Southern Economic Journal 2015, 81(3), 653–678
DOI: 10.4284/0038-4038-2012.243

On Persistent Poverty in a Rich Country

T. M. Tonmoy Islam,* Jenny Minier,† and James P. Ziliak‡

We examine differences in income within the United States, and the regions of persistent poverty that have arisen, using a newly assembled county-level data set linking 19th century Census data with contemporary data. We identify the roles of current differences in aggregate production technologies and factor endowments, together with contributions of historical institutions, culture, geography, and human capital. We allow for possible cross-county factor mobility via a correlated random effects GMM estimator and find evidence of significant regional differences in production technologies. Our decompositions of the poor/nonpoor income gap suggest that at least three-fourths of the gap is explained by differences in productive factors. Persistently poor counties are different (and poorer) primarily because they have lower levels of factors of production, not because they use the factors they have less efficiently. Together, historical and contemporary human capital explain over half of the overall income gap between persistently poor and nonpoor counties.

JEL Classification: O1, R1

1. Introduction

Between 1960 and 2000, real income per capita in the United States increased 175% and aggregate poverty rates fell by half, from 22% to just over 11%. Despite this economic progress, poverty has remained stubbornly high in several regions of the country, as shown in Figure 1. These so-called persistently poor areas, defined as those areas with county poverty rates in excess of 20% since 1960, encompass five distinct regions and 11% of all U.S. counties: Appalachian Kentucky, the "Black Belt" region spanning the Carolinas to Alabama, the Mississippi Delta region, the Texas "colonias" along the Rio Grande River, and Native American reservations in the four corners states and the Dakotas.[1] The regions differ greatly in racial, ethnic, geographic, and economic composition, and thus our aim is to identify why these counties in an otherwise rich nation share the enduring legacy of persistent poverty.

While canonical economic growth models (Solow 1956; Mankiw, Romer, and Weil 1992) have focused on differences in factor accumulation as the primary reason underlying cross-

* Department of Economics, Elon University, 2075 Campus Box, Elon, NC 27244, USA; E-mail: tislam@elon.edu

† Department of Economics, University of Kentucky, 335 Gatton College, Lexington KY 40506, USA; E-mail: jminier@uky.edu

‡ Center for Poverty Research and Department of Economics, University of Kentucky, Mathews Building Suite 300, Lexington KY 40506-0047, USA; E-mail: jziliak@uky.edu; corresponding author.

Received November 2012; accepted March 2014.

[1] The U.S. Department of Agriculture (USDA) defines a county as persistently poor if its poverty rate exceeds 20% in each Census since 1970 (http://www.ers.usda.gov/briefing/rurality/typology/maps/Poverty.htm). We extend this to include the 1960 Census. We note that the typography of persistent poverty is the same if one adopts a more stringent criterion of a 30% poverty rate, though fewer counties meet the criteria in each subregion. In section 5, we include sensitivity analyses on the definitions of poor and non-poor.

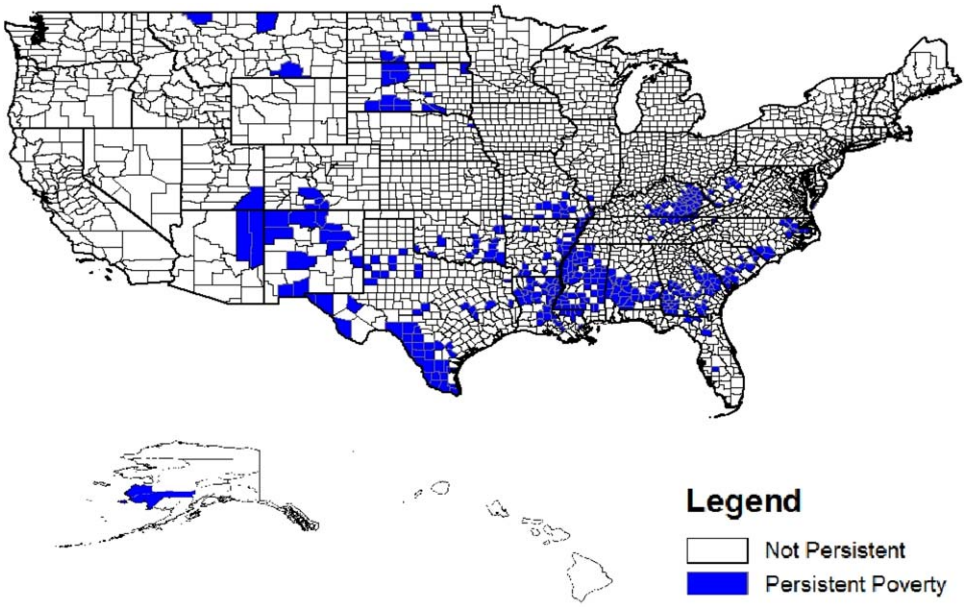

Figure 1. U.S. Counties with Persistent Poverty 1959–1999. Shaded areas represent counties identified as persistently poor, with poverty rates of 20% or higher each decennial Census between 1960 and 2000. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

economy income differences, differences in production technologies (i.e., differences in the efficiency with which a region employs its factors) may also result in persistent income differences across regions.[2] Yet another possibility is that these regions are on divergent growth paths because of different institutions, geography, and culture, as emphasized in some of the more recent cross-country growth literature (Greif 1994; Hall and Jones 1999; Easterly and Levine 2001; Rappaport and Sachs 2003; Acemoglu, Johnson, and Robinson 2005). These considerations lead us to examine three broad explanations why certain regions in the United States might be on a divergent growth path towards persistent poverty: (i) lower levels of contemporaneous factors of production, such as physical and human capital; (ii) less efficient use of those factors[3]; and (iii) lower levels of productivity (residual income), perhaps determined by historical institutions, culture, geography, and endowments of human capital.

We assemble a new data set of U.S. counties that links historical data from the end of the 19th century (1890 and 1900) to contemporaneous data from the end of the 20th century (1960–2000). This permits us to jointly consider current factors of production such as human capital, physical capital, and labor force growth alongside historical factors such as illiteracy, religiosity, urbanicity, temperature and precipitation, and land tenure. Some studies have examined growth in the United States using subnational data at the state level (Barro and Sala-i-Martin 1992; Evans and Karras 1996; Turner et al. 2007; Bauer, Schweitzer, and Shane 2012; Gennaioli, La Porta, Lopez-de-Silanes, and Shleifer 2013), and fewer still at the county level (Rupasingha, Goetz, and

[2] See Azariadis and Drazen (1990) for a theoretical justification, and Durlauf and Johnson (1994) for empirical evidence.

[3] In the textbook case, the aggregate production function is $Y=AK^{\alpha}L^{1-\alpha}$. We refer to different values of the parameters α and $(1-\alpha)$ as "less efficient use of factors," while "lower levels of productivity" refers to different levels of A.

Freshwater 2000; Higgins, Levy, and Young 2006), but to our knowledge none has linked historical county data with contemporaneous data as we do here.[4]

Our starting point is a descriptive examination of links between historical determinants of productivity and the probability of being persistently poor in the late 1900s. The results here point to the primacy of low initial levels of human capital leading to substantially higher odds of persistent poverty. Rates of illiteracy in 1900 among the persistently poor counties are more than three times higher than nonpoor counties, and this initial human capital shortfall dominates geography, culture, and institutions as a reduced-form predictor of long-term poverty.

We then turn to a more formal model of the income process that builds on the dynamic panel data model popularized by Islam (1995), who advanced the empirical cross-country growth literature by explicitly allowing for heterogeneity in aggregate production functions via the inclusion of permanent cross-economy differences. This is appealing because a key determinant of steady-state income in Mankiw, Romer, and Weil (1992) is a country's initial level of productivity, and heterogeneity in productivity seems likely. In our case, we allow productivity to differ across counties by including county-specific intercepts. Whereas many growth articles treat the fixed effects as nuisance parameters, historical human capital, institutions, geography and culture enter via these county-specific intercepts and thus shed important light on the development process. Hence, we adopt the correlated random effects Generalized Method of Moments (GMM) estimator of Arellano and Bover (1995) that identifies parameters on both contemporaneous time-varying and historical time-invariant factors.

The idea of the Arellano–Bover estimator in our context is to parameterize the county fixed effect as a function of observable (historical) factors, and then augment the standard first-difference moment conditions found in dynamic panel models (Anderson and Hsiao 1982; Arellano and Bond 1991; Blundell and Bond 1998) with the additional moment conditions for the historical regressors in levels. In the baseline model, we assume that lagged income and the time-varying factors are predetermined to the current-period error term, but are correlated with the error in first differences, and thus we use further lags as instruments to purge the correlation with the differenced time-varying error. As is typical in growth models (Deaton 2010), the historical variables are exogenous to current income dynamics and they serve as instruments for themselves in the levels equation. An added feature of this estimator is that it is straightforward to relax the assumption that current factors of production are predetermined. For example, if current human capital is mobile across counties, and this migration is related to a county's current income prospects (Greenwood 1997; Kennan and Walker 2011; Gennaioli, La Porta, Lopez-de-Silanes, and Shleifer 2013), then imposing the assumption of predetermined migration would lead to inconsistent estimates. Thus, we test our baseline estimates to alternative identification assumptions.

Since our interest is in understanding what sets the persistently poor counties apart, we extend our baseline model to admit potential heterogeneity in the production function parameters by estimating the dynamic income model separately for counties classified as persistently poor versus those that are not poor. Lee, Pesaran, and Smith (1997) discuss the issue of parameter heterogeneity in the cross-country context, but as noted by Durlauf, Johnson, and Temple (2005), identification in these models is perilous, and a possible alternative is to split the sample into groups likely to share similar parameter values. We estimate separate models based on persistent

[4] Rappaport and Sachs (2003) and Beeson, DeJong, and Troesken (2001) use county data but focus on population change rather than income, while Clifton and Romero-Barrutieta (2006) use county data to focus on poverty rates in 2003 in Appalachia.

Table 1. Summary Statistics of Social Indicators by Persistent Poverty Status

	Not Persistently Poor		Persistently Poor	
	Mean	SD	Mean	SD
Pooled 1960–2000 census data				
Per capita income ($)	12,496	4,879	8,558	3,505
Population	82,831	275,896	21,107	38,171
Fraction in labor force	0.425	0.065	0.352	0.057
Growth in labor force	0.158	0.234	0.037	0.207
Fraction high school graduate	0.585	0.189	0.418	0.171
Capital exp. ($1000 per capita)	0.306	0.554	0.160	0.590
Fraction living in urban area	0.367	0.288	0.217	0.238
Fraction black	0.067	0.114	0.286	0.254
Fraction churched	0.396	0.180	0.388	0.183
Fraction voted in pres. elections	0.597	0.124	0.523	0.135
1960 census data				
Per capita income ($)	6,864	1,872	4,074	1,192
Population	66,094	229,251	21,424	41,973
Fraction in labor force	0.361	0.036	0.301	0.043
Growth in labor force	0.054	0.224	−0.166	0.136
Fraction high school graduate	0.355	0.104	0.212	0.064
Capital exp. ($1000 per capita)	0.132	0.241	0.040	0.054
Fraction living in urban area	0.330	0.278	0.182	0.215
Fraction black	0.073	0.128	0.307	0.267
Fraction churched	0.464	0.159	0.335	0.183
Fraction voted in pres. elections	0.664	0.178	0.467	0.229
2000 census data				
Per capita income ($)	18,664	3,696	13,400	1,897
Population	100,491	322,864	21,865	35,766
Fraction in labor force	0.483	0.047	0.394	0.040
Growth in labor force	0.143	0.153	0.056	0.108
Fraction high school graduate	0.790	0.073	0.637	0.065
Capital exp. ($1000 per capita)	0.319	0.301	0.315	0.308
Fraction living in urban area	0.411	0.301	0.241	0.250
Fraction black	0.065	0.106	0.284	0.257
Fraction churched	0.532	0.170	0.476	0.183
Fraction voted in pres. elections	0.544	0.085	0.493	0.080
Observations	10,830		1,170	
Number of counties	2,166		234	

Notes: "Persistently poor" counties have poverty rates of at least 20% in 1960, 1970, 1980, 1990, and 2000. "Not persistently poor" are all others. Per capita income and capital expenditures are in real 2000 dollars, based on the personal consumption expenditure deflator.

poverty status and test whether production functions differ across regions. A key additional advantage of separate models is that we can use a decomposition method common in labor economics (Oaxaca and Ransom 1994) that permits us to quantify the relative contributions of contemporaneous factor shares, historical determinants of productivity, and the parameters of the production function to the income gap between persistently poor and nonpoor counties.

We find evidence of significant regional differences in aggregate production functions, but our decompositions of the poor/nonpoor income gap suggests that at least three-fourths of the gap is explained by differences in current and past productive factors. Not surprisingly, current factors explain a larger share of the income gap than do historical factors, but historical factors, particularly those measuring human capital, contribute substantially to the gap. Of the income

difference explained by contemporary factors, own past income, human capital, and urban share matter most. While some of the geographic variables are important for income growth (e.g., terrain, temperature, and precipitation), little of the overall income gap is explained by regional differences in geography, culture, and institutions. The historical role of human capital, however, is striking, and combined with contemporaneous differences, accounts for between 50 and 60% of the income gap between persistently poor and nonpoor counties. Persistently poor counties are different (and poorer) because they have lower levels of factors of production, not because they use the factors less efficiently.

2. The Origins of Persistent Poverty: Some Preliminaries

At the dawn of the 1960s, the poverty status of a vast stretch of the United States was bleak. County poverty rates over 50% were common in the South, and poverty rates over 20–30% were the norm in the Midwest and Plains states. With strong economic growth and the expansion of income support due to the Johnson Administration's Great Society programs, the poverty landscape changed dramatically over the next decade, although poverty rates remained over 40% in parts of the South. Poverty rates generally fell during the 1970s, leveled out during the 1980s, and fell further during the expansion of the 1990s. While in an absolute sense poverty rates have improved in almost all of the country, poverty in the counties identified as persistently poor in Figure 1 remains four times higher than the national rate.[5]

To fix ideas, the growth literature generally specifies that output or income is determined by stocks of physical and human capital along with economy-specific "productivity," which encompasses institutions and endowments of natural resources.[6] Table 1 illustrates some descriptive differences in incomes and observed human and physical capital between the persistently poor counties and nonpoor counties pooled across the 1960–2000 Censuses. In our sample, there are 12,000 county-years, consisting of 2400 counties, 234 of which are categorized as persistently poor.[7] Appendix A provides details on the sources and definitions of the variables and sample composition. Not surprisingly, the persistently poor counties have lower real per capita incomes, proportionally fewer people in the labor force, slower growth of the labor force, lower rates of education among the adult population, and lower per capita new capital expenditures in manufacturing. The persistently poor counties are also less urban, smaller (in terms of population), and have higher percentages of African Americans, but are similar in terms of fraction affiliated

[5] The poverty line in the United States was established in the 1960s and has been updated each year thereafter by changes in the CPI. Thus, in real terms it measures a constant standard of living and so is useful in making comparisons over time. It varies by household size, but is common across states. Pre-tax income, not including in-kind transfers such as food stamps and health insurance, is used in computing poverty rates.

[6] While the production function approach, such as we take here, emphasizes the determinants of production rather than income, we use county income as our dependent variable in the empirical analysis due to data availability. A potential drawback to this, of course, is that many people work in one county and live in another. While our Census data include income earned by residents of a county, the Regional Economic Information System (REIS) data from the Bureau of Commerce, which are available for the later part of our sample, measure income earned in a county by residents and non-residents; for 2000, the correlation between the two measures is 0.91.

[7] Today there are 3141 counties in the United States, but many of these did not exist in the late 19th century (e.g., Alaska and Hawaii were not part of United States; other young states still comprises territories; and some other counties were created by splitting larger counties). To abstract from differences owing to changes in composition of counties, our sample comprises 2400 counties with consistently defined boundaries from 1890 to 2000. See the Appendix for details.

Table 2. Summary Statistics of Historical Indicators by Persistent Poverty Status

	Not Poor		Persistently Poor	
Land tenure in 1890	0.800	0.115	0.768	0.145
Share churched in 1890	0.292	0.120	0.308	0.150
Share Baptist in 1890	0.065	0.081	0.134	0.097
Share Calvinist in 1890	0.020	0.022	0.009	0.016
Share Catholic in 1890	0.058	0.086	0.035	0.124
Average temperature 1895-05	52.902	7.417	60.293	6.181
Average precipitation 1895-05	3.073	0.918	3.590	0.746
SD of elevation	0.090	0.132	0.065	0.077
Urban share in 1890	0.127	0.210	0.026	0.106
Share foreign born in 1900	0.095	0.103	0.026	0.076
Illiteracy rate in 1900	0.112	0.135	0.362	0.221
Capital exp. in 1890 ($1000 per capita)	0.400	0.100	0.082	0.200
Number of counties	2166		234	

Notes: "Persistently poor" counties have poverty rates of at least 20% in 1960, 1970, 1980, 1990, and 2000. "Not poor" are all others. The historic Census data include only counties without redefined borders between the relevant year (1890, 1900) and 1960. See Appendix A for detailed variable definitions.

with a church and voting in presidential elections. However, the lower two panels of Table 1 indicate some signs of convergence between the regions: between 1960 and 2000, income per capita in the persistently poor counties increased from 59 to 72% of nonpoor income. Likewise, high school completion rates in the persistently poor counties increased from 60% of the nonpoor rate in 1960 to 81% in 2000, and real capital spending per capita increased from 30% of the nonpoor rate in 1960 to 99% in 2000.

In the cross-country context, many have argued that historical determinants of growth, such as geography, institutions, and culture, explain more of the difference in growth rates than do current differences in factor accumulation (e.g., Easterly and Levine 2001; Acemoglu, Johnson, and Robinson 2005). To proxy for these components of growth at the county level, we use data from the Censuses of the late 19th century and turn of the 20th century. Since 1890 was the cusp of expansions in the logging and coal industries in the United States, we use 1890 data whenever possible (but for several variables that are not available in 1890, we use 1900 data).

Sociologists such as Duncan (1999) and Billings and Blee (2000) argue that the roots of persistent poverty in Appalachia and the Mississippi Delta can be traced to social and economic institutions that bestowed ownership rights of land and resources to a select group of individuals—in the case of Appalachia to absentee coal and timber barons, and in the Delta to plantation owners. Economists such as Engerman and Fogel (1974) and Ransom and Sutch (2001) have made a similar case about the role of institutions on the economic development and growth in the South in the decades following the Civil War, especially the economic organization of sharecropping. Likewise, resettlement of Native Americans in the 19th Century often took the form of removal from productive lands in the South and East to nonproductive, arid lands in the central Plains (Barrington 1999). This suggests that the extent of local ownership of land and natural resources, sometimes referred to as land tenure, likely varies across the United States in response to regional political institutions, and the higher the share of local land tenure, the more productive income remains in the local community.

In the historical Censuses, data were collected on the number of improved and unimproved acres of farmland, distinguishing whether or not the improved acres were owner-occupied. We thus construct a proxy for local institutions as the fraction of farmland in the hands of local

owners (following Clifton and Romero-Barrutieta 2006). As seen in Table 2, average land tenure in 1890 was slightly higher in nonpoor counties (80%) compared to persistently poor counties (77%), and the variance lower. However, because the differences are not striking, the empirical importance of historical land tenure on persistent poverty today is not clear a priori.

Culture is also suggested as a possible source of persistent poverty, both across and within countries (Banfield 1970; Billings 1974; Murray 1984; Greif 1994). Most prominent among these is the role of religion in economic development, especially Max Weber's argument that Protestantism (particularly Calvinism) played a crucial part in the economic success of Northern Europe compared to its southern neighbors who were predominantly Roman Catholic. Barro and McCleary (2003) and Cavalcanti, Parente, and Zhao (2007) provide some evidence in support of the Weber thesis in the cross-country context. In the case of the United States, religion-based cultural influences are determined in part by historical patterns of immigration. Nineteenth Century immigrants in the East, Midwest, and West tended to be dominated by Roman Catholics, while those in the South were primarily Baptists. The early Scots-Irish who settled northern and central Appalachia in the 18th Century tended to be Presbyterian, while later immigrants from Germany tended to adhere to movements within the Baptist faith as well as Catholicism. The 1890 Census of Religious Bodies recorded the number of persons in a county who claimed membership in a church, both overall and by denomination. Table 2 includes the means of the overall share of the population who belonged to an organized church, as well as the share of some major denominations. The table shows that differences in the share of the population counted as church members are very small across poverty groupings, although the distribution across religious affiliation varies considerably between the persistently poor and other counties. For example, the share of residents who were Calvinist is twice as high in the nonpoor counties, providing some prime facie support for the Weber hypothesis. However, the nonpoor counties also have higher historical shares of Roman Catholics (5.8 vs. 3.5%), which is contrary to the progrowth Weberian view. The historical share of Baptists is twice as high in persistently poor counties as in nonpoor counties.

Geography, such as differences in temperature and variation in altitude, is also often considered to contribute to differences in income levels (Gallup, Sachs, and Mellinger 1999; Acemoglu, Johnson, and Robinson 2005; Iyigun 2005; Rappaport and Sachs 2003; Rappaport 2007; Eller 2008). For example, the Appalachian Mountains, which span from Mississippi to Maine, are rugged and densely packed with narrow valleys, making development a challenge. Likewise, the lowlands of the Mississippi Delta were historically prone to flooding, another barrier to development. The hot and humid summers of the South and the arid farmland of Native American reservations also posed challenges to agricultural productivity.

In Table 2, we present three measures of county geography: the standard deviation (SD) of elevation as a measure of how mountainous a county's terrain is; average monthly temperature from 1895 to 1905; and average monthly precipitation from 1895 to 1905. Table 2 shows that persistently poor counties are warmer by just over 7° on average, which is expected given their southerly location depicted in Figure 1, and on average experience nearly one-half inch more precipitation. Nonpoor counties have greater variation in elevation, and also higher dispersion, than poor counties.

Finally, economists have stressed the important roles of human capital endowments and potential agglomeration economies in urban areas (e.g., Glaeser et al. 1992; Moretti 2004; Shapiro 2006). In the 1900 Census, individuals were asked whether they could read or write, which leads to our focal historical measure of human capital, the illiteracy rate. We also include the share of the county's population that was foreign born in human capital, while we use the share of the

Table 3. Linear Probability Estimates of the Probability of Being Persistently Poor

Variables	(1)	(2)	(3)	(4)	(5)
Land tenure in 1890	−0.072 (0.057)	0.050 (0.054)	0.042 (0.056)	−0.137 (0.063)	−0.191 (0.064)
Average temperature	0.009 (0.001)	−0.0002 (0.001)	0.0002 (0.001)	0.009 (0.001)	0.008 (0.001)
Average precipitation	0.015 (0.007)	−0.026 (0.007)	−0.025 (0.007)	0.016 (0.007)	0.008 (0.007)
SD of elevation	−0.069 (0.029)	−0.029 (0.029)	−0.017 (0.030)	−0.054 (0.029)	−0.048 (0.030)
Share foreign-born in 1900		−0.048 (0.081)	−0.115 (0.093)	0.102 (0.084)	−0.004 (0.099)
Urban share 1890		−0.052 (0.024)	−0.065 (0.025)	−0.169 (0.026)	−0.141 (0.026)
Capital 1890	−24.851 (5.888)	−2.845 (4.705)	−1.826 (4.799)	−3.848 (5.215)	−3.942 (5.078)
Illiteracy rate 1900		0.921 (0.064)	0.914 (0.071)		
Church share 1890		−0.181 (0.058)		0.027 (0.063)	
Baptist share 1890			−0.255 (0.120)		0.341 (0.117)
Calvin share 1890			−0.557 (0.263)		−1.413 (0.296)
Catholic share 1890			−0.017 (0.107)		0.191 (0.113)
Constant	−0.380 (0.064)	0.094 (0.077)	0.056 (0.077)	−0.356 (0.074)	−0.166 (0.078)
Adjusted R^2	0.089	0.220	0.219	0.097	0.117
Number of counties	2400	2400	2400	2400	2400

Notes: Standard errors are in parentheses. The dependent variable is a binary variable equal to one for counties that are persistently poor between 1960 and 2000 and zero otherwise.

population that resided in an urban area as our measure of agglomeration economies. Finally, we include a measure of capital spending per capita in 1890 to capture initial endowments in physical capital. Table 2 shows that the most striking differences in the historical data are in the human and physical capital and agglomeration variables: counties that are classified as persistently poor today had illiteracy rates more than three times higher than other counties in 1900 (36% vs. 11%). Persistently poor counties also had urban shares 80% lower, capital spending per capita 80% lower, and shares of foreign-born residents 73% lower, than nonpoor counties.

For a preliminary look at whether these historical differences in institutions, geography, and capital matter for the chances of being persistently poor a century later, in Table 3, we specify a linear probability model of the probability that a county is persistently poor over the period 1960–2000 as a function of historical Census data. Column 1 of Table 3 contains a parsimonious specification of persistent poverty as a function of institutions (land tenure), geography (temperature, precipitation, and terrain), and physical capital.[8] The results in Column 1 show that a 1% point increase in the share of owner-occupied farmland in 1890 lowers the probability of being persistently poor a century later by 7%, being located in a warmer or wetter climate raises the odds of being persistently poor, and more "mountainous" terrain lowers the probability of being poor. In addition, a $1000 increase in per capita capital spending lowers the odds of being persistently poor by 24%. In Columns 2 and 3, we present results from our preferred specifications that add measures of culture and human capital to the regression model. Higher literacy rates, higher church membership (especially Calvinist and Baptist), and higher urban population shares all significantly lower the odds of being persistently poor. At the same time, the addition of these variables negates any role for institutions, geography, and physical capital as we have measured them. In particular, as we show in Columns 4 and 5, omitting illiteracy causes the coefficient estimate on the institution and geography variables to regain statistical significance, and also causes the coefficient estimates on the shares of Baptists and Catholics to become positive and statistically significant, while the coefficient estimate on the share of Calvinists increases in magnitude but remains negative. Illiteracy is highly correlated with the prevalence of some denominations, in particular the share of Baptists in a county,[9] so that the models in Columns 4 and 5 yield the spurious result that a higher share of Baptists increases the probability of being a persistently poor county when in fact the driving force is county rates of illiteracy. This result highlights the long-run importance of human capital to counties' economic status. In the following sections, we turn to the growth literature to examine more formally the mechanisms through which these historical factors affect modern income levels while simultaneously accounting for current levels of factor accumulation.

3. Dynamic Model of Income

In the canonical neoclassical growth model of Solow (1956), an economy converges to a steady state determined by factors such as the economy's rates of saving and population growth,

[8] This is similar to the model estimated by Clifton and Romero-Barrutieta (2006). The dependent variable in their model is the poverty rate in 2003, rather than an indicator variable for persistent poverty for 1960-2000 as we use. In addition, we use three measures of geography-temperature, precipitation and terrainwhereas they use only terrain, and we include a measure of physical capital.

[9] The correlation coefficient between the 1900 illiteracy rate and the share of all church members in 1890 is 0.265; for Baptists, Calvinists, and Catholics, respectively, the correlations are 0.630, −0.189, and −0.163.

where income per capita grows at the rate of technological progress. To illustrate, consider the human capital-augmented version of the Solow model following Mankiw, Romer, and Weil (1992) and Islam (1995). The production function is given by

$$Y(t)=K(t)^{\alpha}H(t)^{\gamma}[A(t)L(t)]^{1-\alpha-\gamma} \tag{1}$$

where Y is aggregate output, K and H are stocks of physical and human capital, L is the labor force, and A is what we refer to as "productivity," which grows at the exogenous rate g. Output is invested in physical and human capital at the constant rates s_k and s_h respectively.

Under standard assumptions, an economy's growth rate as it transitions toward its steady-state level of income per capita ($y = Y/L$) can be derived:

$$\begin{aligned}\ln y_t-\ln y_0=&(1-e^{-\lambda t})\frac{\alpha}{1-\alpha-\gamma}\ln s_k+(1-\mathrm{e}^{-\lambda t})\frac{\gamma}{1-\alpha-\gamma}\ln s_h-(1-\mathrm{e}^{-\lambda t})\frac{\alpha+\gamma}{1-\alpha-\gamma}\\ &\ln(n+g+\delta)-(1-\mathrm{e}^{-\lambda t})\ln y_0+(1-\mathrm{e}^{-\lambda t})\ln A_0+gt\end{aligned} \tag{2}$$

where $\lambda = (n + g + \delta)(1 - \alpha - \gamma)$. Thus, transitional growth in income per capita depends on investment in physical and human capital (s_k and s_h, respectively), a term including population growth, technological progress and depreciation ($n + g + \delta$), initial income (y_0), and the initial level of technology (A_0). Furthermore, with reasonable assumptions about the values of α and γ (the shares of physical and human capital, respectively, from the production function), the coefficient estimate on the log of initial income can be used to infer the speed of convergence toward the steady state (i.e., λ).

Traditionally, this model had been estimated in the cross section, with an assumption not only of identical production functions across economies (most commonly, countries) but also identical rates of technological progress; economies' steady-state levels of income still differ, based on their savings and population growth rates. Under these assumptions, controlling for rates of population growth and savings, an initially poor economy "converges" to the same steady state as an initially richer economy. This test of "conditional convergence" has received wide support in the empirical literature.

Moving to a panel framework allows for the estimation of different initial *levels* of technology across economies (the rate of growth of technology is still assumed constant across economies). The workhorse specification in a panel data setting comes from Islam (1995) (see also eqn. 59 in Durlauf, Johnson, and Temple 2005):

$$y_{it}=(1+\beta)y_{it-1}+X_{it}\varphi+\epsilon_{it} \tag{3}$$

where y_{it} is the natural log of real income per capita for county i ($=1,\ldots, N$) in year t ($=1,\ldots, T$); y_{it-1} is the lag of the dependent variable; and X_{it} are time-varying rates of factor accumulation (new capital investment per capita, labor force growth rate, school attainment rates); and $\epsilon_{it}=\mu_i+\delta_t+\xi_{it}$ is a compound error term that is a function of unobserved, permanent differences across counties in productivity that do not vary over time (μ_i), a time-varying macroeconomic shock (δ_t) that is constant across counties, and an iid error term (ξ_{it}). The parameter identifying the speed of convergence is β.

In the growth literature, μ_i embeds much of what is thought of as "productivity," and can be explained by historical institutions, culture, and other factors. By construction, μ_i is correlated with y_{it-1}. A standard approach is to treat this unobserved heterogeneity as a nuisance parameter

and apply first differences to sweep it out of the model. We, however, have historical data to proxy for these initial productivity factors, and thus we parameterize initial productivity (A_0) by adopting a correlated random effects framework of Hausman and Taylor (1981) as $\mu_i = Z_i\theta + \psi_i$, where Z_i are observed time-invariant factors that may affect initial productivity such as land tenure, church share, weather, and initial human and physical capital endowments, and ψ_i is an error term. Substituting this into Equation 3 yields:

$$y_{it} = (1+\beta)y_{it-1} + X_{it}\varphi + Z_i\theta + \psi_i + \delta_t + \xi_{it} \tag{4}$$

where we still allow lagged income (y_{it-1}) and current factors (X_{it}) to be correlated with ψ_i but assume that $E[Z_i\psi_i] = 0$. That is, our identifying assumption is that the unobserved, random component of the initial conditions, (ψ_i), is uncorrelated with the Z_i factors in Equation 4. As is standard in the dynamic panel data literature, lagged income is treated as predetermined. Much of the growth literature treats the X_{it} as exogenous, but if factors are mobile and that decision is potentially related to unobservables affecting income levels (beyond the possible correlation with the fixed effect), then the assumption of exogeneity may be violated. We instead assume in our baseline model that the X_{it} are predetermined with respect to current income, and then relax that further in a robustness section to permit possible endogeneity.

Appendix B provides details of the correlated random effects Generalized Method of Moments estimator of Arellano and Bover (1995) that we use to identify the unknown parameters in Equation 4. In words, the idea is to stack moment conditions whereby the first ($T - 1$) equations are identified using the first differences transformation to estimate the parameters on time-varying regressors β and Φ, and the Tth equation is specified in levels to identify the parameters on the time-invariant regressors θ. First differences of the first ($T - 1$) equations eliminate the correlation between lagged income and current factors with ψ_i. However, because the lagged dependent variable in first differences is correlated with the first-difference error term ($\Delta\xi_{it}$) owing to the predeterminedness assumption, we use lags of own income dated ($t - 2$) and earlier as instruments as described below and in Appendix B (Anderson and Hsiao 1982). These lagged instruments are correlated with the regressors, but by assumption are uncorrelated with the future error terms. Likewise, since we assume that current factors of production are predetermined with respect to current income, first differences of those factors are also correlated with the first-difference error term. Thus, we use lags of the factors dated ($t - 1$) and earlier as instruments. Furthermore, in the levels equation, we use the historic variables as their own instruments because we are assuming that future shocks to the economy (through the error term) are independent of the historic variables, a standard assumption in the growth and development literature (Deaton 2010). We note that Equation 4 could be estimated in two steps, where in the first step, we apply a first-difference GMM estimator to identify β and Φ, retrieve the county intercepts $\hat{\mu}_i$, and regress the intercepts on the Z_i using least squares to identify θ. The unified correlated random effects GMM estimator we use, however, is more efficient and lends itself more readily to the decomposition below.

The model in Equation 4 assumes common production functions across counties (except for the intercepts), and thus a common speed of convergence. Letting Λ denote the set of unknown coefficients in Equation 4, we test this assumption by applying the correlated random effects GMM estimator to the subsamples of persistently poor counties with parameters ($\hat{\Lambda}^P$) and non-poor counties with parameters ($\hat{\Lambda}^N$). We then conduct a Wald test for whether the production functions differ between the groups as $\text{Wald} = (\hat{\Lambda}^P - \hat{\Lambda}^N)'[\text{var}(\hat{\Lambda}^P) + \text{var}(\hat{\Lambda}^N)]^{-1}(\hat{\Lambda}^P - \hat{\Lambda}^N)$, which is distributed asymptotically chi-square with d.f. equal to the rank of Λ.

Estimating separate models is also important because we are ultimately interested in quantifying how much of the income gap between persistently poor and non-poor counties can be attributed to differences in factor shares and historical endowments, and how much to differences in production function parameters (or the efficiency with which counties use factors to produce output). We use a decomposition technique that pervades labor economics known as the Oaxaca decomposition (Oaxaca and Ransom 1994). Intuitively, if there were no difference in production functions between poor and non-poor counties, the production function would be characterized by the pooled model in Equation 4, implying that differences in income between poor and non-poor counties would be solely due to differences in factor shares and initial productivity. If production functions do differ, however, then the differences in income would be a function both of different contemporaneous and historical factors and different production functions.[10] This suggests the following decomposition based on the estimating model in B2.1:

$$\bar{\hat{y}}^P - \bar{\hat{y}}^N = (\bar{d}^P - \bar{d}^N)\hat{\Lambda} + \bar{d}^P(\hat{\Lambda}^P - \hat{\Lambda}) + \bar{d}^N(\hat{\Lambda} - \hat{\Lambda}^N) \tag{5}$$

where the left hand side is the difference in mean predicted log real income per capita between poor and non-poor counties, the first term on the right hand side reflects differences in factors and determinants of productivity (i.e., d consists of the X_{it}, Z_i), the second term is differences in aggregate production functions (differences in parameter estimates) between the poor counties and the pooled counties, and the third term is differences in aggregate production functions between the pooled counties and the non-poor counties. The decomposition in Equation 5 requires separate estimates of the pooled model and the subsamples of persistently poor and non-poor counties, which we present below.

4. Convergence and Persistent Poverty

In Table 4, we present the estimates of our baseline model following the Arellano and Bover (1995) GMM specification. Our data set is described in detail in Appendix A. The dependent variable is the log of real per capita income, while the independent variables include lagged log real income per capita (y_{it-1}); the percentage of high school graduates; real per capita private capital expenditures in manufacturing; labor force growth; urban share; and the shares of residents who are black, who are churched, and who vote in presidential elections; and the levels of time-invariant historical variables in Table 2 that proxy for culture, institutions, geography, human and physical capital, and agglomeration. The lagged dependent variable is predetermined, but in first differences becomes correlated with the model error and thus is instrumented (Anderson and Hsiao 1982). The other time-varying independent variables are predetermined. To enhance efficiency of the parameter estimates, we use a block diagonal instrument set as proposed in Arellano and Bond (1991), consisting of $(t - 2)$ to $(t - 4)$ lags of log real income per capita (depending on year), $(t - 1)$ to $(t - 4)$ lagged levels of the time-varying variables, and the level of time-invariant historical variables (see Appendix B).

[10] The functional form of the production function is assumed constant for the groups; only the parameters are allowed to vary.

Table 4. Dynamic Panel GMM Estimates of Income

	(1) Pooled Model	(2) Persistently Poor	(3) Nonpoor
Current time-varying factors			
Lag income per capita	0.365 (0.023)	0.361 (0.064)	0.355 0.023)
Fraction high school	0.822 (0.065)	1.159 (0.252)	0.809 (0.063)
Capital spending (per capita)	8.497 (4.054)	7.943 (9.517)	9.530 (4.294)
Labor force growth	0.117 (0.014)	0.142 (0.044)	0.107 (0.014)
Urban share	0.333 (0.042)	0.031 (0.113)	0.349 (0.043)
Black share	0.603 (0.109)	0.353 (0.383)	0.385 (0.097)
Fraction churched	0.071 (0.019)	0.004 (0.048)	0.128 (0.020)
Fraction voted in pres. election	0.255 (0.044)	0.194 (0.115)	0.227 (0.045)
Historic time-invariant factors			
Illiteracy rate 1900	−0.457 (0.081)	−0.408 (0.255)	−0.152 (0.068)
Capital in 1890	0.014 (4.057)	−25.327 (45.576)	4.113 (3.888)
Proportion foreign born 1900	−0.075 (0.037)	0.228 (0.297)	−0.045 (0.037)
Urban share 1890	−0.241 (0.038)	−0.076 (0.131)	−0.230 (0.038)
Proportion churched 1890	0.041 (0.020)	−0.074 (0.107)	0.023 (0.021)
Land tenure 1890	0.013 (0.028)	−0.058 (0.113)	−0.044 (0.027)
SD to area	0.197 (0.021)	0.324 (0.270)	0.179 (0.020)
Average temperature 1895 (x100)	−0.105 (0.078)	0.206 (0.287)	−0.143 (0.076)
Average precipitation 1895	0.036 (0.004)	0.052 (0.020)	0.027 (0.003)
Constant	5.216 (0.177)	5.056 (0.475)	5.392 (0.182)
Convergence rate	0.101	0.102	0.097
Sargan test [d.f., *p*-value]	669.39 [68, 0.000]	82.73 [68, 0.1078]	617.38 [68, 0.000]
Wald test of equal Coef. [d.f., *p*-value]		245.73 [20, 0.000]	

Notes: Robust standard errors are reported in parentheses. The number of county-years in the pooled model is 12,000 (1170 persistently poor and 10,830 nonpoor). Each model controls for time effects. See the text for details.

Pooled Model

Column 1 of Table 4 reports the results of the pooled model for the 12,000 county-year observations in our sample. The results broadly indicate that both current and past levels of human capital accumulation have important effects on current income levels. For example, a 1% point increase in the fraction of high-school graduates implies an increase in income levels of 8.2%, while a 1% point decrease in 1900 illiteracy rates implies a 4.6% increase in income, holding other factors constant. Other variables such as current urban agglomeration, labor force growth, fraction churched, and fraction voting in Presidential elections are also positively correlated with income. The coefficient estimate on the black share is positive and statistically significant, but becomes negative if we drop the human capital controls. This suggests that controlling for differences in human capital (both current and historical), the share of blacks in the county is associated with higher per capita income.

The convergence rate implied by the models is found by the transformation $\text{Convergence} = -1 \times \ln(\hat{\beta})/\text{time}$, where *time* (10 years in our case) is the interval of data measurement (Islam 1995) and $\hat{\beta}$ is the coefficient defined in Equation 4. The estimated convergence rate is around 10%, which is within the range of estimates found in previous research at various levels of aggregation (e.g., Islam 1995 at the cross-country level; Higgins, Levy, and Young 2006 at the county level).

In addition to illiteracy rates, Column 1 of Table 4 includes the other historical factors. The percentage foreign-born in 1900 is negatively correlated with modern income levels, and this effect is statistically significant; the historical share of urban residents is also negatively correlated with income. While counterintuitive, this is explained by the inclusion of the contemporaneous urban share. That is, dropping current urban shares results in the 1890 urban share having no effect on income levels. Culture and institutions, as proxied for by the proportion churched in 1890 and land tenure in 1890, respectively, do not seem substantively correlated with current income levels.[11]

We find that warmer and drier counties grow more slowly than cooler, wetter climates, ceteris paribus. These results are similar to the findings of Iyigun (2005), Masters and McMillan (2001), and Easterly and Levine (2003). Also, the more mountainous a county's terrain, the higher the per capita income, as evidenced by the positive, statistically significant coefficient estimate on the SD of elevation variable. Overall, the results for our historical variables are mixed: historical levels of human capital and measures of geography seem to matter for modern income levels and growth, while historical variables proxying for culture and institutions do not, similar to our preliminary model on persistent poverty in Table 3.

The pooled model thus shows that both past and current human capital accumulation are highly correlated with current income levels and growth. As shown in the last panel of the table, Hansen's *J* test indicates that the model's over-identifying restrictions are not accepted.[12] There is evidence that the *J* test is prone to over reject in the presence of multiple overidentifying restrictions as in our model (Ziliak 1997; Hall and Horowitz 1996; Parente and Silva; 2012); thus, in section 5, we vary the instrument matrix in several ways to examine the sensitivity of our estimates.

Persistently Poor Versus Nonpoor Aggregate Production Technologies

Consistent with the decomposition in Equation 5, we divide the pooled sample into persistently poor and nonpoor samples, and then estimate the models using GMM, as in the previous section, on each of the subsamples. Recall that in our baseline model, a county is considered persistently poor if it has a poverty rate of at least 20% in each decennial Census between 1960 and 2000, and "nonpoor" otherwise. We test this definition of persistent poverty in the robustness section of section 5. Column 2 of Table 4 reports the results for the persistently poor counties, while Column 3 reports those for the nonpoor counties.

For the persistently poor counties, Column 2 indicates that the current percentage of high school graduates, labor force growth, and higher precipitation are positively and statistically significantly correlated with income levels. Other contemporaneous and historical variables are not statistically significant at the 5% level. The convergence rate among the persistently poor counties remains about 10%, and Hansen's *J* test does not reject the null that the overidentifying restrictions are valid.

The results of Column 3 generally follow those of the pooled model in Column 1 (which is not surprising, since nearly 90% of the counties in the pooled sample are nonpoor). Current and past measures of human capital, along with current capital, urbanicity, and culture (fraction churched and voters) all lead to higher income levels and growth. In addition, the geography

[11] Results are unchanged the religion variable is disaggregated by denomination.

[12] The Sargan test for overidentification is derived under the null of homoskedasticity and is known to overreject in the presence of heteroskedasticity (Arellano and Bond 1991). We adopt the heteroskedasticity robust variant proposed in Ahn (1994).

variables are all correlated strongly with growth rates (negatively in the case of temperature), while indicators of historical capital, culture, and institutions are generally not.

As shown in the last panel of Table 4, the Wald test of equal coefficients between the regressions in Columns 2 and 3 rejects the null, indicating that the persistently poor and nonpoor counties have different production technologies. We find that the persistently poor counties have a qualitatively higher return on current human capital (1.16 vs. 0.81, and both are highly statistically significant) than the nonpoor counties, which is expected given the lower levels of human capital, although the *t*-test of equal returns does not reject the null that they are the same (*t*-statistic = 1.34). However, the marginal effect on growth of the 1890 urbanization rate is much higher in absolute value, and the marginal effect of illiteracy is lower, for the nonpoor counties than for the poor counties (recall that the persistently poor counties had much higher illiteracy rates in 1900, and were much less urbanized in 1890, than the nonpoor counties). The marginal effects of the SD of elevation and average temperature are also large in absolute value and statistically significant for the nonpoor sample, but only precipitation is statistically significant in the persistently poor sample.

Persistently Poor/Nonpoor Decomposition

In Table 5, we more formally quantify the underlying sources of the income gap between poor and nonpoor counties into differences in factor levels and productivity (current and historical economic and demographic endowments) and differences in production functions (coefficients) via the Oaxaca decomposition of Equation 5.

As seen in Table 5, the mean predicted per capita income difference between the persistently poor and nonpoor counties is 39.7% ($3938 in per capita income levels in Table 1). Of this, roughly 21% can be attributed to differences in the coefficients, while 79% is due to differences in factor endowments. Of that 79%, three quarters is due to differences in the levels of the current factors, while one quarter is due to differences in the historical factors. Of the historical factors, differences

Table 5. Poor/Nonpoor Decomposition

Predicted Gap in Current Income	−0.397
Proportion difference due to	
Factor endowments	0.791
Coefficients	0.209
Of factor endowments share, proportion due to	
Historical factors	
Human capital (illiteracy 1900; foreign born 1900)	0.346
Agglomeration (urban share 1890)	−0.077
Culture (fraction churched 1890)	−0.002
Institutions (land tenure 1890)	0.001
Geography (terrain; temperature; precipitation)	−0.018
Physical capital (capital 1890)	<0.000
Current factors	
Lag log income	0.479
High school completion	0.453
Capital	0.004
Labor force growth	0.036
Urban share	0.159
Black share	−0.413
Culture (churched; vote for President)	0.029

Notes: The decomposition is based on the results in Table 4, as described in the text.

Table 6. Decomposition Results, Alternative Model Specifications

	(1)	(2)	(3)
Predicted gap in current income	−0.394	−0.394	−0.258
Proportion difference due to			
Factor endowments	0.809	0.796	0.799
Coefficients	0.191	0.204	0.211
Of factor endowments share, proportion due to			
Historical factors			
Human capital (illiteracy 1900; foreign born 1900)	0.339	0.349	0.456
Agglomeration (urban share 1890)	−0.085	−0.079	−0.254
Culture (fraction churched 1890)	−0.001	−0.002	0.003
Institutions (land tenure 1890)	0.001	0.002	0.024
Geography (terrain; temperature; precipitation)	−0.013	−0.017	0.185
Physical capital (capital 1890)	−0.002	<0.000	−0.012
Current factors			
Lag log income	0.481	0.476	0.531
High school completion	0.424	0.460	0.446
Capital	0.018	<0.000	0.005
Labor force growth	0.039	0.036	0.058
Urban share	0.170	0.164	0.372
Black share	−0.415	−0.418	−0.781
Culture (churched; vote for President)	0.044	0.028	−0.034
Number of poor counties	234	234	234
Number of nonpoor counties	2166	2166	2166

Notes: Column 1 is also a dynamic panel GMM estimate, treating current factors as endogenous. In Column 2, current capital stock is used instead of current capital flow as a regressor. Column 3 uses a county-specific price deflator based on the county-level median rent. See text for further details.

in human capital remain extremely important, explaining 35% of the variation in current income, or approximately 27% of the total income gap between the persistently poor and nonpoor samples.[13] Given that the initial human capital variables (illiteracy and the percentage foreign-born) were measured 60 years before the sample starts, and that we traditionally think of human capital as capital embodied in people, it is remarkable that the level of human capital in a geographic area could have such a strong effect decades later.

The negative contribution of historical agglomeration suggests that agglomeration rates of the persistently poor counties predict that they should have *higher* modern incomes (as noted earlier, this is because, we also include the contemporaneous urban share). While the persistently poor counties were more rural in 1890 (urbanization rates of 0.03 for the persistently poor counties compared to 0.13 for the nonpoor), the coefficient estimate on 1890 urbanization in Table 4 is negative and statistically significant in both the pooled and nonpoor samples, suggesting that the overall relationship is negative. Differences in the other historical factors do not seem important in explaining the 1960–2000 income gap between the persistently poor and nonpoor.

Of the contemporary explanatory variables, differences in lagged income and human capital (fraction high school) explain a large amount of the income gap. Contemporary human capital explains 45% of the gap attributed to differences in the explanatory variables, or approximately 36% of the total gap. Differences in current urbanization rates are also important, although do not explain quite as much. Finally, the negative contribution of the difference in black share is

[13] Twenty-seven percent is calculated as 0.346×0.791; that is, the coefficients explain 79% of the income gap, and historical human capital explains 35% of the difference due to coefficients.

Table 7. Decomposition Results, Alternative Definitions of Persistently Poor/Nonpoor

	(1)	(2)	(3)	(4)
Predicted gap in current income	−0.346	−0.419	−0.327	−0.500
Proportion difference due to				
Factor endowments	0.797	0.781	0.876	0.810
Coefficients	0.203	0.219	0.124	0.190
Of factor endowments share, proportion due to				
Historical factors				
Human capital (illiteracy 1900; foreign born 1900)	0.087	0.285	0.264	0.407
Agglomeration (urban share 1890)	−0.047	−0.061	−0.097	−0.047
Culture (fraction churched 1890)	−0.003	0.004	−0.001	0.001
Institutions (land tenure 1890)	−0.007	−0.004	−0.0005	0.003
Geography (terrain; temperature; precipitation)	−0.039	−0.028	−0.018	0.017
Physical capital (capital 1890)	−0.001	−0.006	<0.000	<0.000
Current Factors				
Lag log income	0.457	0.475	0.446	0.499
High school completion	0.507	0.408	0.439	0.402
Capital	0.003	0.0007	0.005	0.005
Labor force growth	0.029	0.049	0.032	0.033
Urban share	0.086	0.137	0.206	0.087
Black share	−0.108	−0.275	−0.296	−0.448
Culture (churched; vote for President)	0.034	0.016	0.019	0.040
Number of poor counties	229	151	579	38
Number of nonpoor counties	1632	1199	1821	2362

Notes: In Column 1, all metro counties (Beale<4) are dropped. Column 2 drops all the counties west of the Mississippi River. In Column 3, persistently poor is defined as having poverty rates at least 20% for at least 30 of the 50 years. Column 4 changes the definition of persistently poor to at least 30% for 50 years. See the text for further details.

primarily due to our result that, when controlling for current and historical human capital, the coefficient estimate on black share is large, positive, and statistically significant in the pooled and nonpoor samples.

To summarize, the decomposition results presented in Table 5 show that most of the income difference between the persistently poor and nonpoor counties is due to differences in factors of production, rather than in differences in the coefficients. In other words, the persistently poor counties are different, and poorer, because they have lower levels of factors of production, rather than because they use the factors they have less efficiently. While much of the income difference is explained by contemporary factors, the contribution of historical levels of human capital is surprisingly large.[14] The combined contribution of historical and contemporary human capital explains 60% of the overall income gap between the persistently poor and nonpoor counties.[15]

5. Sensitivity Analysis

In this section, we consider two broad sets of robustness checks on our model estimates: one set on model specification and a second on the definition of poor and nonpoor counties. The test

[14] Similarly, Simon and Nardinelli (2002) find long-lasting effects of initial human capital in the U.S. cities.

[15] We note that historical data on institutional factors are limited, and to the extent that omitted historical institutions are correlated with (current and historical) factor endowments, our decomposition procedure may attribute some of the difference in incomes to differences in factor endowments and not institutions.

of overidentifying restrictions in Table 4 rejects the null of valid restrictions for the pooled model and nonpoor sample, but not for the persistently poor sample. We thus vary the number of instruments to see how the model estimates change under alternative identification conditions. Specifically, in Column 1 of Table 6, we present estimates of our model under the assumption that current factors of production are contemporaneously correlated with the model time-varying error, as would occur if migration, and factor mobility more generally, were endogenous. This means lagged X instruments dated at time $(t-1)$ are not valid, and thus, we drop them (and use $[t-2]$ and earlier). The baseline results in Table 4 are remarkably robust to the exclusion of the additional instruments as reported in Column 1 of Table 6; namely, lagged income, current and past human capital, and current urban share account for the majority of the income gap. Strictly, the pseudo likelihood ratio test of Newey and West (1987), which compares the J-test from the unrestricted model in Table 4 to the J-test from the restricted model in Column 1 of Table 6, rejects the $(t-1)$ instruments for the pooled model and the nonpoor model, but economically the results are the same.[16]

In results not tabulated (but available on request), we use a more parsimonious standard instrument set rather than the block-diagonal approach of Arellano and Bond (1991), consisting of $(t-2)$ lags of log income, $(t-1)$ and $(t-2)$ lagged levels of the time-varying variables, and the levels of time-invariant variables, as well as a just identified variant.[17] Overall, even though the Hansen test continues to reject in the nonpoor sample with the much smaller instrument set, these estimates tell a similar story to those in Tables 4 and 5. It remains the case that differences in the explanatory variables explain the majority (75%) of the income gap between the persistently poor and nonpoor counties, while differences in the coefficient estimates (reflecting the efficiency with which the factors of production are used) explain 25%. Both historical and contemporary human capital contribute less to the income gap with the standard instrument set, while geography and current lagged income contribute more, but human capital overall continues to be the dominant factor.

In Columns 2 and 3 of Table 6, we return to the specifications of Tables 4 and 5 (with current factors predetermined rather than endogenous) and consider robustness to alternative definitions of key variables in our model. In Column 2, we consider the effect of using capital stock, rather than capital flow, as a regressor, with very little change to the results presented in Table 5.[18] In a related robustness check, we included the value of farms and buildings in the historical capital variable to better approximate the historical capital stock; results were nearly identical.[19]

Finally, in Column 3 of Table 6, we incorporate county-level median rent into the price deflator, creating a county-specific price deflator. Although our controls for county fixed effects capture permanent cost-of-living differences across counties, there could be some geographic cost differences that vary over time. Housing costs are the most likely factor in the consumption bundle that varies across counties, and is the largest component, accounting for between 33 and 45% of total outlays (Renwick 2010; Moretti 2013). We thus construct a county-year specific price index as

[16] For example, the pseudo LR test statistic is 146, and with 28 d.f. distributed asymptotically chi-squared, the p-value is <0.001 in the non-poor sample..

[17] We also restricted the instrument set to only $(t-1)$ and $(t-2)$ lagged levels of the time-varying and time-invariant variables, and $(t-2)$, $(t-3)$, and $(t-4)$ lags of log income. The results were very similar to what we obtained in Table 4, and are available on request.

[18] We use the method outlined in Hall and Jones (1999) to calculate capital stock from capital flow data, assuming a depreciation rate of 6%.

[19] These results are available from the authors.

$\text{index}_{ct} = 0.4 \times \frac{\text{median rent}_{ct}}{\text{median rent}_{t}} \times \text{PCE}_t + 0.6 \times \text{PCE}_t$, where c denotes county, t denotes year, and PCE is the aggregate personal consumption deflator. We assume that housing affects 40% of outlays, and normalize median rent in the county by the average median rent in the nation so that the index captures housing costs differences that are above or below average. We use this deflator for all income and capital variables instead of the aggregate deflator in this exercise. Not surprisingly, this reduces the predicted current income gap (by about one-third); however, the shares due to factor endowments and coefficients are unchanged. There are some differences in the relative contributions of endowments, for example, geography and current urban share matter more, but historical and current human capital *each* explain 36% of the total income gap.

In the second major robustness check, presented in Table 7, we consider alternative samples and alternative definitions of persistently poor counties, and repeat the analysis of Tables 4 and 5. In Column 1 of Table 7, we consider only nonurban counties. Since the persistently poor counties are generally rural, the comparison between poor and nonpoor may be overstated because of the large urban areas located in nonpoor counties.[20] In Column 2, we exclude all counties west of the Mississippi River, since many of these counties were substantially less developed than eastern counties at the start of the twentieth century. In the final two specifications, we consider alternative definitions of persistently poor: in Column 3, a county must have a poverty rate over 20% in at least three of the five Census years from 1960 to 2000, while in Column 4, a county must have a poverty rate of at least 30% in each decennial Census between 1960 and 2000. In all of the estimations, we use the base case block diagonal instrument matrix consisting of $(t-2)$, $(t-3)$, and $(t-4)$ lags of log income per capita and $(t-1)$ through $(t-4)$ lagged levels of the time-varying variables, and the level of time-invariant variables.

Across these different specifications, the decompositions in Table 7 show that the predicted gap of log income between the poor and nonpoor counties ranges from -0.33 to -0.50. In each specification, variation in factors of production explains at least 78% of the income gap, consistent with the results from Tables 5 and 6. Human capital, both historical and contemporary, generally remains important: in Columns 2–4, historical human capital explains between 26 and 41% of the income gap explained by variation in production factors (between 22 and 33% of the total gap), while in all regressions, contemporary human capital explains between 40 and 51% of the gap attributed to differences in factors (and 32–40% of the overall gap). When urban counties are excluded, the contribution of historical human capital decreases markedly because rates of illiteracy were more comparable across counties, but the share attributable to current human capital increases. In sum, across each specification of Table 7, historical and contemporary human capital combine to explain at least half of the total income gap between persistently poor and nonpoor counties.

6. Conclusion

Deep pockets of poverty have remained in several regions of the United States despite the widespread availability of technology and the lack of institutional barriers to labor mobility across county and state borders. To examine the sources of persistent poverty across regions, we estimated a dynamic panel data model of conditional income convergence using county-level data

[20] We use the Economic Research Service's rural-urban continuum code (aka the Beale code) of 1974 to define whether a county is urban or rural. If a county has a Beale score less than 3, it is considered an urban county and is excluded from this specification (whether persistently poor or non-poor). See a description of the coding at http://www.ers.usda.gov/briefing/rurality/ruralurbcon/.

from the past five Censuses. We also incorporated some fundamental determinants of income, including variables proxying for historical culture, geography, human capital, and institutions.

We find evidence of significant regional differences in production function parameters, but our decompositions of the income gap between poor and nonpoor suggests that at least three-fourths of the income gap is explained by differences in productive factors. Persistently poor counties are different (and poorer) because they have lower levels of factors of production, not because they use the factors they have less efficiently. Much of the income difference is explained by contemporary factors—own (recent) past income, human capital, and urban share. We can rule out any direct major role of geography, culture, and institutions as we measure them. However, the combined contribution of historical and contemporary human capital is large, explaining over half of the overall income gap between the persistently poor and nonpoor counties. While geography (captured by SD of elevation, precipitation, and temperature) does not appear to be a barrier to riches—in fact, more rugged terrain counties have higher per capita incomes, controlling for other factors—the importance of current urbanization, or at least proximity to urban areas, is an important determinant of growth, accounting for up to 20% of the income gap.

In short, the most important factors contributing to persistent poverty are low levels of human capital and low urbanization rates. While daunting, these are not insurmountable barriers from a policy perspective. For example, in the case of Appalachia, Bollinger, Ziliak, and Troske (2011) and Black and Sanders (2012) found that human capital shortfalls are greatest at the college level, while Kahn (2012) argues that the region has not adequately developed the urban centers to take advantage of ecotourism or regional universities as incubators for innovation. Similar arguments can be made for the other four persistently poor regions in the nation. This suggests that investments in education, coupled with economic development programs that aim to diversify the economic base around nearby urban centers, may offer a path out of persistent poverty.

Appendix A: Construction and Source of Data

A summary of variable definitions and their measurement units are shown in Table A1, and below we provide a description.

Contemporaneous Data

The contemporaneous county-level variables were collected from the 1960–2000 Decennial Censuses. The U.S. Counties Basic Information database of the Census Bureau provides information on many of the variables for 1980–2000 Census. Included in this database are county per capita income (average income earned by the residents of the county), the total population of the county, civilian labor force residing in the county (defined as the number of people in the labor force over the age of 16 in the county who are not in the armed forces and are not institutionalized), number of people living in urban areas in the county (defined below), number of African-Americans living in each county, the number of adults who voted in a presidential election, persons living under the poverty level in the county according to the official poverty definition of the United States, and the proportion of residents residing in the county who are over the age of 25 and have at least a high school degree. These data are publicly available from the URL: http://censtats.census.gov/.

The corresponding variables for the years 1960 and 1970 were collected from the County and City Data Book of the Census (1962, 1972, and 1977), which are available on the website of the Inter-University Consortium for Political and Social Research (ICPSR). There were no presidential elections in 1970 and 1990, and so voting data of 1972 and 1992 presidential elections, respectively, were used instead for those years. This site is maintained by the University of Michigan, and the data can be obtained from the following link: http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/2896?q=historical+data&searchSource=icpsr-landing.

Table A1. Description of Variables Used in the Regressions

Variable	Description	Measurement Units
Per capita income ($)	Average income earned by the residents of the county	Dollar amounts; values from 1960 to 1990 have been converted to $2000 using the personal consumption expenditure deflator from the BEA
Population	Total number of people living within the boundary of a county	Absolute value
Fraction in labor force	The number of people over the age of 16 in the county who are not used in the armed forces and are not institutionalized, divided by the population of the county.	Between 0 and 1
Growth in labor force	The increase in civilian labor force in a country from one decade to the next	Between 0 and 1
Fraction high school graduate	The number of residents over the age of 25 with at least a high school degree in the county, divided by the population of the county	Between 0 and 1
Capital expenditure	A permanent addition or a major change made by a manufacturing firm and/or the addition or replacement of any machinery or equipment in the plant (and whose depreciation account was maintained), divided by population	Thousands of dollars per capita (values have been converted to real 2000 dollars using the personal consumption expenditure deflator from the BEA)
Fraction living in urban area	The number of residents living in an urban area as defined by the Census, divided by the total population	Between 0 and 1
Fraction black	The number of African-Americans living in a county, divided by the total population	Between 0 and 1
Fraction churched	The number of people in a county who attend a church, divided by the total population of that county	Between 0 and 1
Fraction voted	The proportion of the population in the voting age group who voted in the presidential election	Between 0 and 1
Land tenure in 1890	The total area of farmland farmed by their respective owners in the year 1890, divided by the total area under cultivation in a county in 1890	Between 0 and 1
Share churched in 1890	The total number of people who attended church services in 1890, divided by the total number of people living in the county in 1890	Between 0 and 1
Share Baptist in 1890	The total number of people in a county who identify themselves as Baptists (Regular [North, South and Colored], Freewill, General, Primitive and Old Two-Seed denominations) in 1890, divided by the total population in 1890	Between 0 and 1
Share Catholic in 1890	The total number of people in a county who identify themselves as Catholics in 1890, divided by the total population in 1890	Between 0 and 1
		Between 0 and 1

Table A1 (*Continued*)

Variable	Description	Measurement Units
Share Calvinist in 1890	The total number of people in a county who identify themselves as Calvinists (Welsh Calvinist, Presbyterian (Northern and Southern), Cumberland Presbyterian (Regular and Colored), United Presbyterian, U.S. Reformed Church and American Reformed Church Organizations) in 1890, divided by the total population in 1890	
Average temperature 1895	The average monthly temperature of a county for the years 1895–1905	Degrees Fahrenheit
Average precipitation 1895	The average monthly precipitation of a county for the years 1895–1905	Inches
Terrain (SD to area)	The SD of elevation of a county divided by area of that county	Feet per square mile
Urban share in 1890	The number of residents living in an urban area as defined by the Census, divided by the total population in 1890	Between 0 and 1
Capital in 1890	Manufacturing capital spending, divided by the population of 1890	Thousands of dollars per capita, converted to 2000 real terms using the 1929 PCE
Share foreign born in 1900	The number of residents living in the county who were not born in the United States in 1900, divided by the total population	Between 0 and 1
Illiteracy rate in 1900	The number of people who could not read or write in a county in 1900, divided by the total number of people living in the given county in 1900	Between 0 and 1

The growth rate of the labor force is defined as the percentage change of civilian labor force in a county from one decade to the next. To construct this for 1960, we obtained the county-level civilian labor force population from the 1950 Census to construct the 1950–1960 change.

The definition of what constitutes as an urban area has changed over time. In 2000, the definition of urban areas was a core census block group or census block that had at least 1000 persons per square mile and the surrounding census blocks that have a population density of at least 500 persons per square mile (https://www.census.gov/geo/reference/ua/urban-rural-2000.html). For the years 1960, 1970, 1980, and 1990, the definition of an urban area was less stringent; any area that was one of the Census designated places with more than 2500 people, or was incorporated in an urban area (http://www.census.gov/population/censusdata/urdef.txt) was considered to be an urban area. We use this latter definition for the years 1960–1990.

Contemporaneous church attendance data is obtained from the website of the Association of Religious Data Archives (http://www.thearda.com). This website is maintained by the Department of Sociology of Pennsylvania State University. The fraction who attends church is obtained dividing the total number of adherents with the population of a county.

Private capital expenditure in the manufacturing sector is measured in thousands of dollars and scaled by county population. Data for the years 1960, 1970, and 1980 are taken from the appropriate County and City Data Books. Private capital expenditure is defined as either a permanent addition or a major change made by a manufacturing firm and/or the addition or replacement of any machinery or equipment in the plant (and whose depreciation account was maintained). The data for 1990 are obtained from the 1992 Census of Manufactures Report on each county (https://www.census.gov/prod/www/manufacturing.html). The 1990 data are not in electronic format and thus had to be coded in using a pdf-to-Excel converter. Some counties have missing data for 1990, and were filled in by taking the county-level average capital expenditure for the years 1960–2000 (except 1990). Private capital expenditure data for the year 2000 was obtained from U.S. Counties Basic Information database. The income and capital data from 1960 to 1990 were converted to real 2000 dollars using the personal consumption expenditure deflator from the Bureau of Economic Analysis (BEA).

Historical Data

Historical data on geography, institutions, culture, and human capital were obtained from various sources. Data on the SD of elevation of each county were provided by Jordan Rappaport of the Federal Reserve Bank of Kansas City (Rappaport and Sachs 2003). Elevation data were measured in feet relative to sea level to capture the ruggedness of terrain. The higher the SD, the more extreme the terrain of the county.

Historical temperature and precipitation data were obtained from the website of NOAA (http://www7.ncdc.noaa.gov/CDO/CDODivisionalSelect.jsp). However, weather data is not available by county, only by region within a state. The website (http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/regional_monitoring/CLIM_DIVS/states_counties_climate-divisions.shtml) provided a map that indicated which counties belonged to which region in a state. This data was coded by hand and then the file was merged to the main data set. Historical county data on land tenure, level of capital in each county, illiteracy rates, the fraction foreign-born, and the fraction living in urban areas (areas that have been legally incorporated as cities, towns, or boroughs, http://www.census.gov/population/www/documentation/twps0027/twps0027.html) were obtained from the 1890 and 1900 Census Database. Land tenure is defined as the percentage of total farmland that was farmed by owners. Data on acres of farmland in 1890 were provided in brackets of 0–9 acres of farmland, 10–19 acres, 20–49 acres, 50–99 acres, 100–499 acres, 500–999 acres, and 1000+ acres; the midpoint of each range was used to construct this variable. We obtained data on total church attendance and followers of different denominations, namely Baptists, Calvinists, and Catholics, from the 1890 Census of Religious Bodies. The Baptist denomination includes Regular (North, South, and Colored), Freewill, General, Primitive, and Old Two-Seed denominations. The Calvinist denomination includes Welsh Calvinist, Presbyterian (Northern and Southern), Cumberland Presbyterian (Regular and Colored), United Presbyterian, U.S. Reformed Church, and American Reformed Church Organizations. The religious data were obtained from the ICPSR database: http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/2896?q=historical+data&searchSource=icpsr-landing.

A number of counties changed their area and many new counties formed over time. Large counties were split to form new counties and some counties were merged to form a new county. Only counties with consistent borders since 1890 were used in the analysis. Consistent county borders were determined by comparing by hand current and 1890 county maps of the United States from the website of the Newberry Library (http://www.newberry.org).

Appendix B: Correlated Random Effects Estimator

To implement the correlated random effects estimator for Equation 4 in the text, we adopt the Generalized Method of Moments framework of Arellano and Bover (1995). We begin by rewriting Equation 4 as

$$y_i = d_i\Lambda + \upsilon_i \tag{B2.1}$$

where y_i is the $T \times 1$ vector of log income per capita for county i, $d_i=[y_{i,-1}, X_i, \iota_T Z_i']$ is the T $\times$ $(1 + G + P)$ matrix of regressors for county i, ι_T is a $T \times 1$ vector of ones, $\Lambda=[(1+\beta), \varphi, \theta]$ is a $(1+G+P)\times 1$ vector of unknown parameters to estimate, and $\upsilon_i=\iota_T\psi_i+\delta_t+\xi_i$. For consistent estimates of Λ, we then construct a non-singular transformation, C, and a matrix of instruments, M_i, such that the moment conditions $E[M'_i C\upsilon_i]=0$ are satisfied. The transformation that we adopt is $C=\begin{bmatrix} K \\ \iota'_T/T \end{bmatrix}$ where K is a $(T-1)\times T$ matrix containing the first difference operator and ι'_T/T converts a variable into its time mean. Notice that K eliminates ψ_i from the first $(T - 1)$ rows, thus allowing the identification of the coefficients on time-varying regressors (y_{it-1}, X_{it}). The term ι'_T/T creates an equation in levels (i.e., "between-groups"), and permits identification of the coefficients on time-invariant regressors Z_i.

For the instruments, Arellano and Bover suggest the Arellano–Bond (1991) block-diagonal instrument matrix of the form $M_i=I_T \otimes [g'_i, \ldots, g'_i, Z'_i]$, where I_T is a $T\times T$ identity matrix and $g_i=(y_{i(-2)}, x_{i(-1)}, Z_i)$ are instruments consisting of $(t - 2)$ to $(t - 4)$ lags of log income per capita depending on year, $(t - 1)$ to $(t - 4)$ lagged levels of the time-varying variables (X_{it}), and the level of time-invariant variables (historical variables, Z_i):

$$\begin{bmatrix} 0 & 0 & 0 & X_{it-1} & 0 & 0 & 0 \\ y_{it-2} & 0 & 0 & 0 & X_{it-1} & X_{it-2} & 0 \\ 0 & y_{it-2} & y_{it-3} & \cdots\ 0 & 0 & 0 & \ldots\ 0 \\ & \vdots & & & \vdots & & \vdots \\ & 0 & & & 0 & & Z_i \end{bmatrix}$$

Under first-differencing of the first $(T-1)$ rows in C, the twice-lagged level of the dependent variable is a valid instrumental variable for the lagged first-difference regressor. Moreover, if we assume that the X_{it} in Equation 4 are predetermined, which for variables such as capital and labor force seems reasonable, then we must lag the instruments by one period to maintain consistency. Note also that g_i contains the time-invariant Z's, which drop out due to first differencing in the first $(T-1)$ rows but then serve as instruments for themselves in the level equation in time T. Stacking the observations across all i, the GMM estimator is given as

$$\hat{\Lambda}=[d'\bar{C}'M(M'\bar{C}\hat{\Omega}\bar{C}'M)^{-1}M'\bar{C}d]^{-1}d'\bar{C}'M(M'\bar{C}\hat{\Omega}\bar{C}'M)^{-1}M'\bar{C}y, \tag{B2.2}$$

where $\bar{C}=I_N\otimes C$, I_N is an $N\times N$ identity matrix, and $\hat{\Omega}$ is a conformable matrix. The one-step GMM estimator replaces $\bar{C}\hat{\Omega}\bar{C}'$ with $J=I_N\otimes j$, where $j=\begin{pmatrix} j_d & 0 \\ 0 & j_l \end{pmatrix}$ with j_d a $(T-1)\times(T-1)$ matrix with 2s on the diagonal and -1 on the off-diagonal accounting for the first difference transform, and j_l is equal to the identity matrix (see Arellano and Bond 1991). The two-step GMM estimator uses the residuals from the one-step estimator to form squared residuals on the principal diagonal of $\hat{\Omega}$. As demonstrated in Ziliak (1997) and others, the two-step GMM estimator may be biased in finite samples owing to the correlation between the estimated first-stage residuals and the second stage moments, and thus we use the one-step GMM estimator.

Acknowledgments

We thank John Pepper, two anonymous reviewers, and seminar participants at Clemson University, University of Illinois, and the Southern Growth Conference at the Southern Economic Association conference for helpful comments on earlier versions. We also thank Jordan Rappaport for generously sharing data on county elevation. Part of this research was conducted while Islam was affiliated with the University of Kentucky and the University of Wisconsin-Green Bay. All errors are our own.

References

Acemoglu, D., S. Johnson, and J. Robinson. 2005. Institutions as the fundamental cause of long-run growth. In *Handbook of Economic Growth*, volume 1, edited by P. Aghion and S. N. Durlauf. Amsterdam: North-Holland Elsevier, pp. 385–472.

Ahn, S. 1994. *Robust GMM tests for model specification, with applications to conditional moments testing and structural instability testing*. Mimeo, Arizona State University.

Anderson, T., and C. Hsiao. 1982. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics* 18(1):47–82.

Arellano, M., and B. Stephen. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies*. 58(2):277–297.

Arellano, M., and O. Bover. 1995. Another look at the instrumental variable estimation of error-components model. *Journal of Econometrics*. 68(1):29–51

Azariadis, C., and A. Drazen. 1990. Threshold externalities in economic development. *Quarterly Journal of Economics* 105(2):501–26.

Banfield, E. 1970. *The Unheavenly City*. Boston: Little, Brown and Company.

Barrington, L. 1999. Editor's Introduction: Native Americans and U.S. Economic History. In *The Other Side of the Frontier: Economic Explorations into Native American History*, edited by L. Barrington, Westview Press. https://www.ilr.cornell.edu/directory/ljb239/download/publication/Linda%20Barrington%20Publications.pdf.

Barro, R., and R. M. McCleary. 2003. Religion and economic growth. *American Sociological Review* 68(5):760–81.

Barro, R., and X. Sala-i-Martin. 1992. Convergence. *Journal of Political Economy* 100(2):223–51.

Bauer, P., M. Schweitzer, and S. Shane. 2012. Knowledge matters: The long-run determinants of state income growth. *Journal of Regional Science* 52(2):240–55.

Beeson, P. E., D. N. DeJong, and W. Troesken. 2001. Population growth in U.S. counties, 1840-1990. *Regional Science and Urban Economics* 31:669–99.

Billings, D. 1974. Culture and poverty in Appalachia: A theoretical discussion and empirical analysis. *Social Forces* 53: 315–23.

Bollinger, Christopher, James P. Ziliak, and Kenneth R. Troske. 2011. Down from the mountain: skill upgrading and wages in Appalachia. *Journal of Labor Economics* 29(4):819–857.

Billings, D., and K. Blee. 2000. *The Road to Poverty: The Making of Wealth and Hardship in Appalachia*. New York: Cambridge University Press.

Black, D., and Seth Sanders. 2012. Inequality and Human Capital in Appalachia, 1960–2000. Appalachian Legacy: Economic Opportunity after the War on Poverty, James P. Ziliak (editor). Washington: Brookings Institution, 45–80.

Blundell, R., and S. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87(1):115–43.

Cavalcanti, T., S. Parente, and R. Zhao. 2007. Religion in macroeconomics: A quantitative analysis of Weber's thesis. *Economic Theory* 32:105–32

Clifton, E., and A. Romero-Barrutieta. 2006. Institutions v. geography: Sub-national evidence from the United States, *International Monetary Fund Working Paper*.

Deaton, A. 2010. Instruments, randomization, and learning about development. *Journal of Economic Literature* 48(2): 424–55.

Duncan, C. 1999. *Worlds Apart: Why Poverty Persists in Rural America*, New Haven: Yale University Press.

Durlauf, S. N., and P. A. Johnson. 1994. Multiple regimes and cross-country growth behavior, *Journal of Applied Econometrics* 10:365–84.

Durlauf, S. N., P. Johnson, and J. Temple. 2005. Growth Econometrics. In *Handbook of Economic Growth*, volume 1, edited by P. Aghion and S. N. Durlauf. Amsterdam: North-Holland Elsevier, pp. 555–677.

Easterly, W., and R. Levine. 2001. What have we learned from a decade of empirical research on growth? It's not factor accumulation: Stylized facts and growth models. *The World Bank Economic Review* 15(2):177–219.

Easterly, W., and R. Levine. 2003. Tropics, germs, and crops: How endowments influence economic development. *Journal of monetary economics* 50(1):3–39.

Eller, R. 2008. *Uneven Ground: Appalachia Since 1945*. Lexington: The University Press of Kentucky.

Engerman, S., and R. Fogel. 1974. *Time on the Cross: The Economics of American Negro Slavery*. Boston: Little Brown.

Evans, P., and G. Karras. 1996. Convergence revisited. *Journal of Monetary Economic* 37(2):249–65.

Gallup, J., J. Sachs, and A. Mellinger. 1999. Geography and economic development, *International Regional Science Review* 22(2):179–232.

Gennaioli, N., R. La Porta, F. Lopez-de-Silanes, A. Shleifer. 2013. Human capital and regional development. *Quarterly Journal of Economics* 128(1):105–64.

Glaeser, E., H. D. Kallal, J. A. Scheinkman, and A. Shleifer. 1992. Growth in cities. *The Journal of Political Economy* 100(6):1126–52.

Greif A. 1994. Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *The Journal of Political Economy* 102(5):912–50.

Greenwood, M. 1997. Internal migration in developed countries. In *Handbook of Population and Family Economics*, volume 1B, edited by M. Rosenzweig and O. Stark, Amsterdam: North Holland.

Hall, P., and J. Horowitz. 1996. Bootstrap critical values for tests based on generalized method of moments estimators, *Econometrica* 64:891–916.

Hall, R. E., and C. I. Jones. 1999. Why do some countries produce so much more output than others? *Quarterly Journal of Economics* 114(1):83–116.

Higgins, M., D. Levy, and A. Young. 2006. Growth and convergence across the United States: Evidence from county-level data. *Review of Economics and Statistics* 8(4):671–81.

Islam, N. 1995. Growth empirics: A panel data approach. *The Quarterly Journal of Economics* 110(4):1127–70.

Iyigun, M. 2005. Geography, demography, and early development. *Journal of Population Economics* 18(2):301–21.

Kahn, M. 2012. Cities, Economic Development and the Role of Placed-based Policies: Prospects for Appalachia. Appalachian Legacy: Economic Opportunity after the War on Poverty, James P. Ziliak (editor). Washington: Brookings Institution, 149–168.

Kennan, J., and J. Walker. 2011. The effect of expected income on individual migration decisions, *Econometrica* 79(1): 211–51.

Lee, K., M., H. Perasan, and R. Smith. 1997. Growth and convergence in a multi-country empirical stochastic Solow model. *Journal of Applied Econometrics* 12(4):357–92.

Mankiw, N. G., D. Romer, and D. N. Weil. 1992. A contribution to the empirics of economic growth, *The Quarterly Journal of Economics* 107(2):407–37.

Masters, W. A., and McMillan, M. S. 2001. Climate and scale in economic growth. *Journal of Economic Growth* 6(3), : 167–86.

Moretti, E. 2004. Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *The American Economic Review* 94(3):656–90.

Moretti, E. 2013. Real wage inequality. *American Economic Journal: Applied Economics* 5(1):65–103.

Murray, C. 1984. *Losing ground: American social policy, 1950-1980*. New York: Basic Books.

Newey, Whitney K., and Kenneth D. West. 1987. Hypothesis testing with efficient method of moments estimation. *International Economic Review* 777–787.

Oaxaca, Ronald L., and M. Ransom. 1994. On discrimination and the decomposition of wage differentials. *Journal of Econometrics* 61(1):5–21

Parente, P., and J. Silva. 2012. A cautionary note on tests of overidentifying restrictions. *Economics Letters* 115:314–7.

Ransom, R., and R. Sutch. 2001. *One Kind of Freedom: The Economic Consequences of Emancipation*, 2nd edition. New York: Cambridge University Press.

Rappaport, J. 2007. Moving to nice weather. *Regional Science and Urban Economics* 37(3):375–98.

Rappaport, J., and J. Sachs. 2003. The United States as a coastal nation. *Journal of Economic Growth* 8(1):5–46.

Renwick, T. 2010. *Geographic adjustments of supplemental poverty measure thresholds: Using the American community survey five-year data on housing costs*. U.S. Census Bureau.

Rupasingha, A., S. Goetz, and D. Freshwater. 2000. Social capital and economic growth: A county-level analysis. *Journal of Agricultural and Applied Economics* 32(3):565–72.

Shapiro, T. 2006. Race, homeownership and wealth. *Washington University Journal of Law and Policy* 53.

Simon, C. J., and C. Nardinelli. 2002. Human capital and the rise of American cities, 1900-1990. *Regional Science and Urban Economics* 32:59–96.

Solow, R. 1956. A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70(1):65–94.

Turner, C., R. Tamura, S. E. Mulholland, and S. Baier. 2007. Education and income of the United States: 1840-2000. *Journal of Economic Growth* 12(2):101–58.

Ziliak, J. 1997. Efficient estimation with panel data when instruments are predetermined: An empirical comparison of moment-condition estimators. *Journal of Business and Economic Statistics* 15(4):419–31.