income at time t, $P_{it} =$ price level state i at time t, $P_{ut} =$ U.S. price level at time t, so that $y_{it} = \frac{Y_{it}}{P_{it}}$ and $y_{ut} = \frac{Y_{ut}}{P_{ut}}$. Suppose that D_{it} is an index of prices in state i at time t such that $D_{it} = \frac{P_{it}}{P_{ib}}$ where b is the base year of the index, and $D_{ut}$is an index of prices in the U.S. at time t such that $D_{ut} = \frac{P_{ut}}{P_{ub}}$. Then $y_{it} = \frac{Y_{it}}{P_{ib}D_{it}}$ and $y_{ut} = \frac{Y_{ut}}{P_{ub}D_{ut}}$ so that

$$log\ y_{it} - log\ y_{ut} = log\frac{Y_{it}}{D_{it}} - log\frac{Y_{ut}}{D_{ut}} + log\frac{P_{ub}}{P_{ib}}. \qquad \text{(A-1)}$$

It follows that the left hand side of (3) can be written as

$$log\frac{Y_{it}}{D_{it}} - log\frac{Y_{ut}}{D_{ut}} - \left(log\frac{Y_{i0}}{D_{i0}} - log\frac{Y_{u0}}{D_{u0}}\right)$$

so that using a price index rather that the price level to deflate nominal incomes produces no measurement error on that side of the equation. Substitution of the result in (A-1) into the right-hand side of (3) gives

$$log\frac{Y_{it}}{D_{it}} - log\frac{Y_{ut}}{D_{ut}} - \left(log\frac{Y_{i0}}{D_{i0}} - log\frac{Y_{u0}}{D_{u0}}\right) = -\beta\left(log\frac{Y_{i0}}{D_{i0}} - log\frac{Y_{u0}}{D_{u0}} + log\frac{P_{ub}}{P_{ib}}\right) + \epsilon_i$$

which can be written as

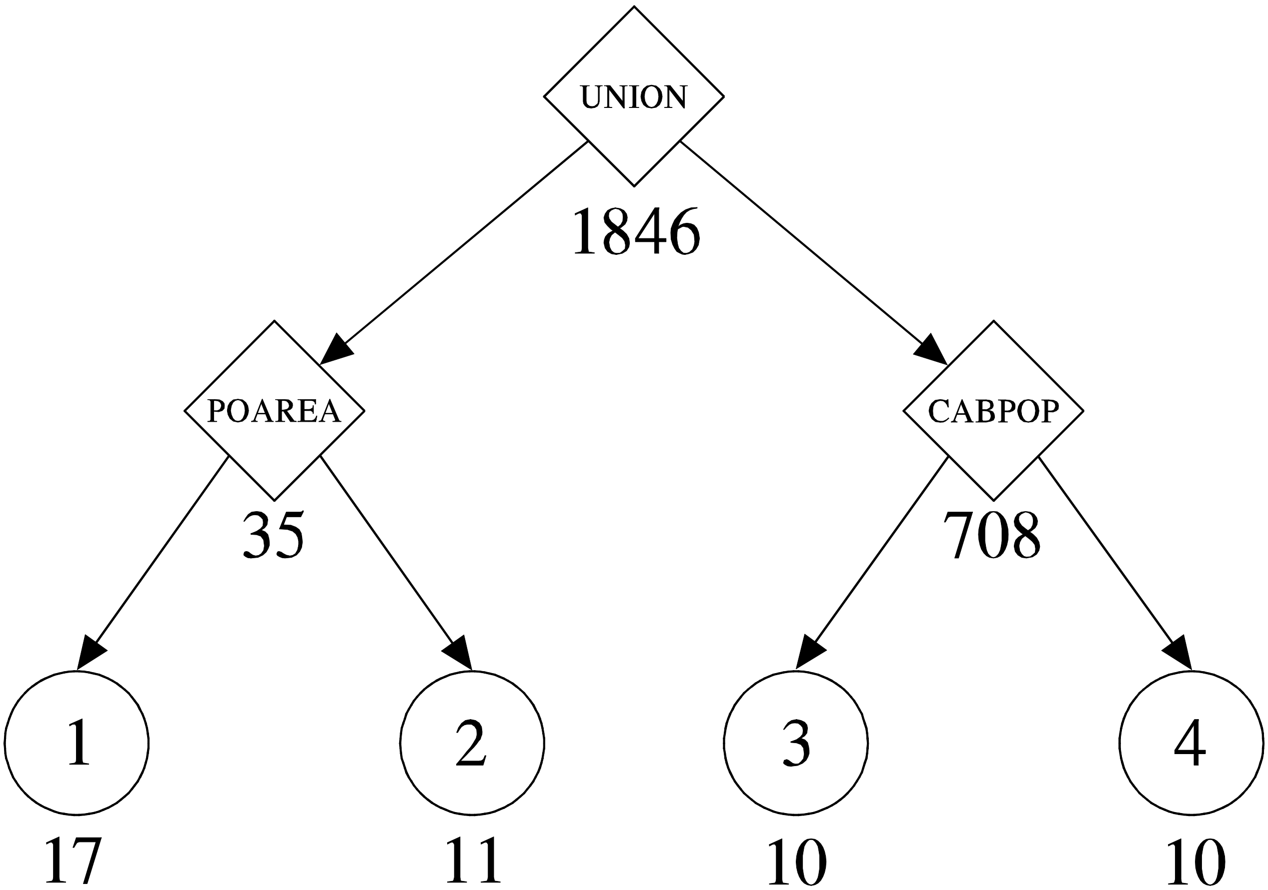
$$log\frac{Y_{it}}{D_{it}} - log\frac{Y_{ut}}{D_{ut}} - \left(log\frac{Y_{i0}}{D_{i0}} - log\frac{Y_{u0}}{D_{u0}}\right) = -\beta\left(log\frac{Y_{i0}}{D_{i0}} - log\frac{Y_{u0}}{D_{u0}}\right) + u_i \qquad \text{(A-2)}$$

where $u_i = \epsilon_i - \beta log\frac{P_{ub}}{P_{ib}}$. Application of OLS to equation (A-2) gives an estimator of the slope coefficient whose probability limit is

$$-\beta + \frac{\beta Var(logP_{ub} - logP_{ib})}{Var(log\ y_{i0} - log\ y_{uo}) + Var(logP_{ub} - logP_{ib})}$$

provided that $log\ \ y_{i0} - log\ \ y_{uo}$ and $logP_{ub} - logP_{ib}$ are uncorrelated and ϵ_i and $logP_{ub} - logP_{ib}$ are uncorrelated. The asymptotic bias in the estimator of β thus depends on the dispersion of prices across the states in the base year.

Figure 1: Regression Tree



The left descendent of each nonterminal node contains those observations for which Split Variable < Split Value. The right contains those for which Split Variable ≥ Split Value.

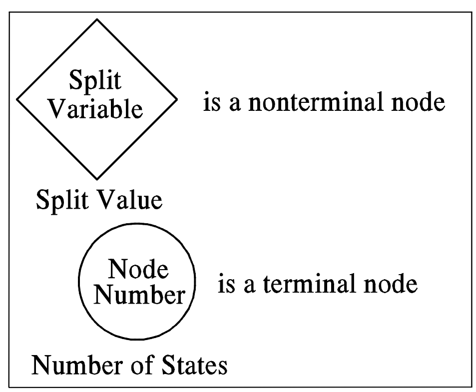


Figure 2: States and Terminal Nodes

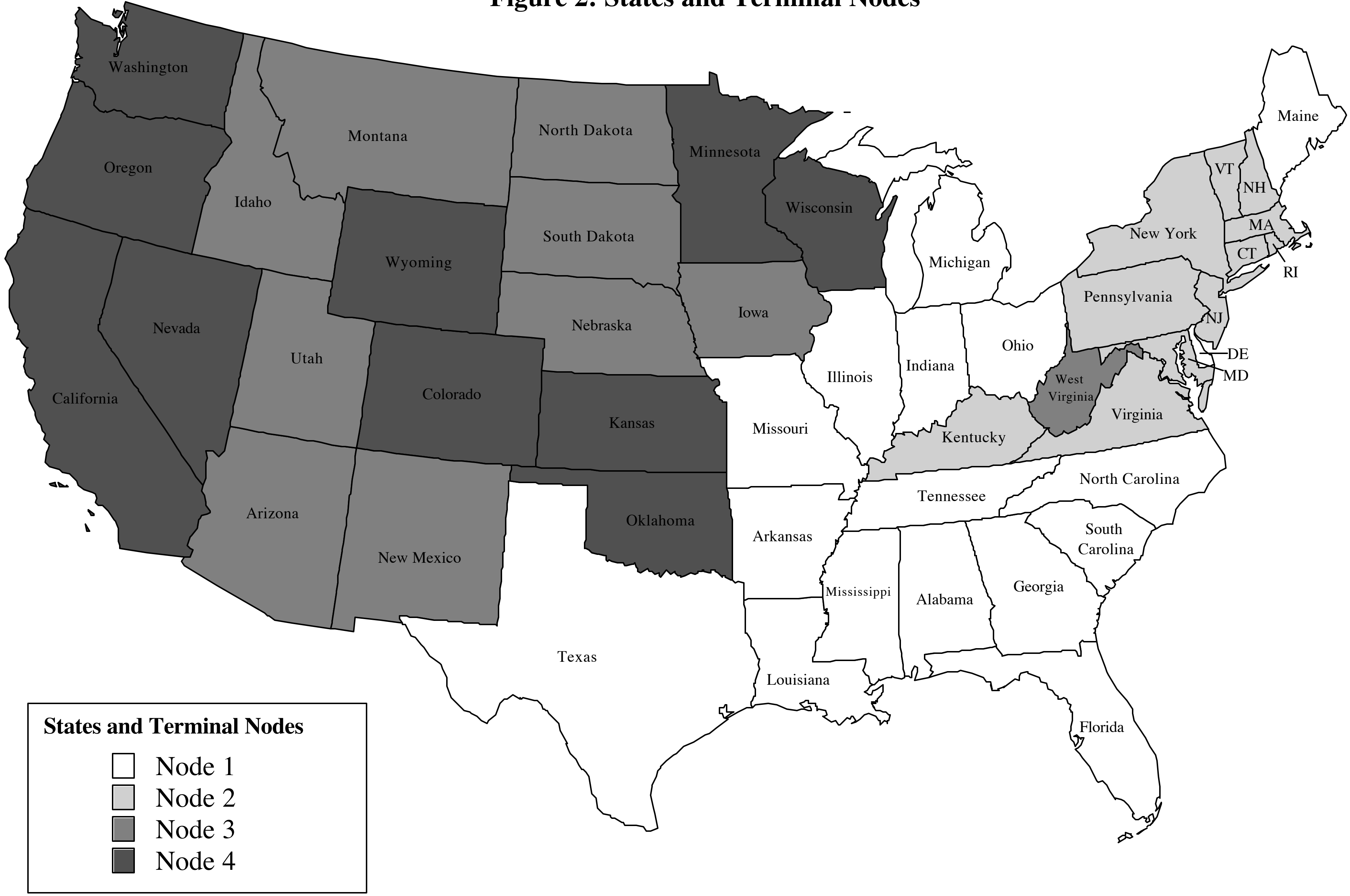


Figure 3: Average Growth Rate versus Log Initial Income

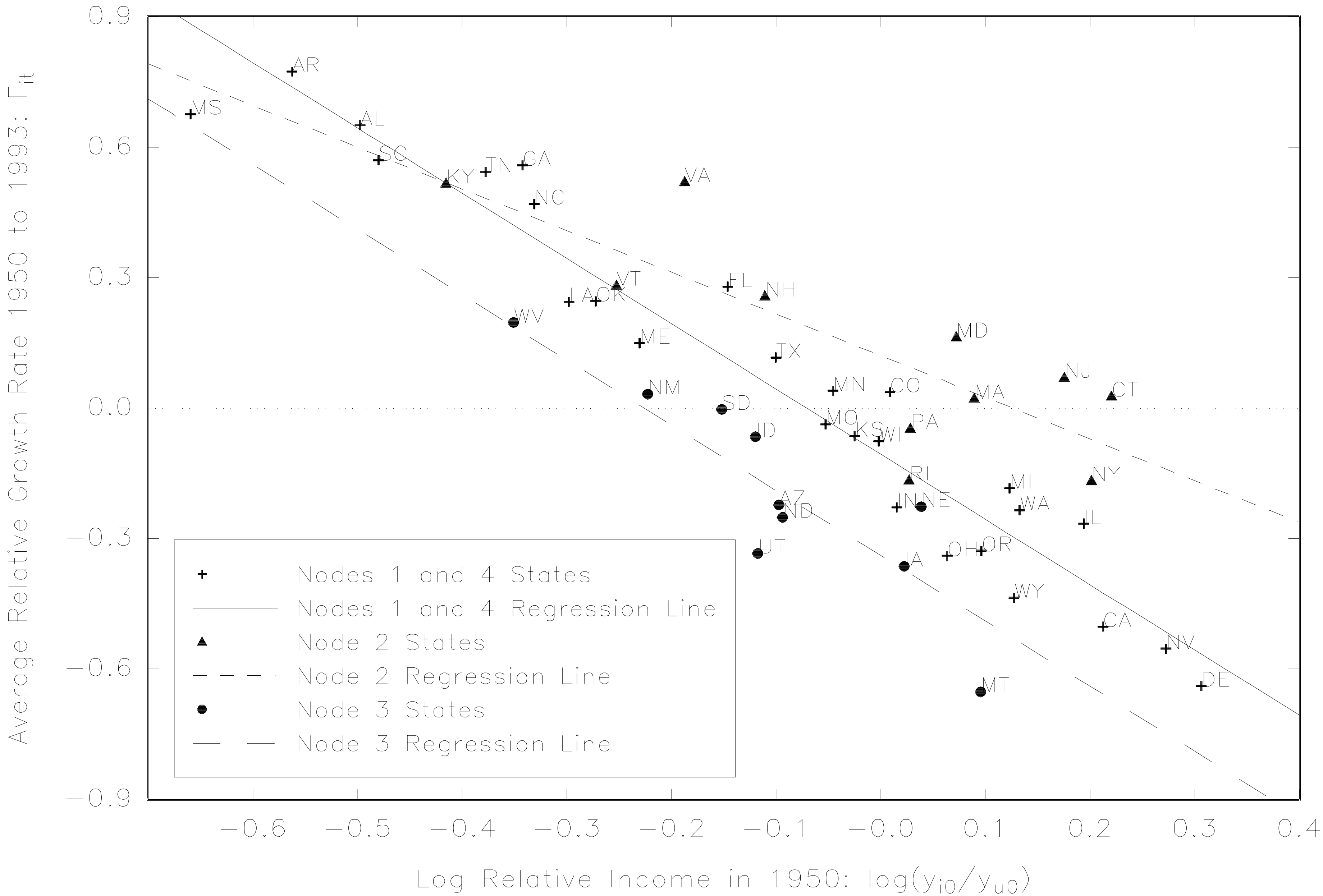


Table 1: City Price Data and Assignments to States

Regional CPI	1993 Level (1950 = 100.0)	States
U.S. City Average	599.6	—
New York, Northern New Jersey, Long Island	628.5	NY, NJ
Philadelphia, Wilmington, Atlantic City	608.1	DE, PA, WV
Boston, Brockton, Nashua	639.7	CT, MA, ME, NH, RI, VT
Chicago, Gary, Kenosha	600.8	IL, IN
Detroit, Ann Arbor, Flint	567.5	MI
Cleveland, Akron	615.4	OH
Los Angeles, Riverside, Orange County	634.2	CA, NV, AZ
Seattle, Tacoma, Bremerton	609.0	WA
Portland, Salem	587.1	OR, ID
Atlanta	613.8	AL, FL, GA, MS, NC, SC, TN
Minneapolis, St. Paul	624.1	MN
Milwaukee, Racine	592.5	WI
Cincinnati, Hamilton	593.6	OH, KY
Kansas City	604.7	MO, KS, CO, IA, MT, NE, ND, SD, UT, WY
Washington DC	599.6	VA
Baltimore	598.7	MD
Dallas, Fort Worth	580.8	AR, LA, NM, OK, TX

This table shows the city price data as described in the text and the assignments to states made in order to estimate the real personal income data used in the paper. Data sources are described in the text.

Table 2: Descriptions of Control/Split Variables

Mnemonic	Description	US Average	Sample Coefficient of Variation
AFDC	Percentage of children receiving AFDC (1951, p. 247)	3.23	0.524
AGE	Median age of the population (1960, p. 27)	30.3	0.098
AGEMP	Fraction of total employed that are employed in agriculture, forestry, and fisheries (1953, pp. 207-8, 211)	0.125	0.625
AIRAREA	Number of airports per 1000 square miles (1951, p. 523)	1.796	0.838
AIRPOP	Number of airports per 100,000 persons (1951, p. 523)	4.216	1.093
BANKAREA	Number of banks, credit unions, and S&L's per 1000 square miles (1952, pp. 395, 414 and 424)	8.763	1.465
BANKPOP	Number of banks, credit unions, and S&L's per 100,000 persons (1952, pp. 395, 414 and 424)	20.57	0.428
BILO	Percentage of persons born in the state living in another state in 1940 (1952, p. 38)	22.5	0.577
BIRTH	Live births per 1000 persons (1960, p. 54)	24.1	0.127
BOLI	Percentage of persons born outside the state living in the state in 1940 (1952, p. 38)	22.5	0.333
BOOKS	Number of volumes in public libraries in 1945 per 1000 persons (1953, p. 133)	821.0	0.823
BORROW	Number of registered borrowers at public libraries in 1945 as a fraction of 1950 population (1953, p. 133)	0.151	0.416
CABAREA	Miles of (aerial and underground) telephone cable (class A & B carriers only) per square mile. (1952, p. 461)	41.35	1.783
CABPOP	Miles of (aerial and underground) telephone cable (class A & B carriers only) per 1000 persons (1952, p. 461)	970.6	0.395
CAPFRAC	Fraction of state government spending on capital items (1952, p. 366)	0.160	0.310
COLLEGE	Fraction of population 25 years of age or older with a college degree (1960, p. 111)	0.060	0.210
DEATH	Deaths per 1000 persons (1960, p. 63)	9.6	0.104
DEMOCRAT	Percentage vote for Democratic Party candidates in 1950 House of Representatives election (1952, p297)	48.9	0.357
FAMSIZE	Mean family (group of related persons living together) size (1960, p. 44)	3.595	0.058
FARM	Value of all farm products sold per farm worker (1952, p. 601)	7088.7	0.423
FARMVAL	Mean value per acre of farm land (1952, p. 575)	66.64	0.706
HIED	Number of institutions of higher education per 1000 persons under 20 years of age (1953, p. 119 and 1960, p. 27)	3.605	0.383
HIWAY	Federal highway projects completed during fiscal year, under construction, or approved but not yet under construction as of June 1, 1951 per 100000 persons (1952, p. 494)	26.87	1.328
HOSAREA	Number of hospital beds per 1000 square miles (1952, p. 82)	408.7	1.725
HOSPOP	Number of hospital beds per 1000 persons (1952, p. 82)	9.59	0.240
INFANT	Infant (children aged less than 1 year) deaths per 1000 live births (1960, p. 67)	29.2	0.213
INSCHO	Fraction of 5 to 19 year olds in primary or secondary school (1960, pp. 27 and 116)	0.718	0.078
IONVA	Manufactures expenditure on new plant and equipment in 1947 as a fraction of value added in manufacturing (1952, pp. 782 and 786)	0.067	0.451

KWPOP	Electrical generation capacity, kW per 1000 persons (1952, p. 480)	545.6	0.831
LUNCH	Percentage of children participating in school lunch programs (1952, p. 91)	30.3	0.296
MALE	Fraction of the population 14 years and older that is male (1960, p. 39)	0.505	0.030
MANEMP	Fraction of total employed that are employed in manufacturing (1953, p. 211)	0.259	0.543
MARRIAGE	Number of marriages per 1000 persons (1952, p. 80)	11.0	2.373
MARRIED	Fraction of males aged 14 years or more who are married (1960, p. 39)	0.675	0.032
MFG	Value added in manufacturing per worker (1952, p. 782)	6152.4	0.202
MURDER	Murders per 100,000 persons (1952, p. 114)	5.11	0.891
NEWBUS	Gross new firms as a fraction of number of existing firms (1952, p443)	0.100	0.220
NOTAXES	Fraction of state government spending financed by means other than taxation (1952, p. 366)	0.400	0.183
PATENTS	Number of patents and designs issued 1941 to 50 (1951, p. 450)	208.46	1.093
PHAREA	Number of telephones per square mile (1952, p.461)	10.635	1.787
PHPOP	Number of telephones per 1000 persons (1952, p. 461)	249.65	0.324
POAREA	Number of post offices per 1000 square miles (1952, p. 475)	11.504	0.867
POPDEN	Population per square mile (1960, p. 13)	42.6	1.507
POPINFAM	Fraction of population living in a family (1960, p. 44)	0.909	0.024
POPOP	Number of post offices per 100,000 persons (1952, p. 475)	27.004	0.614
PROFEMP	Fraction of total employed that are employed in professional or technical occupations (1953, pp. 207-8, 211)	0.087	0.165
PUBEMP	Fraction of total employed that are local, state, or federal (civilian) government employees (1952, p. 364 and 1953, p. 211)	0.111	0.237
PUPTCH	Pupils per teacher in public primary and secondary schools (1953, p. 121)	24.4	0.144
RAILAREA	Miles of railroad per 1000 square miles (1952, p. 507)	62.77	0.587
RAILPOP	Miles of railroad per 1000 persons (1952, p. 507)	1.474	0.893
SAMHSE	Percentage of population living in the same house in 1950 as in 1949 (1960, p. 37)	81.0	0.067
SCHEXP	Per pupil expenditures "chargeable to pupils" in public primary and secondary schools (1953, p. 122)	208.83	0.250
SCHOUP	Percentage increase in SCHOYRS from 1940 to 1950 (1952, p. 114 and 1960, p. 111)	10.71	0.351
SCHOYRS	Median school years completed by those 25 years of age or older (1960, p. 111)	9.3	0.115
TAXES	State government tax revenue (excluding unemployment compensation) as a fraction of total personal income (1952, p. 366)	0.035	0.278
UNION	Year in which the state ratified the US constitution.	1834.7	0.023
URBAN	Fraction of population living in urban areas (1960, p. 16)	0.588	0.302
VECHPOP	Number of registered motor vehicles per 1000 persons (1952, p. 497)	319.81	0.213
VOC	Number of students enrolled in federally funded vocational training programs as a fraction of the population 14 years of age and over (1953, p. 136)	0.031	0.602

This table gives the mnemonic, description, US average and sample coefficient of variation for the variables used as control and split variables. The US average is that for the US as a whole and not simply the average of the figures for the 48 contiguous states. The former includes Alaska and Hawaii, which did not become states until 1959, as well as the District of Columbia. Each variable is measured in 1950 except as noted. The references in parentheses give the year and page number of the edition of the *Statistical Abstract of the United States* where the data used to construct the variable may be found.

Table 3: Tests for Conditional and Club Convergence

	Estimates of Equation (4)				Split Tests	Estimates of Equation (4) with Dummies Added for Regression Tree Splits			
	Levels		*Logs*			*Levels*		*Logs*	
Variable	$\hat{\Pi}$	R^2	$\hat{\Pi}$	R^2		$\hat{\Pi}$	R^2	$\hat{\Pi}$	R^2
$log\,(y_{i0} / y_{u0})$	—	—	—	—	0.47	—	—	—	—
AFDC	−0.005 (0.015)	0.72	−0.007 (0.044)	0.72	0.29	−0.006 (0.012)	0.91	−0.013 (0.029)	0.91
AGE	0.039* (0.011)	0.76	1.018* (0.309)	0.76	0.98	0.0007 (0.0086)	0.91	0.029 (0.225)	0.91
AGEMP	−0.988* (0.267)	0.78	−0.150* (0.042)	0.79	2.84*	0.121 (0.226)	0.91	0.067 (0.044)	0.91
AIRAREA	0.037* (0.013)	0.78	0.150* (0.039)	0.80	3.75*	−0.008 (0.011)	0.91	0.012 (0.033)	0.91
AIRPOP	−0.012* (0.004)	0.79	−0.141* (0.026)	0.82	6.03*	−0.002 (0.001)	0.91	−0.011 (0.022)	0.91
BANKAREA	0.0034* (0.0006)	0.81	0.078* (0.017)	0.82	3.85*	0.0002 (0.0007)	0.91	0.012 (0.012)	0.91
BANKPOP	0.001 (0.002)	0.72	0.001 (0.055)	0.72	1.10	0.001 (0.001)	0.91	0.035 (0.029)	0.91
BILO	−0.004* (0.002)	0.75	−0.114* (0.046)	0.74	1.17	0.0004 (0.0012)	0.91	0.034 (0.034)	0.91
BIRTH	−0.042* (0.009)	0.81	−1.108* (0.219)	0.82	5.40*	−0.003 (0.008)	0.91	−0.032 (0.210)	0.91
BOLI	−0.008* (0.003)	0.77	−0.192* (0.082)	0.76	3.74*	−0.0003 (0.0017)	0.91	−0.002 (0.042)	0.91
BOOKS	0.00003 (0.00003)	0.73	0.009 (0.040)	0.72	0.44	−0.00003 (0.00002)	0.91	−0.040 (0.023)	0.91
BORROW	−0.220 (0.625)	0.72	−0.064 (0.089)	0.73	0.79	−0.030 (0.296)	0.91	−0.054 (0.043)	0.91
CABAREA	0.0005* (0.0001)	0.80	0.079 (0.003)	0.84	5.45*	−0.0001 (0.0002)	0.91	0.009 (0.015)	0.91
CABPOP	0.0005* (0.0001)	0.83	0.423* (0.071)	0.82	5.61*	0.00007 (0.00010)	0.91	0.062 (0.084)	0.91
CAPFRAC	−0.342 (0.384)	0.72	−0.062 (0.079)	0.72	0.73	0.098 (0.311)	0.91	0.034 (0.058)	0.91
COLLEGE	1.188 (3.680)	0.72	0.056 (0.205)	0.72	2.59	−0.255 (2.009)	0.91	−0.010 (0.109)	0.91
DEATH	0.050 (0.027)	0.74	0.489 (0.251)	0.74	3.31*	−0.028 (0.023)	0.91	−0.229 (0.222)	0.91
DEMOCRAT	0.003* (0.001)	0.74	0.175 (0.092)	0.74	1.71	0.0009 (0.0012)	0.91	0.055 (0.071)	0.91
FAMSIZE	−0.052 (0.186)	0.72	−0.166 (0.692)	0.72	0.65	−0.057 (0.100)	0.91	−0.197 (0.374)	0.91
FARM	−0.00009* (0.00002)	0.78	−0.274* (0.073)	0.77	2.62*	−0.00002 (0.00002)	0.91	−0.055 (0.054)	0.91
FARMVAL	0.0017* (0.0004)	0.79	0.142* 0.035)	0.80	2.15	0.0001 (0.0004)	0.91	0.020 (0.026)	0.91
HIED	−0.004 (0.014)	0.72	−0.001 (0.062)	0.72	1.60	−0.003 (0.009)	0.91	0.0002 (0.0401)	0.91
HIWAY	−0.0012* (0.0003)	0.78	−0.080* (0.017)	0.81	3.56*	−0.00007 (0.00021)	0.91	0.012 (0.016)	0.91
HOSAREA	0.0005* (0.0002)	0.78	0.082* (0.018)	0.83	4.60*	−0.00002 (0.00001)	0.91	0.007 (0.015)	0.91

HOSPOP	0.032* (0.013)	0.74	0.308* (0.122)	0.75	2.41	−0.001 (0.010)	0.91	0.005 (0.093)	0.91
INFANT	−0.005 (0.004)	0.73	−0.170 (0.141)	0.73	1.42	0.001 (0.002)	0.91	0.031 (0.073)	0.91
INSCHO	−1.36* (0.375)	0.77	−0.948* (0.272)	0.76	3.70*	−0.168 (0.364)	0.91	−0.075 (0.269)	0.91
IONVA	−0.922* (0.470)	0.73	−0.128 (0.072)	0.73	4.79*	−0.146 (0.460)	0.91	0.017 (0.070)	0.91
KWPOP	−0.00006 (0.00003)	0.73	−0.043 (0.047)	0.73	1.10	−0.00002 (0.00001)	0.91	−0.004 (0.027)	0.91
LUNCH	0.0005 (0.0027)	0.72	0.019 (0.097)	0.72	1.14	0.0003 (0.0021)	0.91	0.020 (0.069)	0.91
MALE	−6.819* (1.382)	0.80	−3.563* (0.710)	0.80	4.04*	−0.732 (1.205)	0.91	−0.354 (0.637)	0.91
MANEMP	0.859* (0.210)	0.79	0.154* 0.030)	0.80	3.93*	−0.136 (0.202)	0.91	−0.007 (0.030)	0.91
MARRIAGE	−0.0004* (0.0002)	0.72	−0.035 (0.026)	0.73	0.84	−0.0001 (0.0001)	0.91	−0.007 (0.014)	0.91
MARRIED	1.100 (1.342)	0.73	0.764 (0.886)	0.73	2.17	0.921 (0.822)	0.91	0.622 (0.540)	0.91
MFG	−0.00005* (0.00002)	0.74	−0.291* (0.107)	0.74	2.13	0.000007 (0.000017)	0.91	0.058 (0.110)	0.91
MURDER	0.016* (0.005)	0.75	—	—	5.33*	0.011* (0.003)	0.92	—	—
NEWBUS	0.164 (1.056)	0.72	0.038 (0.110)	0.72	1.46	0.632 (0.678)	0.91	0.053 (0.068)	0.91
NOTAXES	−0.660 (0.374)	0.74	−0.250 (0.145)	0.74	1.44	−0.381 (0.211)	0.91	−0.150 (0.085)	0.91
PATENTS	0.0005 (0.0003)	0.75	0.153* (0.046)	0.78	4.22*	−0.0001 (0.0001)	0.91	−0.036 (0.037)	0.91
PHAREA	0.0018* (0.0006)	0.79	0.078* (0.017)	0.79	5.30*	−0.0007 (0.0007)	0.91	0.005 (0.014)	0.91
PHPOP	0.0016* (0.0006)	0.76	0.252 (0.135)	0.74	3.32*	−0.0003 (0.0004)	0.91	−0.050 (0.077)	0.91
POAREA	0.005* (0.001)	0.81	0.113* (0.025)	0.82	4.68*	−0.00004 (0.00110)	0.91	0.015 (0.018)	0.91
POPDEN	0.0005* (0.0002)	0.79	0.084* (0.018)	0.83	4.88*	−0.0002 (0.0002)	0.91	0.007 (0.016)	0.91
POPINFAM	1.133 (2.05)	0.72	1.041 (1.824)	0.72	1.07	−0.329 (1.295)	0.91	−0.264 (1.157)	0.91
POPOP	−0.0035* (0.0009)	0.78	−0.147* (0.091)	0.78	3.07*	−0.0007 (0.0008)	0.91	0.0006 (0.0414)	0.91
PROFEMP	1.054 (3.32)	0.72	0.040 (0.273)	0.72	3.44*	−0.237 (1.898)	0.91	−0.013 (0.156)	0.91
PUBEMP	−3.188* (0.797)	0.78	−0.393* (0.110)	0.78	1.74	−0.566 (0.662)	0.91	−0.061 (0.082)	0.91
PUPTCH	0.012 (0.009)	0.73	0.293 (0.185)	0.73	1.31	−0.003 (0.006)	0.91	−0.061 (0.124)	0.91
RAILAREA	−0.048* (0.011)	0.81	−0.140* (0.030)	0.83	5.96*	0.00004 (0.00033)	0.91	0.024 (0.023)	0.91
RAILPOP	0.0015* (0.0004)	0.78	0.140* (0.033)	0.80	6.17*	−0.008 (0.007)	0.91	0.005 (0.031)	0.91
SAMHSE	0.010* (0.004)	0.74	0.778* (0.343)	0.74	2.71*	−0.003 (0.003)	0.91	−0.217 (0.245)	0.91
SCHEXP	−0.002 (0.001)	0.74	−0.369* (0.178)	0.74	1.43	−0.001 (0.001)	0.91	−0.123 (0.135)	0.91
SCHOUP	−0.008 (0.006)	0.73	−0.063 (0.072)	0.73	1.74	−0.005 (0.003)	0.91	−0.060 (0.036)	0.91
SCHOYRS	−0.092* (0.029)	0.77	−0.890* (0.293)	0.76	2.20	−0.034 (0.019)	0.87	−0.323 (0.189)	0.91

TAXES	−0.065* (0.024)	0.75	−0.316* (0.093)	0.76	2.29	−2.156 (1.304)	0.91	−0.100 (0.062)	0.91
UNION	−0.0029* (0.0005)	0.84	−5.375* (0.992)	0.84	11.21*	−0.0001 (0.0005)	0.91	−0.218 (0.970)	0.91
URBAN	0.671* (0.226)	0.77	0.359* (0.110)	0.77	2.15	−0.206 (0.202)	0.91	−0.060 (0.100)	0.91
VECHPOP	−0.0017* (0.0003)	0.80	−0.567* (0.117)	0.80	6.72*	0.00002 (0.00032)	0.91	0.026 (0.107)	0.91
VOC	−0.015 (0.016)	0.73	−0.061 (0.052)	0.73	3.78*	0.001 (0.011)	0.91	0.041 (0.044)	0.91

The left-hand columns of this table show estimates of the parameter Π in equation (4) with each of the control variables included separately both in level and log form. MURDER is not included in log form as the murder rate in two states was zero in 1950. Below each parameter estimate is shown the estimated heteroskedasticity-consistent standard error. The middle column shows the test statistic for tests of parameter constancy when the variable in question is used to split the sample as described in the text. The right-hand columns repeat the tests in the left-hand columns after dummy variables as defined by the regression tree procedure are added to the regression. A * indicates significance at the 5% level. Data sources and definitions are given in the text.

Table 4: Growth Regressions

	Baseline $(\alpha^2=\alpha^3=\alpha^4=0)$ $(\beta_t^2=\beta_t^3=\beta_t^4=0)$	**Intercept Dummies** $(\beta_t^2=\beta_t^3=\beta_t^4=0)$	**Slope and Intercept Dummies**	**Nodes 1 & 4 Combined** $(\alpha^4=\beta_t^4=0)$	**Restricted** $(\alpha^4=\beta_t^4=0)$ $(\beta_t^3=0)$
α	-0.0900^* (0.0316)	-0.0792^* (0.0316)	-0.0928^* (0.0325)	-0.1051^* (0.0200)	-0.1063^* (0.0198)
β_t	1.3283^* (0.1045)	1.3828^* (0.0862)	1.4511^* (0.1027)	1.4900^* (0.0808)	1.5010^* (0.0782)
α^2	—	0.1940^* (0.0566)	0.2135^* (0.0517)	0.2258^* (0.0449)	0.2270^* (0.0448)
β_t^2	—	—	-0.4925^* (0.1794)	-0.5314^* (0.1678)	-0.5424^* (0.1666)
α^3	—	-0.2482^* (0.0436)	-0.2597^* (0.0617)	-0.2473^* (0.0562)	-0.2329^* (0.0386)
β_t^3	—	—	0.1823 (0.2772)	0.1434 (0.2698)	—
α^4	—	-0.0391 (0.0444)	-0.0132 (0.0435)	—	—
β_t^4	—	—	0.1762 (0.1914)	—	—
R^2	0.72	0.89	0.91	0.91	0.91

This table shows estimates of the parameters of the model $\Gamma_{it} = \alpha + \alpha^2 D_i^2 + \alpha^3 D_i^3 + \alpha^4 D_i^4 - (\beta_t + \beta_t^2 D_i^2 + \beta_t^3 D_i^3 + \beta_t^4 D_i^4)(log\, y_{i0} - log\, y_{u0}) + \epsilon_i$ where $D_i^n = 1$ if state i is in node n, for $n = 2, 3, 4$. Numbers in parentheses below parameter estimates are estimated heteroskedasticity-consistent standard errors. A * indicates significance at the nominal 5% level but, as discussed in the text, conventional significance values apply to only the first column. Data definitions and sources are given in the text.