Inclusive Institutions and Long-Run Misallocation

Galor, Oded and Munshi, Kaivan and Wilson, Nicolas

Brown University, University of Cambridge, Reed College

1 October 2013

Inclusive Institutions and Long-Run Misallocation

Oded Galor[*], Kaivan Munshi[†], and Nicholas Wilson[‡]

October 30, 2013

Abstract

This research advances the hypothesis that resource abundant economies characterized by a socially cohesive workforce and network externalities triggered the emergence of efficiency-enhancing inclusive institutions designed to restrict mobility and to enhance the attachment of community members to the local labor market. However, the persistence of these institutions, and the inter-generational transmission of their value, ultimately resulted in the misallocation of talents across occupations and a reduction in the long-run level of income per capita in the economy as a whole. Exploiting variation in resource intensity across the American Midwest during its initial development, the empirical analysis establishes that higher initial resource-intensity in 1860 is indeed associated with greater community participation over the subsequent 150 years, and reduced mobility and labor misallocation in the contemporary period.

Keywords: Inclusive institutions, Exclusive institutions, Growth, Networks, Labor misallocation, Persistence.
JEL classification Numbers: I12, J13, N3, O10.

[*]Brown University and NBER
[†]University of Cambridge
[‡]Reed College

1 Introduction

The transition from an epoch of stagnation to an era of sustained economic growth has triggered one of the most remarkable transformations in the course of human history. While living standards in the world economy stagnated during the millennia preceding the Industrial Revolution, income per capita has undergone an unparalleled fourteen-fold increase over the past two centuries, profoundly altering the level and distribution of education, health, and wealth across the globe. Nevertheless, the rise in the standard of living has not been universally shared among individuals and societies. Inequality, which had been modest until the nineteenth century, has widened considerably and the ratio of income per capita between the richest and the poorest regions of the world has been magnified from a moderate 3:1 ratio in 1820 to a staggering 18:1 ratio in 2000 (Maddison (2008)).

Theories of comparative economic development have highlighted multiple proximate and ultimate factors that have contributed to the vast inequity in living standards across the globe. The role of geography, culture, and institutions, the contribution of the level and diversity of innate and acquired human capital, as well as the impact of the forces of colonization and globalization have been central to a debate about the origin of the differential timing of the transition from stagnation to growth and the remarkable transformation in the world income distribution.

The effects of institutional factors on the growth process have been largely segmented into two opposing forces (Acemoglu and Robinson (2012)). The emergence of inclusive institutions, typically in homogeneous socially cohesive populations, has been widely viewed as a significant catalyst in the process of economic development. These institutions facilitated the protection of property rights, and provided economic incentives that stimulated the entrepreneurial spirit and the process of accumulation in the economy as a whole. Inclusive institutions therefore fostered economic activity, technological development, and knowledge creation and diffusion, and variation in their presence has contributed to divergence in productivity across the globe (North (1981), Landes (1998), Acemoglu and Robinson (2012)).

In contrast, the implementation of exclusive institutions in environments characterized by a conflict of interest between the elite and the masses has been regarded as highly detrimental to development. These institutions, which are often associated with colonization, were designed to maintain the political power of the elite and to preserve existing inequality. They stifled the incentives for investment in physical and human capital in the economy as a whole, and limited the scope for innovations, leading to inferior development trajectories and reduced productivity in the long run (Engerman and Sokoloff (1997, 2000), Acemoglu, Johnson and Robinson (2001, 2002, 2012)).

A particular type of exclusive institution, designed to restrict the mobility of workers from the agricultural to the industrial sector, emerged in economies marked by a conflict of interest between the industrial and the landed elites. While complementarity between physical and human capital provided the industrial elite with an incentive to implement inclusive institutions that would induce workers to invest in human capital (Galor and Moav (2006)), landowners, whose livelihood was endangered by the mobility of workers from the local (rural) to the global (urban) environment, had the incentive to block the provision of education (Galor et al., (2009)) and to implement mobility restricting coercive institutions (Dell (2012), Naidu and Yuchtman (2012)), leading to delayed industrialization, reduced human capital formation, and lowered productivity in the long-run.

This research advances the hypothesis that resource abundant economies characterized by a socially cohesive workforce and network externalities triggered the emergence of efficiency-enhancing inclusive institutions designed to restrict mobility and to enhance the attachment of community members to the local labor market. Nevertheless, the persistence of these inclusive institutions that were designed to internalize the positive externalities in the local community, and the inter-generational transmission of their value, may have ultimately led to labor misallocation and a reduction in the long-run level of income per capita in the economy as a whole. Thus, as the process of development has altered economic opportunity in the global economy, while reducing the importance of historical networks, the persistence of inclusive institutions that were optimally designed for an earlier stage of development, has inadvertently led to labor misallocation in the contemporary era.[1]

The test of the proposed theory is based on the experience of the American Midwest over the past 150 years. This region provides an ideal setting for an empirical assessment of the theory for several reasons. First, the Midwest is a geographically and culturally homogeneous

[1] This is consistent with the observation from diverse locales that close-knit working class and farming communities often stubbornly resist economic change (Gans (1962), Elder and Conger (2000), Munshi and Rosenzweig (2006), Willis (1974), Kornblum (1977)).

region in which extractive institutions have not emerged. Second, the rapid development of the Midwest, in the aftermath of the railroad expansion of the 1850's, had a significant community component. In line with a fundamental element of the theory, ethnic churches in this region played a pivotal role in forming communities, characterized by economic networks and social support systems based on nationality and residential location. Third, conditions that affected economic opportunities and the initial level of investment in community services 150 years ago are unlikely to be relevant in the contemporary period.

In the early stages of the development of the Midwest, counties that were located near a canal, Great Lakes harbor, or railroad line enjoyed a significant economic advantage. Complementarity between these resources and labor inputs, would have led according to the theory to investments in the community services designed to restrict mobility in these resource-rich counties. This resource advantage dissipated, however, once the railroad network had penetrated the entire region and the canals had lost their importance. Thus differences in labor market choices across counties in the contemporary period can be plausibly attributed to the persistent attachment to community institutions, rather than the persistent effect of initial resources.

The theory suggests that inclusive institutions in historically resource-rich counties provided a persistently higher level of community services, designed to restrict mobility and to enhance the complementarity between human and natural and physical resources in the local community. Church participation is indeed greater in historically resource-rich counties and this difference has persisted since 1870. Moreover, a significantly larger fraction of students in those counties are enrolled in parochial schools.

The theory further indicates that individuals born in the contemporary period in historically resource-rich counties, long after the ethnic networks that gave rise to the inclusive community institutions dissipated, are nevertheless less mobile. The evidence suggests that individuals born in historically resource-rich Midwestern counties are indeed represented disproportionately in less mobile occupations, and are significantly less likely to migrate from their county of birth, despite the fact that the available composition of jobs, government expenditures on education and other public goods, educational attainment, and test scores do not differ between resource rich and poor counties in the contemporary period. A one standard deviation increase in the measure of resource intensity in the past would result in as much as a 14 percent decline in migration. Moreover, the income of individuals born in historically resource-rich counties is significantly lower despite the fact that income in their county is not significantly lower. Thus, inclusive institutions, which were designed specifically to increase efficiency by internalizing network externalities may generate labor misallocation in the long-run.[2]

2 The Basic Structure of the Model

Consider an overlapping-generations economy in the process of development. In every period the economy produces a single homogeneous good that can be used for either consumption or investment. The good is produced in a local, geographically isolated, community as well as in a global perfectly competitive environment, using labor and a location-specific non-producible resource.

The local economy consists of a large number of individuals. They are homogeneous in their labor endowment, and a subset of them owns the non-producible resource. Labor productivity is higher in the local community, reflecting the positive externalities of a community-based economic network, which is absent in the global, anonymous, market. However, individual resource owners are too small to account for these externalities when making their profit-maximizing employment decisions, and a competitive allocation would result in a sub-optimal level of employment in the local economy and excessive exit to the global economy. This market failure leads to the formation of inclusive institutions that provide community services, designed to increase the attachment of workers to the local community and to increase the aggregate well being of the local population.

2.0.1 Production of Final Output

The aggregate output produced in the global market, Y_t^G, is governed by a linear production function using labor as the only factor of production. In particular, $Y_t^G = aL_t^G$, where $a > 0$, and L_t^G is the number of individuals employed in the global economy. The wage in the global economy is therefore constant over time, i.e., $w_t^G = a \equiv \bar{w}^G$, for all t.

The local economy consists of a large number of perfectly competitive firms that produce the final output with a neoclassical, constant-returns-to-scale production technology, using labor and the immobile, location-specific, resource.

[2] From a welfare perspective, the proposed theory indicates that individuals born in historically resource-rich counties are as well off staying versus moving. From an economic perspective, however, the intergenerational transmission of local attachment appears to have generated a substantial labor misallocation.

The aggregate output produced by resource owners in the local community at time t, Y_t^c, is

$$Y_t^c = A_t F(X, L_t^c), \tag{1}$$

where X is the aggregate amount of the location-specific resource, L_t^c is the number of individuals employed in the local community, and A_t is the level of technology in the local environment in period t.

The production function, $F(X, L_t^c)$, is neoclassical. The marginal productivity of both factors of production are positive and diminishing, satisfying the neoclassical boundary conditions that assure the existence of an interior solution to the producers' profit-maximization problem. Moreover, the two factors of production are complements in the production process, i.e., $F_{XL}(X, L_t^c) > 0$. In particular,

$$Y_t^c = A_t X^{1-\alpha}(L_t^c)^{\alpha}; \qquad \alpha \in (0,1). \tag{2}$$

Labor productivity in the local community reflects a positive externality from the local community-based economic network. In particular, the level of technology in the local environment in period t increases at a decreasing rate with the level of employment in the local community, L_t^c, in period t, i.e.,

$$A_t = A(L_t^c), \tag{3}$$

where

$$A'(L_t^c) > 0, \quad A''(L_t^c) < 0. \tag{4}$$

In particular,

$$A_t = (L_t^c)^{\beta}; \qquad \beta \in (0,1). \tag{5}$$

Resource owners in the local community operate in a perfectly competitive environment. In particular, they are too small to internalize their employment decisions on the aggregate productivity, A_t. In every period t, given the wage rate in the local community, w_t^c, and the level of productivity, A_t, they choose the level of employment of labor, L_t^c, so as to maximize profits. That is,

$$\{L_t^c\} = \arg\max[A_t F(X, L_t^c) - w_t^c L_t^c]. \tag{6}$$

Given the fixed supply of the resource in the local community, X, producers' aggregate inverse demand for labor is therefore,

$$w_t^c = A(L_t^c)F_L(X, L_t^c) = \alpha X^{1-\alpha}(L_t^c)^{\alpha+\beta-1}. \tag{7}$$

The positive externality from local employment is assumed to be limited so as to assure that the marginal productivity of workers diminishes, implying that

$$\alpha + \beta < 1. \tag{A1}$$

The aggregate rental rate on the location-specific resource, r_t^c, is therefore,

$$r_t^c = A_t F(X, L_t^c) - w_t^c L_t^c = (1-\alpha)X^{1-\alpha}(L_t^c)^{\alpha+\beta}. \tag{8}$$

2.0.2 Local Institutions

The local community cooperates to provide services designed to enhance local ties and thus to reduce the incentive of workers to migrate from the local market into the global one. These services can be viewed as a social infrastructure, provided by the church and other inclusive community institutions, that facilitate social interactions and provide social support. All individuals who are employed in the local economy participate in the community institutions and benefit from their provision of these services.

Each member of the local community derives utility from community services. The net value of these services for each community member in period $t+1$ is, g_{t+1}, in terms of the final output. This value is affected positively by: (i) the resources invested in the provision of the services in period $t+1$, v_{t+1}, and (ii) the value of the services in the past generation, g_t, which is transmitted imperfectly across generations.[3]

In particular, the value of community services for each local worker of generation $t+1$, g_{t+1}, is

$$g_{t+1} = g(v_{t+1}) + \rho g_t. \tag{9}$$

where $\rho \in (0,1)$.

Resources invested in the provision of the community services are productive but their effect is diminishing, i.e., for all $v_{t+1} > 0$, $g(v_{t+1}) > 0$, $g'(v_{t+1}) > 0$, and $g''(v_{t+1}) \leq 0$. Furthermore, $\lim_{v_{t+1}\to 0} g'(v_{t+1}) = \infty$ and $\lim_{v_{t+1}\to\infty} g'(v_{t+1}) = 0$, assuring the direct profitability of

[3]This last term reflects the idea that social infrastructure established by the preceding generation affects social interactions and social support in the current period, but not as strongly as current infrastructure.

positive, but finite, investment in the community services. In particular,

$$g(v_{t+1}) = v_{t+1}^{\gamma}; \qquad \gamma \in (0,1). \tag{10}$$

Investments in community services are financed by taxes imposed on the community as a whole, prior to the allocation of labor between the local and the global market. Hence this tax rates does not affect the mobility of workers from the local to the global economy.

2.0.3 Workers

In every period t, workers in the local community face the choice of employment in either the local or the global market. Those who choose to remain in the local community earn the competitive local wage, w_t^c, and benefit from the value of community services, g_t. However, if they choose employment in the global market they relinquish the benefits of these services, while earning the competitive global wage, $\bar{w}^G$.

Hence, as depicted in Figure 1, labor allocation between the local and the global market will be determined such that

$$w_t^c + g_t = \bar{w}^G, \tag{11}$$

where the size of the local population L is sufficiently large so as to assure that in equilibrium some individuals will choose employment in the global market. Equivalently, this implies that if the entire local population were to be employed in the local market, the competitive wage that would emerge in the local market, $w_{\min}^c \equiv A(L)F_L(X,L)$, would be such that for all $g_t \geq 0$

$$w_{\min}^c + g_t < \bar{w}^G. \tag{A2}$$

As follows from the competitiveness of the local labor market, $w_t^c = A(L_t^c)F_L(X, L_t^c)$, and thus as depicted in Figure 1, the level of employment in the local market in period t, L_t^c, is determined by

$$A(L_t^c)F_L(X, L_t^c) + g_t = \bar{w}^G. \tag{12}$$

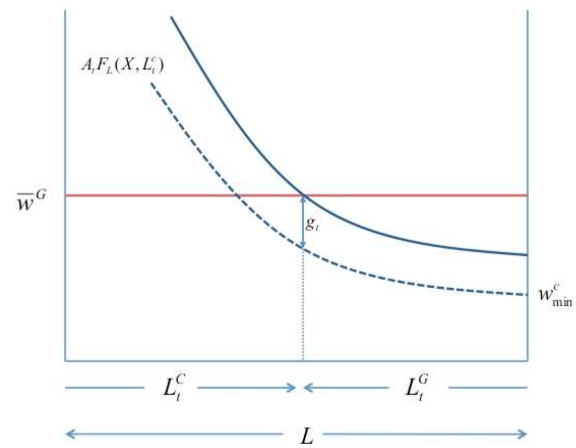

Figure 1. Labor allocation between the local and the global markets

Given the labor endowment in the local community in every time period, L, the level of employment in the local market in period t, L_t^c, and the level of employment in global market, L_t^G, are such that,

$$L_t^c + L_t^G = L. \tag{13}$$

Given the properties of the production function, it follows from (12), that for any given level of location-specific resource, $X > 0$, the level of employment in the local community, L_t^c, is a single-valued function of the value of community services, g_t. Moreover, given the existence of an immobile, location-specific resource, $X > 0$, and the neoclassical properties of the production function (i.e., $\lim_{L_t^c \to 0} A(L_t^c) F_L(X_t, L_t^c) = \infty > \bar{w}^G$), employment in the local market is strictly positive, even if the local community does not provide any services.

Thus, for all $g_t \geq 0$,

$$L_t^c = \alpha^{\frac{1}{1-\alpha-\beta}} \left[\bar{w}^G - g_t\right]^{\frac{1}{\alpha+\beta-1}} X^{\frac{1-\alpha}{1-\alpha-\beta}} \equiv L(g_t; X) > 0, \tag{14}$$

where $\partial L_t^c / \partial g_t > 0$ and $\partial L_t^c / \partial X > 0$. Holding the value of community services constant, employment in the local community increases with the amount of the location-specific resource, X, and holding the amount of the location-specific resource, X, constant, employment in the local community increases with the value of the services, g_t.

2.0.4 Optimal Provision of the Community Services

Suppose that in every period the aggregate level of investment in community services is chosen so as to maximize the well being of the community as whole.[4] The aggregate level of investment in community services at time t, v_t, will be the level that maximizes the sum of output in the local community, Y_t^c, the wage income of local workers in the global market, $L_t^G \bar{w}^G$, and the value of the services for the local workers, $g_t L_t^c$, net of resources invested in community services, v_t, i.e.,

$$v_t = \arg\max Y_t^c + L_t^G \bar{w}^G + g_t L_t^c - v_t. \tag{15}$$

Noting that $L_t^G = L - L_t^c$ and $w_t^c = \bar{w}^G - g_t$, and using (8), it follows that

$$v_t = \arg\max (1-\alpha) X^{1-\alpha} (L_t^c)^{\alpha+\beta} + L\bar{w}^G - v_t, \tag{16}$$

where $L_t^c = L(g_t; X)$, and $g_t = g(v_t) + \rho g_{t-1}$.

Lemma 1 *The aggregate level of investment in the provision of community services, v_t, is:*

- *a positive, increasing function*

  $$v_t = \omega(L_t^c) > 0,$$

  of the level of employment in the local community, $L_t^c > 0$, where $\omega'(L_t^c) > 0$;

- *a positive, increasing function,*

  $$v_t = v(X, \rho g_{t-1}) > 0,$$

  of the amount of the location-specific resource, X, and a concave function of the value of community services transmitted from the previous generation, $\rho g_{t-1} \geq 0$, where

  $$v_x(X, \rho g_{t-1}) > 0, \;\; v_g(X, \rho g_{t-1}) > 0, \;\; \text{and } v_{gg}(X, \rho g_{t-1}) < 0$$

  if for all t,

  $$g(v_t) < \mu w_{\min}^c,$$

[4] Community members are not altruistic and therefore they do not consider the effect of their decisions on the well being of future generations. Moreover, the community does not account for the effect of services that it currently provides and thus the level of employment on the future rental rate of the local resource. Given the positive effect of local employment on the rental rate, if this effect were accounted for, investment in community services would be even higher at each point in time, leaving the dynamic analysis that follows qualitatively unchanged.

where $\mu \equiv (1-\alpha-\beta)(1-\gamma)/\gamma < 1$.[5]

Proof. (see Appendix).

2.1 Evolution of Local Employment

The evolution of employment in the local community $\{L_t^c\}_{t=0}^{\infty}$, (which is synonymous with participation in the inclusive institution), is governed by the evolution of the value of community services for each community member, $\{g_t\}_{t=0}^{\infty}$, as well as by the endowment of the location-specific resource, X.

2.1.1 Evolution of the Value of the Community Services

The time path of the value of community services,$\{g_t\}_{t=0}^{\infty}$, is governed by

$$g_{t+1} = g(v_{t+1}) + \rho g_t, \tag{17}$$

where as follows from Lemma 1, $v_{t+1} = v(X, \rho g_t)$, and thus the evolution of g_t is governed by the one-dimensional first-order difference equation,

$$g_{t+1} = g(v(X, \rho g_t)) + \rho g_t \equiv \psi(g_t; X). \tag{18}$$

Lemma 2 *The evolution of the value of the services in the local community, $\{g_t\}_{t=0}^{\infty}$, as depicted in Figure 2, is governed by the one-dimensional, first-order difference equation*

$$g_{t+1} = \psi(g_t; X).$$

- $\psi(g_t; X)$ *is positive, increasing and strictly concave in g_t for all $X > 0$,*

$$\psi(0; X) > 0;$$

$$\psi_g(g_t; X) > 0.$$

$$\psi_{gg}(g_t; X) < 0$$

- $\psi(g_t; X)$ *is positive and increasing in $X > 0$ for all $g_t \geq 0$,*

$$\psi_x(g_t; X) > 0,$$

[5] This condition is satisfied if the value of the additional resources invested in community services, $g(v_t)$, does not exceed the minimal competitive wage defined in A2.

- *The system has a unique and globally stable steady-state equilibrium, $\bar{g}(X)$, i.e.,*

$$\lim_{t \to \infty} g_t = \bar{g}(X) > 0, \quad \text{for all } g_0 \geq 0.$$

where

$$\bar{g}'(X) > 0.$$

Proof. (see Appendix).

As follow from Lemma 2 and as depicted in Figures 2 and 3, the evolution of community services approaches a steady-state level $\bar{g}(X) > 0$. The higher is the amount of the local resources the higher will be the level of community services in the steady state.

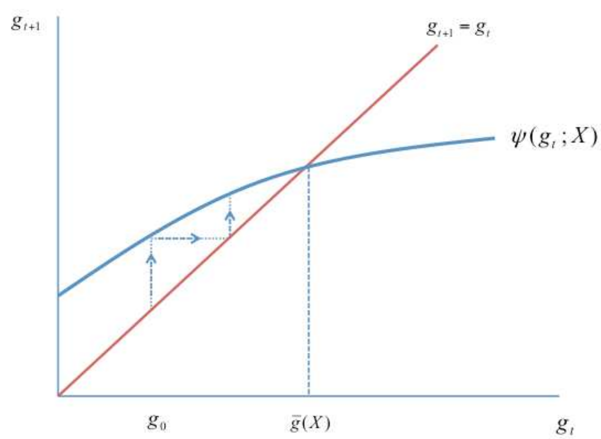

Figure 2. The evolution of the value of the local community services

Corollary 1 *Consider two local communities, a resource abundant community, with resource endowment, X^H, and a resource poor community with resource endowment, X^L. The gap in the value of community services will persist over time.*

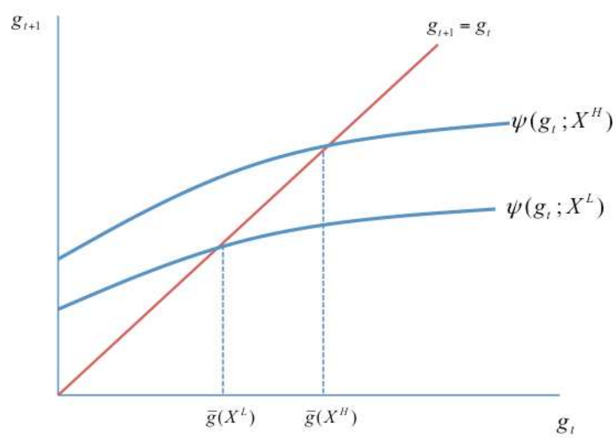

Figure 3. Persistent gap in the value of community services in resource rich, X^H, and resource poor, X^L, areas

2.1.2 Evolution of Employment in the Local Community

The evolution of employment in the local community and the accompanying participation in the inclusive institutions, $\{L_t^c\}_{t=0}^{\infty}$, is governed by by the evolution of the value of community services for each community member, $\{g_t\}_{t=0}^{\infty}$..

As follows from (14) and (27), $L_t^c = L(g_t, X)$ is monotonically increasing in both g_t and X. Hence, given the monotonically increasing evolution of g_t towards steady-state equilibrium, $\bar{g}(X)$, the level of employment in the local community, L_t^c, increases monotonically as well and it has a higher level the higher is the resource endowment, X.

Furthermore,, the evolution of the level of employment has a unique and globally stable steady-state equilibrium, $\bar{L}^c(X)$, i.e.,

$$\lim_{t\to\infty} L_t^c = L(\bar{g}(X), X) \equiv \bar{L}^c(X) \text{ for all } L_0^c > 0, \tag{19}$$

where as follows from (27) and Lemma 2,

$$\frac{d\bar{L}^c(X)}{dX} = \frac{\partial L(\bar{g}(X), X)}{\partial \bar{g}}\bar{g}'(X) + \frac{\partial L(\bar{g}(X), X)}{\partial X} > 0.$$

Corollary 2 • *The level of employment in the local community, $L_t^c = L(g_t, X)$, increases monotonically over time*

- *The evolution of the level of employment in the local community has a unique and globally stable steady-state equilibrium,* $\bar{L}^c(X)$*, i.e.,*

$$\lim_{t\to\infty} L_t^c = \bar{L}^c(X) \quad \textit{for all } L_0^c > 0.$$

- *Consider two local communities, a resource abundant community, with resource endowment,* X^H*, and a resource poor community with resource endowment,* X^L*. Employment will be persistently higher in the resource abundant community, i.e., for every period* t*,*

$$L_t^c(X^H) > L_t^c(X^L).$$

2.2 Persistence of Institutions and Long-Run Misallocation

This section establishes that while the emergence of local institutions that internalize the positive externalities from employment in the local community is beneficial in the short-run, the persistence of these institutions and the intergenerational transmission of local affinity over time could lead to a reduction in the level of income per capita in the long-run in the economy as a whole.

Proposition 1 establishes that the formation of local institutions in period 0, that internalize the positive externalities from employment in the local community, increases aggregate output in the economy as a whole,

Proposition 1 *An increase in the equilibrium level of employment in the local community in period* 0*,* L_0^c*, increases aggregate output in the economy in period* 0*,* $Y_0^G + Y_0^c$*.*

Proof. As follows from (1), noting that $Y_t^G = \bar{w}^G L_t^G = \bar{w}^G(L - L_t^C)$,

$$\frac{\partial[Y_0^c + Y_0^G]}{\partial L_0^c} = [A(L_0^c)F_L(X, L_0^c) + A'(L_0^c)F(X, L_0^c)] - \bar{w}^G. \tag{20}$$

Furthermore, as follows from (12), in the absence of community services, $w_0^c = A(L_0^c)F_L(X, L_0^c) = \bar{w}^G$ and therefore

$$\frac{\partial[Y_0^c + Y_0^G]}{\partial L_0^c} = A'(L_0^c)F(X, L_0^c) > 0. \tag{21}$$

Although the establishment of inclusive institutions increases output in the short run, Proposition 2 establishes that it may reduce the long-run level of output in the economy. In particular, inclusive institutions will have an adverse effect on long-run output if the steady-state value of community services, $\bar{g}(X)$, exceeds the contribution of an additional worker to the output in the local community via its affect on aggregate productivity, $A'(\bar{L}^c)$, i.e.,

$$A'(\bar{L}^c(X))F(X, \bar{L}^c(X)) < \bar{g}(X). \tag{A3}$$

Proposition 2 *Under A3, a movement of worker from the local to the global economy will increase the steady-state level of aggregate output:*

Proof. As follows from (1), noting that $Y_t^G = \bar{w}^G L_t^G = \bar{w}^G(L - L_t^C)$,

$$\frac{\partial[\bar{Y}^c + \bar{Y}^G]}{\partial \bar{L}^c} = [A(\bar{L}^c)F_L(X, \bar{L}^c) + A'(\bar{L}^c)F(X, \bar{L}^c)] - \bar{w}^G. \tag{22}$$

Furthermore, as follows from (12), $A(\bar{L}^c)F_L(X, \bar{L}^c) = \bar{w}^G - \bar{g}(X)$ and therefore

$$\frac{\partial[\bar{Y}^c + \bar{Y}^G]}{\partial \bar{L}^c} = A'(\bar{L}^c)F(X, \bar{L}^c) - \bar{g}(X) < 0. \tag{23}$$

Corollary 3 *If the value of the local network dissipates in the long run (i.e., if $A(\bar{L}^c) = 1$), then the persistence of the inclusive institution will generate a larger reduction in the level of steady-state output in resource rich communities.*

Proof. As follows from (5), $A(\bar{L}^c) = 1$, and thus $A'(\bar{L}^c) = 0$, if and only if $\beta = 0$. Hence, substituting into (23),

$$\frac{\partial[\bar{Y}^c + \bar{Y}^G]}{\partial \bar{L}^c} = -\left.\bar{g}(X)\right|_{\beta=0} < 0. \tag{24}$$

Moreover, as follows from the proof of Lemma 2,

$$\frac{\partial^2[\bar{Y}^c + \bar{Y}^G]}{\partial \bar{L}^c \partial X} = -\left.\bar{g}'(X)\right|_{\beta=0} < 0. \tag{25}$$

Hence, inclusive institutions, which were designed specifically to increase efficiency by internalizing network externalities may generate labor misallocation in the long-run. This misallocation will be larger in resource-rich communities.

3 Testing the Theory

The point of departure for the empirical analysis is the settlement and rapid development of the American Midwest in the middle of the nineteenth century. The theory is tested by exploiting variation in initial resource abundance across counties in the Midwest and its effect on the formation of local inclusive institutions and the emergence of labor misallocation in the long-run.

3.1 Institutional Setting

The Midwest began to be settled in the early nineteenth century with the expansion of the national canal system (Fishlow (2000)). However, it was only with the arrival of the railroad that the Midwest took off on a steeper growth trajectory. Before 1850 the Midwest had less than one thousand miles of track, but almost ten thousand were added by 1857 (Meyer (1989)). Improved rail transportation spurred industrialization and this region's share of national manufacturing increased rapidly between 1860 and 1920, with almost half of this increase occurring in the 1860's (Meyer (1989)). This increase in economic activity led, in turn, to an increase in the demand for labor. In 1810, approximately 6 percent of the labor force (outside the southern states) resided in the Midwest. By 1860, this share had increased to 41 percent, with a further increase to 51 percent by 1880, after which regional growth converged to the national average (Margo (1999)).

The preceding discussion suggests that 1860, just after the railroad boom, would serve as an appropriate point in time at which initial conditions in the Midwest were established. Settlement patterns and railroad expansion over the 1850-1900 period presented in Figure 4 are consistent with this view. Restricting attention to states that had been settled by 1860, the Midwest consists of Illinois, Indiana, Iowa, Michigan, Minnesota, Ohio, and Wisconsin.[6] Using county-level census data it is apparent that the number of incorporated counties in those states increases sharply from 1850 to 1860 and then flattens out. Moreover, the number of these counties with a direct access to the railroad system also increases steeply over the 1850-1870 period, growing thereafter at a slower rate.[7]

[6] Missouri, the only pre-Civil War slave state in the Midwest, is excluded from the analysis.

[7] Information on railroads, obtained from the Historical Map Archive at the University of Alabama. Railroad maps were used to construct a county-level binary variable indicating whether any part of a railroad ran through the county in a given year. Railroad maps were unavailable in some census years in which case maps that were closest in vintage to those census years are used (the discrepancy never exceeded three years).

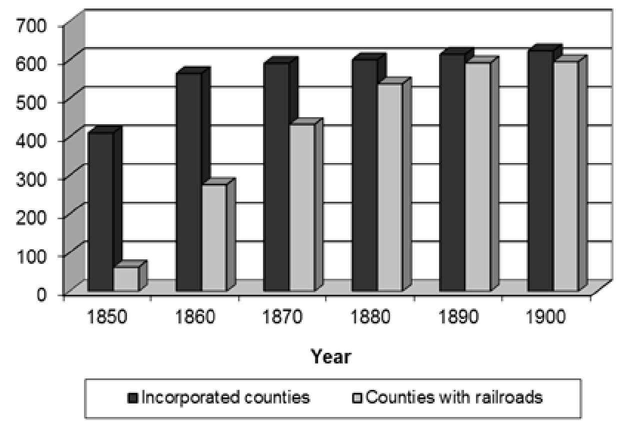

Figure 4. Early growth in the Midwest

The rapid expansion of the railroad system and the economic activity that accompanied it led to a steep increase in the population of the Midwest as well as an influx of foreign migrants. Using county-level census statistics, the total population in the seven Midwestern states in the sample grew from less than 5 million in 1850 to 20 million in 1900. The number of foreign-born migrants nearly tripled between 1850 and 1860, reaching close to 15 percent of the population. Labor markets in the nineteenth century could be divided into three segments: a stable segment with permanent employment, an unstable segment with periodic short-term unemployment, and a marginal but highly flexible segment characterized by spells of long-term and short-term unemployment (Gordon, Edwards, and Reich (1982)). Migrants being newcomers to the U.S. market typically ended up in the unstable and marginal segments, where the uncertain labor demand and the lack of information about their ability and diligence naturally provided an impetus for the formation of ethnic job networks (Conzen (1976), Hoerder (1991)).

Accounts by contemporary observers and an extensive social history literature indicate that friends and kin from the origin community in Europe played an important role in securing jobs for migrants in the Midwest in the nineteenth century and the first quarter of the twentieth century. As an immigrant to the Midwest put it, “The only way you got a job [was] through somebody at work who got you in” (Bodnar, Simon, and Weber (1982): 56). Early historical studies used census data, which provide occupations and country of birth, to identify ethnic clusters in particular locations and occupations (Hutchinson (1956), Gordon, Edwards, and Reich (1982)). More recently, social historians have linked parish registers and county data in specific European sending communities to census and church records in the United States to construct the entire chain of migration from those communities as it unfolded over time. This research has documented the formation of new settlements in the Midwest by pioneering migrants, the subsequent channeling of migrants from the origin community in Europe to these settlements, as well as the movement of groups from the original settlement to new satellite colonies elsewhere in the United States (Gjerde (1985), Kamphoefner (1987), Bodnar (1985)).

Once an ethnic group had established a "toe-hold" (Thistlethwaite (1991)) or a "beach-head" (Bodnar, Simon, and Weber (1982)) in a particular industry or location, it was essential to maintain and even consolidate that presence. Given the variety of economic opportunities in the United States, individuals and small groups drawn from the same parish in Europe often had an incentive to move and seek employment in new locations. The stability of the *local* community in the United States, based on a common national origin rather than narrower social affiliations, was thus essential for the viability of the labor market network. Because *ex post* social sanctions would have little impact once the individual or small group had moved, a more effective strategy would have been to discourage exit *ex ante*. As described in the previous section, one strategy to discourage exit would have been to establish inclusive community institutions, which provided an infrastructure that facilitated social interactions among their members. These social interactions would, in turn, have supported economic networks and social support systems, which members of the community would forego when they left.

3.2 Measuring Resource Abundance

The theory indicates that investments in the community institution and, hence, the level of community services would have been greater in resource-rich locations. However, resources in a county are not directly observed and they may reflect multiple attributes, including the physical infrastructure and natural endowments. The construction of a single resource statistic takes advantage of the fact that a resource-rich county would have typically supported the entry of a larger number of ethnic groups, each with its own network.

To pin down the number of groups, suppose that each ethnic group faces a fixed cost of entry that is increasing in the number of groups. Assume that resources, X, are non-excludable.

This is a reasonable assumption, given that X in the initial period, when the Midwest was being settled, would have been largely determined by transportation infrastructure that was available to everyone. Because larger X is associated with a larger surplus, it follows that more groups would have entered resource-rich counties. As the number of groups increases, the population share of each group i, S_i, mechanically declines. A standard measure of market entry is one minus the Herfindahl Index of market concentration, $1-\sum_i S_i^2$, and this will be used as a proxy measure of resource-intensity in the empirical analysis.[8]

Alternatively, the *level* of entry, measured by the population density, rather than the *heterogeneity* of entry could have been used to capture resource abundance in each county. One major advantage of the heterogeneity-based measure of resource abundance is that it is based on the ethnic groups around which labor networks, and the community institutions that supported these networks, were organized during this period. A second advantage of using the heterogeneity measure is that it allows us to estimate the long-run effect of initial resources in each county, net of the scale of production, measured by the initial population, which could have directly determined the contemporary structure of the local economy through an independent dynamic process associated with agglomeration externalities. As will become apparent, the population of a county in 1860 has a significant effect on the contemporary population level. Thus the estimates of the effect of resource-intensity in 1860 on contemporary outcomes must account for the potentially confounding effect of the initial population level in each county.[9]

Table 1 describes the ethnic shares used in the construction of the resource measure in the initial period. The 1-in-100 individual sample from the Population Census (IPUMS) is used to compute the ethnic share in each Midwestern county. In the absence of data on ethnic affiliation among native born, ethnic shares are computed using first-generation migrants whose country of birth determines their ethnic affiliation. In light of the intergenerational persistence in ethnic migrant networks over this period, these shares are very likely to reflect the ethnic composition in the county as a whole. The English (13 percent), the Irish (25 percent), and the Germans

[8] When there is a single group in a resource-poor county, the value of this measure is zero. When there are many groups, this measure increases in size and in the limit approaches one. Greater resource-intensity is associated with a higher level of community services because of its complementarity with labor and the presence of network externalities. If networks compete directly with each other, the cost of an individual's exit on members that remain will be even larger in resource rich (ethnically heterogenous) locations, reinforcing this results.

[9] In contrast, a measure based on population density, (i.e., the ratio of population to total county area) will not permit us to disentangle the long-run effect of initial resources from the population level.

(32 percent), dominated the migrant labor force in the Midwest in 1860.[10] Subsequently, over the 1860-1900 period, the dominating ethnic groups are the Germans and the Scandinavians.

Given the variety of occupations in this period, as reported in Table 2, and in light of the differential complementarity between resources and labor across occupations, the resource measure ought to capture the heterogeneity of entry within each occupation.[11] The 1860 resource measure in each county is thus constructed as the weighted average of the heterogeneity of entry within each occupational category (where the weight is the fraction of migrant workers in that occupation in the county). However, the results are robust to the use of alternative initial years (i.e., 1850 and 1870) and to constructing this measure without regard to occupation (Table A1).[12] The resource measure across the Midwest in 1860 is depicted in Figure 5. Evidently, there exists substantial variation in this measure, as reflected by a mean of 0.5 and a standard deviation of 0.2 across these Midwestern counties.[13]

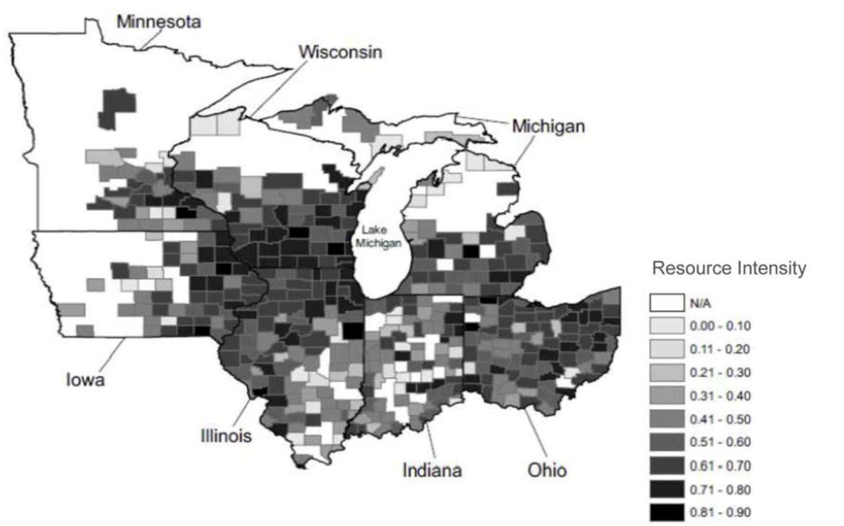

Figure 5. Resource intensity in 1860

The resource measure is designed to capture multiple dimensions of resource intensity. Reassuringly, the constructed resource measure captures resource intensity in each county, as

[10] No other ethnic group accounted for more than a 3 percent share.

[11] Although heavy manufacturing and the growth of large cities came later, 40 percent of the migrant jobs in the 1860 census are already outside the agricultural sector.

[12] The main results are also robust to the inclusion of a full set of ethnic dummies, excluding the possibility that settlement of particular ethnic groups in certain counties had a long run effect (Table A1).

[13] Counties that were not incorporated and those without foreign-born workers in the 1860 sample are unshaded in the figure.

reflected by transportation infrastructure and land quality. Given the importance of transportation links in the development of the Midwest, counties with access to the railroad system, as well as counties situated closer to a canal or Great Lakes harbor would have been associated with a higher level of productive resources.[14] Similarly, superior soil quality would have been associated with higher resource intensity. As established in Column 1 of Table 3, the constructed resource measure is indeed positively associated with each component of the transportation infrastructure as well as soil quality.[15] Furthermore, consistent with the hypothesis that resources and labor are complements, the relationship between resource intensity and its determinants (e.g., transportation infrastructure and soil quality) as reported in Column 1, is qualitatively identical to the relationship between the scale of production in the county, measured by population, and the same determinants in Column 2.

3.3 Identifying Labor Misallocation

Although the provision of community services by inclusive institutions can be efficiency enhancing at the outset, the theory suggests that this can lead to labor misallocation in the long run. Labor misallocation is identified by exploiting the fact that variation in resource-intensity across counties was temporary in nature.[16] In particular, as will become apparent, while resource-intensity was conducive for employment in 1860, it is uncorrelated with standard determinants of contemporary occupational choice. Thus, if one observes that individuals born in historically resource-rich counties are nevertheless disproportionately represented in less mobile occupations characterized by lower human capita intensity, without facing a relative disadvantage with respect to occupation and education opportunities, then this would be indicative of a labor misallocation.

In a standard model of occupational choice, the occupation that an individual chooses

[14] Data on the distance to the nearest canal (or navigable river) and the nearest Great Lakes harbor is obtained from Jordan Rappaport's website at the Kansas City Federal Reserve Bank. The distance is computed in each case from the county centroid.

[15] The transportation infrastructure variables, which are evidently no longer directly salient, are highly jointly-significant in Column 1. Underlining the fundamental role of resource-intensity, the effect of historical resource-intensity on contemporary community participation and occupational choice would grow even larger in (absolute) magnitude if only the part of the variation in resource-intensity that could be explained by the transportation infrastructure variables was used in the analysis.

[16] If natural land productivity is associated with the emergence of generalized social capital rather than community-specific social capital, then it may foster mobility and efficiency (Litina (2013)). The results, however, suggest that this is not the dominating force in the Midwest. Relatedly, historical land inequality could have led land owners to implement persistent mobility reducing institutions (Galor et al., (2009)). As noted, extractive institutions have not emerged in the Midwest. Hence, not surprisingly, as reported in Table A1, the result are unaffected once the potentially confounding effects of historical land inequality are accounted for.

will depend on the supply of occupations in the local area and the cost of investing in human capital, reflecting individual's ability and school quality. To identify a labor misallocation, 1860 resource-intensity must be therefore uncorrelated with each of these determinants of occupational choice today. Initial support for the claim that reduced mobility in historically resource-rich counties cannot be attributed to standard determinants of occupational choice is provided in Tables 4 and 5, using county-level data. Additional validation tests, using individual-level data from the NLSY, are reported in the subsequent section.

First, reduced labor mobility in historically resource rich counties cannot be attributed to superior contemporary economic conditions. As established in Table 4, resource intensity in 1860 is uncorrelated with 1990 manufacturing share (Column 1), agriculture share (Column 2) and population (Column 3).[17] In contrast, recall from Table 3 that resource intensity in 1860 is associated with superior economic conditions at that time, as reflected by the population of the county and the sectoral composition of the local economy. Since these economic attributes could have had persistent effects on economic activity in the long-run, the entire empirical analysis estimates the effect of resource intensity in 1860 *conditional* on 1860 population, agriculture share, and manufacturing share.[18]

Second, reduced labor mobility in historically resource-rich counties cannot be attributed to the differential provision of government-provided public goods. As established in Tables 4 and 5, historically resource-rich counties are not characterized by superior government-provided local amenities or inferior schools. While lower provision of public goods has been attributed to higher fractionalization in society, in the United States, racial, rather than ethnic or religious fractionalization in society is associated with lower provision of public goods (Alesina, Baqir, and Easterly (1999)). As established in Table 4, 1860 resources are uncorrelated with contemporary racial (Column 4), and ethnic (Column 5) fractionalization and have a negative association with religious fractionalization (Column 6).[19] Hence, as expected, there is no direct

[17] The manufacturing share in 1990 is defined as the share of the civilian labor force employed in manufacturing in that year. The agriculture share in 1990 is computed using the farm population and the total population in the county in that year. All these statistics, as well as the area of each county used to compute the population density, are obtained from the 1994 County Data Book, compiled by the U.S. Bureau of the Census.

[18] As established in Figure A1, the main results are unaffected if these controls are removed.

[19] Ethnic fractionalization is computed from the 1990 IPUMS as one minus the Herfindahl index of ethnic concentration, using the same 16 white ethnic groups as in Alesina, Baqir, and Hoxby (2004). Racial fractionalization and religious fractionalization are computed from the 1990 IPUMS and the 1990 Census of Religious Bodies, respectively, using the same five racial groups and the same 18 religious groups as in Alesina, Baqir, and Hoxby.

evidence that historical resource-abundance has resulted in superior government-provided local amenities or inferior schools. In particular, Table 5 establishes that there is no statistically significant association between historical resource-intensity and both public expenditures per capita and education expenditures per student in 1967, 1977, and 1990.[20]

Third, reduced labor mobility in historically rich-resource counties cannot be attributed to lower individual ability in those counties. As established in Column 1 of Table 9, standardized test scores are not associated with historical resources in the county. The absence of a significant statistical association between contemporary ability and initial resources in 1860 removes this potential explanation for reduced mobility.[21]

4 Empirical Results

The theory generates two testable predictions. First, differences in the provision of community services between resource-abundant and resource-scarce counties will persist over time. Second, individuals born in historically high-resource counties will be over-represented in less mobile occupations, resulting in a misallocation of labor in the long-run. These predictions are tested by estimating the relationship between the measure of resource intensity in 1860 and (i) participation over time in community institutions, and (ii) the propensity for mobility and the associated labor misallocation, as reflected by contemporary occupational and educational choices.

4.1 Church Participation

As predicted by the theory, the empirical analysis establishes that resource abundance in 1860 has a positive and persistent effect on church participation, encouraging members of the local community to stay. The focus on church participation reflects the fact that churches were among the first community institutions to be established when European immigrants arrived in the Midwest (Hoerder (1991), Bodnar, Simon, and Weber (1982)). "[Migrants] from varying regions formed a community based on common nationality and religion centered on the central cultural institution – the church" (Gjerde (1991): 176).

[20]Data on expenditures is obtained from the Annual Survey of Governments. School enrollment is obtained from the population census.

[21]Reduced labor mobility in historically-rich resource counties could potentially be attributed to the sorting of individuals with greater inate attachment to local communities (with superior community institutions). This potential sorting, however, has no bearing on the presence of labor misallocation as long as this trait is uncorrelated with ability. These individuals are less mobile despite being equally able.

Historically, the church congregation provided various forms of mutual assistance including credit, insurance, job referrals, business information, and social support (Gjerde (1985), Alexander (1991)). Indeed, it has been argued that the participation of early immigrants to the Midwest in their local churches was primarily motivated by the economic and social services they provided, rather than by belief or ideology (Bodnar (1985)). Midwestern churches have continued to provide important forms of social support for their members over the subsequent 150 years. Church activities have included Sunday school service, youth groups, pot-lucks, informal home parties, and food, visits, and other forms of support when members of the congregation are ailing or infirm. The church also lies at the center of a cluster of inter-linked civic institutions, including the school and various voluntary organizations, that emerged over time. Life in a Midwest community revolves around these institutions, bringing families and friends together on a regular basis (Elder and Conger (2000)). Although other institutions, such as parochial schools, play a role in building an attachment to the local community, the initial focus on the church reflects its central position in community life and the availability of data on church participation at the county level, by denomination, from the Census of Religious Bodies (CRB) at roughly ten-year intervals from 1860 till the present.[22]

In the presence of heterogeneity in the cost of participation, the theory implies that the participation rate will reflect the level of services provided by the community institution. Thus church participation-rates are expected to be persistently higher in resource-rich counties. Consistent with this prediction, as depicted in Figure 6, resource intensity in 1860 has a positive and significant effect on church participation (by 1890) growing steadily thereafter and remaining statistically significant all the way through 2000.[23]

Church participation in the Midwest has remained at roughly 55 percent of the population over the 1860-2000 period. However, this stability masks substantial variation in the mix of denominations over time. Five denominations – Baptist, Catholic, Lutheran, Methodist, and Presbyterian – account for nearly 80 percent of church participants over the 1860-2000 period. Among these denominations, the Catholics and Lutherans grew substantially in popularity,

[22]The CRB was conducted as part of the population census from 1860 to 1890, with census enumerators collecting information from individual churches in each county. Subsequently, the U.S. Bureau of the Census conducted the CRB separately from the population census in ten-year intervals from 1906 to 1936. Starting from 1952, the National Council of Churches of Christ undertook the responsibility of conducting the CRB, with subsequent census rounds in 1972, 1980, 1990 and 2000.

[23]Figure 6 and 7 depict the relationship between 1860 resource-abundance and church participation in each census year between 1870 to 2000, net of population, manufacturing share, and agriculture share in 1860.

accounting for 33 percent and 20 percent of all church participants by 2000, while the other denominations (especially the Methodists and Presbyterians) faced a corresponding decline.

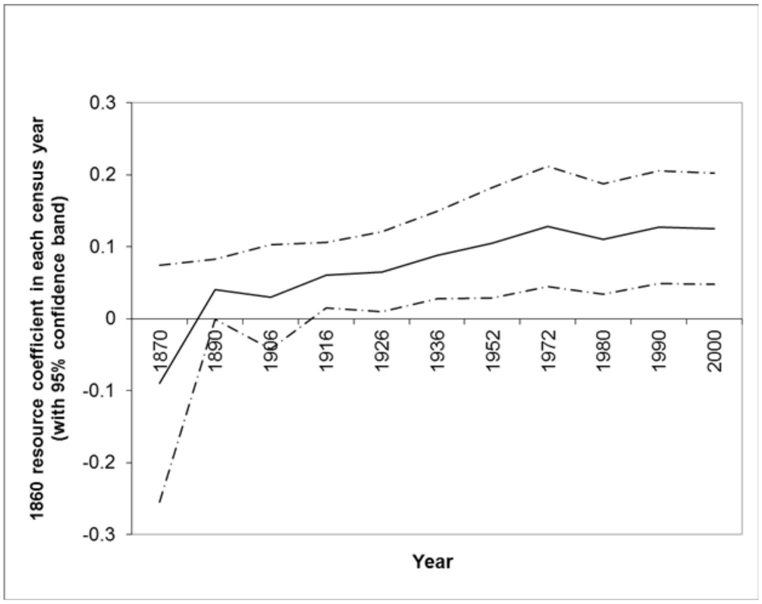

Figure 6. The effect of resource-abundance in 1860 on the time path of church participation

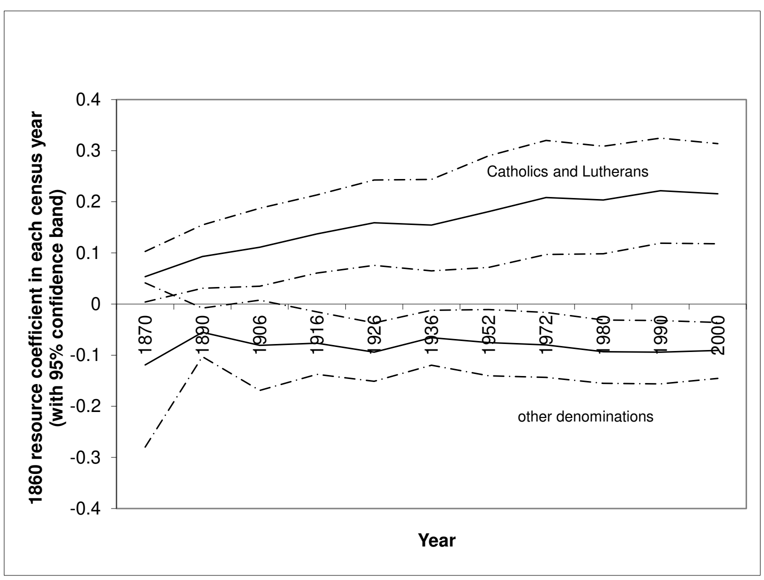

Figure 7. The effect of resource-abundance in 1860 on the time path of church participation by denomination

Based on the country of origin of the first-generation migrants, reported in Table 1, most migrant churches would have been Catholic or Lutheran.[24] Economic networks and social support systems are especially important for new entrants in an economy. Since this is the motivation for the provision of community services, the theory should be most applicable for migrant communities, who were disproportionately Catholic and the Lutheran in 1860. The persistently higher church participation in historically resource-rich counties observed in Figure 6 is indeed driven by participation in Catholic and Lutheran churches in those counties. As established in Column 1 of Table 6 for the year 2000, and as depicted for the entire time period in Figure 7, resource-abundance in 1860 had a positive and significant effect on the share of Catholics and Lutherans in the population; an effect that remains at least as large over time.[25] A one standard deviation increase in 1860 resource intensity would increase the population share of Catholics and Lutherans in the county by four percentage points in 2000 (an increase of 22 percent). In contrast, as established in Column 2 of Table 6 for the year 2000, and as depicted for the entire time period in Figure 7, the effect of resource-abundance in 1860 is much smaller and *negative*, but remains significant and stable over time for other denominations.

The CRB provides information on church participation among residents of the county. To study the effect of historical resource intensity on church participation among individuals *born* in the county, we turn to the National Longitudinal Survey of Youth (NLSY). The NLSY consists of a nationally representative sample of American high school seniors in 1979 who were interviewed annually from 1979 to 1994 and biennially thereafter. A unique feature of the NLSY, which is especially relevant for the analysis, is the availability of the county of birth for each respondent. Matching the patterns established in Columns 1-2, and in Figure 7, with the CRB, the results from the NLSY indicate that individuals born in historically resource-rich counties are significantly more likely to be raised as Lutherans or Catholics (Column 3), whereas 1860 resource-abundance has a negative effect of a much smaller magnitude for other

[24] Apart from the Germans and the Irish, the English were also an important migrant group in 1860. Although it is possible that many of the English were Anglican (Episcopalian), this denomination accounts for just 3 percent of church participants in 1860 and never has a significant presence in the Midwest.

[25] This pattern is obtained for Catholics and Lutherans individually. The 1860-1890 census rounds collected information on the number of church seats by denomination in each county. From 1890 onwards, information was collected on the number of members directly, and from 1972 onwards the number of adherents was collected as well. We use church seats rather than members to measure participation in 1890 because it is more in line with trends in participation over time and we use adherents rather than members to measure participation in the 1972-2000 census rounds because membership information for the Catholics is unavailable in that period. The 1860 resource coefficient increases steadily over time and not in three distinct steps, so our results are unlikely to be driven by changes in the participation measure.

denominations (Column 4).[26]

4.2 Parochial School Enrollment

The focus on the church in the analysis of community participation reflects the role that this institution has occupied in community life from the time the Midwest was developed. Indeed, the importance of the church is a common thread running through historical accounts of the settlement of this region. Other institutions, such as parochial schools, that subsequently formed around the church would also have contributed to building community ties. The prediction of the theory is that individuals born in historically resource-rich counties are more likely to have been exposed to these socializing institutions in childhood as well. This section thus begins by estimating the relationship between enrollment in parochial schools (which are primarily Catholic and Lutheran in the Midwest) and the indicator of resource-abundance in 1860.

Consistent with the prediction of the theory, as established in Table 7, resource-abundance in 1860 has a positive and significant effect on the share of students enrolled in parochial schools in 1970 (Column 1) and in 1980 (Column 2). A standard deviation increase in resource intensity in 1860 would increase the parochial-school share in each of the periods by 0.014 (20 percent).[27] While the 1970 and 1980 Population Censuses report parochial school enrolment at the county level, this information was not collected on the 1990 round. However, all three rounds collected information on private school enrolment. Parochial schools account for the bulk of private schooling in the Midwest (73% of private school enrolment in 1970 and 74% in 1980). The share of private school enrollment in 1990 can thus be used as a proxy for parochial schools enrolment in that year. As established in Column 3-5 of Table 7, resource-abundance in 1860 had a positive, significant and stable effect on the share of students enrolled in private school in the years 1970, 1980, and 1990. Moreover, the coefficients in Columns 3 and 4 are nearly identical in magnitude to those in Columns 1 and 2, respectively, indicating that the effect of resource-abundance in 1860 on private school enrollment is indeed a good proxy for its effect on parochial school enrolment.

The results on parochial schooling, together with the results presented on church par-

[26] The NLSY collects information on the respondents' county of birth and the religious denomination they were raised in. To be consistent with the results using the CRB, we aggregate the individual data in the NLSY to the county level. The regressors in all columns include our 1860 resource-measure and the usual county characteristics.

[27] The 1970 and 1980 population censuses provide county-level information on the total number of children enrolled in school (grades K-12) as well as the number of children enrolled in parochial schools.

ticipation, indicate that individuals growing up in historically resource-rich counties have been more likely to participate in specific community institutions associated with the provision of community services. The consistency of the results on church participation and parochial school enrollment, obtained with three independent data sources (i.e., the CRB, NLSY, and Population Census) further increases confidence in the persistent effects of initial differences in community institutions.

4.3 Reduced Mobility and Labor Misallocation

The second prediction of the theory is that in the long run, individuals born in historically resource-rich counties are more likely to select into occupations associated with low mobility, that are less demanding in terms of human capital accumulation, giving rise to a labor misallocation. This prediction is tested using the NLSY, which specifies the county of birth for each individual as well as residential location in the year 2000.[28]

As established in Table 8, individuals born in historically resource-rich counties are indeed less likely to migrate from their county of birth by the year 2000 and sort into less mobile occupations. As reported in Column 1, a standard deviation increase in historical resource intensity reduces the migration rate by 8 percentage points to fifty percent (i.e., a fourteen percent decline). Moreover, the effect of resource-abundance in 1860 on the average migration rate in the individual's occupation (Column 2) and on the average college completion in that occupation (Column 3) is indeed negative and significant.[29]

The documented selection into less mobile, less demanding occupations is only indicative of a labor misallocation if the availability of jobs in mobile occupations and the effectiveness of investments in human capital do not vary across counties with different historical resource intensities. As has been shown in Table 4 and 5, historical resource-intensity is uncorrelated with local economic conditions, public goods expenditures per capita, and public education expenditures per student. Moreover, previous research indicates that Catholic schools, which make up the bulk of parochial schools, provide higher quality education than public schools (Evans and Schwab (1995), Altonji, Elder, and Taber (2005)). Since students in historically resource-rich counties are significantly more likely to be enrolled in parochial schools, this would

[28] Importantly, the focus on the year 2000 assures that the respondents who were first sampled in the year 1979 are old enough to be settled in their careers and to potentially experience occupation related migration.

[29] Migration and college completion are strongly correlated across occupations (the correlation is 0.6).

suggest that they actually have access to schools of superior quality.

As an additional confirmatory test we establish that individuals residing in historically resource-rich counties in 2000, regardless of their place of birth, are employed in occupations that are as mobile than those of their counterparts residing in historically resource-poor counties. Reassuringly, as reported in Column 5, focusing on the county of residence, the effect of resource-abundance in 1860 on the average migration rate in the individual's occupation is statistically insignificant and much smaller in magnitude than the corresponding coefficient in Column 2. In addition, the effect of resource-abundance in 1860 on average college completion in the individual's occupation changes sign from Column 3, and is positive and marginally significant in Column 6. These results establish that individuals residing and working in historically resource-rich counties are engaged in occupations that are at least as mobile and educationally demanding as the occupations available to individuals in historically resource-poor counties. Moreover, as reported in Column 4 and 7, individuals born in historically resource-rich counties earn significantly lower income, despite the fact that income in their county of birth is not significantly lower. The fact that individuals born in those counties nevertheless select into less mobile occupations is thus indicative of labor misallocation.

Finally, to dismiss the possibility that differences in occupational choice across counties are driven by differences in the ability distribution, the analysis proceeds to examine the relationship between measures of educational attainment and schooling choices obtained from the NLSY and 1860 resources intensity. As reported in Columns 1-3 of Table 9, 1860 resource-abundance does not have a statistically significant effect on AFQT scores, high school completion, and college completion. Apart from being statistically insignificant, the point estimates are extremely small in magnitude.[30] Interestingly, individuals born in historically resource-rich counties are significantly less likely to attend college out of state (Column 4). This choice keeps them closer to home while in college and most likely reduces their mobility when they do graduate. Reinforcing this pattern, individuals born in historically resource abundant counties also select into majors associated with significantly lower migration, conditional on completing college (Column 5).[31] As established in Table A2 and A3, the effects on migration, occupa-

[30]Based on the imprecisely measured point estimates, a one standard deviation increase in 1860 resource intensity would reduce AFQT scores by 0.6 points. As a basis for comparison, the black-white gap is 25 points. The same change in 1860 resource intensity would lower the probability of high school completion by 0.01 and the probability of college completion by 0.02.

[31]These individuals could select into majors associated with higher migration and then forego the benefits

tional choice and educational attainment (Table 8 and 9) are robust to the inclusion of a AFQT scores as a proxy for individual ability.

5 Conclusion

This research advances the hypothesis that economies characterized by a socially cohesive workforce and network externalities gave rise to a particular variety of efficiency-enhancing inclusive institutions, designed to restrict mobility and to increase the attachment of community members to the local labor market. However, their persistence, and the inter-generational transmission of their value, ultimately led to the misallocation of talents across occupations and a reduction in the long-run level of income per capita in the economy as a whole.

Exploiting variation in resource intensity across the American Midwest during its initial development process, the empirical analysis establishes that higher initial resource-intensity in 1860 is indeed associated with greater community participation over the subsequent 150 years, and reduced mobility and labor misallocation in the contemporary period.

that come with higher mobility over the course of their careers. By the same argument, individuals could select into more mobile occupations and subsequently choose not to move as new opportunities arise. The returns to these relative remunerative careers could nevertheless be dominated by the returns in alternative less mobile occupations if the training costs in human-capital intensive careers are relatively high.

6 Appendix

Proof of Lemma 1. Maximizing (16), the optimal level of investment in the provision of public goods must satisfy the condition

$$(1-\alpha)(\alpha+\beta)X^{1-\alpha}(L_t^c)^{\alpha+\beta-1}\frac{\partial L_t^c}{\partial g_t}g'(v_t)=1. \tag{26}$$

As follows from (A1), noting from (7) that $\bar{w}^G - g_t = w_t^c = \alpha X^{1-\alpha}(L_t^c)^{\alpha+\beta-1}$,

$$\begin{aligned}\tfrac{\partial L_t^c}{\partial g_t} &= \tfrac{\partial L(g_t;X)}{\partial g_t} = \tfrac{(L_t^c)^{2-\alpha-\beta}}{(1-\alpha-\beta)\alpha X^{1-\alpha}} > 0;\\ \tfrac{\partial L_t^c}{\partial X} &= \tfrac{\partial L(g_t;X)}{\partial X} = \tfrac{(1-\alpha)L_t^c}{(1-\alpha-\beta)X} > 0.\end{aligned} \tag{27}$$

Using (27), it follows that the optimal level of v_t is given by

$$g'(v_t)=\frac{\alpha(1-\alpha-\beta)}{(1-\alpha)(\alpha+\beta)L_t^c}. \tag{28}$$

As established in (14), $L_t^c = L(g_t;X) > 0$ for all $g_t \geq 0$, and therefore

$$v_t = \omega(L_t^c) > 0, \quad \text{for all } g_t \geq 0, \tag{29}$$

where

$$\omega'(L_t^c) = -\frac{\alpha(1-\alpha-\beta)}{(1-\alpha)(\alpha+\beta)(L_t^c)^2 g''(v_t)} > 0. \tag{30}$$

To derive the second part of the Lemma, note from (14) that,

$$L^c = \left[\frac{\bar{w}^G - g_t}{\alpha X^{1-\alpha}}\right]^{\frac{1}{\alpha+\beta-1}}. \tag{31}$$

It thus follows from (28) that

$$g'(v_t) = \frac{(1-\alpha-\beta)\left[\bar{w}^G - g_t\right]^{\frac{1}{1-\alpha-\beta}}}{(1-\alpha)(\alpha+\beta)\alpha^{\frac{\alpha+\beta}{1-\alpha-\beta}}X^{\frac{1-\alpha}{1-\alpha-\beta}}}. \tag{32}$$

Noting that $g_t = g(v_t) + \rho g_{t-1}$,

$$[(1-\alpha)(\alpha+\beta)]\alpha^{\frac{\alpha+\beta}{1-\alpha-\beta}}g'(v_t)\left[\bar{w}^G - g(v_t) - \rho g_{t-1}\right]^{\frac{1}{\alpha+\beta-1}} = (1-\alpha-\beta)X^{\frac{1-\alpha}{\alpha+\beta-1}}, \tag{33}$$

and therefore

$$\frac{\partial v_t}{\partial X} = -\frac{X^{\frac{2-2\alpha-\beta}{\alpha+\beta-1}}}{(\alpha+\beta)\alpha^{\frac{\alpha+\beta}{1-\alpha-\beta}}\left[\bar{w}^G - g(v_t) - \rho g_{t-1}\right]^{\frac{2-\alpha-\beta}{\alpha+\beta-1}}\left\{g''(v_t)\left[\bar{w}^G - g(v_t) - \rho g_{t-1}\right] + \frac{1}{1-\alpha-\beta}[g'(v_t)]^2\right\}}. \tag{34}$$

Noting that $\bar{w}^G - g_t > 0$ (as follows from (A2)), it follows that

$$\frac{\partial v_t}{\partial X} > 0 \quad \text{if and only if} \quad [g"(v_t) \left[\bar{w}^G - g(v_t) - \rho g_{t-1}\right] + \frac{1}{1-\alpha-\beta}[g'(v_t)]^2] < 0. \tag{35}$$

Hence, since $g(v_t) = v_t^{\gamma}$, where $\gamma \in (0,1)$, noting that $\bar{w}^G - g(v_t) - \rho g_{t-1} = w_t^c$,

$$\frac{\partial v_t}{\partial X} > 0 \quad \text{if} \quad g(v_t) < \frac{(1-\alpha-\beta)(1-\gamma)}{\gamma} w_t^c. \tag{36}$$

Thus $dv/dX > 0$ if $g(v_t) < \mu w_{\min}^c$, where $\mu \equiv (1-\alpha-\beta)(1-\gamma)/\gamma < 1$.

Similarly, noting that $\bar{w}^G - g_t > 0$,

$$\frac{\partial v_t}{\partial g_{t-1}} = -\frac{\frac{\rho}{1-\alpha-\beta} g'(v_t)}{g"(v_t)\left[\bar{w}^G - g_t\right] + \frac{1}{1-\alpha-\beta}[g'(v_t)]^2} > 0, \tag{37}$$

if and only if $[g"(v_t)\left[\bar{w}^G - g_t\right] + \frac{1}{1-\alpha-\beta}[g'(v_t)]^2] < 0$, and therefore if if $g(v_t) < \mu w_{\min}^c$. Moreover,

$$\frac{\partial^2 v_t}{\partial g_{t-1}^2} = \frac{\left[\frac{\rho}{1-\alpha-\beta} g"(v_t)\right]\left[-\{g"(v_t)\left[\bar{w}^G - g_t\right] + \frac{1}{1-\alpha-\beta}[g'(v_t)]^2\}\right] + \left[\frac{\rho}{1-\alpha-\beta} g'(v_t)\right][\rho g"(v_t)]}{\left[-\{g"(v_t)\left[\bar{w}^G - g_t\right] + \frac{1}{1-\alpha-\beta}[g'(v_t)]^2\}\right]^2} < 0. \tag{38}$$

Proof of Lemma 2. As derived in (18), $g_{t+1} = g(v(X, \rho g_t)) + \rho g_t \equiv \psi(g_t; X)$, and

$$\psi(0; X) = g(v(X, 0)) > 0, \tag{39}$$

as established in Lemma 1. Moreover, as follows from (9), (27), and Lemma 1,

$$\begin{aligned} &\psi_g(g_t; X) = \rho[g'(v_{t+1}) v_g(X, \rho g_t) + 1] > 0; \\ &\psi_x(g_t; X) = g'(v_{t+1}) v_x(X, \rho g_t) > 0; \\ &\psi_{gg}(g_t; X) = \rho[g"(v_{t+1}) v_g(X, \rho g_t) + \rho g'(v_{t+1}) v_{gg}(X, \rho g_t)] < 0. \end{aligned} \tag{40}$$

Hence, as depicted in *Figure* 2, $\psi(g_t; X)$ has a unique intersection with the 45^0 line and thus there exists a unique and globally stable steady-state equilibrium,

$$\bar{g} = g(v(X, \rho\bar{g})) + \rho\bar{g} \equiv \psi(\bar{g}; X). \tag{41}$$

where

$$\frac{d\bar{g}}{dX} = \frac{\rho g'(v) v_g(X, \rho\bar{g}))dg + g'(v) v_x(X, \rho\bar{g}))}{(1-\rho)} > 0. \tag{42}$$

6.1 Appendix: Robustness Tests

Each of the robustness test includes separate regressions with migration in 2000 (Table 6, Column 1), the share of Lutherans and Catholics in the population in 2000 (Table 5, Column 1) and the share of children enrolled in parochial schools in 1980 (Table 7, Column 2).[32]

Panel A demonstrates the robustness of the results for alternative construction of the historical resource variable. Columns 1-6 report estimates based on resource intensity measured in the year 1850 and 1870, rather than 1860. Subsequently, Columns 7-9 provide that estimates for a more accurate resource measure, excluding all counties with less than 20 observations in the 1860 IPUMS (approximately 10 percent of the sample).

Panel B further demonstrates the robustness of the results for alternative construction of the historical resource variable.. The statistic that we use in the regressions reported thus far is constructed by averaging across occupations. Results with the corresponding resource measure, constructed without regard to occupation, are reported in Columns 1-3. Table 9 concludes with specifications that include historical characteristics that could potentially determine the outcomes of interest and are correlated with 1860 resource intensity as additional regressors. The population share of each nationality in the 1860 census (computed at the county level) is included in Columns 4-6 to allow for the possibility that particular ethnicities rather than overall resource intensity determine the outcomes. Although the focus is on the role played by inclusive community institutions in discouraging mobility, an alternative explanation is that large landowners in historically resource-rich counties restricted labor mobility to maximize their rents, with these restrictions persisting over time with the support of extractive local institutions. The Midwest, unlike the American South, is not typically associated with such institutions and, not surprisingly, the Gini coefficient included as a regressor to measure land inequality in 1860 has no effect on long-run outcomes in Columns 7-9.

The results in Table 9 indicate that the 1860 resource coefficient remains significant at the 5 percent level and is very stable across the alternative specifications. As a final robustness test, conditional and unconditional estimates of the nonparametric relationship between the three outcomes and 1860 resources are depicted in Figure A3.[33] The individuals in the NLSY

[32] To preserve space, we only report the 1860 resource coefficient for each test in Table 9.

[33] The solid line in the figure is the conditional estimate, with additional regressors partialled out nonparametrically, while the dashed line is the unconditional estimate. The Epanechnikov kernel function is used in Figure 8.

counties are drawn from 150 of our approximately 400 counties, so it is unlikely that a few outlying counties are driving any of the results. It is nevertheless reassuring to observe that the relationships reported earlier with the linear parametric regressions hold up across the entire range of the 1860 resource variable. The population share of Catholics and Lutherans and the share of students enrolled in parochial schools double over this range, while the migration rate declines by nearly 50 percent, emphasizing the importance of the effects uncovered.

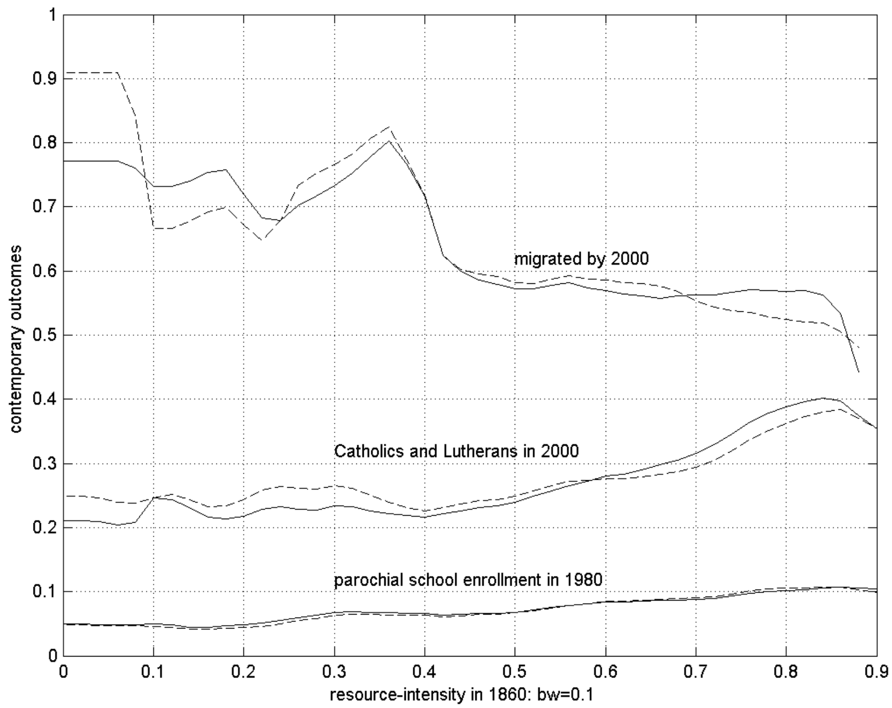

Figure A1. Historical Resource Intensity and Contemporary Outcomes

References

[1] Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. The Colonial Origins of Comparative Development: An Empirical Investigation, *American Economic Review* 91: 1369-1401.

[2] Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2002. Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution, *Quarterly Journal of Economics* 118: 1231-1294.

[3] Acemoglu, Daron and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, New York: Crown Publishers.

[4] Alba, Richard D. 1990. *Ethnic Identity: The Transformation of White America*. New Haven and London: Yale University Press.

[5] Alesina, Alberto, Reza Baqir, and William Easterly. 1999. Public Goods and Ethnic Division, *Quarterly Journal of Economics* 114(4): 1243-84.

[6] Alesina, Alberto, Reza Baqir, and Caroline Hoxby. 2004. Political Jurisdictions in Heterogeneous Communities, *Journal of Political Economy* 112(2): 348-396.

[7] Alexander, June Granatir. 1991. Moving into and out of Pittsburgh: Ongoing Chain Migration, in *A Century of European Migrants, 1830-1930*, eds. Rudolph J. Vecoli and Suzanne M. Sinke, Urbana: University of Illinois Press.

[8] Altonji, Joseph, Todd Elder, and Christopher Taber. 2005. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy* 113(1): 151-184.

[9] Bodnar, John. 1980. Immigration, Kinship, and the Rise of Working-Class Realism in Industrial America, *Journal of Social History*. 45-65.

[10] Bodnar, John. 1985. *The Transplanted: A History of Immigrants in Urban America*. Bloomington: Indiana University Press.

[11] Bodnar, John, Roger Simon, and Michael P. Weber. 1982. *Lives of Their Own: Blacks, Italians, and Poles in Pittsburgh, 1900-1960*. Urbana: University of Illinois Press.

[12] Conzen, Kathleen Neils. 1976. *Immigrant Milwaukee 1836-1860: Accomodation and Community in a Frontier City*. Cambridge, Massachusetts: Harvard University Press.

[13] Dell, Melissa. 2012. "The Persistent Effects of Peru's Mining *Mita*," *Econometrica* 78(6): 1863-1903.

[14] Elder Jr., Glen H. and Rand D. Conger. 2000. *Children of the Land: Adversity and Success in Rural America*. Chicago and London: University of Chicago Press.

[15] Engerman, Stanley L. and Kenneth L. Sokoloff. 1997. Factor Endowments, Institutions, and Differential Growth Paths Among New World Economies, in *How Latin America Fell Behind*, ed. Stephen Haber, Stanford, CA: Stanford University Press.

[16] Evans, William N. and Robert M. Schwab. 1995. Finishing High School and Starting College: Do Catholic Schools Make a Difference? *Quarterly Journal of Economics* 110(3): 947-974.

[17] Fishlow, Albert. 2000. Internal Transportation in the Nineteenth and Early Twentieth Centuries, in *The Cambridge Economic History of the United States*, eds. Stanley L. Engerman and Robert E. Gallman, Cambridge, England: Cambridge University Press.

[18] Galor, Oded and Omer Moav. 2006. "Das Human Kapital: A Theory of the Demise of the Class Structure," *Review of Economic Studies* 73: 85-117.

[19] Galor, Oded, Omer Moav, and Dietrich Vollrath. 2009. "Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence," *Review of Economic Studies* 75: 143-179.

[20] Gans, Herbert. 1962. *The Urban Villagers: Group and Class in the Life of Italian-Americans*. New York: The Free Press of Glencoe.

[21] Gans, Herbert. 1979. Symbolic Ethnicity: The Future of Ethnic Groups and Cultures in America, *Ethnic and Racial Studies* 2(January): 1-20.

[22] Gjerde, Jon. 1985. *From Peasants to Farmers: The Migration from Balestrand, Norway to the Upper Middle West*. Cambridge, England: Cambridge University Press.

[23] Gjerde, Jon. 1991. Chain Migrations from the West Coast of Norway, in *A Century of European Migrants, 1830-1930*, eds. Rudolph J. Vecoli and Suzanne M. Sinke, Urbana: University of Illinois Press.

[24] Gordon, David M., Richard Edwards, and Michael Reich. 1982. *Segmented Work, Divided Workers: The Historical Transformation of Labor in the United States*. Cambridge England: Cambridge University Press.

[25] Hoerder, Dirk. 1991. International Labor Markets and Community Building by Migrant Workers in the Atlantic Economies, in *A Century of European Migrants, 1830-1930*, eds. Rudolph J. Vecoli and Suzanne M. Sinke, Urbana: University of Illinois Press.

[26] Hutchinson, E. P. 1956. *Immigrants and Their Children 1850-1950* New York: John Wiley and Sons, Inc.

[27] Kamphoefner, Walter D. 1987. *The Westfalians: From Germany to Missouri*. Princeton: Princeton University Press.

[28] Kornblum, William. 1974. *Blue Collar Community* Chicago and London: University of Chicago Press.

[29] Landes, David. 1998. *The Wealth and Poverty of Nations: Why Some are so Rich and Some so Poor* New York: W.W. Norton.

[30] Litina, Anastasia. 2013. Natural Land Productivity, Cooperation and Comparative Development. University of Luxembourg.

[31] Maddison, Angus. 2008. "Shares of the Rich and the Rest in the World Economy: Income Divergence Across Nations, 1820-2030," *Asian Economic Policy Review*, June: 67-82.

[32] Margo, Robert A. 1999. Regional Wage Gaps and the Settlement of the Midwest, *Explorations in Economic History* 36: 128-143.

[33] Meyer, David R. 1989. Midwestern Industrialization and the American Manufacturing Belt in the Nineteenth Century, *Journal of Economic History* 49(4): 921-937.

[34] Munshi, Kaivan and Mark Rosenzweig. 2006. "Traditional Institutions Meet the Modern World: Caste, Gender and Schooling Choice in a Globalizing Economy," *American Economic Review* 96(4):1225-1252.

[35] Naidu, Suresh and Noam Yuchtman. 2013. "Coercive Contract Enforcement: Law and the Labor Market in 19th Century Industrial Britain," *American Economic Review*103(1):107-144.

[36] North, Douglas C. 1981. *Structure and Change in Economic History* New York: W.W. Norton.

[37] Sokoloff, Kenneth L. and Stanley L. Engerman. 2000. "History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives* 14(3): 217-232.

[38] Thistlethwaite, Frank. 1991. Migration from Europe Overseas in the Nineteenth and Twentieth Centuries, in *A Century of European Migrants, 1830-1930*, eds. Rudolph J. Vecoli and Suzanne M. Sinke, Urbana: University of Illinois Press.

[39] Willis, Paul. 1977. *Learning to Labor: How Working Class Kids Get Working Class Jobs* New York: Columbia University

Table 1: Ethnic Distribution, 1860-1900

Census year:	1860	1880	1900
	(1)	(2)	(3)
Scandinavia			
Danish	0.01	0.02	0.02
Norwegian	0.03	0.07	0.07
Swedish	0.02	0.06	0.10
British Isles			
English	0.13	0.11	0.09
Irish	0.25	0.19	0.11
Scottish	0.03	0.03	0.02
Welsh	0.01	0.01	0.01
Western Europe			
Dutch	0.01	0.01	0.01
French	0.03	0.02	0.01
German	0.32	0.37	0.41
Swiss	0.02	0.02	0.02
Eastern Europe			
Czech	0.00	0.01	0.02
Polish	0.00	0.01	0.02
Other	0.14	0.07	0.09
Total	1.00	1.00	1.00

Source: IPUMS 1:100 sample, including all foreign-born individuals.

Table 2: Occupational Distribution, 1860-1900

Census year:	1860	1880	1900
	(1)	(2)	(3)
White collar			
Professional	0.04	0.04	0.05
Manager	0.04	0.04	0.06
Clerical	0.00	0.01	0.02
Sales	0.01	0.02	0.03
Farm			
Farmer	0.50	0.41	0.31
Laborer, Farm	0.12	0.17	0.17
Blue collar, nonfarm			
Craftsman	0.10	0.08	0.09
Operative	0.05	0.08	0.09
Household Service	0.05	0.05	0.05
Service	0.00	0.01	0.02
Laborer, Non-Farm	0.09	0.10	0.12
Total	1.00	1.00	1.00

Source: IPUMS 1:100 sample, including all foreign-born individuals
who report that they are employed and report an occupational category.

Table 3: County Characteristics, 1860

Dependent variable:	resource intensity	population	manufacturing share	agriculture share
	(1)	(2)	(3)	(4)
Railroad through county, 1860	0.038	0.098	0.009	-0.026
	(0.021)	(0.012)	(0.006)	(0.012)
Distance to canal, 1890	-0.498	-0.418	0.060	0.157
	(0.191)	(0.112)	(0.081)	(0.113)
Distance to Great Lakes harbor	-0.291	-0.148	-0.041	0.057
	(0.069)	(0.057)	(0.023)	(0.046)
Land quality	0.157	0.047	-0.102	0.176
	(0.069)	(0.033)	(0.035)	(0.037)
Joint sig. of transport variables				
F statistic	11.13	34.99	3.38	3.94
p-value	0.00	0.00	0.02	0.01
Observations	401	401	401	401

Note: Robust standard errors in parentheses.
Distance to canal and distance to Great Lakes harbor measured in thousands of kilometers.
Land quality is county mean of 0.5 degree by 0.5 degree soil suitability measure from Ramankutty et al (2002).
Resource intensity is measured by the heterogeneity of (ethnic) entry: one minus the sum of squared ethnic shares, averaged across occupations.
Manufacturing share and agriculture share in 1860 computed using IPUMS.
Population divided by 100,000.

Table 4: County Characteristics, 1990

Year:	1990					
Dependent variable:	agric share	manufac share	pop	racial frac	ethnic frac	religious frac
	(1)	(2)	(3)	(4)	(5)	(6)
Resource intensity, 1860	0.003	0.018	-0.269	0.011	-0.027	-0.124
	(0.015)	(0.020)	(0.854)	(0.017)	(0.113)	(0.034)
Manufacturing share, 1860	-0.090	-0.056	1.636	0.084	0.176	-0.257
	(0.038)	(0.053)	(1.063)	(0.055)	(0.309)	(0.110)
Agriculture share, 1860	0.136	0.057	-0.962	-0.080	0.049	-0.119
	(0.027)	(0.037)	(0.746)	(0.030)	(0.225)	(0.074)
Population, 1860	-0.087	0.030	11.716	0.246	0.326	0.124
	(0.031)	(0.026)	(5.372)	(0.052)	(0.110)	(0.076)
Observations	437	437	437	437	437	437

Note: Robust standard errors in parentheses.
Resource intensity is measured by the heterogeneity of (ethnic) entry: one minus the sum of squared ethnic shares, averaged across occupations.
Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.
Population is divided by 100,000.
Manufacturing share in 1990 defined as share of civilian labor force employed in manufacturing.
Agriculture share in 1990 is computed using farm population and total population in county.
Ethnic fractionalization in 1990 is one minus the Herfindahl index of (white) ethnic concentration based on 16 ethnicities.
Racial fractionalization in 1990 is one minus the Herfindahl index of racial concentration based on 5 racial groups.
Religious fractionalization in 1990 is one minus the Herfindahl index of religious concentration based on 18 denominations.

Table 5: Public Goods Expenditures

Dependent variable:	local public expenditures per capita			education expenditures per student		
Year:	1967	1977	1990	1967	1977	1990
	(1)	(2)	(3)	(4)	(5)	(6)
Resource intensity, 1860	0.059	0.168	0.143	-0.072	0.274	0.379
	(0.059)	(0.091)	(0.145)	(0.119)	(0.187)	(0.279)
Manufacturing share, 1860	0.082	0.168	0.431	0.027	1.058	1.669
	(0.167)	(0.311)	(0.469)	(0.443)	(0.594)	(0.848)
Agriculture share, 1860	0.081	-0.291	-0.042	0.317	0.044	-0.452
	(0.146)	(0.215)	(0.410)	(0.343)	(0.425)	(0.617)
Population, 1860	-0.121	0.055	0.168	0.013	0.781	0.032
	(0.132)	(0.175)	(0.160)	(0.235)	(0.388)	(0.312)
Observations	437	437	437	437	437	437

Note: Robust standard errors in parentheses.
Resource intensity is measured by the heterogeneity of (ethnic) entry: one minus the sum of squared
Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.
Population is divided by 100,000.
Education expenditure per student is total public spending on grades K-12 (measured in thousands of 1990 dollars) divided by total public school enrollment grades K-12.

Table 6: Church participation in 2000

Data set:	CRB		NLSY	
Dependent variable:	share of county participating in church denomination		share of sample raised in church denomination	
1860 variables measured in:	county of residence		county of birth	
Church denomination:	Catholic/Lutheran	other	Catholic/Lutheran	other
	(1)	(2)	(3)	(4)
Resource intensity, 1860	0.216	-0.091	0.224	-0.132
	(0.049)	(0.027)	(0.113)	(0.139)
Manufacturing share, 1860	0.464	-0.182	0.341	-0.014
	(0.134)	(0.066)	(0.317)	(0.381)
Agriculture share, 1860	0.314	0.056	0.125	-0.207
	(0.103)	(0.065)	(0.229)	(0.255)
Population, 1860	-0.229	0.063	-0.111	0.119
	(0.103)	(0.031)	(0.089)	(0.089)
Observations	437	437	222	222

Note: Standard errors in parentheses.
Resource intensity is measured by the heterogeneity of (ethnic) entry: one minus the sum of squared
Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.
Population divided by 100,000.

Table 7: Education (Population Census)

Dependent variable:	parochial school share		private school share		
Year:	1970	1980	1970	1980	1990
	(1)	(2)	(3)	(4)	(5)
Resource intensity, 1860	0.071	0.070	0.077	0.080	0.075
	(0.016)	(0.012)	(0.016)	(0.013)	(0.012)
Manufacturing share, 1860	0.126	0.072	0.138	0.088	0.075
	(0.049)	(0.044)	(0.049)	(0.047)	(0.040)
Agriculture share, 1860	-0.014	-0.016	-0.025	-0.015	0.001
	(0.039)	(0.032)	(0.040)	(0.034)	(0.029)
Population, 1860	0.025	0.050	0.046	0.064	0.064
	(0.021)	(0.015)	(0.020)	(0.016)	(0.013)
Observations	437	437	437	437	437

Note: Robust standard errors in parentheses.

Resource intensity is measured by the heterogeneity of (ethnic) entry: one minus the sum of squared ethnic shares, averaged across occupations.

Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.

Population is divided by 100,000.

Parochial school share is parochial school enrollment grades K-12 divided by school enrollment grades K-12.

Private school share is private school enrollment grades K-12 divided by school enrollment grades K-12.

Table 8: Mobility in 2000 (NLSY)

1860 variables measured in:	county of birth				county of residence		
Dependent variable:	migrated	avg migration in occupation	avg college completion in occupation	income	avg migration in occupation	avg college completion in occupation	income
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Resource intensity, 1860	-0.414	-0.127	-0.188	-9.513	-0.010	0.123	-4.444
	(0.138)	(0.062)	(0.084)	(5.547)	(0.058)	(0.076)	(5.201)
Manufacturing share, 1860	-0.242	-0.178	-0.407	-21.01	0.159	0.272	-18.41
	(0.260)	(0.120)	(0.190)	(11.30)	(0.096)	(0.218)	(9.293)
Agriculture share, 1860	-0.317	-0.127	-0.377	-14.63	0.179	0.232	-1.964
	(0.222)	(0.089)	(0.134)	(8.670)	(0.078)	(0.122)	(7.507)
Population, 1860	0.065	0.030	0.030	1.544	0.002	0.037	3.761
	(0.040)	(0.014)	(0.024)	(1.501)	(0.013)	(0.017)	(1.344)
White	0.284	0.089	0.091	5.284	0.016	0.094	5.239
	(0.052)	(0.018)	(0.013)	(1.454)	(0.021)	(0.020)	(1.400)
Female	-0.014	-0.002	0.002	-13.60	-0.012	0.010	-13.99
	(0.024)	(0.013)	(0.014)	(1.072)	(0.012)	(0.013)	(1.072)
Age	0.015	0.001	0.004	0.181	0.003	0.005	0.384
	(0.005)	(0.003)	(0.004)	(0.292)	(0.003)	(0.003)	(0.246)
Observations	1437	1216	1216	1096	1237	1237	1113

Note: Standard errors in parentheses are clustered at the county level.

Resource intensity is measured by the heterogeneity of (ethnic) entry: one minus the sum of squared ethnic shares, averaged across occupations.

White, female, and age are individual-level characteristics.

Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.

Population divided by 100,000.

Annual income in 2000 in thousands of dollars. Top 5% of the income distribution is truncated.

Table 9: Education (NLSY)

Dependent variable:	AFQT score	high school completion	college completion	college out of state	average migration in college major
	(1)	(2)	(3)	(4)	(5)
Resource intensity, 1860	-3.185	-0.056	-0.120	-0.386	-0.242
	(8.695)	(0.082)	(0.126)	(0.187)	(0.094)
Manufacturing share, 1860	-14.421	-0.185	-0.347	-0.686	-0.077
	(19.737)	(0.180)	(0.241)	(0.338)	(0.143)
Agriculture share, 1860	-17.311	-0.269	-0.477	-0.665	0.110
	(17.015)	(0.178)	(0.198)	(0.306)	(0.165)
Population, 1860	-3.028	-0.061	-0.009	0.018	0.026
	(2.039)	(0.027)	(0.026)	(0.036)	(0.033)
White	25.469	0.043	0.128	0.021	0.041
	(1.724)	(0.024)	(0.021)	(0.039)	(0.056)
Female	-3.314	0.005	-0.041	0.019	0.004
	(1.893)	(0.020)	(0.026)	(0.027)	(0.023)
Age	2.642	0.009	0.010	0.011	0.0009
	(0.366)	(0.003)	(0.004)	(0.006)	(0.006)
Observations	1390	1437	1437	814	289

Note: Standard errors in parentheses are clustered at the county level.
Resource intensity is measured by the heterogeneity of (ethnic) entry: one minus the sum of squared ethnic shares, averaged across occupations.
Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.
Population is divided by 100,000.
Ethnic competition, manufacturing share, and agricultural share measured in county of birth.
White, female, and age are individual-level characteristics.
AFQT is the score on the Armed Forces Qualification Test.
High school completion is a binary variable indicating whether the individual completed high school, including GED.
College completion is a binary variable indicating whether the individual completed a four-year college/university degree.
Average migation in college major is the proportion of individuals in the respondents' college major who reside outside of their counties of birth.

Table A1: Robustness Tests

Dependent variable:	a	b	c	a	b	c	a	b	c
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Alternative construction of resource measure, 1860									
Resource intensity, 1860	0.356	-0.368	0.102	0.361	-0.292	0.075	0.288	-0.410	0.081
	(0.053)	(0.121)	(0.015)	(0.050)	(0.195)	(0.012)	(0.051)	(0.169)	(0.012)
Observations	290	1157	290	517	1471	517	415	1417	415
Robustness test	resource intensity, 1850			resource intensity, 1870			exclude counties w/ less than 20 obs, 1860		
Panel B: Alternative construction of resource measure and additional regressors, 1860									
Resource intensity, 1860	0.120	-0.422	0.055	0.095	-0.482	0.034	0.127	-0.459	0.056
	(0.047)	(0.136)	(0.012)	(0.049)	(0.178)	(0.013)	(0.048)	(0.129)	(0.012)
Observations	437	1437	437	437	1437	437	432	1437	432
Robustness test	no occupation categories			pop share of each ethnicity			Gini as additional regressor		

Notes: Dependent variable "a" is share of Catholics and Lutherans in the population in 2000 (CRB data at the county level).
Dependent variable "b" is migrated by 2000 (NLSY data at the individual level).
Dependent variable "c" is share of students in parochial schools in 1980 (population census at the county level).
In individual-level regressions, standard errors in parentheses clustered at the county level.
In county-level regressions, robust standard errors in parentheses.
All specifications include 1860 manufacturing share, 1860 agricultural share, and 1860 population as regressors.
Individual-level regressions with migrated as the dependent variable include white, female, and age as additional regressors.
Population share of each ethnicity in 1860 and Gini coefficient of land inequality in 1860 are included as regressors in Panel B, Columns 4-9.

Table A2: Mobility in 2000 (Controlling for AFQT)

1860 variables measured in:	county of birth				county of residence		
Dependent variable:	migrated	avg migration in occupation	avg college completion in occupation	income	avg migration in occupation	avg college completion in occupation	income
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Resource intensity, 1860	-0.410	-0.116	-0.178	-10.19	-0.025	0.079	-6.448
	(0.147)	(0.070)	(0.089)	(4.836)	(0.059)	(0.061)	(4.515)
Manufacturing share, 1860	-0.225	-0.130	-0.358	-19.54	0.168	0.282	-16.08
	(0.270)	(0.136)	(0.186)	(9.449)	(0.086)	(0.149)	(6.723)
Agriculture share, 1860	-0.297	-0.064	-0.286	-9.65	0.196	0.257	-0.611
	(0.209)	(0.098)	(0.135)	(7.211)	(0.072)	(0.101)	(6.257)
Population, 1860	0.069	0.040	0.057	2.942	0.003	0.050	4.608
	(0.040)	(0.015)	(0.023)	(1.096)	(0.013)	(0.015)	(1.059)
White	0.221	0.054	-0.001	-0.538	-0.007	0.003	0.044
	(0.052)	(0.020)	(0.014)	(1.344)	(0.021)	(0.018)	(1.380)
Female	-0.006	0.002	0.022	-12.67	-0.011	0.022	-13.45
	(0.023)	(0.013)	(0.013)	(1.109)	(0.013)	(0.012)	(1.063)
Age	0.008	-0.003	-0.007	-0.488	0.00001	-0.004	-0.129
	(0.006)	(0.003)	(0.003)	(0.260)	(0.003)	(0.003)	(0.245)
AFQT score	0.002	0.002	0.004	0.245	0.001	0.004	0.213
	(0.001)	(0.0002)	(0.0003)	(0.023)	(0.0002)	(0.0003)	(0.019)
Observations	1390	1176	1176	1062	1193	1193	1074

Note: Standard errors in parentheses are clustered at the county level.

Resource intensity is measured by the heterogeneity of (ethnic) entry: one minus the sum of squared ethnic shares, averaged across occupations.

White, female, and age are individual-level characteristics.

Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.

Population divided by 100,000.

AFQT is the score on the Armed Forces Qualification Test.

Annual income in 2000 in thousands of dollars. Top 5% of the income distribution is truncated.

Table A3: Education (Controlling for AFQT)

Dependent variable:	high school completion	college completion	college out of state	average migration in college major
	(1)	(2)	(3)	(4)
Resource intensity, 1860	-0.044	-0.074	-0.422	-0.263
	(0.075)	(0.123)	(0.179)	(0.103)
Manufacturing share, 1860	-0.089	-0.219	-0.703	-0.061
	(0.151)	(0.256)	(0.320)	(0.157)
Agriculture share, 1860	-0.176	-0.346	-0.653	0.260
	(0.148)	(0.199)	(0.303)	(0.179)
Population, 1860	-0.044	0.011	0.031	0.059
	(0.020)	(0.027)	(0.038)	(0.029)
White	-0.069	-0.056	0.013	0.046
	(0.023)	(0.021)	(0.040)	(0.063)
Female	0.027	-0.016	0.026	0.011
	(0.018)	(0.017)	(0.029)	(0.021)
Age	-0.002	-0.009	0.013	0.006
	(0.004)	(0.004)	(0.006)	(0.006)
AFQT score	0.005	0.007	0.0003	0.0004
	(0.0004)	(0.0005)	(0.001)	(0.001)
Observations	1390	1390	795	283

Note: Standard errors in parentheses are clustered at the county level.
Resource intensity is measured by the heterogeneity of (ethnic) entry: one minus the sum of
averaged across occupations.
Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.
Population is divided by 100,000.
Ethnic competition, manufacturing share, and agricultural share measured in county of birth.
White, female, and age are individual-level characteristics.
AFQT is the score on the Armed Forces Qualification Test.
High school completion is a binary variable indicating whether the individual completed high school, including GED.
College completion is a binary variable indicating whether the individual completed a four-year college/university degree.
Average migation in college major is the proportion of individuals in the respondents' college major who reside
outside of their counties of birth.