A Contribution to the Theory & Empirics of Regional Economic Growth

Campbell J, Low K, Rybak C

Department of Economics

New York University

Introduction

This paper has a modest goal of studying the long run variation in growth of output per worker across states of the United States of America (hereafter, USA) after the mid-70s. Specifically, can it be explained by the variations in population growth and savings across states in accordance with the neoclassical model of economic growth initiated by Robert Solow in his classic 1956 article. Using the standard ordinary least squares estimation, we extend this investigation to the human capital-augmented model by Mankiw, Romer & Weil (1992) and are unable to find any validity for both models at the state level in the 1976-2010 dataset constructed. We believe this is likely to be related to the choice of proxy variables in our regressions but cannot conclusively confirm this observation through the econometric techniques available to us.

The evolution of economic thought about economic growth traces its roots from Malthus (1798) and the more recent Domar (1946) but neither had the impact of the subsequent work of Solow (1956). The Solow model improves on Domar (1946), founded on the basis that capital accumulation puts economies on to a balanced growth path. Despite the prominence of the Solow model in the field of economic development, its success appears to be confined to its ability to explain the past growth, grounded on the process of industrialization and propagation of manufactured goods across economies. We thus look into this by examining the relevance of this growth model on the output data during 1976-2010 across the states of the USA and for the

This paper is done at part of a econometrics project for a Core Applied Statistics and Econometrics course on the MA Economics Programme

purpose of this paper, we focus solely on the key variables of output per worker and savings rate. We also assume that the technology or productivity parameter should not differ across states.

I. Literature Review

Solow (1956) initiated the neoclassical model of economic growth, and through this model economists are able to argue that steady state output per worker depends positively on the savings rate, and negatively on population growth and depreciation. Barro & Sala-i-Martin (1991) first applied the Solow model to regional economies rigorously, inspecting convergence of state economies in the USA. They determined that the neoclassical model could only work if diminishing returns to capital set in very slowly.

The Solow model was subsequently updated by Mankiw, Romer & Weil (hereafter, MRW) in A Contribution to the Empirics of Economic Growth in 1992. MRW tested the Solow model on the cross-country dataset covering the period 1960-1985. They argued that although the model predicts the direction of the effects on savings and the population growth rate, the magnitudes of coefficients on the variables were grossly off the mark. MRW (1992) believe this disparity is due to an omitted variable, namely human capital accumulation. MRW believe that including the omitted variable would help explain the cross-country differences in steady state output per worker. When MRW augmented Solows model with human capital and tested on the same dataset it generated mixed results. Notably, the MRW augmented model performed well among developing countries while doing a poor job explaining growth in OECD countries.

Garofalo & Yamarik (2002) then took the model from MRW (1992) and attempted to apply it across state economies in the USA. They focused on the period of 1977-1996, investigating if different levels of output per worker could be explained by the different savings, population growth rates, as well as human capital accumulation. Without any data on state-level investments, they apportioned the national capital stock among the states. They adopted the MRW (1992) methodology of incorporating human capital accumulation. They were able to find evidence that the Solow Model explained statewide growth during the specified time period.

Our approach utilizes the same techniques on a different dataset; using actual manufacturing capital expenditure as a proxy for investments rather than estimates generated from apportioned national level data. We obtained results that indicated that the Solow model does a poor job of explaining output per worker variations across states; a negative coefficient on the capital accumulation variable indicates different mechanics at work with the interaction between capital accumulation rate and that of output per worker. The result casts questions on our assumptions of steady state rates of growth, and the use of capital expenditure as a proxy for capital accumulation.

It also suggests the presence of an omitted variable strongly correlated with output per worker that is missing from the neoclassical growth model. We attempt to augment the model with human capital adopting the ideas of MRW, again with different data, and at state level. In this paper, we report our findings.

II. Theoretical Modelling

The relevance of the basic Solow growth model is not confined to national economies and can be applied to state level economies. The advantage of applying the model to state-level economies in the USA is that we could more easily argue that the technology available to all states is the same and the productivity parameter should not vary across states.

Theory of Growth

The model assumes the following Cobb-Douglas production function:

$$Y_t = A_t' K_t^{\alpha} L_t^{1-\alpha} \tag{1}$$

with Y being output, K being capital, L being labor and A is the transformed productivity parameter. α is a parameter between 0 and 1 that denotes the share of income attributable to capital. In per effective worker terms,

$$\frac{Y_t}{A_t L_t} = (\frac{K_t}{A_t L_t})^{\alpha} \Longrightarrow y_t = k_t^{\alpha} \tag{2}$$

Steady state is achieved when capital per effective worker in each period is constant; and this is given by:

$$(\delta + n + g) k_{ss} = s y_{ss} = s k_{ss}^{\alpha} \tag{3}$$

$$k_{ss} = (\frac{s}{\delta + n + g})^{\frac{1}{1-\alpha}} \Longrightarrow y_{ss} = (\frac{s}{\delta + n + g})^{\frac{\alpha}{1-\alpha}} \tag{4}$$

where s is an exogenously decided rate of capital accumulation, and δ is the rate of depreciation, n is the population growth and g is the technology growth.

$$\frac{Y_t}{A_t L_t} = (\frac{s}{\delta + n + g})^{\frac{\alpha}{1-\alpha}} \tag{5}$$

This result and the process by which we arrive at this result is identical to that of MRW (1992); the model implies that at steady state, the output per effective worker is stationary since the right-hand side variables are all constants in the model. Barro and Sala-i-Martin make the legitimate argument that the rate of depreciation, δ and the technology growth rate g should not vary significantly across states. Therefore, the major implication of this result is that output per worker should be positively correlated with savings rate across states; and negatively correlated with population growth, n.

Theory of Human Capital

To incorporate human capital into the basic Solow model according to MRW, we rewrite the production function:

$$Y_t = K_t^{\alpha} H_t^{\beta} (A_t L_t)^{1-\alpha-\beta} \tag{6}$$

Rewriting the function in per effective worker terms,

$$y_t = k_t^{\alpha} h_t^{\beta} \tag{7}$$

where $y = \frac{Y}{AL}$, $k = \frac{K}{AL}$ and $h = \frac{H}{AL}$. Using the same steady-state dynamics as in the previous manipulation, we obtain the following steady state levels of physical and human capital;

$$k_{ss} = \left(\frac{s_k^{1-\beta} s_h^{\beta}}{n+g+\delta} \right)^{\frac{1}{1-\alpha-\beta}} \quad h_{ss} = \left(\frac{s_k^{1-\alpha} s_h^{\alpha}}{n+g+\delta} \right)^{\frac{1}{1-\alpha-\beta}} \tag{8}$$

And substituting these values back into the production function, we get:

$$y_t = \left(\frac{s_k}{n+g+\delta} \right)^{\frac{\alpha}{1-\alpha-\beta}} \left(\frac{s_h}{n+g+\delta} \right)^{\frac{\beta}{1-\alpha-\beta}} \tag{9}$$

We observe that this is similar to the results from the basic model, with the additional term for the human capital accumulation contribution to the output per effective worker.

III. Empirical Model, Econometric Specification & Estimation Technique

Basic Solow Model and Augmented Solow Model

For the basic Solow Model, we use equation (5) to produce our econometric specification. After taking logs of the variables and rearranging, we obtain:

$$\log\left(\frac{Y_t}{L_t}\right) - \log A_t = \frac{\alpha}{1-\alpha} \log s - \frac{\alpha}{1-\alpha} \log(\delta + n + g) \tag{10}$$

Since we know that $A_t = A_0 e^{gt}$, we can thus obtain the theoretical basis for our regression:

$$\log\left(\frac{Y_t}{L_t}\right) = \log A_0 + gt + \frac{\alpha}{1-\alpha}\log s - \frac{\alpha}{1-\alpha}\log\left(\delta + n + g\right) \tag{11}$$

The first 2 terms constitute the constant of the regression; and the last 2 terms comprise of the regressors of interest and their respective coefficients. This basic econometric specification can be written in terms of variables in our dataset as:

$$lnopw = \beta_0 + \beta_1 * lnngd + \beta_2 * lninv + \epsilon \tag{12}$$

Based on the theoretical foundations of the model, we should expect a positive coefficient on the savings rate term and a negative coefficient on the $n+g+\delta$ term. The results we report will be based on this regression of interest and further investigation branches out from this specification. As for the model augmented with human capital, we have the corresponding theoretical equation:

$$\ln\left(\frac{Y_t}{L_t}\right) = \ln A(0) + gt - \frac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta) + \frac{\alpha}{1-\alpha-\beta}\ln(s_k) + \frac{\beta}{1-\alpha-\beta}\ln(s_h) \tag{13}$$

Taking logs as we did previously, we obtain our econometric specification that incorporates human capital. We now have 3 regressors; the original 2 and the additional human capital explanatory variable. We rewrite the econometric specification in terms of our dataset variables:

$$lnopw = \beta_0 + \beta_1 * lnngd + \beta_2 * lninv + \beta_3 * lnhcap + \epsilon \tag{14}$$

Our dependant variable, $lnopw$, is our proxy for steady state output per worker. Our first independent variable, $lnngd$, is the log of the $n + g + \delta$ term. The second independent variable used, $lninv$, is our proxy for state level investment, or savings. The final variable, $lnhcap$, is our proxy for human capital accumulation. As with the coefficient on $lninv$, we should expect the coefficient on $lnhcap$, the variable for human capital accumulation, to be positive. In the human capital section of our paper, we report results on this regression specification based on the data gathered and interpret our findings.

We used an Ordinary Least Squares (OLS) regression in STATA to test the models described above. In particular, we hope to test the impact of the log of the investment rate variable on the log of the output per worker.

IV. Data & Methodology

For every variable used, we had 51 observations, the 50 states of USA and the District of Columbia. More details on the data used can be found in Appendix II.

Due to the lack of a single database or survey of the data required to test the Solow Model, we had to put together our own dataset. We calculated output per worker, our independent variable, by finding the Real GDP of each state in 2010, and dividing it by the labor force for the same year. Our Real GDP figures were found by taking the Nominal GDP numbers from the Bureau of Economic Analysis and then applying a Price Deflator [1] to turn the Nominal GDP figure into a Real GDP figure. The labor force figure we used was taken from the Bureau of Labor Statistics. We took the natural log of this number to get our final lnopw variable used in the regression. The reason we only used 2010 instead of an average from 1976-2010 was because our regression is attempting to explain the final state output per worker.

We had two independent variables that we used in our first regression, $lninv$, and $lnngd$. The first variable, $lninv$, is our proxy for state investment, or savings. The second one, $lnngd$, takes into account three different growth rates: n, population growth, g, technology growth, and δ, depreciation. In order to calculate our investment numbers by state, we found the Capital Expenditure (CapEx) numbers in the Annual Survey of Manufacturers. We are able to find the figures for 23 years from 1976-2010, and we omitted the missing years [2] . We then found the Nominal GDP for each state to match up with the years we had CapEx figures for. These nominal GDP figures were taken from the Bureau of Economic Analysis. The CapEx for each year was divided by the Nominal GDP figures for each year, giving us 23 total Investment rate numbers for each state. The average of these figures was calculated, and then we took the natural log of these percentages to obtain the final variable, $lninv$, that we used in our regression. In order to calculate our $lnngd$ variable, we needed to obtain estimates for the following: Population growth, Depreciation, and Technology growth. As in the MRW paper we based our model off of, we took g and ?,, to be constant for each state, at .02 and .05 respectively [3]. To find the growth rate by state from 1976-2010 we used population figures from the Census Bureau. We

[1] Years 1979-1981,1986,1996-2001,2003

[2] Base year 2009; so all the real GDP figures are listed in 2009 prices, obtained from the Federal Reserve Bank of St. Louis

[3] Mankiw, Romer, Weil used .03 for their estimate of depreciation

added this number to our .07 (technology + depreciation) to form $n + g + \delta$ value. We took the natural log of this figure to obtain the final variable, $lnngd$, which we used in our regression.

Adding a Manufacturing variable

In order to scrutinize our variables we ran a second regression, adding in the percentage of the labor force that was employed in the manufacturing sector in the year 2010. The number of employees in the manufacturing sector is obtained from the Annual Survey of Manufactures 2011 and the labor force statistics were obtained from the National Bureau of Labor statistics. This variable is used to help explain the results from the first regression we ran. We took the log of the percentage we calculated, and this variable, $lnman$, was used in our subsequent regression. Since the variable is not in the theoretical model, using the proportion of manufacturing employment in each state for the year 2010 would eliminate idiosyncratic variations in output per worker in 2010, our dependent variable, associated with manufacturing.

Adding a Human Capital Variable

To determine our human capital accumulation variable we started by taking the population of each state aged between 14 and 35. In order to determine the rate during our period we extrapolated data figures between 1990 and 2010 since they are only collected in the per-decade census. These figures were found using the Center for Disease Controls website. We divided this by the total state population for the same years. This percentage was multiplied by the total school enrollment for the same age groups and years. The total school enrollment figures for these ages were taken from the Current Population Survey of the Population Census Bureau and were only available as far back as 1990. These enrollment figures were then divided by the state-level labor force data for the respective years (1990 and 2010); the labor force data was obtained from the Local Area Unemployment Statistics by the Bureau of Labor Statistics. We took the natural log of these numbers to form our variable, lnhcap. The raw figures before taking logs had a correlation coefficient with the state-level population in 2010 of 0.18. This quelled concerns that it may act as a proxy for population and its high correlation with our dependent variable (reported below) gave us confidence.

V. Results & Discussion

Our preliminary regression showed heteroskedasticity with 95% level of significance based on both a White-test and a Breusch-Pagan test. The results of both heteroskedasticity tests on our final regression are reported in I. In order to ensure the heteroskedasticity did not create inaccurate levels of significance in our variable we ran our regression using robust standard errors. The results in this section will primarily refer to regression one on Table II, with Regressions (2) through (4) of Table II discussed in the manufacturing decline methodological section below.

TABLE I

Heteroscedasticity Tests for basic Regression

	Test-Statistic	p-value
White Test	11.21	0.0474
Breusch-Pagan Test	12.47	0.0004

We begin by investigating the signs on the coefficients of $lninv$ and $lnngd$. Immediately some clear inconsistencies between the application of the Solow model at the state level and its application at the international level arise. While the sign of $lnngd$ is correct based on the Solow model the sign of $lninv$ is not. The negative sign on $lnngd$ is interpreted to mean that the sum of technology growth, population growth and depreciation have a negative effect on state level growth as Solow predicted. However, the sign of $lninv$ does not fall in line with the Solow predictions. The negative sign on $lninv$ would portend that savings actually cause growth to decrease as they increase. This conclusion is both illogical, as practically, investment is necessary for growth, and does not agree with the hypothesis of the model.

The magnitude of our coefficients illustrated the strengths and weaknesses of the model. The model suggests that a one percent increase in savings should be accompanied by a roughly quarter percent decrease in growth of output per worker. Simply put, by adding capital to their economies states are causing slower growth. The coefficient on $lnngd$ is much more sensible. That coefficient, of -0.263, suggests that a portion of growth is being offset by depreciation, technology and population growth over time. Thus our econometric result mirrors the theoretical assertion of the Solow model.

Our regression was estimated using an OLS regression in which we regressed the log of output per worker on the log savings and the log of the sum of n, g and δ. Our results are provided in Table II Regression (1). Both the coefficients, $lninv$ and $lnngd$ showed some level of significance, with $lninv$ having a p-value below .01 and $lnngd$ having a p-value below 0.1. Based solely on the significance of the variables, the log of savings appears to be an excellent predictor of overall growth at the state level while the log of the sum of n, g and δ provides some explanatory power as well. Also, the model seems to do a good job describing the overall variation of the model with an R-Squared of 0.52.

The mixture of positive and negative results in our model leaves us in a position where the overall model can neither be soundly rejected nor accepted. The primary issue is the high significance of savings, despite a nonsensical coefficient. At this point we discussed possible strategies for addressing this inconsistency (see appendix I). Ultimately, we concluded our use of a sectoral variable, manufacturing capital expenditures, was likely the root of the inconsistency.

TABLE II

BASIC REGRESSIONS WITH 51 OBSERVATIONS (50 STATES AND DISTRICT OF COLUMBIA)

	(1)	(2)	(3)	(4)
VARIABLES	lnopw	lnopw	lnopw	lnopw
lninv	-0.263***	-0.148***	-0.195**	-0.165*
	(0.0666)	(0.0441)	(0.0879)	(0.0828)
lnngd	-0.443*	-0.0760	-0.454*	-0.0600
	(0.259)	(0.195)	(0.246)	(0.212)
dc		0.800***		0.828***
		(0.147)		(0.187)
lnman			-0.0629	0.0194
			(0.0819)	(0.0778)
Constant	9.275***	10.68***	9.425***	10.68***
	(0.853)	(0.585)	(0.742)	(0.589)
Observations	51	51	51	51
R-squared	0.520	0.636	0.530	0.637

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Basic Model - Manufacturing Decline and DC Dummy

Of the many possible causes of the negative coefficient on savings (See Appendix I) the most plausible revolved around the idea that our data for savings only explained one sector of the economy. We began investigating our hypothesis: did states, which are heavily manufacturing-focused suffer a decline due to the growing dominance of the service sector in the overall economy? A strong negative correlation between $lnopw$ and $lninv$ indicates the $lninv$ that we are using is proxy-ing for the effect of manufacturing decline. This is reported in Table III. We used $lnman$ as a variable to test for the possibility of spurious correlation.

TABLE III

CORRELATION TABLE (INCLUDING DC)

	lnopw	lninv	lnman
lnopw	1		
lninv	-0.6975	1	
lnman	-0.6563	0.8770	1

As seen in the correlation Table III, the high degree of negative correlation between our dependent variable and the explanatory variable parallels that between $lnman$ and $lnopw$. There is a strong indication that the negative correlation between investment rate and output per worker is driven largely by the manufacturing decline across states. If this manufacturing decline hits manufacturing-oriented states more than the service-oriented states there would be a spurious negative correlation between the investment rate and the output per worker.

Appending a dummy for DC reduced the coefficient of our $lninv$ variable significantly and DC is the archetypical non-manufacturing state that has extremely high output per worker. We tested our hypothesis further by running the original regression with $lnman$ as an explanatory variable. This reduced both the magnitude and the significance of $lninv$ although the coefficient of the $lnman$ variable itself is not significant.

These results do not allow us to reject the null hypothesis that higher investment rates are producing lower output per worker even controlling for the proportion of labor force in manufacturing. Whilst we initially believed that the negative correlation between output per worker and investment rate is spurious and due to the varying sizes of the manufacturing sector across states, the regression shows this is not the case. It is likely that this manufacturing decline explanation is only part of the full story of why capital expenditures are inversely related to output per worker. The fact that service sector is dominating the economys output does suggest that the capital expenditure figure we are using for physical capital accumulation may not provide the full picture of capital accumulation in the production function.

Human Capital Model

The strong correlation reported in Table IV between our human capital variable and the dependent variable indicated some support for the augmented model.

TABLE IV

CORRELATION TABLE (INCLUDING DC)

	lnopw	lnhcap
lnopw	1	
lnhcap	0.9827	1

Regression (1) of Table V presents our results from our second least squares regression in which we added the log of our human capital variable to the regression. Our new regression included our original $lnngd$ and lninv along with our new variable, the log of our human capital proxy, regressed against the log of output per worker. The positive sign of our new variable supports the idea that human capital investment spurs economic growth. Specifically, the coefficient of 0.662 indicates that a one percent increase in the proportion of population with advanced degrees leads to a 0.662 percent increase in the growth. Further, the variables enter the model significantly with a p-value below 0.05. It also provides some additional explanation for the variation in the model as it increases the R-squared by roughly 4 percent.

Overall the addition of our new variable has little impact on the coefficients of our original variables. Both the signs and magnitudes of $lninv$ and $lnngd$ stay roughly the same as does the significance of both $lninv$, with the significance of $lnngd$ improving slightly. This would suggest that the log of savings ability to explain state level growth is not due to correlation with human capital growth.

TABLE V

HUMAN CAPITAL AUGMENTED REGRESSIONS WITH 51 OBSERVATIONS (50 STATES AND DC)

	(1)	(2)
VARIABLES	lnopw	lnopw
lninv	-0.258***	-0.157***
	(0.0558)	(0.0431)
lnngd	-0.508**	-0.149
	(0.234)	(0.203)
lnhcap	0.662**	0.354
	(0.304)	(0.295)
dc		0.717***
		(0.159)
Constant	14.92***	13.55***
	(2.440)	(2.465)
Observations	51	51
R-squared	0.563	0.647

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

VI. CONCLUSION

The results of our regression analysis provided inconclusive results about the role of capital accumulation in explaining the variation in output per worker across the states in the USA. The problem may stem from the empirics as much as from the theory. It is likely that our available measures of human and physical capital are unable to capture the concept of capital in these models. Perhaps more importantly, the quantitative measures of capital fail to encapsulate the qualitative nature of capital that is required for these models to work. The fact that our analysis is restricted to a cross-section dataset may have limited our ability to draw insights. Additionally, it is plausible that a panel data analysis would give us more insights into the validity of the model.

It is also possible that the theory is unable to provide a valid explanation for the variation in output per worker because the production function has changed over time, despite having worked for earlier time periods. That would not have been possible to test or demonstrate using traditional econometric techniques, especially when we do not have any theoretically grounded means of breaking up the time period of the dataset. The availability of a better proxy for the investment rate at the state-level would perhaps provide a better test of the model. Nevertheless, the negative relationship between state-level output and capital expenditure figures would be an interesting area of research to look into for future work.

BIBLIOGRAPHY

Barro, Robert J., and Xavier Sala-i-Martin (1991). Convergence across U.S. States and Regions, Brookings Papers on Economic Activity, 1, 107-182,

Domar, Evsey D (1946). Capital Expansion, Rate of Growth, and Employment Econometrica, Vol. 14, No. 2, pp. 137-147

Garofalo, Gasper A., and Steven Yamarik (2002). Regional Convergence: Evidence From A New State-By-State Capital Stock Series, The Review of Economics and Statistics, LXXIV, 316-323

Mankiw, Gregory N., David Romer, and David N. Weil (1992).A Contribution to the Empirics of Economic Growth, The Quarterly Journal of Economics, CVII, 407-437

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