Output Convergence in the United States: Evidence from State-Level Data

Chan Wang

qwang1@macalester.edu

Macalester College

Introduction to Econometrics

Gary Krueger Advisor

Abstract

The classical Solow model of economics development posits a negative relationship between GDP per capita and economics growth rate and therefore predicts output convergence. In this paper, I use Barro and Sala-i-Martin's (1990) specification to estimate the output convergence rate among 50 U.S states by 10 year subperiods from 1930 to 2015. The convergence rate significantly declines after 1980s but improves when I control for regional shocks, sectoral shocks and education attainment. I conclude that while the Solow model is still applicable to the current U.S economy, including these additional controls helps improve its significance and robustness.

I. INTRODUCTION

In his seminal paper, Solow (1956) ushered in a new field in neoclassical economics, income convergence. The principles behind his model are strikingly simple: steady state output is determined by capital per labor and exogenous factors – the savings rate, the capital depreciation rate, and the population growth rate. Additionally, capital per labor exhibits diminishing marginal returns. His model predicts that capital will flow from capital-intensive to capital-scarce economies, inducing output convergence in the long run. Research by Barro and Sala-i-Martin (1990) as well as Mankiw (1992) has empirically tested the convergence hypothesis. Barro and Sala-i-Martin (1990) studied convergence among US states. They included a regional dummy and a sector composition variable in their specifications to control for regional and sectoral shocks. Mankiw (1992) introduced human capital per labor into the Solow model, thereby improving its statistical fit and reducing the implausibly high capital share of income it implied.

This paper examines output convergence among the US states using Barro and Sala-i-Martin's (1990) specification for β-convergence. It contributes to the existing literature by testing the effectiveness of β-convergence models with data up to 2015, and by suggesting additional controls for Barro and Sala-i-Martin's (1990) specification. This paper is organized as the following. In section II, I discuss the classical Solow model, Barro and Sala-i-Martin's (1990) model, and past empirical research. In section III, I introduce the conceptual model. In section IV, I discuss the ideal data and measurement of concepts in the model. Section V describes my data and estimation equation. In section VI, I discuss my results. Section VII concludes.

II. LITERATURE REVIEW

II.a. Classical Solow Model

The Solow model begins by assuming a Cobb-Douglas production function with constant returns to scale and diminishing marginal returns to capital:

$$Y = F(K,L) = K^a L^{1-a} \ (0<a<1)^1 \qquad (1)$$

Since output per labor is $y = Y/L$ and capital per labor is $k = K/L$, dividing both side of (1) by L, we obtain:

$$y = k^a \ (0<a<1) \qquad (2)$$

[1] The production function omits total factor productivity (A) because it's expressed as the labor augmenting technology (E) in the equation $Y = F(K,EL)$. I left E out of my discussion to simplify the model.

The second assumption of Solow model is: income (Y) is saved and reinvested at the savings rate (s) and capital stock depreciates at rate δ. Both rates are exogenously given. Therefore, the change in capital (K') is given by:

$$K' = sY - \delta K \tag{3}$$

Solow then assumes that labor (L) is the entire population and grows at a fixed rate n:

$$L_t = L_0 e^{nt} \tag{4}$$

Taking log on both sides of the capital per labor equation $k = K/L$, we obtain:

$$\log k = \log K - \log L \tag{5}$$

Taking the derivative on both sides of (5) and substituting L for (4), we obtain:

$$k'/k = K'/K - L'/L = K'/K - n \tag{6}$$

Substituting K' for (3), we obtain:

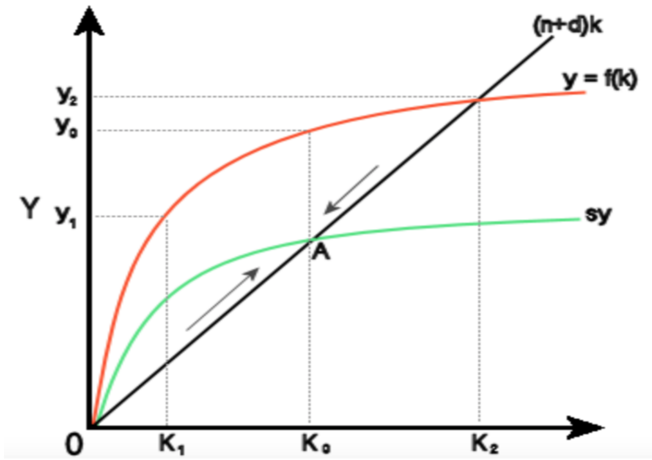
$$k'/k = (sY - \delta K)/K - n \tag{7}$$

Rearranging (7), we obtain:

$$k' = sy - (\delta+n)k \tag{8}$$

Therefore, the change in capital per labor (k') is a function of the savings rate (s), capital depreciation rate (δ) and population growth rate (n). If $k'=0$, the inflow of new capital (sy) equals the outflow of capital $(\delta+n)k$. We arrive at the steady state capital per labor ($k0$) and output per labor ($y0$) according to Figure 1. If an economy starts below the steady state at $k1$ and $y1$, $k' = sy - (\delta+n)k > 0$, capital and output per labor grow to approach $k0$ and $y0$. If an economy starts above the equilibrium at $k2$ and $y2$, $k' = sy - (\delta+n)k < 0$, capital and output per labor decline to approach k0 and y0. Therefore, equation (8) predicts that holding all else constant (s, δ and n), poorer economies grow faster than richer economies and income converges.

Figure 1. Classical Solow Growth Model and the Steady State Output (y_0)



II.b. Barro and Sala-i-Martin's (1990) Specification

To model the inverse relationship between initial output and growth rate implied by the classical Solow model, Barro and Sala-i-Martin (1990) derived the following equation:

$$(1/T)log(y_{i,t+T}/y_{i,t}) = a - (1-e^{-\beta t})/T*log(y_{i,t}) + (other\ variables) + u_{i,t,t+T} \qquad (9)$$

where $y_{i,t}$ is the output per labor of state i at time t; $y_{i,t+T}$ is the output per labor of state i at time $t+T$; and $u_{i,t,t+T}$ is the error term between time t and $t+T$. Therefore, T is the period that measures the convergence.

On the left side of equation (9), $(1/T)log(y_{i,t+T}/y_{i,t})$ is the growth rate of output per labor between time t and $t+T$. On the right side of equation (9), a is the intercept; $-(1-e^{-\beta t})/T$ is the coefficient of $log(y_{i,t})$, the initial output; and β is the convergence coefficient.

II.c. Empirical Research Using Barro and Sala-i-Martin's (1990) Specification

Barro and Sala-i-Martin (1991) observed a robust convergence rate of 2% among the US states and European countries when controlling for regional and sectoral shocks.

Hart (2001) replicated Barro and Sala-i-Martin (1991) using updated data from the 2000 census and observed a robust convergence rate of 1.7% among the US states when controlling for regional and sectoral shocks. He attributed the low convergence rate in 1930s and 1980s to the great depression and stagflation, respectively.

Higgins (2006) tested convergence among the US states using county level data and 41 control variables. He observed a robust convergence rate of 1.6% when controlling for geographical region, state government size, education attainment and the size of the finance, insurance, real estate and entertainment sectors.

When using county level data from 1970 to 1998, Young (2008) found that convergence among the US states is significant but convergence among counties within each individual state is not always significant. Nevertheless, the convergence rates are always positive when significant.

III. CONCEPTUAL MODEL

In this paper, I use Barro and Sala-i-Martin's (1990) specification:

$$(1/T)log(y_{i,t+T}/y_{i,t}) = a - (1-e^{-\beta t})/T*log(y_{i,t}) + (other\ variables) + u_{i,t,t+T} \qquad (9)$$

If we assume all states have the same steady state output, we do not need to control for the exogenous determinants of the steady state: the savings rate (*s*), the population growth rate (*n*) and the capital depreciation rate (δ). This assumption is tenable among the US states, where these exogenous factors are assumed to be identical under the same institutions, social norms, and culture.

However, if we assume that each state has a different steady state output, we need to control for *s*, *n* and δ by using the *other variables* option on the right side of equation (9).

IV. IDEAL DATA AND CONCEPT MEASUREMENT

The output growth rate $(1/T)log(yi,t+T/yi,t)$ and the initial output $log(yi,t)$ can be calculated from annual GDP per capita by state, or approximated with annual personal income per capita by state. The savings rate s is equal to the physical capital investment rate plus the human capital investment rate (Mankiw 1992). Unfortunately, neither are available by state as time series data. Besides, it is hard to define human capital investment, which includes education, job training and government programs among others. Therefore, past literature typically uses the percentage of population holding a bachelor's degree as a proxy for the human capital investment rate.

Capital depreciation δ is the rate at which the economic value of the capital stock decreases. Unfortunately, it is also not available by state as time series. Past literature typically fixes capital depreciation at 5% per year. Since it is assumed to be a constant, I do not control for capital depreciation in my specifications.

The population growth rate can be calculated from the census in 10-year intervals or approximated with population estimates from the Bureau of Economic Analysis (BEA).

V. DATA AND ESTIMATION

The equation I estimate is:

$$Output\ Growth\ Rate = \beta_0 + \beta_1\ Initial\ Output + \beta_2\ Regional\ Dummy + \beta_3\ Sectoral\ Composition + \beta_4\ Population\ Growth\ Rate + \beta_5\ Percentage\ of\ Population\ with\ Bachelor's\ Degree + \mu$$

My dependent variable is *Output Growth Rate*, calculated from BEA annual personal income per capita by state. My key independent variable is *Initial Output*, also calculated from BEA personal income by state.

Regional Dummy is a vector of four dummy variables that divides the states into four BEA regions: Northeast, Midwest, South and West.[2] Each dummy variable has a value of one for states that belong to the region and zero otherwise. This variable controls for regional shocks.

Sectoral Composition is calculated by:

$$S_{it} = \sum_{j=1}^{9} \frac{W_{ijt} \log\left(\frac{y_{i,t+T}}{y_{it}}\right)}{T}$$

where W_{ijt} is the total income share of sector j in state i at time t.[3] $log(y_{i,t+T}/y_{it})/T$ is the national growth rate of sector j from time t to $t+T$. There are 9 sectors. Therefore, S_{it} is the output growth rate of state i at time t if each of its sectors grows at the national rate. *Sectoral Composition* controls for sectoral shocks. For example, if sector j suffers a shock between time t and $t+T$, $log(y_{i,t+T}/y_{it})/T$ will be low. States with high total income shares of sector j at time t will have a low S_{it}. All the inputs of *Sectoral Composition* are from BEA personal income by state and by industry.

Population Growth Rate and *Percentage of Population with Bachelor's Degree* are obtained from the decennial census. They control for steady state output.

---

[2] The regions are made up of the following states:
Northeast: ME, NH, VT, MA, RI, CT, NY, NJ, PA.
South: DE, MD. VA, WV, NC, SC, GA, FL, KY, TN, AL, MS, AR, LA, OK, TX.
Midwest: MN, IA, MO, ND, SD, NE, KS, OH, IN, IL, MI, WI.
West: MT, ID, WY, CO, NM, AZ, UT, NV, WA, OR, CA

[3] From 1930 to 2001, BEA personal income by state by industry has 9 industrial sectors: agriculture, mining, construction, manufacturing, wholesale and retail trade, finance and insurance, services, government, and transportation and communication. After 2001, these 9 sectors are split into 21. I aggregated these 21 sectors back into 9. My industrial classification is therefore consistent in this paper from 1930 to 2015.

Table 1 – Summary Statistics of Variables

Variable Name	**Mean**	**Std. Deviation**	**Minimum**	**Maximum**
Output Growth Rate	5.32%	0.38%	4.64%	6.11%
Initial Output	6.22	0.38	5.36	6.94
Percentage of Population with Bachelor's degree	15.3%	2.7%	9.9%	21.2%
Population Growth Rate	1.12%	0.83%	0.07%	4.05%
Sectoral Composition	0.062	0.002	0.058	0.066

Note: output growth rate, percentage of population with bachelor's degree, population growth rate and sectoral composition are aggregated over the full sample period 1930 – 2015. Initial output refers to 1930.

Table 2 – Summary of Independent Variables by Availability, Source, Expected Sign and Economic Concept

Dependent Variable is Output Growth Rate					
Variable Name	**Availability**	**Source**	**Expected Sign**	**Justification of Expected Sign**	**Economic Concept**
Initial Output	1930 - 2010	BEA	Negative	Convergence predicts that the higher the initial output is, the lower the growth rate is	Initial Output
Percentage of Population with Bachelor's Degree	1940 - 2010	Census	Positive	The higher the human capital investment rate is, the higher the steady state output is and the higher the growth rate is	Human Capital Investment Rate
Population Growth Rate	1930 - 2010	Census	Negative	The higher the population growth rate is, the lower the steady state output is and the lower the growth rate is	Population Growth Rate
Regional Dummy	1930 - 2010	BEA	Indeterminate	The signs correspond to the directions of the regional shocks	Regional Shock
Sectoral Composition	1930 - 2010	BEA	Positive	The larger the sectoral composition variable is, the faster the state's sectors grow and the higher the growth rate is	Sectoral Shock

VI. RESULTS

VI.a. Assuming the Same Steady State Output for All States

If we assume all states have the same steady state output, we do not need to control for the exogenous determinants of the steady state. Therefore, I estimate three models. Model 1 is the basic model with initial output as the only independent variable. Model 2 is the basic model plus the regional dummy. Model 3 is the basic model plus the regional dummy and the sectoral composition variable. Each model is estimated in eleven periods: seven 10-year periods from 1930 to 2000, three 5-year periods from 2000 to 2015, and one full sample period from 1930 to 2015.

Each regression estimates $-(1-e^{-\beta t})/T$, the coefficient of initial output, from which I calculate the implied convergence rate β. Intriguingly, the coefficient of initial output is inversely related to the convergence rate. If $-(1-e^{-\beta t})/T=0$, then $\beta=0$, implying that the initial output has no effect on the growth rate and thus no convergence. If $-(1-e^{-\beta t})/T<0$, then $\beta>0$, implying that states with higher initial output grow more slowly and thus output *converges*. If $-(1-e^{-\beta t})/T>0$, then $\beta<0$, implying that states with higher initial outputs grow faster and thus output *diverges*.

Table 3 – Coefficients of Initial Output, T-Statistics, Implied Convergence Rate ß and Adjusted R-squared by Models and Estimation Periods

	Model 1 Basic Equation				**Model 2 Basic with Regional Dummy**				**Model 3 Basic with Regional Dummy and Sectoral Composition**			
	Coef.	*T*	*ß*	R^2	*Coef.*	*T*	*ß*	R^2	*Coef.*	*T*	*ß*	R^2
1930-1940	-0.012***	-4.31	1.3%	0.27	-0.004	-1.25	0.4%	0.42	-0.007***	-2.18	0.7%	0.52
1940-1950	-0.041***	-11.21	5.2%	0.73	-0.040***	-13.02	5.2%	0.87	0.039***	-11.91	5.0%	0.87
1950-1960	-0.019***	-6.45	2.1%	0.46	-0.019***	-5.51	2.1%	0.56	-0.027***	-8.17	3.1%	0.70
1960-1970	-0.025***	-7.76	2.9%	0.56	-0.018***	-5.09	2.0%	0.67	-0.022***	-5.64	2.5%	0.70
1970-1980	-0.016***	-4.00	1.7%	0.24	-0.013**	-2.79	1.3%	0.28	-0.015**	-3.26	1.7%	0.31
1980-1990	-0.003	-0.34	0.3%	-0.02	-0.003	-0.57	0.3%	0.55	-0.007	-1.39	0.7%	0.67
1990-2000	0.003	1.02	-0.3%	0.00	0.002	0.58	-0.2%	-0.05	0.008	1.59	-0.8%	0.00
2000-2005	-0.022**	-3.23	2.3%	0.17	-0.023**	-2.99	2.5%	0.23	-0.023***	-4.00	2.4%	0.59
2005-2010	-0.007	-0.59	0.7%	-0.01	-0.022*	-1.75	2.4%	0.16	-0.031**	-2.38	3.3%	0.22
2010-2015	-0.005	-0.73	0.5%	-0.01	0.001	0.09	-0.1%	0.23	-0.002	-0.30	0.2%	0.29
Full Period	-0.009***	-18.37	0.9%	0.88	-0.009***	-14.36	0.9%	0.88	-0.009***	-10.44	0.9%	0.88
No. of Obs	48				48				48			

R-squared is Adjusted
Significance level: *p<0.1, **p<0.05, ***p<0.001

Note: Hawaii, Alaska and District of Columbia are excluded according to Barro and Sala-i-Martin (1990)

Table 3 shows that over the full sample period from 1930 to 2015, US states exhibit a robust and significant convergence rate of 0.9% across all three models, lower than the 2% consensus among earlier literature. The adjusted R-squared of the basic model is 0.88, indicating that the initial output in 1930 can explain most output growth rate variations in the subsequent 85 years. The fitted line in Figure 2 shows the strong negative linear relationship between the output growth rate and the initial output. The residual plot in Figure 3 shows that the basic model is homoscedastic.

Figure 2. Regression of Output Growth Rate from 1930 to 2015 Against Output in 1930

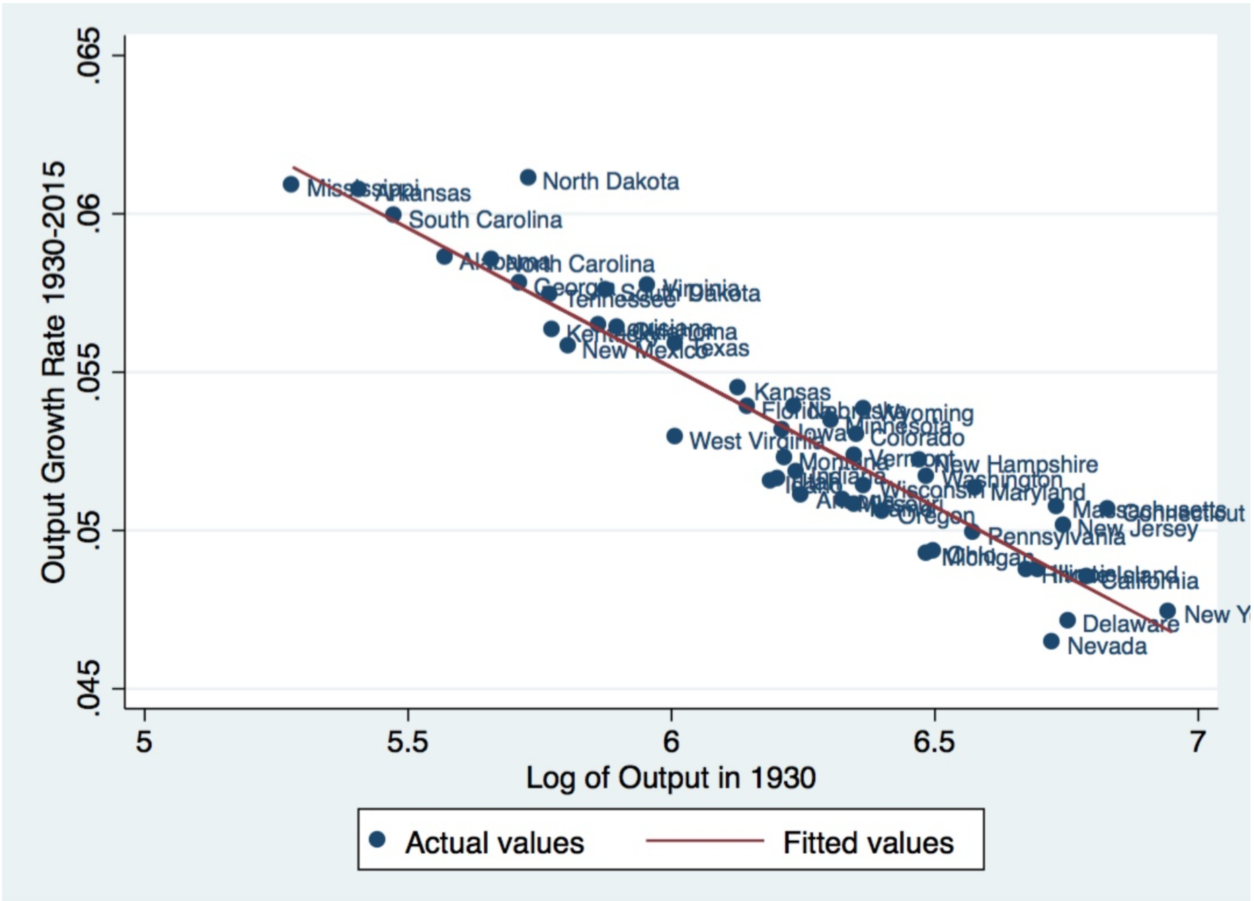
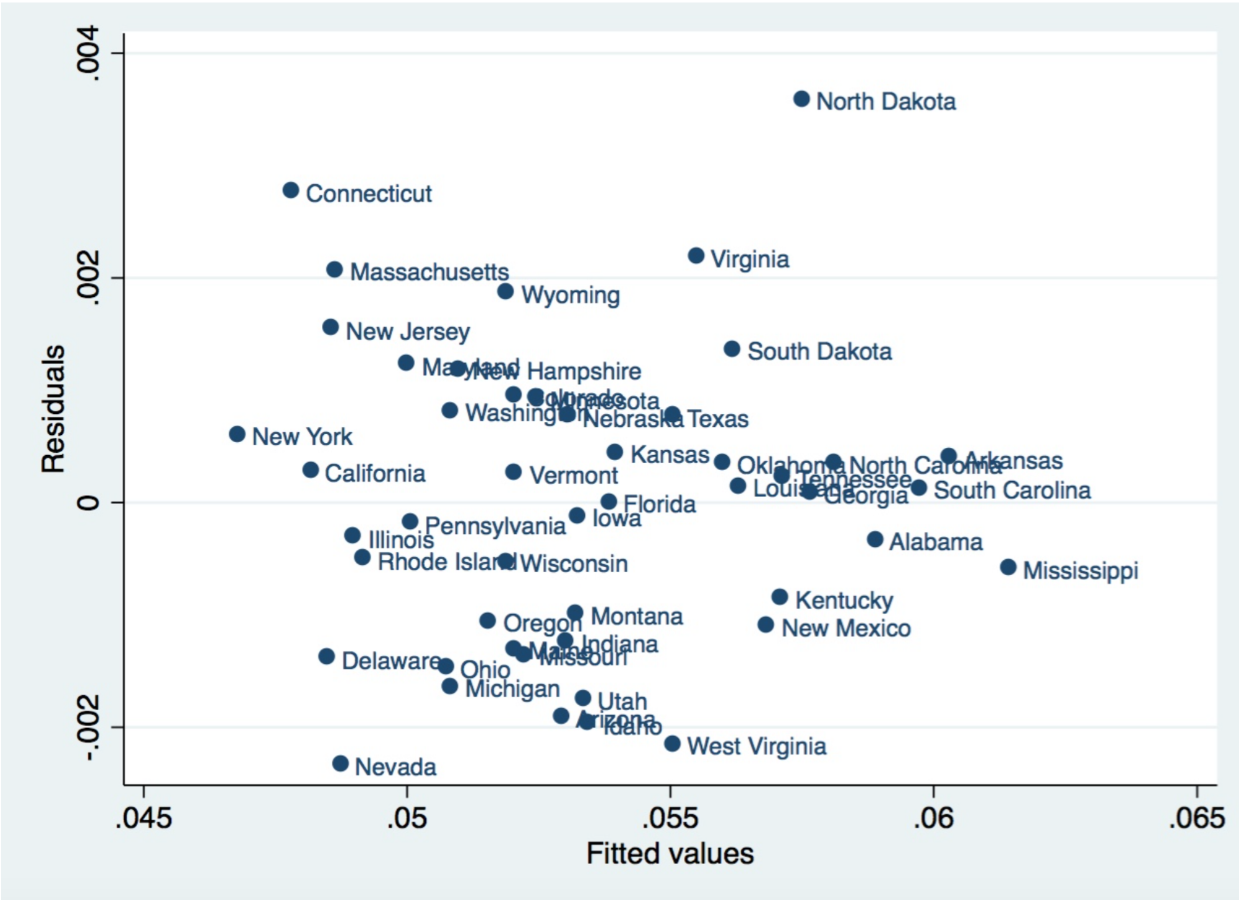


Figure 3. Rvfplot of the Regression in Figure 2



Unfortunately, output convergence in the 10-years and 5-year sub-periods is not consistent. From 1930 to 1980, β is nearly always positive and significant at 0.001 across all three models. In the 1980s, however, β suddenly drops from about 2% to about 0.5% and becomes insignificant across all three models. In the 1990s, β becomes negative, implying output *divergence*.

Figure 4. Fitted Line and Rvfplot of the 1940s and the 1990s (the Basic Model)

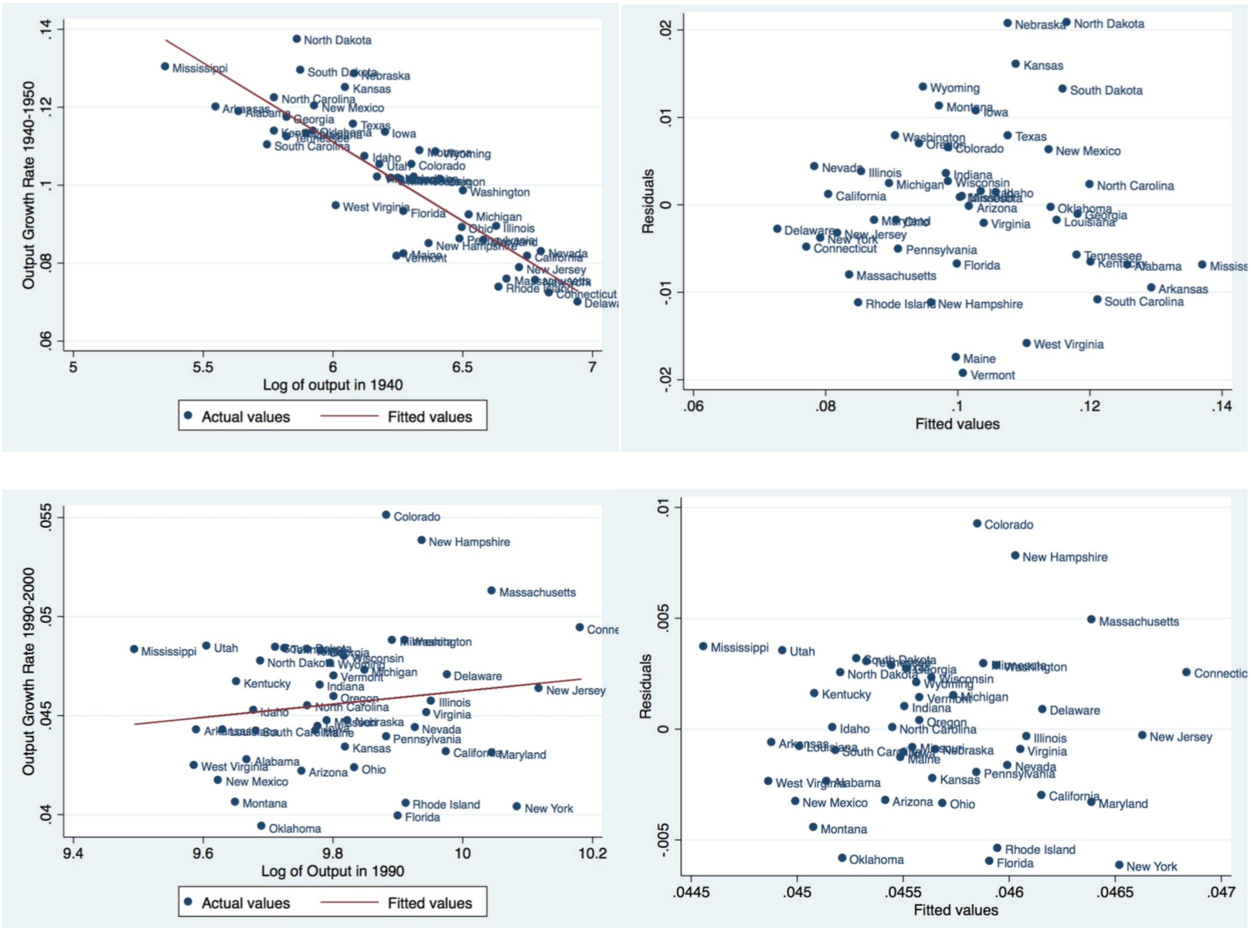


Figure 4 compares the 1940s (the most significant convergence) with the 1990s (divergence). The residual plots show that both sub-periods are homoscedastic. In fact, all the sub-periods from 1930 to 2015 are homoscedastic (Appendix). Furthermore, to examine whether weak convergence results from regional and sectoral shocks, we can compare the adjusted R-squared and T-statistics of Model 1 with those of Model 3. By adding the regional dummy and the sectoral composition variable, Model 3 improves the adjusted R-squared from -0.02 to 0.67 and the T-statistics from -0.34 to -1.39 in the 1980s. This indicates that regional and sectoral shocks can explain most output growth rate variations in the 1980s. The significance of *β* improved after controlling for these shocks but is still insignificant. In the 1990s, Model 3 improves neither the adjusted R-squared nor the T-statistics. The output *divergence* becomes more significant when controlling for regional and sectoral shocks. These imply that during the 1980s and the 1990s, output convergence was disrupted by factors other than regional and sectoral shocks. Intensive technological evolution, the emerging information technology sector, and the expanding service sector might be responsible factors. These transformations correspond with BEA industrial sector reclassification in 2001, which added the information sector and split the service sector into 9 subsectors.

β is positive and significant from 2000 to 2005 but becomes insignificant following the Great Recession. Fortunately, regional and sectoral shocks can explain the weak convergence from 2005 to 2010. When controlling for these shocks, β rises from 0.7% to 3.3% and becomes significant. The adjusted R-squared improves from -0.01 to 0.22. However, the same is not true for the period 2010 to 2015. When controlling for regional and sectoral shocks, β declines from 0.5% to 0.2% and becomes less significant.

Consequently, the full period convergence rate of 0.9% is lower than the 2% consensus among earlier literature, because the magnitude and significance of β falls after 1980. Intensive technological evolution, uncontrolled by my models, might be responsible for the weak convergence since 1980.

Figure 5. Implied ß of Model 1 (the basic equation) and Model 3 (Controlled for Regional and Sectoral Shocks)

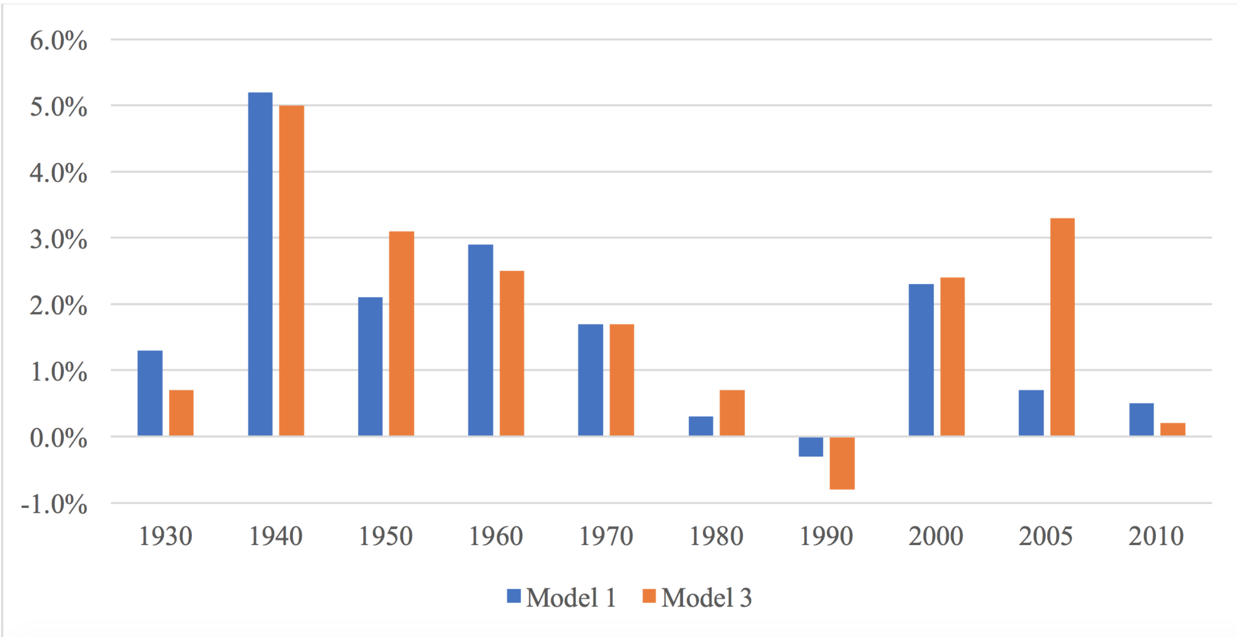


Figure 5 compares the implied ß of the basic equation and that of Model 3, which controls for regional and sectoral shocks. It shows that for the sub-periods after 1980, the implied ß becomes larger after controlling for these shocks. Besides, Table 3 shows that for the sub-periods after 1980, Model 3 improves the significance of ß and the adjusted R-squared 4 out of 5 times. Therefore, Model 3 is a good starting point for the next subsection, which relieves the assumption of identical steady state outputs.

VI.b. Assuming Different Steady State Outputs

If we assume that each state has a different steady state output, we need to control for the exogenous determinants of the steady state. In this section, I control for the population growth rate and the percentage of population with bachelor's degree, a proxy for human capital investment rate. All three models in this section control for the regional and sectoral shocks.

They are identical to Model 3 except that Model 4 controls for educational attainment (bachelor's degree), Model 5 controls for population growth and Model 6 controls for both.

Table 4 – Coefficients of Initial Output, T-Statistics, Implied Convergence Rate ß and Adjusted R-squared by Models and Estimation Periods

	Model 4 Education Attainment				**Model 5 Population Growth**				**Model 6 Education Attainment and Population Growth**			
	Coef.	*T*	*ß*	R^2	*Coef.*	*T*	*ß*	R^2	*Coef.*	*T*	*ß*	R^2
1930-1940					-0.008 88	-2.31	0.8%	0.51				
1940-1950	-0.043 ***	-9.03	5.7%	0.87	-0.037 ***	-9.35	4.6%	0.87	-0.042 ***	-8.62	5.4%	0.88
1950-1960	-0.035 ***	-8.33	4.2%	0.74	-0.030 ***	-8.66	3.5%	0.73	-0.035 ***	-8.46	4.3%	0.75
1960-1970	-0.024 ***	-4.55	2.7%	0.69	-0.026 ***	-6.05	3.0%	0.71	-0.028 ***	-5.07	3.3%	0.71
1970-1980	-0.020 **	-3.14	2.2%	0.31	-0.015 **	-3.15	1.6%	0.29	-0.020 **	-3.07	2.2%	0.29
1980-1990	-0.014 **	-2.27	1.5%	0.69	-0.007	-1.37	0.8%	0.66	-0.014 **	-2.22	1.5%	0.68
1990-2000	0.000	-0.08	0.0%	0.11	0.009 *	1.90	-0.9%	0.09	-0.001	-0.13	0.1%	0.26
2000-2005	-0.025 **	-3.02	2.7%	0.58	-0.024 ***	-3.92	1.5%	0.58	-0.027 **	-3.01	2.9%	0.57
2005-2010	-0.040 **	-2.22	4.5%	0.21	-0.028 **	-2.16	3.1%	0.21	-0.039 **	-2.15	4.4%	0.21
2010-2015	-0.027 **	-2.65	2.9%	0.42	-0.009	-1.25	0.9%	0.42	-0.025 **	-2.57	2.7%	0.48
Full Period	-0.010 ***	-14.80	1.1%	0.93	-0.009 ***	-10.37	0.9%	0.88	-0.010 ***	-14.80	1.1%	0.93
No. of Obs	48				48				48			

R-squared is Adjusted
Significance level: *p<0.1, **p<0.05, ***p<0.001

Table 4 shows that over the full sample period from 1930 to 2015, US states exhibit a robust and significant convergence rate of about 1% across all three models. The adjusted R-squared for all three models is above 0.88. When controlling for educational attainment, Model 4 and Model 6 improve the T-Statistics from -10.44 (Model 3) to -14.80 and the adjusted R-squared from 0.88 (Model 3) to 0.93. *ß* rises from 0.9% to 1.1%. In contrast, when controlling for the population growth rate, Model 5 has the same adjusted R-squared as and a lower T-Statistics than Model 3. This evidence suggests that educational attainment *might be* necessary in the specification, whereas the population growth rate *might not be*.

Table 5 – Coefficients and T-Statistics of Educational Attainment in Model 4 and the Population Growth Rate in Model 5

	Educational Attainment (Model 4)		**Population Growth Rate (Model 5)**	
	Coefficient	*T*	*Coefficient*	*T*
1930-1940	0.15	0.77		
1940-1950	-0.13	-1.03	0.0021	1.14
1950-1960	0.11 **	2.10	0.0022 **	2.73
1960-1970	0.16 *	1.88	0.0003	0.45
1970-1980	0.02	0.20	0.0005	1.10
1980-1990	0.02	0.17	0.0006 *	1.92
1990-2000	0.19 **	2.28	0.0005 **	2.51
2000-2005	0.08	0.53	0.0001	0.38
2005-2010	-0.27	-0.79	0.0004	0.76
2010-2015	0.62 **	3.27	0.0010 **	3.20
Full Period	-0.02	-0.59	0.0004 ***	5.69
No. of Obs	48		48	

Significance level: *p<0.1, **p<0.05, ***p<0.001

To better determine whether educational attainment and the population growth rate are necessary in the specification, Table 5 summarizes the coefficients and t-statistics of educational attainment in Model 4 and population growth rate in Model 5. The coefficient of educational attainment is positive and significant over the full sample period and is positive in 10 out of 10 sub-periods. This aligns with our expectation that higher human capital investment rate implies higher steady state output and higher output growth rate. In contrast, the population growth rate coefficient is negative but insignificant over the full sample period, and is positive in 8 out of 10 sub-periods. This contradicts our expectation that higher population growth rate implies lower steady state output and lower output growth rate. This anomaly might result from interstate migration: people migrate from low-growth states to high-growth states seeking better opportunities, thereby driving up population growth in states with higher output growth rate. This evidence confirms that I should keep educational attainment and drop population growth rate in my specification. Therefore, I only use Model 4 for the rest of my discussion.

Compared to Model 1, 2 and 3, Model 4 has three major improvements. First, as discussed above, Model 4 improves the overall fit and the significance of β over the full sample period. Second, Model 4 has significant β in all sub-periods except the 1990s. Third, Model 4 observes a significant convergence rate in the 1980s and the 2010s, while Model 1, 2 and 3 all fail to do so. The improved significance of β is especially obvious in the 2010s. The t-statistic rises from -0.73, 0.09 and -0.30 to -2.65. β rises from 0.5%, -0.1% and 0.2% to 2.9%. The adjusted R- squared rises from -0.01, 0.23 and 0.29 to 0.42. Therefore, educational attainment is necessary in the specification as it makes β more significant, improves the overall fit, and has theoretically justified signs.

Figure 6 – Implied β of Model 3 and Model 4

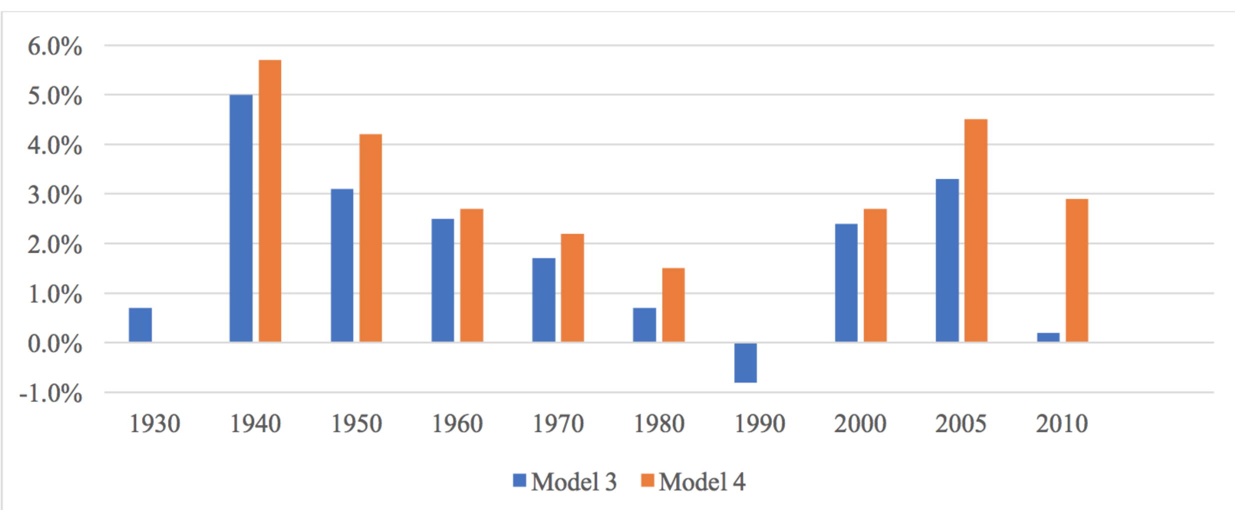


Figure 6 compares the implied β of Model 3 and Model 4, whose only difference is that

Model 4 has educational attainment in its specification. Figure 6 shows that output convergence is more pronounced when Model 4 controls for educational attainment, especially in the 2010s. This indicates that human capital, especially investment in higher education (bachelor's degree), has become more important in driving output growth recently.

VII. CONCLUSION

This paper uses Barro and Sala-i-Martin's (1990) specification and estimates a robust and significant output convergence rate of 1% among US states over the full sample period 1930 – 2015. This is markedly lower than the 2% consensus among earlier literature, because convergence is inconsistent over the sub-periods and declines after 1980. When controlling for regional and sectoral shocks, Model 3 improves the overall fit and the significance of β, but still fails to isolate a significant β from 1980 to 2000, which I attribute to intensive technological evolution after 1980. When controlling for human capital investment rates, Model 4 isolates a significant β in the 1980s and the 2010s whereas all other models fail to do so. Therefore, I conclude that educational attainment is necessary in Barro and Sala-i-Martin's (1990) specification if we assume different steady state outputs.

My research, however, is limited in three ways. First, the ten-years and five-years sub-periods are arbitrarily set up to accommodate for the availability of decennial census data. Different convergence rates might arise simply if the sub-periods are divided differently. Second, my model does not control for technological evolution, which I suspect is responsible for the weak convergence since 1980. Third, to ensure consistency, I use BEA's 9 industry classifications to calculate the sectoral composition variables throughout the entire period.

However, this approach risks losing the nuances of sectoral shocks. For example, 9 subsectors are collapsed into one service sector, which now represents 80% of US GDP.

Therefore, future research can explore the effects of technological evolution on convergence by controlling for the employment share or total income share of technology sectors by state. Besides, since controlling for educational attainment helps isolate more pronounced convergence effects and improves the overall fit, future research can also include proxies for physical capital investment rate in the specification.

REFERENCES

Barro, Robert J., and Sala-i-Martin, Xavier. "Economic Growth and Convergence across the United States." *NBER Working Paper Series*, no.3419 (1990)

Barro, Robert J., and Sala-i-Martin, Xavier. "Convergence across States and Regions." *Brookings Papers Econ. Activity*, no.1 (1991), 107-182.

Hart, James. "Government size and Convergence across the United States." *Macalester College Economics Honors Paper Series,* (2001)

Higgins, Matthew J., Levy, Daniel., Young, Andrew T. "Growth and Convergence across the United States: Evidence from County-Level Data." *The Review of Economics and Statistics*, no 88(4), (2006), 671-681.

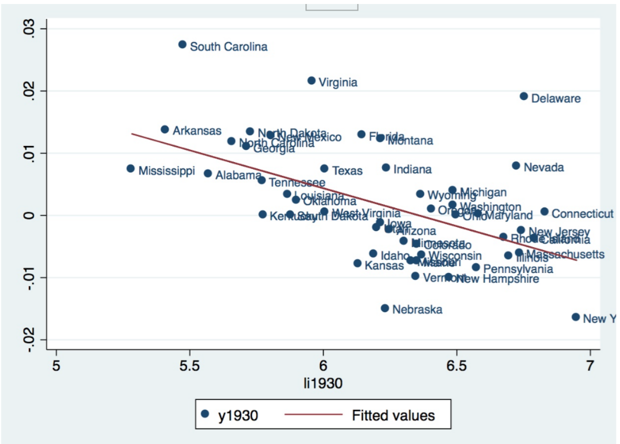
Mankiw, N.G., Romer, D., Weil, D.N. "A contribution to the empirics of economic growth." *Quarterly Journal of Economics,* 107 (1992), 407–437.

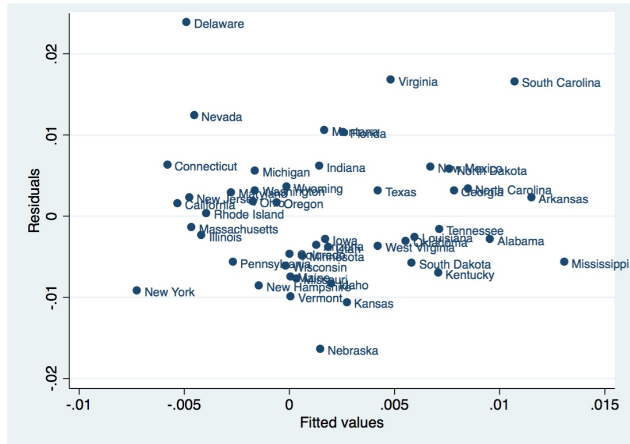
Solow, Robert M. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics*, LXX (1956), 65-94.

Young, Andrew T., Higgins, Matthew J., Levy, Daniel., "Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data." *Journal of Money, Credit and Banking*, Vol. 40, No.5 (2008), 1083-1093.

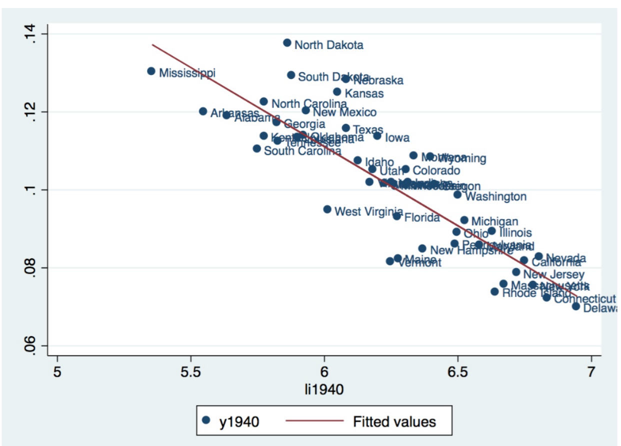
APPENDIX: Fitted Lines and Residual Plots for Each Sub-period (the Basic Model)

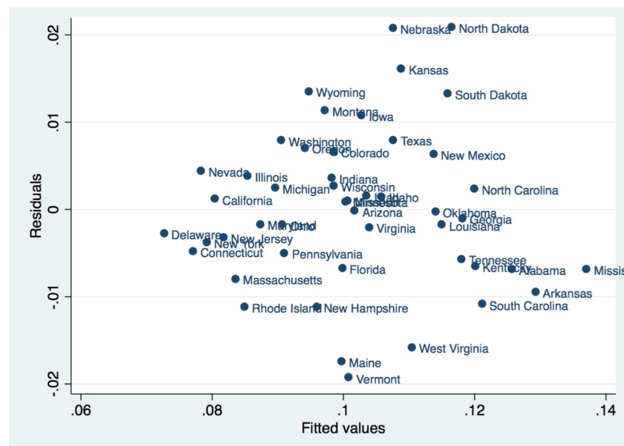
1930 – 1940



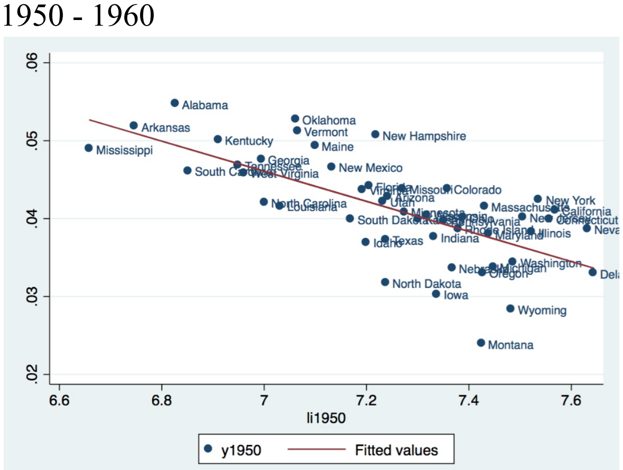


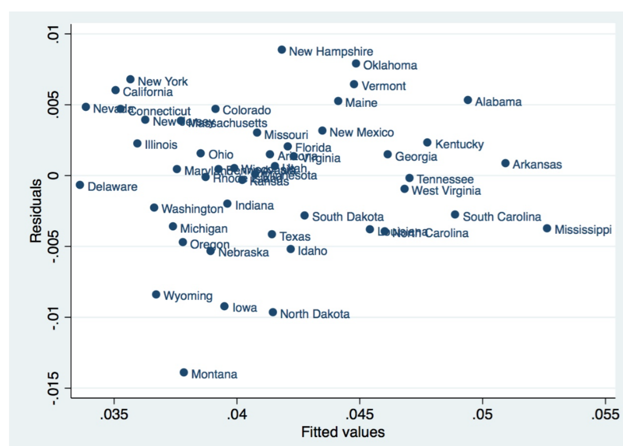
1940 - 1950



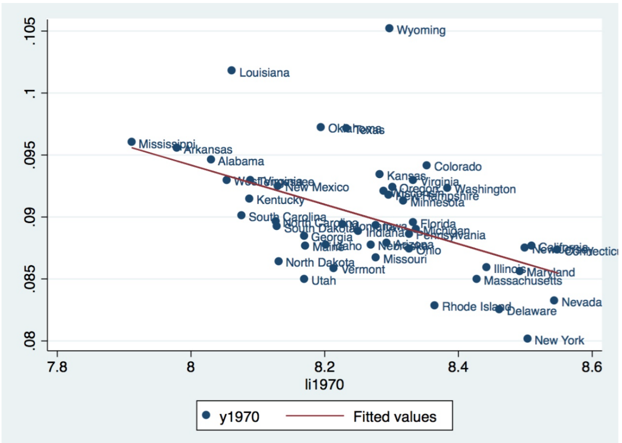


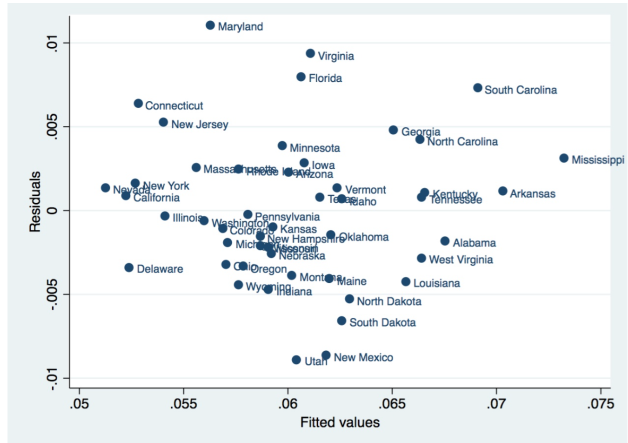
1950 - 1960



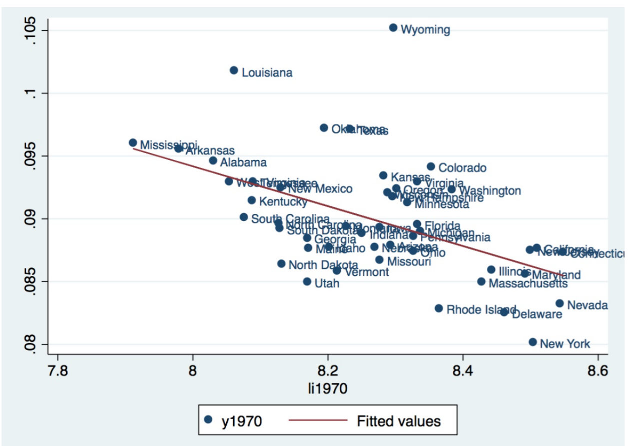


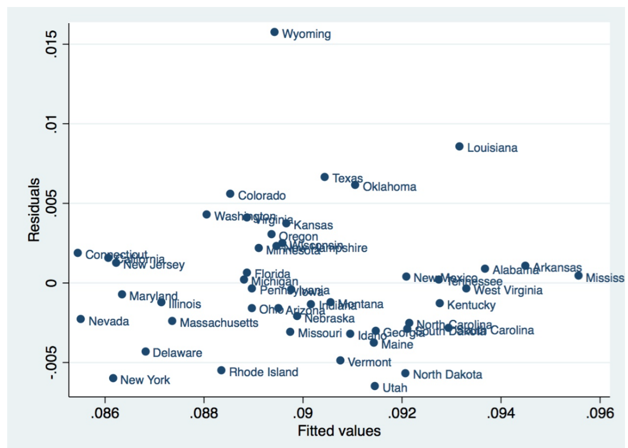
1960 – 1970



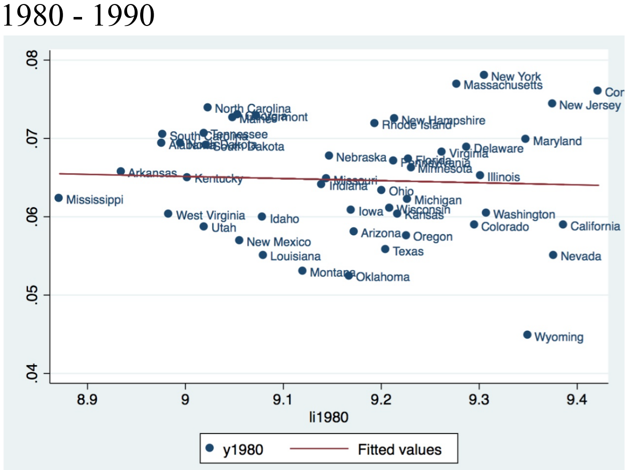


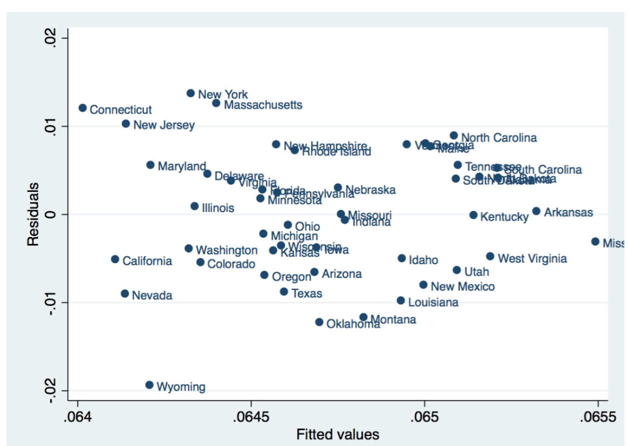
1970 - 1980



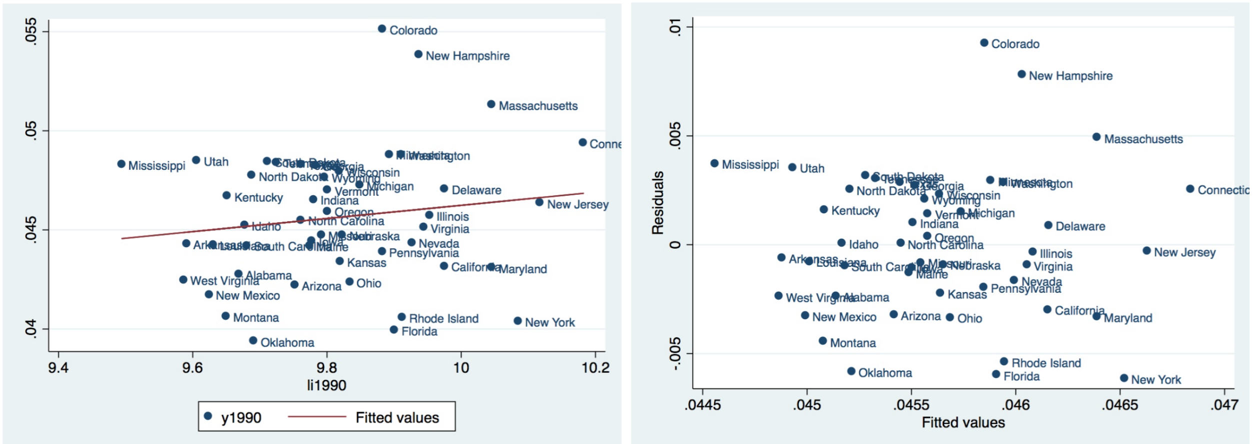


1980 - 1990

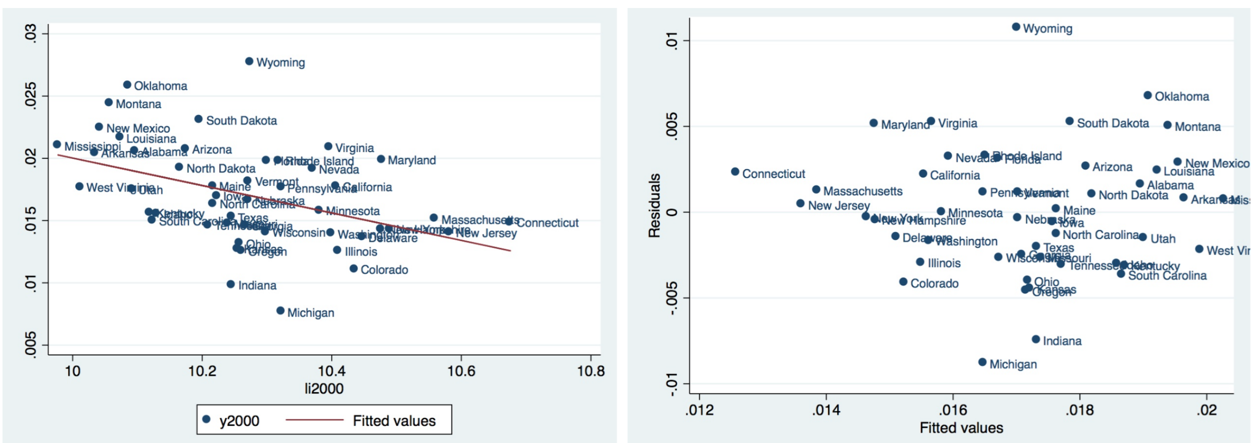




1990 – 2000

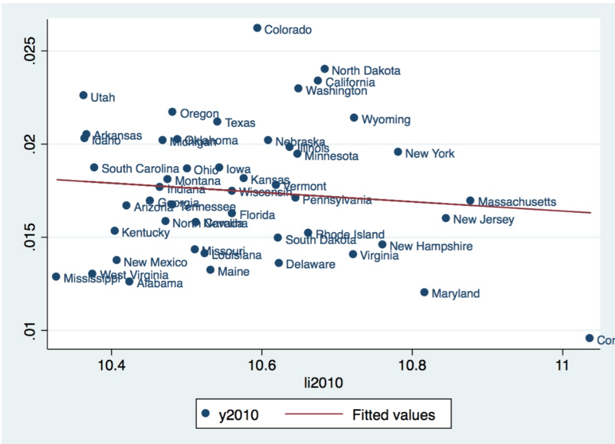


2000 – 2005

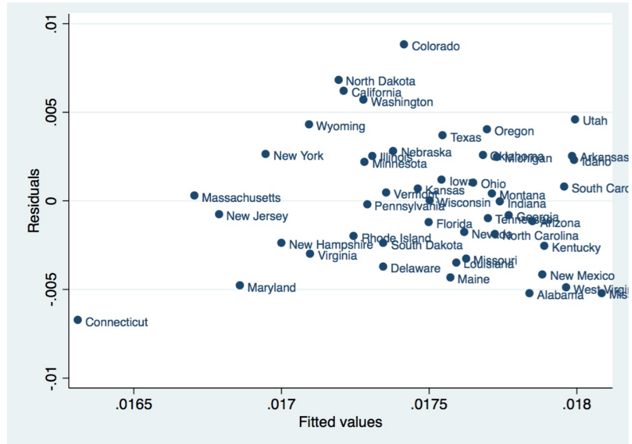


2005 – 2010

2010 – 2015



.025
.02
.015
.01
10.4
10.6
10.8
11
li2010
y2010
Fitted values
Colorado
North Dakota
California
Washington
Utah
Oregon
Texas
Wyoming
Arkansas
Idaho
Nebraska
Minnesota
New York
South Carolina
Ohio
Iowa
Montana
Indiana
Kansas
Vermont
Pennsylvania
Massachusetts
Arizona
Tennessee
Florida
New Jersey
Kentucky
Rhode Island
South Dakota
New Hampshire
Missouri
Louisiana
Virginia
New Mexico
Delaware
West Virginia
Mississippi
Alabama
Maine
Maryland



.01
.005
0
-.005
-.01
Residuals
.0165
.017
.0175
.018
Fitted values
Colorado
North Dakota
California
Washington
Utah
Wyoming
Texas
Oregon
New York
Nebraska
Minnesota
Arkansas
Idaho
Iowa
Ohio
South Carolina
Massachusetts
Kansas
Montana
Pennsylvania
Wisconsin
Indiana
New Jersey
Florida
Georgia
Arizona
North Carolina
Kentucky
New Hampshire
Rhode Island
South Dakota
Virginia
Missouri
Louisiana
Delaware
Maine
New Mexico
Maryland
Alabama
Connecticut