CONVERGENCE OF INCOME ACROSS PENNSYLVANIA COUNTIES

David A. Latzko
Business and Economics Division
Pennsylvania State University, York Campus
1031 Edgecomb Ave., York, PA 17403

e-mail: DXL31@psu.edu
phone: 717-771-4115
fax: 717-771-4062

Abstract

The neoclassical growth model implies that if two economies have the same preferences and technology, the poorer economy will tend to grow faster in per capita terms. The relatively homogenous counties of Pennsylvania provide an excellent test of the model’s prediction of convergence in living standards. I apply a generalized method of moments estimator to a set of panel data to properly account for individual effects and endogenous explanatory variables, and find a convergence rate of 35 percent. This implies that differences in income levels among Pennsylvania counties are mainly due to differences in their steady-state income levels.
CONVERGENCE OF INCOME ACROSS PENNSYLVANIA COUNTIES

The neoclassical growth model predicts convergence: poor regions will grow faster than rich regions so that living standards across all regions will eventually be the same, even though some may start out way behind. Although there is little empirical evidence for this sort of unconditional convergence among heterogeneous economies (Mankiw, Romer, and Weil 1992; Romer 1994), both Dowrick and Nguyen (1989) and Mankiw, Romer, and Weil (1992) found that there has been a significant tendency towards convergence of per capita income among a cross-section of OECD countries, and Barro and Sala-i-Martin (1992; 1995) found evidence of convergence across U.S. states and across Japanese prefectures. The purpose of this paper is to examine whether income has converged in accordance with the prediction of the neoclassical growth model across a set of even smaller economic units, the 67 counties of the state of Pennsylvania.

I. Cross-Sectional Evidence

I test for convergence of per capita income across the 67 Pennsylvania counties over the period 1972 to 1997. The appropriate income measure is not available: a county-level version of gross domestic product. Personal income data is available on a county basis, but use of personal income is problematic if people work in one county and live in another or if people tend to own capital in other counties because the personal income accounts reported by the Bureau of Economic Analysis assign income to the county in which the owner of the inputs resides not to the county in which the income was earned. For instance, Lackawanna County received 7033 commuters from Luzerne County in 1990 while sending there 5175 commuters, and nearly
60,000 Delaware County residents worked in Philadelphia County with an even larger number of Philadelphia commuters coming from out of state (Pennsylvania State Data Center web site). Also, the personal income measure includes transfer payments. Therefore, personal income is not a good measure of county-level economic activity because it includes both unearned income and income earned outside the county.

I use ‘total earnings by place of work’ per capita as the measure of county income because it attributes income to the county where it was earned. Total earnings includes wages and salaries, other labor income, contributions for social insurance, and proprietors’ income. It excludes dividends, interest, rent, and transfer payments. Data for total earnings and population are taken from the Regional Economic Information System web page.¹

¹ <http://fisher.lib.virginia.edu/reis>. I am indebted to participants at the 1998 Pennsylvania Economic Association conference for pointing me towards this web site.
Figure 1 plots the average annual rate of per capita real total earnings growth over 1972-1997 against the log of 1972 real total earnings per capita for all 67 Pennsylvania counties. Unconditional convergence implies a negative relationship between the rate of income growth and initial income: poorer counties grow more rapidly than richer counties. However, the graph reveals no such obvious relationship.

I use ordinary least squares to more rigorously test the convergence prediction of the neoclassical model. The first column of Table 1 reports on the estimation of

\[(1) \frac{1}{25}[y_{1997} - y_{1972}] = \text{constant} + y_{1972},\]

where \(y_t\) is the natural logarithm of real total earnings per capita in year \(t\). The coefficient on the log of initial per capita income is positive, but both the coefficient and the adjusted \(R^2\) are essentially zero. There is no tendency for poor Pennsylvania counties to unconditionally grow faster on average than rich counties.

Early convergence studies utilized OLS to estimate

\[(2) g_{nT,0} = \alpha + \beta y_{n0} + \gamma x_n + e_n,\]

where \(g_{nT,0}\) is the average rate of growth of income per capita for economy \(n\) between periods 0 and T, \(y_{n0}\) is the natural logarithm of income per capita for economy \(n\) in period 0, \(x_n\) is a vector of variables that control for cross-economy heterogeneity, \(\alpha\) and \(\beta\) are parameters, \(\gamma\) is a vector of parameters, and \(e_n\) is an error term.\(^2\) If \(\beta\) is less than zero, economies that are initially poor after controlling for permanent differences associated with their \(x\)’s and \(e\)’s grow more quickly than economies that are initially rich controlling for their \(x\)’s and \(e\)’s. This is conditional

\(^2\) Baumol (1986), De Long (1988), Barro (1991), and Levine and Renelt (1992) are the most prominent examples not previously cited in the text.
convergence: incomes per capita converge holding all other variables influencing growth constant.

I utilize five variables in vector \( x \) which parallel research (Latzko 1999) found were robustly correlated in the sense of Sala-i-Martin (1997) with county economic growth in Pennsylvania: (1) the fraction of adults with a college degree (positively related to growth), (2) the growth rate of county population (negatively related to growth), (3) the fraction of total earnings in agriculture (positively related to growth), (4) the fraction of voters registered Democratic (negatively related to growth), and (5) the percentage of the county population living in urban areas (positively related to growth).

If income growth influences the independent variables, then estimates of equation (2) using OLS are potentially inconsistent. To reduce the simultaneity problem, I take the values of the explanatory variables from the beginning of the study period, 1972. The second column of Table 1 summarizes an estimate of equation (2). The coefficient of the log of initial income is negative but not significantly different from zero. The implied speed of convergence, derived from \( \beta \), is just 0.5 percent a year.

II. Panel Data Evidence

Ordinary least squares estimations of equation (2) produce consistent estimates only if the variables in vector \( x \) account for all of the permanent cross-county variation in income growth rates. Otherwise, omission of the individual effect biases estimates of \( \beta \) and the \( \gamma \)'s towards zero and, therefore, understates the rate of convergence (Caselli, Esquivel, and Lefort 1996; Evans 1997b). A solution is a panel data approach because it allows for control of unobservable,

The first column of Table 2 presents the results from a pooled fixed-effects, or within, regression of equation (3):

\[ g_{n,t+1,t} = \alpha_n + \beta y_{n0} + \gamma x_{n,t} + \epsilon_{n,t}. \]

The fixed-effects estimator takes deviations from individual means to eliminate the individual effects and assumes that each county has its own intercept. The coefficient of the log of initial income is significantly negative. The implied speed of convergence is 0.33.

The fixed-effects regression leads to consistent estimates only if all of the explanatory variables are exogenous. It’s unlikely that county population growth, for instance, is unaffected by county income growth. Caselli, Esquivel, and Lefort (1996) advocate use of a generalized method of moments (GMM) estimator to simultaneously address the problems of correlated individual effects and endogeneity. Their estimator takes all variables as deviations from period means. Taking deviations from period means reduces any bias generated by the business cycle. The growth model is first differenced to remove the individual effects, and then all of the past values of the explanatory variables are used as instruments to alleviate the endogeneity bias. The second column of Table 2 provides the results of a GMM estimation of equation (3) employing the Caselli, Esquivel, and Lefort (1996) procedure. Once again, the coefficient on the log of initial income is significantly less than zero. The implied speed of convergence is 0.35.
typical Pennsylvania county covers 35 percent of the distance between its current and steady-state income levels in one year.

Although the similarity of the convergence rates from the two panel estimations indicates that properly accounting for the individual fixed effects is more critical than dealing with the endogeneity bias, it is important to test the identifying assumption that lagged values of income and the other explanatory variables are valid instruments. The Sargan test statistic of overidentifying restrictions has a value of 58.8; its 0.24 p-value is in the acceptance region. I also perform a Hausman test comparing the GMM estimates in Table 2 to those obtained by adding the current and future values of the explanatory variables to the set of instruments. The test statistic equals 146.8. The p-value of less than 0.001 rejects the null hypothesis that the two estimates are not significantly different. This is strong evidence of the endogeneity of the explanatory variables and suggests that GMM is the appropriate method of estimation.

Figure 2 - Conditional Convergence of Income Across Pennsylvania Counties

![Graph showing conditional convergence of income across Pennsylvania counties.](image)
Figure 2 presents a graphical demonstration of the conditional convergence of income across Pennsylvania counties. The average values over the study period of the proportion of the county population of persons 25 years and over who have a bachelor’s degree or higher, the growth rate of county population, the fraction of total earnings in agriculture, the fraction of voters registered Democratic, and the percentage of the county population living in urban areas are partialled out of the growth rate variable using the coefficients from the GMM estimation to yield each county’s conditional growth rate of total earnings per capita. The figure plots this conditional growth rate against the log total earnings per capita in 1972. The tendency toward convergence is clear.

III. Discussion

The 67 Pennsylvania counties constitute a set of small open economies with similar political and social institutions and access to identical technology. Barro, Mankiw, and Sala-i-Martin (1995) demonstrate that if human capital cannot be completely financed by outside borrowing, an open economy will have a rapid but finite convergence rate. The 35 percent per year convergence rate detected in this paper is not inconsistent with rates found in other panel data studies: 10 percent (Caselli, Esquivel, and Lefort 1996) to 30 percent (Lee, Pesaran, and Smith 1997) for a broad sample of countries and 15.5 percent for a sample of U.S. states (Evans 1997a). To the extent that the economies of Pennsylvania counties are more similar than are state and national economies, the faster rate of convergence would be expected.

The rapid rate of convergence implies that Pennsylvania county economies are usually very near to their steady-states. This means that differences in income per capita are attributable to differences in steady-state levels of income. Counties that were relatively rich in 1972,
Dauphin and Montgomery, were still relatively rich in 1997; counties that were relatively poor in 1972, Perry and Susquehanna, remained relatively poor in 1997.³

³ Dauphin, Philadelphia, and Montgomery Counties had the highest total earnings per capita in 1972; the top three in 1997 were Montgomery, Dauphin, and Montour Counties. Fulton, Susquehanna, and Perry Counties were the three poorest counties in 1972; the bottom three counties in 1997 were Susquehanna, Pike, and Perry. The gap between rich and poor counties is widening. In 1972, the standard deviation of the log of total earnings per capita across Pennsylvania counties was 0.26; in 1997, it was 0.34. Barro and Sala-i-Martin (1995, p. 383) call this σ-divergence.
Table 1  
Tests for Convergence Using Cross Sectional Data

Dependent variable: average annual growth rate of total earnings per capita 1972-1997

<table>
<thead>
<tr>
<th></th>
<th>Unconditional</th>
<th>Conditional</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>-0.013</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.374)</td>
<td>(0.884)</td>
</tr>
<tr>
<td>Log $y_{1972}$</td>
<td>0.002</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.531)</td>
<td>(1.041)</td>
</tr>
<tr>
<td>fraction of population with a college degree</td>
<td>0.142</td>
<td>(4.904)</td>
</tr>
<tr>
<td>population growth rate</td>
<td>-0.383</td>
<td>(3.550)</td>
</tr>
<tr>
<td>farm income as a fraction of total earnings</td>
<td>20.824</td>
<td>(1.708)</td>
</tr>
<tr>
<td>fraction of voters registered Democratic</td>
<td>-0.001</td>
<td>(0.165)</td>
</tr>
<tr>
<td>percent urban population</td>
<td>0.003</td>
<td>(0.505)</td>
</tr>
<tr>
<td>Adjusted R$^2$</td>
<td>-0.01</td>
<td>0.34</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.01</td>
<td>0.007</td>
</tr>
<tr>
<td>Observations</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>Implied Speed of Convergence</td>
<td>-0.002</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Note: Absolute values of t-statistics are in parentheses. $y_{1972}$ is total earnings per capita in 1972. The values of the other explanatory variables are for 1972.
Table 2  
Tests for Convergence Using Panel Data

Dependent variable: average annual growth rate of total earnings per capita

<table>
<thead>
<tr>
<th></th>
<th>pooled fixed effects</th>
<th>GMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log of initial year income</td>
<td>-0.162 (-11.353)</td>
<td>-0.165 (27.223)</td>
</tr>
<tr>
<td>fraction of population with a college degree</td>
<td>0.307 (6.454)</td>
<td>0.039 (3.542)</td>
</tr>
<tr>
<td>population growth rate</td>
<td>0.110 (0.470)</td>
<td>-0.106 (3.127)</td>
</tr>
<tr>
<td>farm income as a fraction of total earnings</td>
<td>-12.167 (0.196)</td>
<td>42.235 (1.967)</td>
</tr>
<tr>
<td>fraction of voters registered Democratic</td>
<td>-0.071 (1.952)</td>
<td>0.017 (3.850)</td>
</tr>
<tr>
<td>percent urban population</td>
<td>0.205 (3.354)</td>
<td>0.217 (7.058)</td>
</tr>
</tbody>
</table>

Observations 335 335  
Implied Speed of Convergence 0.332 0.351

Note: Absolute values of t-statistics are in parentheses. Growth rates are over 5-year intervals: 1972-77, 1977-82, 1982-87, 1987-92, and 1992-97. Values of the explanatory variables are for the beginning of each interval.


Evans, Paul, “Consistent Estimation of Growth Regressions,” Manuscript, Ohio State University, September 1997b.


