

UKCPR
University of Kentucky
Center for
Poverty Research

Discussion Paper Series

DP 2012-01

ISSN: 1936-9379

On Persistent Poverty in a Rich Country

T.M. Tonmoy Islam

Department of Urban and Regional Studies and Economics
University of Wisconsin -Green Bay
islamt@uwgb.edu

Jenny Minier

Department of Economics
University of Kentucky
jminier@uky.edu

James P. Ziliak*

Center for Poverty Research
and
Department of Economics
University of Kentucky
jziliak@uky.edu

Preferred citation

Islam, T.M., Minier, J., & Ziliak, . On Persistent Poverty in a Rich Country. *University of Kentucky Center for Poverty Research Discussion Paper Series, DP2012-01*. Retrieved [Date] from <http://www.ukcpr.org/Publications/DP2012-01.pdf>.

Author correspondence

James P. Ziliak, Center for Poverty Research, University of Kentucky, Lexington, KY 40506-0047; Email: jziliak@uky.edu; Phone: 859-257-6902.

University of Kentucky Center for Poverty Research, 302D Mathews Building, Lexington, KY, 40506-0047
Phone: 859-257-7641; Fax: 859-257-6959; E-mail: ukcpr@uky.edu

www.ukcpr.org

EO/AA

Abstract: We examine differences in income within the U.S., and the regions of persistent poverty that have arisen, using a newly assembled dataset of counties that links historical 19th century Census data with contemporaneous data. The data, along with an augmented human capital growth model, permit us to identify the roles of contemporaneous differences in aggregate production technologies and factor endowments, in conjunction with the historical roles of institutions, culture, geography, and human capital. We allow for possible cross-county factor mobility via a correlated random effects GMM estimator that identifies simultaneously the coefficients on time varying and time-invariant determinants of income. We find evidence of significant regional differences in production technologies, but our decompositions of the poor/non-poor income gap suggests 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. While much of the income difference is explained by contemporary factors, the contribution of historical levels of human capital is surprisingly large. The combined contribution of historical and contemporary human capital is striking: together, they explain nearly 60 percent of the overall income gap between the persistently poor and non-poor counties.

JEL Codes: O1, R1

Between 1960 and 2000, real income per capita in the United States increased 175 percent and aggregate poverty rates fell by half, from 22 percent to just over 11 percent. 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 percent since 1960, encompass five distinct regions and 11 percent 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-economy income differences, differences in production technologies (that is, differences in the efficiency with which a region employs its factors) may also result in persistent income differences across regions.² In the context of this paper, to what extent are persistently poor counties poor because they have lower levels of inputs like physical and human capital, and to what extent because they use the capital they have less efficiently than other counties? 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. (See, among many others, Hall and Jones 1999; Grief 1994; Easterly and Levine 2001;

¹ The U.S. Department of Agriculture (USDA) defines a county as persistently poor if its poverty rate exceeds 20 percent 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 percent poverty rate, though fewer counties meet the criteria in each subregion.

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

Rappaport and Sachs 2003; Acemoglu, Johnson, and Robinson 2005.) These considerations lead us to examine three broad explanations why certain regions in the U.S. might be on a divergent growth path towards persistent poverty: (1) lower levels of contemporaneous factors of production, such as physical and human capital; (2) less efficient use of those factors; and (3) lower levels of productivity (residual income), perhaps determined by the historical roles of institutions, culture, geography, and endowments of human capital.

We begin by assembling a new dataset of U.S. counties that links historical data from the end of the 19th century to contemporaneous data from the end of the 20th century. 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 U.S. using subnational data at the state level (Barro and Sala-i-Martin 1992; Evans and Karras 1996; Bauer, Schweitzer, and Shane 2006; Gennaioli, La Porta, Lopez-de-Silanes, and Shleifer forthcoming), and fewer still at the county level (Higgins, Levy, and Young 2006), but to our knowledge none have linked historical county data with contemporaneous data as we do here.³

Our starting point with the data is a simple descriptive examination of links between initial productivity from the late 19th century and the probability of being persistently poor at the close of the 20th century. 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 non-poor counties, and this initial human capital shortfall dominates geography, culture, and institutions as a reduced-form predictor of long-term poverty.

³ Rappaport and Sachs (2003) 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.

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 attractive because a key determinant of steady-state income in Mankiw, Romer, and Weil (1992) is a country's initial level of productive efficiency (technology/productivity; we will follow Hall and Jones (1999) in using the term "productivity"), and heterogeneity in productivity seems likely. In our case, U.S. counties are the unit of analysis, and so we allow productivity to differ across counties by including county-specific intercepts. Whereas many growth papers treat the latter as nuisance parameters, in our case it is the historical human capital, institutions, geography and culture that enter via fixed county intercepts and thus are fundamental. Hence, we adopt the correlated random effects GMM estimator of Arellano and Bover (1995) that identifies parameters on both time-varying and time-invariant factors. That is, we can simultaneously identify the contributions of current factor accumulation and historical productivity to income levels and growth. An added feature of this estimator is that it is straightforward to treat current factors of production as predetermined, or possibly endogenous, to the income process. For example, if human capital is mobile across counties, and this migration is related directly or indirectly to a county's income prospects as we might expect (Greenwood 1997; Kennan and Walker 2011; Gennaioli, et al. forthcoming), then imposing the assumption of exogenous migration would lead to biased estimates. Our GMM estimator permits us to relax exogeneity of contemporaneous factors.

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

We find evidence of significant regional differences in production technologies, but our decompositions of the poor-non-poor income gap suggests that at least three fourths of the gap is explained by differences in current and past 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 past income, human capital, and urban share—and while some of the geographic variables matter (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 is striking, and combined with contemporaneous differences, accounts for nearly 60 percent the income gap between persistently poor and nonpoor counties.

II. 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. In Figure 2, we present county poverty rates for 1959 based on income reported in the

1960 Census.⁴ County poverty rates in excess of 50 percent were common in the South, and those in excess of 20-30 percent were the norm in the Midwest and Plains states. With strong economic growth and expansion of income support as part of the Johnson Administration's Great Society programs, the poverty landscape changed dramatically over the next decade. Figure 3 shows that county poverty rates in the 1970 Census were considerably lower throughout much of the nation, especially in the South and Midwest. However, shades of the five regions of persistent poverty already emerge as poverty rates remained in excess of 40 percent in many of the counties. Further progress against poverty continued through the 1970s, as seen in Figure 4, but Figure 5 suggests that gains against poverty abated in the 1980s. The economic expansion of the 1990s was similar in strength to the 1960s, and as seen in Figure 6, poverty fell compared to 1989, but the expansion was not enough to lift the poorest areas up to levels found elsewhere. Examining Figures 2 to 6, it is clear that in an absolute sense the poverty status of many counties that are identified as persistently poor in Figure 1 improved over the past four decades, but poverty in many of those counties has remained four times higher than the national rate.

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 technology, institutions, and endowments of natural resources.⁵ Table 1 illustrates some descriptive differences in incomes and observed human and physical capital between the persistently poor counties and non-poor counties pooled across the 1960 to 2000 Censuses. In

⁴ The United States did not produce its first estimates of poverty until the 1960s, but in the special tabulation the Economic Research Service of the USDA produced estimates for the 1960 Census. We thank Robert Gibbs of USDA for providing these data.

⁵ 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 at the county level, many people may work in one county and live in another. While our Census data measures income earned by residents of a county, the Regional Economic Information System (REIS) data from the Bureau of Commerce, which is available for the later part of our sample, measures income earned in a county by residents and non-residents; for 2000, the correlation between the two measures is 0.91.

our sample, there are 12,000 county-years, consisting of 2,400 counties, 234 of which are categorized as persistently poor.⁶ Appendix 1 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 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. 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 percent of non-poor income. Likewise, high school completion rates in the persistently poor counties increased from 60 percent of the non-poor rate in 1960 to 81 percent in 2000; and real capital spending increased from 6.5 percent of the non-poor rate in 1960 to 11 percent 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 et al. 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 U.S., 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

⁶ Today there are 3,141 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 were comprised of territories; and some other counties were created by splitting larger counties). To abstract from differences owing to changes in composition of counties, our sample is comprised of 2,400 counties with consistently defined boundaries from 1890-2000. See the Appendix for details.

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 Fogel and Engerman (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 non-productive, 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 U.S. 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 non-poor counties (80 percent) compared to persistently poor counties (77 percent), 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; Grief 1994). Most prominent among these is the role of religion in economic development, especially the argument made by Max Weber that Protestantism (and Calvinism in particular) 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 U.S., religion-based cultural influences are determined in part by historic patterns of immigration. 19th 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 various 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 non-poor counties, providing some *prime facie* support for the Weber hypothesis. However, the non-poor counties also have higher historical shares of Roman Catholics (5.8 versus 3.4 percent), which is contrary to the pro-growth Weberian view. The historical share of Baptists, however, is twice as high in persistently poor counties as in non-poor 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, et al. 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, which make 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 of elevation as a measure of how mountainous a county's terrain is; average monthly temperature from 1895-1905; and average monthly precipitation from 1895-1905. Table 2 shows that persistently poor counties are warmer by just over 7 degrees on average, which is expected given their southerly location depicted in Figure 1, and on average experience nearly one-half inch more precipitation. Non-poor 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 (for example, Glaeser, Kallal, Scheinkman, and Shleifer 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, while we use the share of the population that resided in an urban area as our measure of agglomeration economies. Table 2 shows that the most striking differences in the historical data are in the human 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 percent versus 11 percent). Persistently poor counties also had urban shares 80 percent lower, and shares of foreign-born residents 75 percent lower, than non-poor counties.

For a preliminary look at whether these historical differences in institutions, geography, and human 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) and geography (temperature, precipitation, and terrain).⁷ The results in column 1 show that a one percentage point increase in the share of owner-occupied farmland in 1890 lowers the probability of being persistently poor a century later by seven percent, 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 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 and geography. 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 county⁸, so that the models in columns 4 and 5 yield the spurious result

⁷ This is akin 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 terrain—whereas they use only terrain.

⁸ The correlation coefficient between the 1900 illiteracy rate and the share of church members in 1890 is 0.265; for Baptists, Calvinists, and Catholics, respectively, the correlations are 0.630, -0.189, and -0.163.

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.

III. 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, 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

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

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:

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

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 of identical rates of technological progress; economies' steady states 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 et al. 2005):

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

where y_{it} is the natural log of real income per capita for county i ($=1, \dots, N$) in year t ($=1, \dots, T$); y_{it-1} is the lag of the dependent variable; and X_{it} are time-varying rates of factor accumulation (new capital investment, labor force growth rate, school attainment rates); and $\varepsilon_{it} = \mu_i + \delta_t + \xi_{it}$ 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” or “technology,” 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 the 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 capital endowments, and ψ_i is an error term. Substituting this into equation (3) yields:

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

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 2 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 T th 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 pre-determinedness assumption, we use lags of own income dated $(t-2)$ and earlier as instruments as described below and in Appendix 2 (Anderson and Hsiao 1982).

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. 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 $Wald = (\hat{\Lambda}^P - \hat{\Lambda}^N)'[Var(\hat{\Lambda}^P) + Var(\hat{\Lambda}^N)]^{-1}(\hat{\Lambda}^P - \hat{\Lambda}^N)$, which is distributed asymptotically chi-square with degrees of freedom 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 functions (or the efficiency with which counties use factors to produce output). We utilize 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.⁹ This suggests the following decomposition based on the estimating model in A2.1:

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

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 factor shares and initial productivity endowments (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 (5) requires separate estimates of the pooled model and the subsamples of persistently poor and non-poor counties, which we present below.

IV. Convergence and Persistent Poverty

In Table 4, we present the estimates of our baseline model following the Arellano and Bover (1995) GMM specification. The source and construction of our dataset is explained in

⁹ The functional form of the production function is assumed constant for the groups; only the parameters are allowed to vary.

Appendix 1. To recap, the dependent variable is the log of real per capita income. The independent variables include lagged log real income per capita (y_{it-1}); the percentage of high school graduates; real private capital expenditures in manufacturing; labor force growth; urban share; the share of residents who are black; and the levels of time-invariant historical variables in Table 2 that proxy for culture, institutions, geography, and human capital/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 2).

A. *Pooled Model*

Column 1 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 one percentage point increase in the fraction of high-school graduates implies an increase in income levels of 7.4 percent, while a one percentage point decrease in 1900 illiteracy rates implies a 4.3 percent increase in income, holding other factors constant. Other variables such as current urban agglomeration and labor force growth 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 $Convergence = -1 \times \ln(\hat{\beta}) / time$, 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 percent, 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 et al 2006 at the county level).

In addition to illiteracy rates, column 2 of Table 4 includes the other historical factors. The percentage foreign-born in 1900 is marginally statistically significant and negatively correlated with modern income levels, while the historical share of urban residents is more strongly, but 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 significantly correlated with current income levels.¹⁰

We find that warmer and drier counties grow more slowly than cooler, wetter climates, *ceteris paribus*. 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 standard deviation 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 models on persistent poverty in Table 3.

¹⁰ We only tabulate results using the aggregated church share as the results are unchanged with the disaggregated by denomination.

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. 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); thus, we vary the instrument matrix in several ways to examine the sensitivity of our estimates, as we discuss in Section V.

B. Persistently-Poor versus Non-poor Aggregate Production Technologies

Consistent with the decomposition in equation (5), we divide the pooled sample into persistently-poor and non-poor 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 percent in each decennial Census between 1960 and 2000, and "non-poor" otherwise. We test this definition of persistent poverty in the robustness section of Section V. Column 2 of Table 4 reports the results for the persistently poor counties, while Column 3 reports those for the non-poor 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 percent level, although the signs of the current factors and the historical human capital/agglomeration variables are generally the same as those of the pooled model. The convergence rate among the persistently poor counties is about 10 percent, 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 percent of the counties in the pooled sample are non-poor).

Current and past measures of human capital, as well as the geography variables, are correlated with growth rates, while indicators of 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 non-poor counties have different production technologies. We find that the persistently poor counties have a qualitatively higher return on human capital (1.045 vs. 0.704, both are highly statistically significant) than the non-poor 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=1.05). However, the marginal effects on growth of the historical illiteracy rate and urbanization rate are much higher in absolute value for the non-poor counties than for the poor counties (but recall that the persistently poor counties had much higher illiteracy rates in 1900, and were much less urbanized in 1890, than the non-poor counties). The marginal effects of the standard deviation of elevation and average temperature are also larger in absolute value and statistically significant for the non-poor sample, but have little effect in the persistently poor sample.

C. Persistently Poor/Non-Poor Decomposition

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

As was seen in Table 1, the mean income difference between the persistently poor and non-poor counties is \$3,985. Of this, roughly 19 percent can be attributed to differences in the coefficients, while 81 percent is due to differences in endowments. Of that 81 percent, 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 in human capital remain extremely important, explaining 31 percent of the variation in current income due to variation in the explanatory variables, or approximately 25 percent of the total income gap between the persistently poor and non-poor samples. 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 non-poor), the coefficient estimate on 1890 urbanization in Table 4 is negative and statistically significant in both the pooled and non-poor 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 non-poor.

Of the contemporary explanatory variables, differences in lagged income and human capital explain a large amount of the income gap. Contemporary human capital explains 40 percent of the gap attributed to differences in the explanatory variables, or approximately 32 percent 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 primarily due to our result that, when county fixed effects are included, the coefficient estimate

on black share is large, positive, and statistically significant in the pooled and non-poor samples (see Table 4, and recall that this result stems from controlling for current and historical human capital).

To summarize, the decomposition results presented in Table 5 show that most of the income difference between the persistently poor and non-poor 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. The combined contribution of historical and contemporary human capital explains 57 percent of the overall income gap between the persistently poor and non-poor counties.

V. Sensitivity Analysis

In this section we consider two broad sets of robustness checks on our model estimates, one set on instrument selection and a second on the definition of poor and non-poor counties. The test of overidentifying restrictions in Table 4 rejects the null of valid restrictions for the pooled model and non-poor 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 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 Tables 6 and 7; 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 Table 6, rejects the (t-1) instruments for the pooled model and the non-poor model, but economically the results are the same.¹¹

In results not tabulated (but available upon 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.¹² 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 6. It remains the case that differences in the explanatory variables explain the majority (74 percent) of the income gap between the persistently poor and non-poor counties, while differences in the coefficient estimates (reflecting the efficiency with which the factors of production are employed) explain 26 percent. 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 continues to be the dominant factor.

In the second major robustness check, we modify the definition of persistently poor (or the non-poor comparison) in several ways, and repeat the preceding analysis. For brevity, we only report the results of the decompositions in Table 8 and suppress the individual regression estimates. In Column 1, we retain the definition of persistently poor (a poverty rate over 20% in

¹¹ The pseudo LR test statistic is 53.71 (=378.92-325.21), and with 20 degrees of freedom distributed asymptotically chi-squared, the p-value is < 0.00 in the pooled model (and non-poor sample).

¹² 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. These results are available upon request.

each decennial Census between 1960 and 2000), but use only non-urban, non-poor counties as the comparison group.¹³ In Column 2, the persistently poor definition is also retained, but non-poor counties in states with no persistently poor counties are omitted (that is, the non-poor comparison group is only counties in states with at least one persistently poor county; there are 26 states with no persistently poor county). Column 3 excludes the persistently poor counties that are primarily *colonias* or Native American reservations (i.e., the persistently poor counties included here are only from Appalachia, the “Black Belt,” and the Mississippi Delta).¹⁴ In Column 4, we relax the definition of persistently poor to require that a county have a poverty rate over 20% in three of the five Census years. Finally, in Column 5, our definition of persistently poor is that a county have a poverty rate of at least 30% (rather than 20%) 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 show that the predicted gap of log income between the poor and non-poor counties ranges from -0.32 to -0.48. In each specification, variation in factors of production explains at least 80 percent of the income gap, consistent with the results from Tables 5 and 7. Human capital, both historical and contemporary, remains important: historical human capital explains between 23 and 38 percent of the income gap explained by variation in production factors (between 19 and 32 percent of the

¹³ 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 of less than or equal to 5, it is considered as an urban county; a score of higher than 5 is considered to be a rural county. See a description of the coding at <http://www.ers.usda.gov/briefing/rurality/ruralurbcon/>.

¹⁴ Technically, we exclude states between the Mississippi Delta and the Pacific Coast states. The non-poor sample includes counties from these states.

total gap), while contemporary human capital explains between 36 and 42 percent of the gap attributed to differences in factors (and 30-34 percent of the overall gap). In each specification of Table 8, historical and contemporary human capital combine to explain over half of the total income gap between persistently poor and non-poor counties.

VI. Conclusion

Deep pockets of poverty persist 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 from the past five Censuses. We also incorporated the fundamental determinants of income; namely, the roles of culture, geography, human capital, and institutions based on measures from historical Censuses of 1890 and 1900.

We find evidence of significant regional differences in production technologies, but our decompositions of the income gap between poor and non-poor suggests that at least three fourths of the 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 past income, human capital, and urban share. We can rule out any major role of geography, culture, and institutions. However, the combined contribution of historical and contemporary human capital is large, explaining almost 60 percent of the overall income gap between the persistently poor and non-poor counties. While geography 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 urbanization, or at least proximity to

urban areas, is an important determinant of growth, accounting for over 15 percent of the income gap.

Together, the shortfalls in current human capital and urbanicity as key determinants of persistent poverty, while daunting to overcome, are surmountable barriers from a policy perspective. For example, in the context of Appalachia, Bollinger, Ziliak, and Troske (2011) and Black and Sanders (2012) find 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 eco-tourism 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.

References

- Acemoglu, D., S. Johnson, and J. Robinson (2005), "Institutions as the fundamental cause of long-run growth," in P. Aghion and S.N. Durlauf, eds., *Handbook of Economic Growth*, vol. 1. Amsterdam: North-Holland Elsevier: 385-472.
- 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 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 L. Barrington, ed., *The Other Side of the Frontier: Economic Explorations into Native American History*, Westview Press.
- Barro, R., and R. M. McCleary (2003), "Religion and Economic Growth," *American Sociological Review* 68 (5): 760-781.
- Barro, R., and X. Sala-i-Martin (1992), "Convergence," *Journal of Political Economy* 100 (2): 223-251.
- Bauer, P., M. Schweitzer, and S. Shane (2006), "State Growth Empirics: The Long-Run Determinants of State Income Growth," *Social Science Research Network*.
- Billings, D. (1974), "Culture and Poverty in Appalachia: A Theoretical Discussion and Empirical Analysis," *Social Forces*, 53, 2, 315-323.
- Billings, D., and K. Blee (2000), *The Road to Poverty: The Making of Wealth and Hardship in Appalachia*. New York: Cambridge University Press.
- Blundell, R., and S. Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics* 87(1): 115-143.
- Cavalcanti, T., S. Parente, and R. Zhao (2007), "Religion in macroeconomics: a quantitative analysis of Weber's thesis," *Economic Theory* 32: 105-132
- Clifton, E., and A. Romero-Barrutieta (2006), "Institutions v. geography: sub-national evidence from the United States," *International Monetary Fund Working Paper*.
- 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 P. Aghion and S.N. Durlauf, eds., *Handbook of Economic Growth* vol. 1.. Amsterdam: North-Holland Elsevier: 555-677.
- Durlauf, S. N., and D. Quah (1999), "The new empirics of economic growth," in J. Taylor and M. Woodford, eds, *Handbook of Macroeconomics* vol. 1. Amsterdam: North-Holland: 235-308.
- 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.
- Eller, R (2008), *Uneven ground: Appalachia since 1945*. Lexington: The University Press of Kentucky.
- Engerman, S. (1966), "The Economic Impact of the Civil War," *Explorations in Economic History* 3: 176-99.
- 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): 249-265.
- 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 (forthcoming), "Human Capital and Regional Development," *Quarterly Journal of Economics*.
- Glaeser, E., H. D. Kallal, J. A. Scheinkman and A. Shleifer (1992), "Growth in Cities," *The Journal of Political Economy* 100(6): 1126-1152.
- 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*, Vol. 1B, M. Rosenzweig and O. Stark (eds), Amsterdam: North Holland.

- Griliches, Z, and J. Hausman (1986), "Errors in Variables in Panel Data," *Journal of Econometrics* 31: 93-118.
- 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.
- Hansen, L. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50(4): 1029-1054.
- Higgins, M., D. Levy, and A. Young (2006), "Growth and Convergence across the United States: Evidence from County-Level Data," *The Review of Economics and Statistics* 8(4): 671-681.
- Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge, MA: Cambridge University Press.
- Islam, N. (1995), "Growth Empirics: A Panel Data Approach," *The Quarterly Journal of Economics* 110(4): 1127-1170.
- Iyigun, M. (2005), "Geography, Demography, and Early Development," *Journal of Population Economics* 18(2): 301-321.
- Kennan, J., and J. Walker (2011), "The Effect of Expected Income on Individual Migration Decisions," *Econometrica* 79(1): 211-251.
- 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): 357-392.
- 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): 407-437.
- Moretti, E. (2004), "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," *The American Economic Review* 94(3): 656-690.
- Murray, C. (1984), *Losing ground: American social policy, 1950-1980*. New York: Basic Books.
- Oaxaca, Ronald L. and M. Ransom (1994) "On discrimination and the decomposition of wage differentials," *Journal of Econometrics*. 61(1): 5-21
- 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): 375-398.

- Rappaport J., and J. Sachs (2003), "The United States as a Coastal Nation," *Journal of Economic Growth*. 8(1): 5-46.
- Shapiro, T. (2006), "Race, homeownership and wealth," *Washington University Journal of Law and Policy* 53.
- Solow, R. (1956), "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70(1): 65-94.
- 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): 419-431.

Appendix 1: Construction and Source of Data

A summary of variable definitions and their measurement units are shown in Appendix Table 1, and below we provide a description.

Contemporaneous Data

The contemporaneous county-level variables were collected from the 1960-2000 Decennial Censuses. The USA 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, persons living under the poverty level in the county according to the official poverty definition of the US, 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://www.census.gov/support/DataDownload.htm>

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). 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/system>

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 (http://www.census.gov/geo/www/ua/ua_2k.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 1960-90.

We obtained private capital expenditure in the manufacturing sector (measured in millions of US dollars) of each county for the years 1960, 1970 and 1980 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 (<http://www.census.gov/prod/1/manmin/92area/92manufa.htm>). The 1990 data are not in electronic format and thus had to be coded in using a pdf-to-Excel converter. Private capital expenditure data for the year 2000 was obtained from US Counties Basic Information database. The 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 standard deviation 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 in order to capture the ruggedness of terrain. The higher the standard deviation, 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 dataset. Historical county data on land tenure, 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, US 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/system>.

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 US from the website of the Newberry Library (<http://www.newberry.org>).

Appendix Table 1: Description of Variables Used in the Regressions

Variable	Description	Measurement Units
Per Capita Income (\$)	Average income earned by the residents of the county	In dollar amounts; values from 1960 to 1990 have been converted to 2000 dollars using the personal consumption expenditure deflator from the Bureau of Economic Analysis (BEA)
Population	Total number of people living within the boundary of a county	In absolute value
Fraction in Labor Force	The number of people over the age of 16 in the county who are not employed 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 county 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).	In millions of dollars, and the values have been converted to real 2000 dollars using the personal consumption expenditure deflator from the Bureau of Economic Analysis
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
Land Tenure in 1890	The total area of farmland farmed by their respective owners in the year	Between 0 and 1

	1890, divided by the total area under cultivation in a county in 1890	
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 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, US Reformed Church and American Reformed Church Organizations) 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
Average Temperature 1895	The average monthly temperature of a county for the years 1895 to 1905	In Fahrenheit
Average Precipitation 1985	The average monthly precipitation of a county for the years 1895 to 1905	In inches
Terrain (Std Dev to area)	The standard deviation of elevation of a county divided by area of that county	In feet per square mile
Urban Share in 1890	The number of residents living in an urban area as defined by the Census,	Between 0 and 1

	divided by the total population in 1890	
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

Appendix 2: 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

$$y_i = d_i\Lambda + v_i \quad (\text{A2.1})$$

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 $v_i = \iota_T \psi_i + \delta_t + \xi_i$. For consistent estimates of Λ , we then construct a nonsingular transformation, C , and a matrix of instruments, M_i , such that the moment conditions $E[M_i' C v_i] = 0$ 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 $M_i = I_T \otimes [g_i', \dots, g_i', Z_i']$, 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):

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

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

$$\hat{\Lambda} = [d' \bar{C}' M (M' \bar{C} \hat{\Omega} \bar{C}' M)^{-1} M' \bar{C} d]^{-1} d' \bar{C}' M (M' \bar{C} \hat{\Omega} \bar{C}' M)^{-1} M' \bar{C} y, \quad (\text{A2.2})$$

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.

Figure 1: U.S. Counties with Persistent Poverty from 1959 - 1999

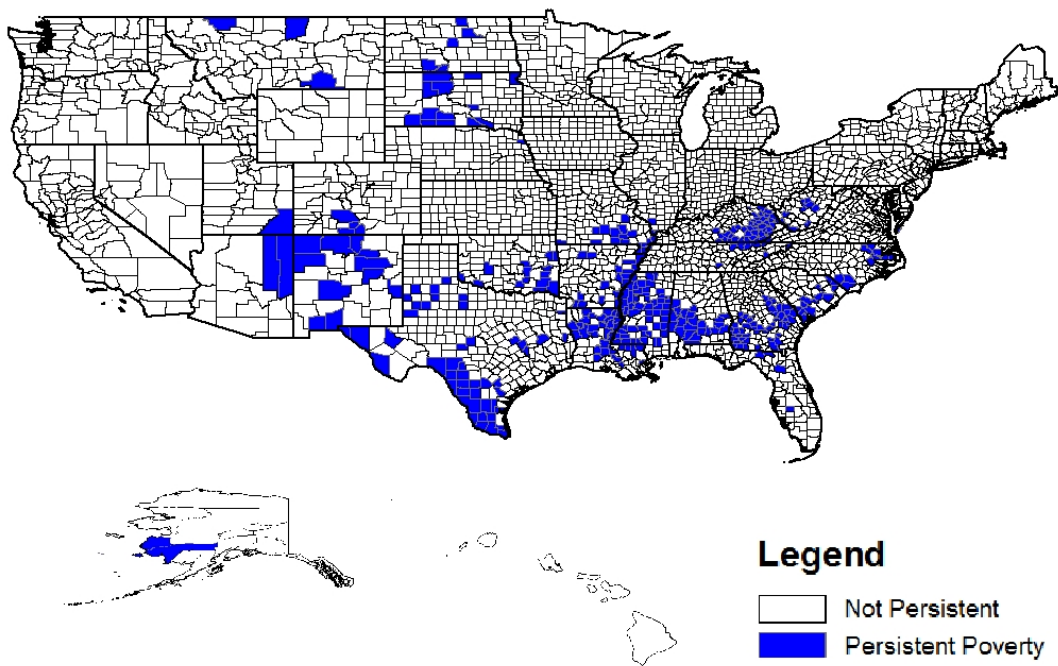


Figure 2: U.S. County Poverty Rates in 1959

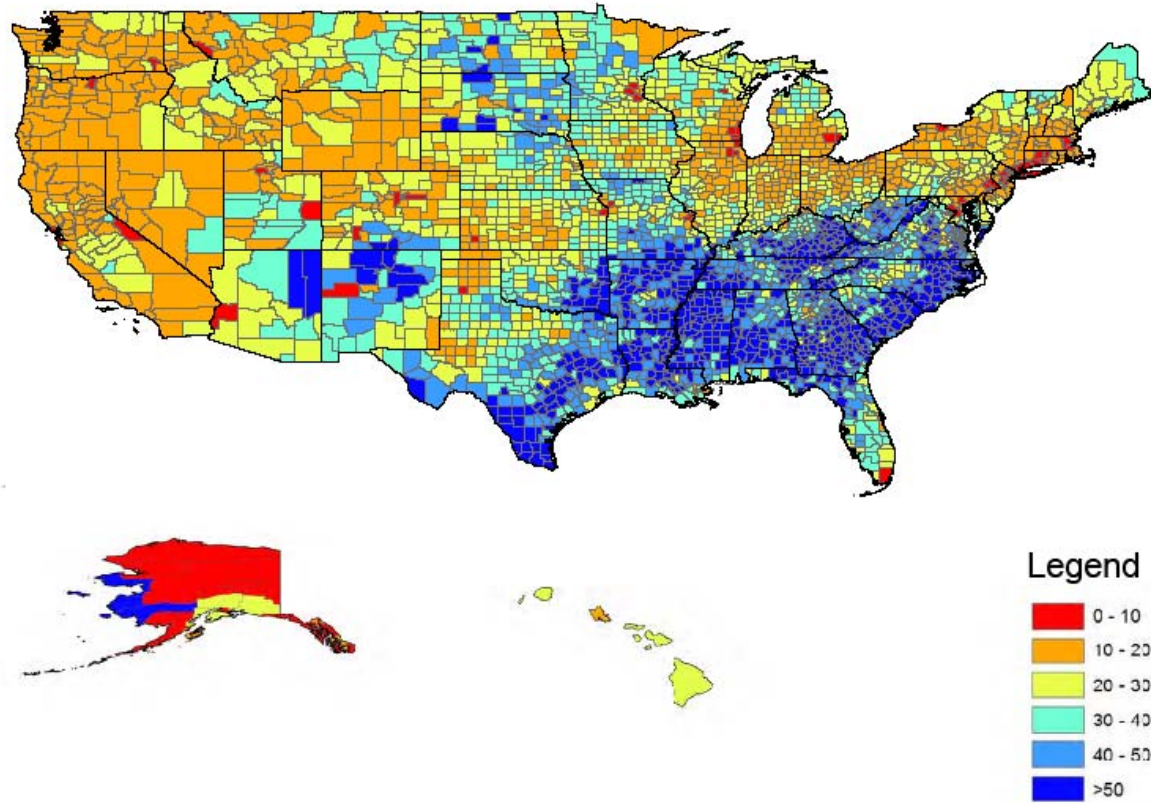


Figure 3: U.S. County Poverty Rates in 1969

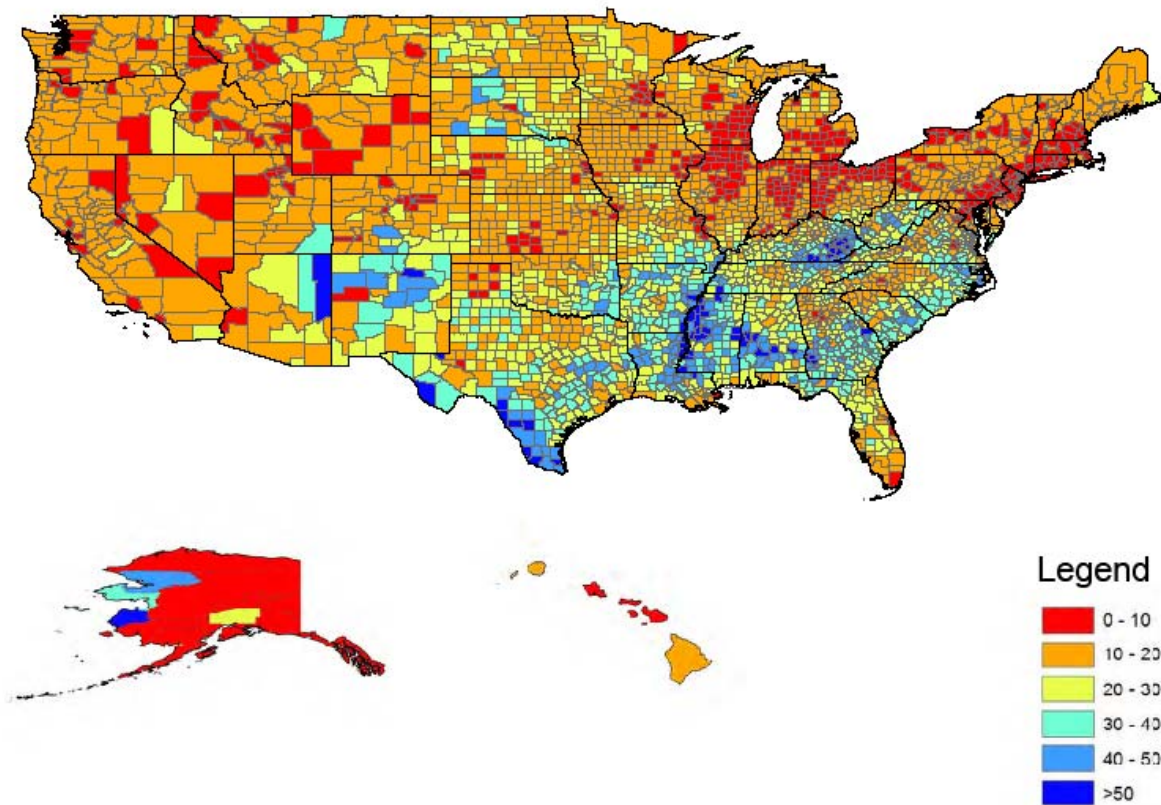


Figure 4: U.S. County Poverty Rates in 1979

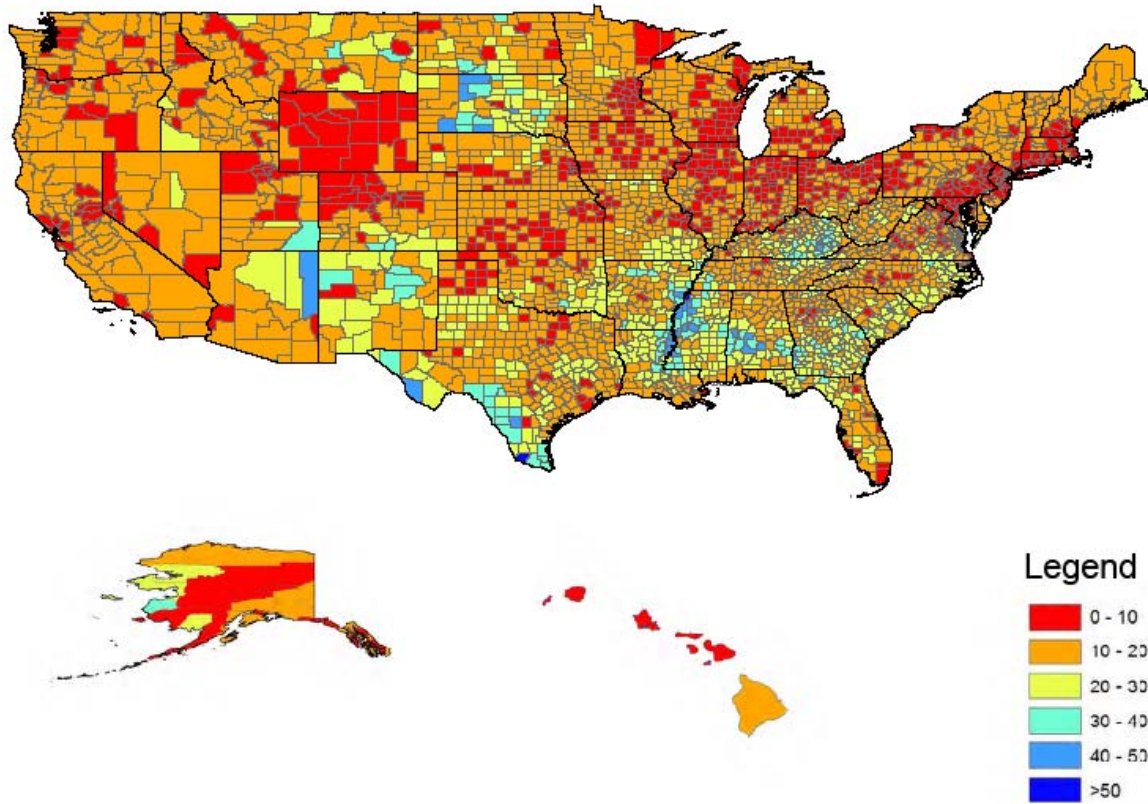


Figure 5: U.S. County Poverty Rates in 1989

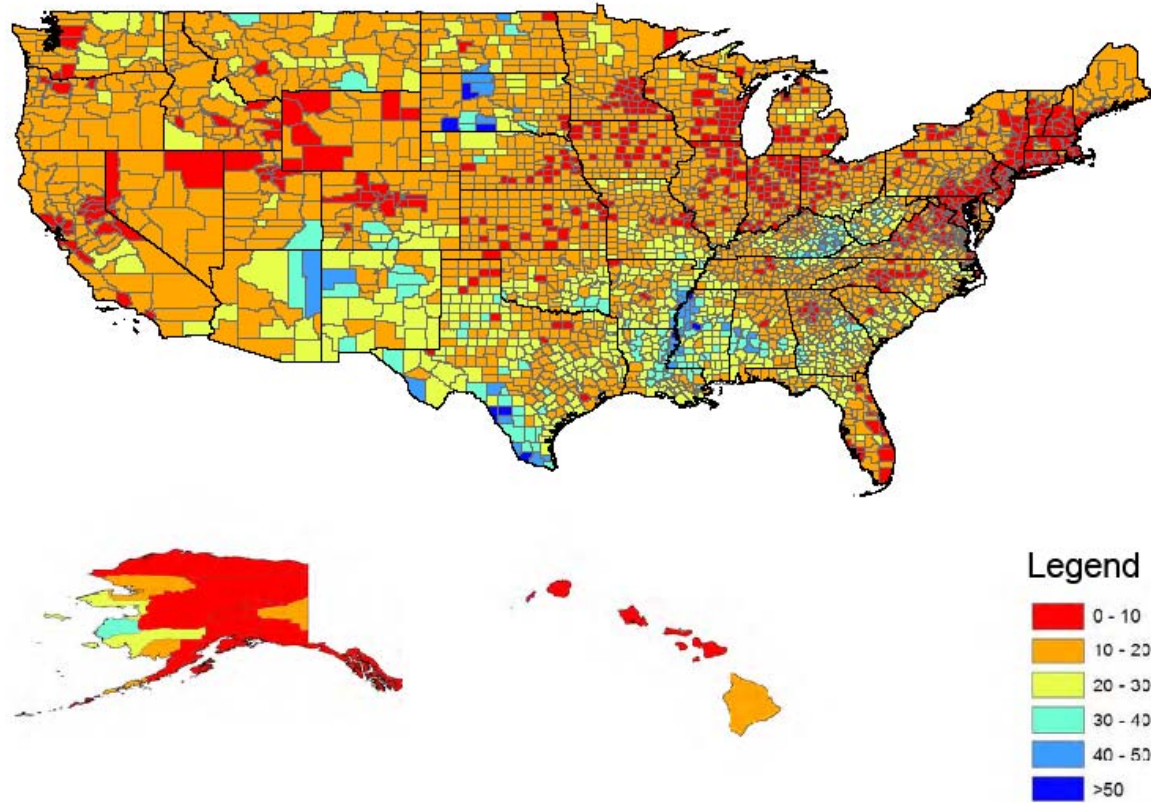


Figure 6: U.S. County Poverty Rates in 1999

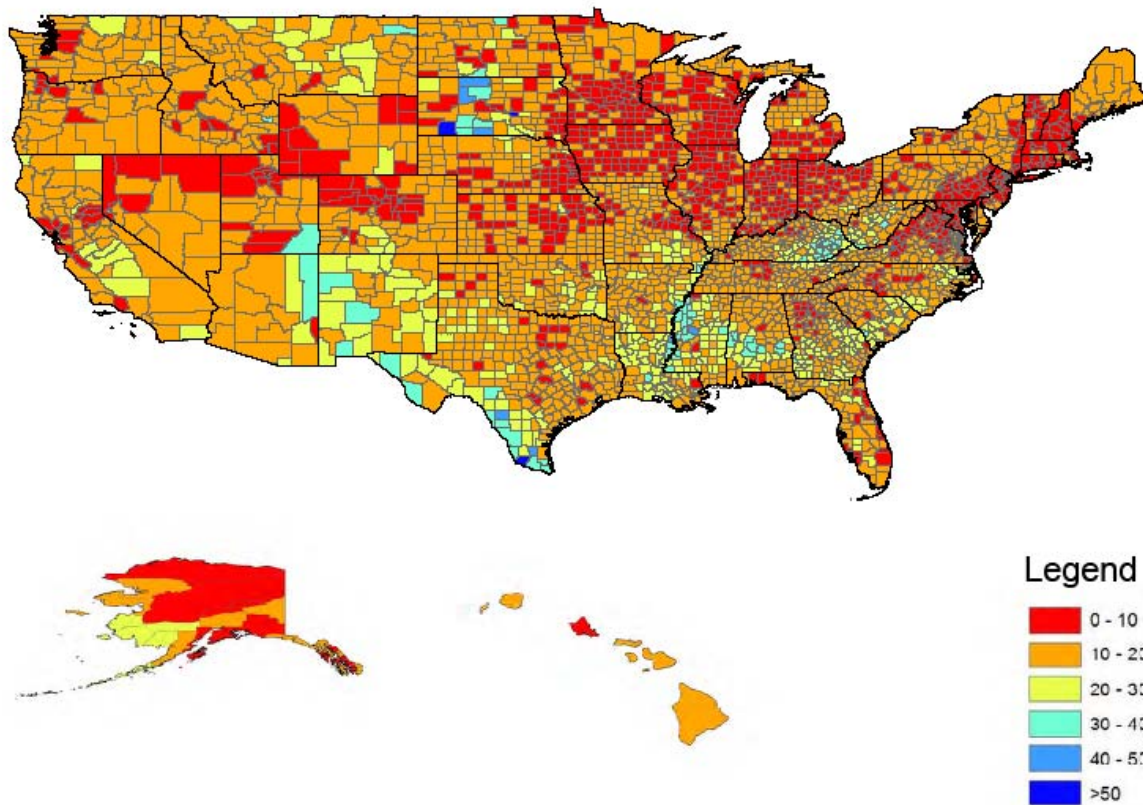


Table 1: Summary Statistics of Social Indicators for Counties by Persistent Poverty Status

	Not Persistently Poor		Persistently Poor	
	Mean	Standard Deviation	Mean	Standard Deviation
<u>Pooled 1960-2000 Census Data</u>				
Per Capita Income (\$)	12,436	4,952	8,451	3,585
Population	83,741	261,370	21,749	39,192
Fraction in Labor Force	0.423	0.065	0.349	0.057
Growth in Labor Force	0.161	0.235	0.036	0.211
Fraction High School Graduate	0.579	0.190	0.409	0.173
Capital Expenditure (\$millions)	38.916	155.244	4.718	18.158
Fraction Living in Urban Area	0.373	0.287	0.217	0.235
Fraction Black	0.068	0.114	0.288	0.254
<u>1960 Census Data</u>				
Per Capita Income (\$)	6,864	1,872	4,074	1,192
Population	66,094	229,251	21,424	41,972
Fraction in Labor Force	0.360	0.036	0.300	0.042
Growth in Labor Force	0.053	0.224	-0.166	0.136
Fraction High School Graduate	0.355	0.103	0.212	0.064
Capital Expenditure (\$millions)	17.221	80.589	1.117	3.324
Fraction Living in Urban Area	0.330	0.277	0.182	0.215
Fraction Black	0.073	0.128	0.307	0.267
<u>2000 Census Data</u>				
Per Capita Income (\$)	18,664	3,695	13,399	1,896
Population	100,491	322,866	21,865	35,766
Fraction in Labor Force	0.482	0.046	0.394	0.040
Growth in Labor Force	0.142	0.152	0.056	0.108
Fraction High School Graduate	0.790	0.073	0.636	0.065
Capital Expenditure (\$millions)	43.997	180.537	4.892	16.898
Fraction Living in Urban Area	0.410	0.301	0.241	0.241
Fraction Black	0.064	0.105	0.283	0.283
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.

Table 2: Summary Statistics of Historical Indicators for Counties by Persistent Poverty Status

	Not Persistently Poor		Persistently Poor	
<i>Institutions</i>				
Land Tenure in 1890	0.798	0.115	0.769	0.143
<i>Culture</i>				
Share Churched in 1890	0.293	0.118	0.309	0.148
Share Baptist in 1890	0.065	0.081	0.135	0.097
Share Calvinist in 1890	0.019	0.022	0.009	0.015
Share Catholic in 1890	0.058	0.086	0.034	0.123
<i>Geography</i>				
Average Temperature	52.90	7.405	60.27	6.037
Average Precipitation	3.089	0.910	3.611	0.721
Std Dev of Elevation	0.089	0.129	0.064	0.076
<i>Human Capital/Agglomeration</i>				
Urban Share in 1890	0.129	0.211	0.026	0.107
Share Foreign Born in 1900	0.094	0.103	0.023	0.072
Illiteracy Rate in 1900	0.112	0.135	0.360	0.217
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. The historic Census data include only counties without redefined borders between the relevant year (1890, 1900) and 1960. See the Data Appendix for detailed variable definitions.

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)	0.049 (0.054)	0.041 (0.055)	-0.138 (0.063)	-0.193 (0.063)
Average Temperature	0.010 (0.001)	-0.0001 (0.001)	0.0002 (0.001)	0.009 (0.001)	0.007 (0.001)
Average Precipitation	0.011 (0.006)	-0.027 (0.007)	-0.025 (0.006)	0.015 (0.007)	0.007 (0.006)
Std. Dev of Elevation	-0.072 (0.029)	-0.028 (0.029)	-0.016 (0.029)	-0.053 (0.029)	-0.048 (0.029)
Share Foreign-born in 1900		-0.053 (0.078)	-0.118 (0.092)	0.094 (0.081)	-0.011 (0.097)
Urban Share 1890		-0.057 (0.021)	-0.067 (0.022)	-0.175 (0.023)	-0.147 (0.024)
Illiteracy Rate 1900		0.921 (0.063)	0.913 (0.071)		
Church Share 1890		-0.181 (0.057)		0.027 (0.063)	
Baptist Share 1890			-0.254 (0.119)		0.340 (0.117)
Calvin Share 1890			-0.557 (0.263)		-1.414 (0.296)
Catholic Share 1890			-0.017 (0.106)		0.189 (0.113)
Constant	-0.422 (0.063)	0.094 (0.077)	0.057 (0.077)	-0.355 (0.074)	-0.164 (0.078)
Adjusted R ²	0.085	0.220	0.219	0.097	0.117
Number of Counties	2,400	2,400	2,400	2,400	2,400

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.

Table 4: Dynamic Panel GMM Estimates of Income, Current Factors Predetermined

	(1)	(2)	(3)
	Pooled Model	Persistently-Poor	Non-Poor
<i>Current Factors</i>			
Lag Income per Capita	0.376 (0.027)	0.367 (0.083)	0.373 (0.029)
Fraction High School	0.742 (0.073)	1.045 (0.313)	0.705 (0.079)
Capital Spending (x1,000,000)	3.795 (41.777)	253.82 (378.66)	20.983 (40.392)
Labor Force Growth	0.135 (0.017)	0.170 (0.040)	0.135 (0.018)
Urban Share	0.434 (0.060)	0.071 (0.142)	0.446 (0.062)
Black Share	0.559 (0.138)	0.042 (0.393)	0.504 (0.129)
<i>Human Capital/Agglomeration</i>			
Illiteracy Rate 1900	-0.427 (0.104)	-0.171 (0.252)	-0.265 (0.090)
Proportion Foreign Born 1900	-0.094 (0.047)	-0.038 (0.295)	-0.068 (0.047)
Urban Share 1890	-0.340 (0.051)	-0.083 (0.129)	-0.348 (0.052)
<i>Culture</i>			
Proportion Churched 1890	0.035 (0.028)	-0.012 (0.098)	0.028 (0.031)
<i>Institutions</i>			
Land Tenure 1890	0.045 (0.034)	-0.083 (0.099)	0.003 (0.033)
<i>Geography</i>			
Standard Dev. to Area	0.176 (0.027)	0.088 (0.246)	0.161 (0.026)
Average Temperature 1895 (x100)	-0.348 (0.080)	0.131 (0.314)	-0.372 (0.081)
Average Precipitation 1895	0.032 (0.005)	0.051 (0.020)	0.025 (0.005)
Constant	5.440 (0.201)	5.217 (0.575)	5.542 (0.216)
Convergence Rate	0.0979	0.100	0.099
Hansen's J (df., p-value)	378.92 [50, 0.000]	33.76 [50, 0.962]	380.47 [50, 0.000]
Wald Test of Equal Coef. (df., p-value)			197.97 [17, 0.000]

Notes: Robust standard errors are reported in parentheses. The number of county-years in the pooled model is 12,000 (1,170 persistently poor and 10,830 non-poor). Each model controls for time effects. The instrument matrix is block diagonal and consists of (t-2), (t-3) and (t-4) lag of log income, and (t-1), (t-2), (t-3) and (t-4) lagged levels of the time-varying variables, and the levels of time-invariant variables.

Table 5: Poor/Non-Poor Decomposition

Predicted Gap in Current Income	-0.398
<i>Proportion Difference due to</i>	
Factor Endowments	0.809
Coefficients	0.191
<i>Of Factor Endowments Share, Proportion due to:</i>	
<i>Historical Factors</i>	
Human Capital	0.311
Agglomeration	-0.107
Culture	-0.002
Institutions	0.004
Geography	0.042
<i>Current Factors</i>	
Lag Log Income	0.482
Human Capital	0.400
Capital	<0.001
Labor Force Growth	0.040
Urban Share	0.203
Black Share	-0.374

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

Table 6: Dynamic Panel GMM Estimates of Income, Current Factors Endogenous

	(1)	(2)	(3)
	Pooled Model	Persistently-Poor	Non-Poor
<i>Current Factors</i>			
Lag Income per Capita	0.404 (0.030)	0.385 (0.108)	0.386 (0.030)
Fraction High School	0.615 (0.081)	0.839 (0.381)	0.607 (0.086)
Capital Spending (x1,000,000)	107.92 (62.41)	979.87 (578.34)	154.06 (78.07)
Labor Force Growth	0.164 (0.027)	0.226 (0.058)	0.167 (0.028)
Urban Share	0.446 (0.071)	0.067 (0.158)	0.481 (0.072)
Black Share	0.545 (0.194)	-0.155 (0.518)	0.405 (0.162)
<i>Human Capital/Agglomeration</i>			
Illiteracy Rate 1900	-0.412 (0.144)	-0.033 (0.265)	-0.203 (0.110)
Proportion Foreign Born 1900	-0.124 (0.051)	-0.231 (0.317)	-0.117 (0.055)
Urban Share 1890	-0.371 (0.061)	-0.063 (0.156)	-0.397 (0.061)
<i>Culture</i>			
Proportion Churched 1890	0.025 (0.028)	0.038 (0.102)	0.017 (0.032)
<i>Institutions</i>			
Land Tenure 1890	0.046 (0.040)	-0.122 (0.110)	0.003 (0.037)
<i>Geography</i>			
Standard Dev. to Area	0.162 (0.032)	-0.084 (0.258)	0.147 (0.029)
Average Temperature 1895 (x100)	-0.462 (0.108)	0.151 (0.348)	-0.514 (0.117)
Average Precipitation 1895	0.030 (0.005)	0.043 (0.023)	0.023 (0.006)
Constant	5.322 (0.232)	5.201 (0.739)	5.561 (0.230)
Convergence Rate	0.091	0.096	0.095
Hansen's J (df., p-value)	325.21 [30, 0.000]	23.43 [30, 0.797]	307.02 [30, 0.000]
Wald Test of Equal Coef.			171.301 [17, 0.000]

Notes: Robust standard errors are reported in parentheses. The number of county-years in the pooled model is 12,000 (1,170 persistently poor and 10,830 non-poor). Each model controls for time effects. The instrument matrix is block diagonal and consists of (t-2), (t-3) and (t-4) lag of log income, and (t-2), (t-3) and (t-4) lagged levels of the time-varying variables, and the levels of time-invariant variables.

Table 7: Poor/Non-Poor Decomposition

Predicted Gap in Current Income	-0.387
<i>Proportion Difference due to</i>	
Factor Endowments	0.837
Coefficients	0.163
<i>Of Factor Endowments Share, Proportion due to:</i>	
<i>Historical Factors</i>	
Human Capital	0.290
Agglomeration	-0.116
Culture	-0.001
Institutions	0.005
Geography	0.071
<i>Current Factors</i>	
Lagged Log Income	0.515
Human Capital	0.329
Capital	0.013
Labor Force Growth	0.049
Urban Share	0.207
Black Share	-0.362

Notes: The decomposition is based on the results from Table 6.

Table 8: Decomposition Results from the Sensitivity Analysis

<u>Poor - Non-Poor Decomposition</u>					
	<u>Using Different Definitions to Define Persistently-Poor/Non-Poor Counties</u>				
	(1)	(2)	(3)	(4)	(5)
Predicted Gap in Current Income	-0.379	-0.319	-0.410	-0.335	-0.477
<i>Proportion Difference due to</i>					
Factor Endowments	0.816	0.830	0.800	0.886	0.853
Coefficients	0.184	0.170	0.191	0.114	0.147
<i>Of Factor Endowments Share, Proportion due to:</i>					
<i>Historical Factors</i>					
Human Capital	0.231	0.308	0.260	0.233	0.377
Agglomeration	-0.107	-0.056	-0.107	-0.132	-0.067
Culture	-0.002	-0.00002	0.003	-0.001	0.001
Institutions	0.002	0.004	0.002	-0.002	0.011
Geography	0.032	-0.022	0.101	0.045	0.067
<i>Current Factors</i>					
Lag Log Income	0.468	0.520	0.476	0.443	0.512
Human Capital	0.421	0.398	0.379	0.383	0.362
Capital	-0.001	0.007	-0.003	0.001	0.0003
Labor Force Growth	0.036	0.041	0.046	0.036	0.038
Urban Share	0.189	0.116	0.218	0.259	0.113
Black Share	-0.268	-0.316	-0.376	-0.265	-0.415
Number of Poor Counties	234	234	202	579	38
Number of Non-poor Counties	2043	1210	1658	1821	2362

Notes: Column (1) compares between persistently-poor counties with non-urban, non-poor counties, according to the Beale System. Column (2) removes all the counties located in states that do not have any persistently-poor counties within their borders. Column (3) keeps the non-poor set the same, but removes poor

counties between Mississippi Delta and the western states of Washington, Oregon and California. Column (4) considers a county to be persistently poor if it has 20% or higher poverty rates for 30 out of 50 years. Column (5) considers a county to be persistently-poor if it has 30% or higher poverty rates for 50 years. The instrument matrix is block diagonal and consists of (t-2), (t-3) and (t-4) lags of log income, and (t-1), (t-2), (t-3) and (t-4) lagged levels of the time-varying variables, and the levels of time-invariant variables.