Identity and Mobility: Historical Fractionalization, Parochial Institutions, and Occupational Choice in the American Midwest *

Kaivan Munshi† Nicholas Wilson‡

December 2010

Abstract

This paper examines the role played by a specific identity, defined as the attachment to a hometown, in determining occupational choice and mobility. The analysis links competition between ethnic networks in the Midwest when it was first developing, and the in-group identity that emerged endogenously to support these networks, to institutional participation and occupational choice today. Individuals born in counties with greater ethnic fractionalization in 1860 are today – 150 years later – (i) significantly more likely to participate in institutions such as churches and parochial schools that transmit identity from one generation to the next, and (ii) significantly less likely to select into mobile skilled occupations. The effect of historical fractionalization on participation in these socializing institutions actually grows stronger over the course of the twentieth century, emphasizing the idea that small differences in initial conditions can have large long-term effects on institutions and economic choices.

Keywords. Identity. Institutional persistence. Networks. Occupational choice. Mobility.
JEL. D85. J62. L14. L22.

*We thank Roland Bénabou, Andrew Foster, Dennis Hogan, Ashley Lester, Nancy Luke, David Weil, and, especially, Oded Galor and Rachel Kranton for helpful comments and discussions. Seminar participants at Brown, Duke, Georgia Tech, Harvard-MIT, Indiana University, NBER, Stanford, Toronto, UCLA, Warwick, and York provided many useful comments. Ben Feigenberg and Jordan Ou provided excellent research assistance. Research support from the National Science Foundation through grant SES-0617847 is gratefully acknowledged. We are responsible for any errors that may remain.

†Brown University and NBER
‡Williams College

1 Introduction

Modern economic theories of growth and development assign individuals to occupations on the basis of their ability, the opportunities they have to invest in human capital, and the types of jobs that are subsequently made available to them in the labor market (Galor and Zeira 1993, Bénabou 1996, Fernandez and Rogerson 1996, Durlauf 1996). Although these theories go beyond the classical model by allowing for credit market imperfections and complementarities in human capital investments, social scientists outside of economics have long argued that factors unrelated to economic opportunities and constraints can also determine career choices (eg., Weber 1930, Gans 1962). In response to this critique and in an attempt to broaden the scope of existing economic models, this paper examines theoretically and empirically the role played by a specific identity, defined as the attachment to a hometown, in determining occupational choice and mobility.

Consider a town in the American Midwest, the setting for this paper. Although we do not attempt to explicitly characterize the process through which individuals come to identify with their hometown, this attachment is presumably a consequence of the activities and social interactions they are exposed to in church, the school, and other community institutions in childhood (Elder and Conger 2000). Attachment to a hometown can also influence economic decisions. In a series of influential papers, Akerlof and Kranton (2000, 2005) develop a theory that describes the effect of group identity on individual behavior. The insight from their theory is that individuals suffer disutility when they choose actions that deviate from customary behavior or are not in the best interest of the group they identify with and will consequently avoid those actions. Individuals who are attached to their hometown will thus be more likely to stay back and contribute to their community, selecting into less mobile occupations independently of the economic incentives and opportunities that they might face.

This decision to stay is well described by the following quote from a resident of a small Midwestern city, "Such a life as I came to understand it ... meant belonging to a place rather than to a profession ... as I came to understand my children's need and my own need for a firm home place, I came to understand my community's need for citizens who stay put" (Sanders 2007: 67). In the sections that follow we will investigate (i) why a stronger sense of attachment emerged in some local areas but not others when the Midwest was first developing, (ii) how this identity was transmitted from one generation to the next with the support of permanent community institutions, and (iii) what consequences it has for occupational choice and mobility today.

The Midwest developed extremely rapidly following the railroad expansion of the 1850's. Economic growth led to an influx of European migrants. These migrants typically ended up

in unstable and marginal segments of the labor market, where the uncertain labor demand provided the impetus for the formation of ethnic networks (Conzen 1976, Hoerder 1991). Most theories of identity formation are based on the in-group versus out-group dichotomy first proposed by Sumner (1906), with competition between groups giving rise to in-group identity. Our explanation in Section 2 for why stronger identity, as we have defined it, would have emerged in some local areas or counties when the Midwest was being settled follows this tradition and complements new research that seeks to model the instilling of values as an endogenous response to economic incentives (Tabellini 2008). When ethnic groups are competing in a local labor market, the cost imposed on a group by the migration of one of its members is increasing in the degree of competition, which is conveniently measured by a standard fractionalization statistic (one minus the Herfindahl index of group concentration). Ethnic groups in highly fractionalized locations would thus have had a greater incentive to establish community institutions that created identity and tied individuals to their hometown when the Midwest was developing. While the presence of ethnic networks in the labor market and subsequent competition between these networks is not particular to the Midwest, what makes this region especially well suited for our analysis is a well defined starting point – around 1860 – and the availability of individual-level data from the population census in that year. Using the 1-in-100 sample from the 1860 IPUMS, ethnic fractionalization in the labor market can be computed in each Midwestern county that was incorporated and had foreign-born workers.

Based on the preceding discussion, a stronger identity or attachment would have emerged in historically fractionalized counties.[1] If identity subsequently persisted over time, the first prediction of the model developed in Section 3 is that individuals born in historically fractionalized counties should be less likely to select into mobile skilled occupations today. But what is the mechanism underlying such persistence? Parents exert costly effort to socialize their children in Bisin and Verdier's (2000) pioneering model of cultural transmission. In an economy where internal migration is relatively high on average, the persistence of identity within families (or dynasties) will not explain differences across counties over many generations. Our dynamic model thus assumes that socialization occurs within permanent community institutions instead. There is a fixed cost to coordinating the efforts of many individuals and so only some communities find it worthwhile to invest in the activities and social interactions that are needed for identity formation in the initial period of the model (for reasons described above). Once this initial investment is made, children who participate in socializing institutions such as churches and schools in those communities develop identity

[1] Attachment to the home community could have subsequently emerged in a county for other reasons. 1860 fractionalization thus explains some but not all of the variation in identity across counties. We focus on the initial condition because this allows us to trace the dynamic process through which identity evolved and because it will be less correlated with other determinants of occupational choice today (as discussed below).

with fixed probability in all the generations that follow. As in Bisin and Verdier, we assume that individuals with identity want their children to share these values. This generates the second prediction of the model, which is that participation in these socializing institutions will be greater in historically fractionalized counties, with this effect growing stronger over time as identity propagates through the population.

A common thread running through historical accounts of the settling of the Midwest is the role of the church in forming new communities based on common nationality and residential location in the United States. The church continues to lie at the center of a cluster of inter-linked civic institutions around which life in a Midwestern town revolves. We thus expect churches and other related institutions in historically fractionalized counties to have been oriented toward instilling values that attach their members to their hometown, with this orientation persisting over the generations. Data on church participation are available at roughly ten-year intervals from 1870 to 2000 from the Census of Religious Bodies (CRB). As predicted, participation in these socializing institutions is greater in historically fractionalized counties, with the participation-gap growing wider over time. Matching this striking persistence over more than a century, population-census data from 1970 and 1980 indicate that a significantly larger fraction of students in historically fractionalized counties are enrolled in parochial schools, which will support the church in building community ties. To test the key prediction of the model for occupational choice, we take advantage of the fact that the National Longitudinal Survey of Youth (NLSY) collects information on the respondents' county of birth. Complementing the results on church participation and parochial school enrollment, individuals born in historically fractionalized counties select into less mobile occupations and are significantly less likely to have migrated from their county of birth by 2000 (when they are on average 39 years old). It is reassuring to observe that results consistent with the model are obtained with three outcomes and three independent data sets.

Our interpretation of these results is that individuals born in historically fractionalized counties are exposed to socializing institutions in childhood, which increases the level of participation in these institutions and discourages them from leaving in adulthood. The fact that identity or attachment is not observed should not be a cause for concern. There is, after all, a well established tradition in economics testing demand theory without observing preferences. The more relevant concern is that 1860 fractionalization could be correlated with economic opportunities or constraints today. While fractionalization might generate in-group identity, the negative effects of fractionalization on inter-group cooperation and public good provision have also been extensively documented. If individuals face a greater cost of investing in human capital in historically fractionalized counties because school quality is lower today, for example, then we are back to a standard explanation for differences in occupational choice across locations. Our strategy to deal with the conflicting effects of

fractionalization is to go back to the mid-nineteenth century when the Midwest was being settled, which is when community institutions were being established and in-group identity would have formed but is long before major expenditures on public schooling and local infrastructure occurred. As we show, theoretically and empirically in the paper, identity could persist long after the economic circumstances that gave rise to it had ceased to be relevant. We nevertheless take care in Section 4 to rule out the possibility that individuals born in historically fractionalized counties select into less mobile occupations, which are associated with lower levels of college completion, simply because the availability of skilled jobs is lower or the cost of investing in human capital is higher in those counties today. Historical fractionalization is uncorrelated with the skill-mix of locally available jobs, expenditures on public education, test scores, and educational attainment.

Can persistent differences across communities be explained by differences in institutional quality alone? Examination of the model indicates that such persistence could be obtained without a role for identity if there are increasing returns to scale in the benefits that community institutions provide. As discussed below, such economies of scale are unlikely to be relevant in our Midwestern setting. Alternatively, suppose that there is a fixed cost to investing in the activities and social interactions that give rise to identity in the initial period. Once these investments were made in historically fractionalized counties, the superior institutions that emerged could have provided greater benefits to all their residents in the generations that followed. We show formally that this cannot explain the key empirical result, which is that the supply of mobile skilled labor is lower in historically fractionalized counties, despite the fact that the availability of skilled jobs does not vary systematically with fractionalization. While wages will adjust for differences in institutional quality across counties in a competitive labor market, this would not affect the allocation of individuals to jobs. What identity does is to place a premium on a specific *location* – the hometown – and this is what generates the mismatch between the supply and the availability of skilled jobs in equilibrium.

How large are the effects we have uncovered? Our estimates indicate that a one standard deviation increase in 1860 fractionalization would reduce migration out of the birth-county in 2000 by 14 percent. The foreign-born workers that are used to compute the fractionalization statistic in 1860 account for less than 15 percent of the population in that year. Although a persistent (and increasing) role for historical fractionalization is consistent with a model in which identity is propagating through the population over many generations, it is nevertheless quite remarkable that county-level differences 150 years ago based on such a small fraction of the population could continue to have large effects on migration and career decisions across counties that are currently indistinguishable with respect to standard determinants of occupational choice.

2 Initial Conditions

We begin this section by describing the early development of the Midwest and the competition between migrant ethnic groups in local labor markets that occurred at the time. We then proceed to derive the relationship between this competition, measured by ethnic fractionalization at the level of the county, and the strength of the hometown ties that would have emerged. This section concludes with a description of the fractionalization statistic that is used to measure initial identity across our Midwestern counties.

2.1 Institutional Setting

The Midwest first 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). Indeed, more than half the track-miles built in the United States between 1853 and 1856 were added in the Midwest (Atack et al., forthcoming). 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 diffusion over the 1850-1900 period presented in Figure 1 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 (Missouri, the only pre-Civil War slave state in the Midwest, is excluded from the analysis). Using county-level census data we see that the number of incorporated counties in those states increases sharply from 1850 to 1860 and then flattens out in Figure 1. Information on railroads, obtained from the Historical Map Archive at the University of Alabama, indicates that the number of these counties with a railroad also increases steeply over the 1850-1870 period, growing thereafter at a slower rate.[2] Although we treat 1860 as the initial period in

[2]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 we used maps that were closest in vintage to those census years (the discrepancy never exceeded three years).

the analysis that follows, the results are robust to using 1850 or 1870 instead.

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 our seven Midwestern states 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 naturally provided an impetus for the formation of ethnic networks that helped their members find jobs (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. Conzen (1976) describes the competition (and conflict) between Irish and German immigrants when Milwaukee was being settled. And the following quote from a Polish immigrant well describes the ethnic competition that very likely was a feature of most labor markets at the time, "You take in the erection department – it was mostly all Slavs ... Not Slovaks, it was Polish ... We didn't have Lithuanians there and the Russians were not involved there ... Now if a Russian got his job in a shear department ... he's looking for a buddy, a Russian buddy. He's not going to look for a

Croatian buddy. And if he see the boss looking for a man he says, 'Look, I have a good man,' and he's picking out his friends" (Bodnar, Simon, and Weber 1982:62). 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 discussed below, the incentive to invest in the activities and social interactions that were needed to build community ties would have been greater in fractionalized locations characterized by greater labor market competition between ethnic groups.

2.2 Identity Formation

Theories of group identity in the social sciences can be broadly divided into two categories: (1) *contact* theory in which diversity fosters inter-group tolerance and weakens in-group identity, and (2) *conflict* theory in which diversity strengthens in-group identity, while at the same time increasing prejudice towards other groups (Putnam 2007). Among these theories, conflict theory has received a disproportionate share of research attention and empirical support in the social science literature. Central to this theory is the in-group versus out-group dichotomy first proposed by Sumner (1906). Competition *between* groups over scarce resources is seen to give rise to greater in-group identification (Ragin 1979, Giles and Evans 1986).

Why does inter-group competition give rise to in-group identity? Much of the social psychology literature appears to treat in-group identity formation as a by-product of competition between groups (Brewer and Brown 1998). In contrast, Bisin et al. (2008) develop a model in which identity is instilled purposefully in response to inter-group competition. In their model, there is a cost to interacting with or competing with members of other groups, which is decreasing in the share of the individual's own group in the population. In-group identity reduces this cost, which implies that smaller groups will have a greater incentive to invest in these values.

Although the motivation for identity formation in Bisin et al. (2008) is psychological, economic incentives can also give rise to identity. Based on the preceding description of the early development of the Midwest, consider a local labor market with N migrant workers drawn from M ethnic groups. These workers are competing for a limited number of permanent jobs, which provide surplus R. Because members of an ethnic network provide referrals for each other, larger networks will capture a greater share of the surplus. Tullock's (1980) canonical model of rent seeking can be conveniently adapted to this setting to describe the

share of the surplus captured by group i

$$S_i = \frac{n_i^\alpha}{n_i^\alpha + \sum_{j=1}^{M-1} n_j^\alpha}$$

where n_i is the number of workers belonging to group i, n_j is the corresponding number of workers from each group $j \neq i$, and $\alpha > 0$.[3] Differentiating with respect to n_i, and assuming that all group sizes are equal, $n_i = n_j = N/M$, the cost to group i from the exit of its members at the margin can be described by the expression

$$R\frac{dS_i}{dn_i} = \frac{R\alpha}{N}\left(1 - \frac{1}{M}\right),$$

which is increasing in the number of groups, M. Conditional on the number of workers N and the surplus R, a decline in the share of each group, $1/M$, is associated with a greater marginal cost of exit regardless of whether the returns to group-size are increasing ($\alpha > 1$) or decreasing ($\alpha < 1$) at the margin.

Self-interested individuals do not internalize the cost imposed on the rest of the group by their exit from the local market. One way for recently established local communities to have aligned such exit more closely with the socially optimal level when the Midwest was first developing would have been to invest in the activities and social interactions that give rise to identity, since attachment to the home community increases the individual's exit-cost. If there is a fixed cost to coordinating the efforts of many individuals when making these investments, the probability that identity would have emerged would have been greater in locations where each group accounted for a relatively small share of the population.

If groups were of equal size, in-group identity would have been more likely to emerge in locations or counties with more ethnic groups and, hence, greater inter-group competition. If group sizes varied, the Herfindahl index, which is the (weighted) average of the group shares $\sum_{j=1}^{M}\left(\frac{n_j}{N}\right)^2$ would measure concentration, which is inversely related to competition. One minus the Herfindahl index yields a well known measure of fractionalization, which would be associated with greater inter-group competition and in-group identity in the county as a whole.

Recent empirical evidence from diverse settings is consistent with the hypothesis that in-group identity is negatively correlated with a group's share of the population. Bisin, Topa, and Verdier (2004) use the General Social Survey to show that parents belonging to religious minority communities in a geographical area put more effort into socializing their children. Bisin et al. (2008) use data from the Fourth National Survey of Ethnic Minorities (FESNEM) in the United Kingdom, which contains detailed information on ethnic identity, to document the same negative correlation between identity and the share of the respondents'

[3] Tullock's specification is identical to the equation above except that the number of individuals is replaced by the investment in rent seeking.

ethnic group in the local population. Abramitzky, Einav and Rigbi (forthcoming) show that Jewish-related expenditure on Hanukkah is higher in U.S. counties with a lower share of Jews. Finally, Fryer (forthcoming), using data from the National Longitudinal Study of Adolescent Health (Add Health) finds that black and Hispanic students in schools that are less than 20 percent black have the strongest sense of minority identity, measured by the disapproval of behaviors associated with 'Acting White' and academic achievement. In a related analysis, blacks in less segregated schools incur a greater tradeoff between peer approval and academic achievement. Identity may have emerged in minority groups in these applications to prevent acculturation, whereas competition *between* migrant (minority) groups may have been more salient in our historical Midwestern setting. The important message from these previous studies is that in-group identity can emerge endogenously in response to external circumstances, as we assume in our analysis.

2.3 Measuring Initial Identity

The preceding description of the development of the Midwest establishes a point of departure – around 1860 – for the empirical analysis. Based on the discussion above, identity as we have defined it is more likely to have emerged in this initial period in counties where the labor market was ethnically fractionalized. To measure fractionalization in 1860, we need information on ethnic shares in the occupation categories within which labor market competition occurred. Individual-level data, including characteristics such as age, sex, occupation, and country of birth, are publicly available from the Population Census each decade from 1850 to 1930. Using the 1-in-100 sample from the 1860 IPUMS we can compute ethnic fractionalization within broad occupational categories in each Midwestern county.

Table 1 describes the overall ethnic shares upon which the fractionalization measure is based in 1860, 1880, and 1900. The English (13 percent), the Irish (25 percent), and the Germans (32 percent), dominated the migrant population in the Midwest in 1860, when initial conditions were being established, with no other ethnic group accounting for more than a 3 percent share of the migrants in that year. Subsequently, the English and the Irish were displaced by the Germans and the Scandinavians over the 1860-1900 period. Notice that the Italians, Poles, and Slavs continue to be insignificant in 1900, although they would display a substantial presence in Midwestern cities such as Cleveland, Chicago, and Pittsburgh by the first quarter of the twentieth century. Table 2 reports the occupational distribution of the migrants from the IPUMS sample in the same census years. Although agriculture was the dominant sector in this period, the share of farm employment declines from 62 percent in 1860 to 48 percent in 1900, with manufacturing operatives and laborers accounting for much of the increase in non-farm employment.

Figure 2 plots the fractionalization statistic based on the broad occupational categories in

Table 2. Counties that were not incorporated and those without foreign-born workers in 1860 are unshaded in the Figure. The implicit assumption when constructing the fractionalization statistic is that individuals can switch occupations within but not across categories and that migrant workers are restricted to a distinct segment of the labor market. The weighted average of fractionalization within each occupational category, where the weight is measured by the share of migrants in that category, then provides us with an overall measure of ethnic fractionalization in the county. For example, suppose that there are two occupations and two ethnicities in the county, with complete occupational segregation along ethnic lines. Our measure of fractionalization will be zero in this case, correctly reflecting the absence of ethnic competition in this labor market. In contrast, if the two ethnicities are of equal size and evenly distributed across the two occupations, our measure of fractionalization will increase to 0.5. The fractionalization statistic that we will use in the empirical analysis has a mean of 0.5 and a standard deviation of 0.2 across the seven Midwest states in Figure 2.

While there may be a connection between historical fractionalization and initial identity in theory, is there sufficient variation in 1860 fractionalization to predict outcomes many years in the future? To answer this question we regressed church participation at the county-level from 1870 to 2000, available at roughly ten-year intervals from the CRB, on 1860 fractionalization and a parsimonious set of variables meant to capture relevant economic conditions in 1860 (discussed below). The coefficient on 1860 fractionalization, together with the 95 percent confidence band, from these regressions is plotted in each year in Figure 3. This coefficient is less precisely estimated in the early census years, but is positive and significant by 1890. Subsequently, it grows steadily larger, while remaining statistically significant, all the way through to 2000. The 1860 fractionalization statistic is based on migrant workers who comprised less than 15 percent of the population in that year. Nevertheless, we see that historical fractionalization has a strong and persistent effect on an outcome – church participation – that is closely associated with identity in our analysis.

3 The Model

The model developed in this section uses identity, defined as the attachment to the birth community, to (i) explain the persistent differences in church participation observed in Figure 3, and (ii) generate predictions for variation in occupational choice across spatial communities that are otherwise indistinguishable.

3.1 Individuals and Communities

There are two types of communities, type-1 communities in which a fraction λ_t of the population is instilled with identity in period t, and type-2 communities without identity. In

our Midwestern context, the probability that a community is a type-1 community would be increasing in historical fractionalization. There are π type-1 communities and $1-\pi$ type-2 communities in this economy, with a continuum of individuals with measure one in each community. Each individual lives for two periods and has a single child at the beginning of the second period, preserving the size of the population over time. Each individual is endowed with ability $\omega \sim U[0,1]$. Ability does not vary systematically across communities and within families is uncorrelated across generations.

Each individual must make two choices in her lifetime. At the beginning of the second period she must decide whether or not to participate in the single collective institution, such as the church, that is active in her community. The cost of participation $c_p\omega$ is increasing linearly in the individual's ability. This could be because more able (intelligent) individuals are less sociable, or because the opportunity cost of participation is higher for them. Larger collective institutions provide more services for their members. The direct benefit from participation for any individual in period t, $\theta + g\underline{\omega}_t$, is thus assumed to be increasing linearly in the measure of the community $\underline{\omega}_t$ that chooses to participate. In addition, the collective institution serves as a vehicle through which individuals are socialized. The child of any individual who participates in this institution in type-1 communities will develop identity with probability $\tau \in (0,1]$. As in Bisin and Verdier (2000) and Tabellini (2008) we assume that an individual with identity would like her child to share these values. She thus derives additional utility U_I from participation.[4]

Individuals must also make career choices at the beginning of the second period of their lives. Two types of jobs are available in this economy: skilled jobs and unskilled jobs. Individuals who select into the skilled occupation must invest in human capital. The cost of this investment, $c_e(1-\omega)$, declines linearly in ability and is incurred instantaneously before commencing work. Apart from training costs and wages, the two types of occupations are distinguished by the level of spatial mobility. There is an idiosyncratic aspect to skilled jobs, which implies that skilled workers must move to specific locations to fill match-specific positions. For example, university professors or management consultants move throughout their careers, across local and regional labor markets, as their ability is revealed to the market and new opportunities arise. In contrast, medical technicians or janitors are less likely to move to further their careers (conditional on remaining in the same occupation). Skilled workers therefore move with probability P_S during the second period of their lives in the model, whereas unskilled workers move with probability $P_{NS} < P_S$. We will verify

[4]Bisin and Verdier assume that τ is a choice variable, determined at the level of the family in each generation. We assume instead that the decision to invest in the activities and the social interactions that are needed for identity formation is made at the level of the community in the initial period (outside the model). Once this decision is made, identity is transmitted with fixed probability τ in all the generations that follow in type-1 communities.

below that occupations with higher levels of college completion are indeed associated with greater migration.[5] Individuals with identity suffer disutility when they leave their birth community, and we will see momentarily that this affects the allocation of individuals to jobs in equilibrium.

3.2 Labor Market Equilibrium

While we allow for uncertainty in job location at the individual level, there is no uncertainty at the level of the local labor market in this economy. There are ω_{NS} unskilled jobs and $1-\omega_{NS}$ skilled jobs to be filled in each spatial community in each period. Because the cost of investing in human capital is declining in ability, there is an ability threshold above which individuals select into the skilled occupation in each community. Let this threshold be $\underline{\omega}_t^e$ in type-2 communities and $\underline{\omega}_t^{e,I}$, $\underline{\omega}_t^{e,NI}$ for individuals with and without identity in type-1 communities. These thresholds will be set so that the supply of skilled and unskilled labor in the economy matches the demand in each period. The market clearing condition for unskilled labor in period t is

$$\pi\left[\lambda_t\underline{\omega}_t^{e,I}+(1-\lambda_t)\underline{\omega}_t^{e,NI}\right]+(1-\pi)\underline{\omega}_t^{e}=\omega_{NS}.$$

The supply and the demand for skilled labor will match by construction if this condition is satisfied.[6]

To derive the allocation of individuals to jobs across communities, we assume that the labor market functions competitively in this economy. Aggregate labor demand and, hence, total output is constant in each period. The First Welfare Theorem can thus be applied to derive the ability threshold in each type of community, without characterizing individual decisions or solving explicitly for wages, as the solution to the Central Planner's Problem of minimizing moving costs and training costs (for skilled workers). The solution to this problem yields our first result:

Proposition 1. *A smaller share of the population born in type-1 communities selects into the skilled occupation and, therefore, migrates in each period despite the fact that the availability of skilled jobs does not vary across communities.*

[5] These results are obtained in 2000 and it is possible that the correlation between skill and mobility was reversed (negative) in the nineteenth century when new areas in the United States were developing. Because the labor market equilibrium is derived independently in each period in the model, this does not affect its predictions for current occupational choice.

[6] Although the market-clearing condition ensures that the supply matches the demand for each type of worker in this economy, the Central Planner must in addition sort the migrants so that ω_{NS} unskilled workers and $(1-\omega_{NS})$ skilled workers end up *ex post* in each community. We will see below that the supply of unskilled (skilled) workers exceeds the demand *ex ante* in type-1 (type-2) communities. To sort completely across all communities, the supply net of migration must therefore not exceed the demand: (i) $(1-P_{NS})[\lambda_t\underline{\omega}_t^{e,I}+(1-\lambda_t)\underline{\omega}_t^{e,NI}]\leq\omega_{NS}$ (ii) $(1-P_S)(1-\underline{\omega}_t^e)\leq(1-\omega_{NS})$. These conditions will be satisfied if P_S, P_{NS} are sufficiently large.

Let c_I be the cost of moving for individuals with identity. Normalize so that the cost of moving is zero for individuals without identity. Total costs are then described by the expression

$$c_I \pi \lambda_t \left[P_{NS} \underline{\omega}_t^{e,I} + P_S (1 - \underline{\omega}_t^{e,I}) \right] + \frac{c_e}{2} \pi \left[\lambda_t (1 - \underline{\omega}_t^{e,I})^2 + (1 - \lambda_t)(1 - \underline{\omega}_t^{e,NI})^2 \right] + \frac{c_e}{2} (1-\pi)(1-\underline{\omega}_t^{e})^2.$$

The first term in the expression above represents the cost of moving for individuals with identity in type-1 communities. The second and third terms represent training costs in type-1 and type-2 communities, respectively. Minimizing this expression with respect to $\underline{\omega}_t^{e,I}$, $\underline{\omega}_t^{e,NI}$, subject to the market-clearing condition (which effectively reduces the problem to two dimensions) we obtain,

$$\underline{\omega}_t^{e,I} = \underline{\omega}_t^{e} + \frac{c_I}{c_e}(P_S - P_{NS})$$
$$\underline{\omega}_t^{e,NI} = \underline{\omega}_t^{e}.$$

From the expressions above, the measure of skilled workers supplied by type-1 communities $\lambda_t(1 - \underline{\omega}_t^{e,I}) + (1 - \lambda_t)(1 - \underline{\omega}_t^{e,NI})$ will be strictly smaller than the measure of skilled workers supplied by type-2 communities $(1 - \underline{\omega}_t^{e})$ if $\lambda_t > 0$. We will show below that the fraction of the population instilled with identity in type-1 communities λ_t is indeed positive in all periods to complete the proof.

3.3 Institutional Participation

The cost of participation in the collective institution is declining in ability and so there is an ability threshold below which individuals choose to participate. There is a strategic element to the individual's participation decision since it depends on the participation decisions of other individuals in the community. The threshold is thus the solution to a fixed point problem, which we solve below. Let this threshold be $\underline{\omega}_t^{p}$ in type-2 communities and $\underline{\omega}_t^{p,I}$, $\underline{\omega}_t^{p,NI}$ for individuals with and without identity in type-1 communities. We make two assumptions to simplify the analysis that follows:

A1. First generation migrants, who arrive in a community as a consequence of job turnover, do not participate in the collective institution. This assumption emphasizes the cost to a community from migration.

A2. The *participation thresholds* $\underline{\omega}_t^{p}$, $\underline{\omega}_t^{p,I}$, $\underline{\omega}_t^{p,NI}$ lie to the left of the corresponding *job thresholds* $\underline{\omega}_t^{e}$, $\underline{\omega}_t^{e,I}$, $\underline{\omega}_t^{e,NI}$ in each period. This implies that only unskilled workers participate in the collective institution, allowing us to solve for the labor market equilibrium and the participation equilibrium independently. We will discuss the consequences of relaxing this assumption below.

Given these assumptions, it is straightforward to derive the participation threshold in each type of community at each point in time. Starting with type-2 communities, the

participation threshold $\underline{\omega}_t^p$ must satisfy the following equality,

$$\theta + g(1 - P_{NS})\underline{\omega}_t^p - c_p\underline{\omega}_t^p = 0. \tag{1}$$

The left hand side of equation (1) describes the utility from participation for the marginal individual with ability $\underline{\omega}_t^p$ who is indifferent between participating and not participating. For such indifference to be obtained, the following condition must be satisfied:

Condition 1. $c_p - g(1 - P_{NS}) > 0$.

The $(1 - P_{NS})$ multiplier in equation (1) reflects the fact, from assumptions A1 and A2, that only individuals who selected into the unskilled occupation and remain in their birth community participate in the collective institution. Collecting terms, the participation threshold in type-2 communities $\underline{\omega}_t^p$ is constant over time,

$$\underline{\omega}_t = \frac{\theta}{c_p - g(1 - P_{NS})}. \tag{2}$$

The corresponding thresholds for individuals without and with identity in type-1 communities, $\underline{\omega}_t^{p,NI}$, $\underline{\omega}_t^{p,I}$ must satisfy the following equalities,

$$\theta + g(1 - P_{NS})\left[\lambda_t\underline{\omega}_t^{p,I} + (1 - \lambda_t)\underline{\omega}_t^{p,NI}\right] - c_p\underline{\omega}_t^{p,NI} = 0 \tag{3}$$

$$U_I + \theta + g(1 - P_{NS})\left[\lambda_t\underline{\omega}_t^{p,I} + (1 - \lambda_t)\underline{\omega}_t^{p,NI}\right] - c_p\underline{\omega}_t^{p,I} = 0. \tag{4}$$

The thresholds $\underline{\omega}_t^{p,NI}$, $\underline{\omega}_t^{p,I}$ only differ because individuals instilled with identity derive additional utility U_I from participation, since their children will then share their values with probability τ. Differencing equation (3) from equation (4), and then substituting the resulting expression for $\underline{\omega}_t^{p,I}$ in equation (3),

$$\underline{\omega}_t^{p,I} = \underline{\omega}_t^{p,NI} + \frac{U_I}{c} \tag{5}$$

$$\underline{\omega}_t^{p,NI} = \frac{\theta}{c_p - g(1 - P_{NS})} + \left(\frac{U_I}{c_p}\right)\frac{g(1 - P_{NS})}{c_p - g(1 - P_{NS})}\lambda_t. \tag{6}$$

The thresholds derived above fully characterize participation decisions in type-1 and type-2 communities.

3.4 Identity and Participation Dynamics

Ability is uncorrelated across generations within the family and the benefit derived from the collective institution is determined by contemporaneous participation decisions. What links one generation to the next in a community is identity and the utility that individuals with identity derive when their children share their values. Because children of participating

individuals in type-1 communities develop identity with probability τ, identity evolves in those communities according to the following law of motion:

$$\lambda_{t+1} = \tau(1 - P_{NS}) \left[\lambda_t \underline{\omega}_t^{p,I} + (1 - \lambda_t) \underline{\omega}_t^{p,NI} \right]. \tag{7}$$

The $(1 - P_{NS})$ multiplier in equation (7) reflects the fact that participation is restricted to individuals who remain int their birth community and select into the unskilled occupation (from assumptions A1 and A2). The result that follows describes how the evolution of identity, as derived above, shapes participation in type-1 and type-2 communities.

Proposition 2. *There exists a $\overline{\tau}$ and a $\overline{\lambda}$, such that for $\tau < \overline{\tau}$ and $\lambda_0 < \overline{\lambda}$, (a) the share of the population instilled with identity in type-1 communities grows monotonically over time until it converges to its steady-state level. (b) Participation in the collective institution is always greater in type-1 communities and the gap between type-1 and type-2 communities widens over time until the steady-state is reached.*

Substituting the expression for $\underline{\omega}_t^{p,I}$ from equation (5), and subsequently the expression for $\underline{\omega}_t^{p,NI}$ from equation (6), in equation (7),

$$\lambda_{t+1} = \tau(1 - P_{NS}) \left[\frac{U_I}{c_p - g(1 - P_{NS})} \lambda_t + \frac{\theta}{c_p - g(1 - P_{NS})} \right]. \tag{8}$$

Equation (8) describes a first-order linear one-dimensional dynamical system. The properties of such a system are well known and the system will converge to a unique steady-state equilibrium $\overline{\lambda}$, regardless of the initial condition λ_0, if Condition 1 and the following condition are satisfied:

$$\textbf{Condition 2}. \ \frac{U_I \tau (1 - P_{NS})}{c_p - g(1 - P_{NS})} < 1$$

Condition 2 will be satisfied if τ is less than a critical $\overline{\tau}$, given the other parameter values, as stated in the Proposition. Intuitively, a smaller transmission probability τ slows down the diffusion of identity through the population, ensuring that the dynamical system is well behaved. For the steady-state $\overline{\lambda}$ to lie in the unit interval, we need a stronger condition, which is derived by setting $\lambda_t = \lambda_{t+1} = \overline{\lambda}$ in equation (8):

$$\textbf{Condition 3}. \ \frac{(U_I + \theta) \tau (1 - P_{NS})}{c_p - g(1 - P_{NS})} < 1$$

Starting with a small λ_0 also ensures that identity will increase over time. If $\lambda_0 < \overline{\lambda}$, as specified in Proposition 2, then it follows from equation (8) that the state variable λ_t will increase monotonically over time as it converges to the steady-state $\overline{\lambda}$ to complete part (a) of Proposition 2. This also implies that λ_t is positive at each point in time to complete the proof of Proposition 1.

Having established that λ_t increases monotonically over time before converging to its steady-state level if τ and λ_0 are sufficiently small, the next step is to compare participation levels in type-1 and type-2 communities. $\underline{\omega}_t^{p,I} > \underline{\omega}_t^{p,NI}$ from equation (5). Because $\lambda_t > 0$, $\underline{\omega}_t^{p,NI} > \underline{\omega}_t^{p}$ from equation (2) and equation (6). Participation in type-1 communities $(1 - P_{NS}) \left[\lambda_t \underline{\omega}_t^{p,I} + (1 - \lambda_t)\omega_t^{p,NI}\right]$ is the weighted average of two terms that are both larger than participation in type-2 communities $(1 - P_{NS})\underline{\omega}_t^{p}$. It follows that participation in type-1 communities will be greater than participation in type-2 communities at each point in time. This difference in participation is effectively driven by the U_I term in equation (4), which induces greater participation by individuals with identity and, by extension, individuals without identity in type-1 communities. Moreover, while participation in type-2 communities is constant over time from equation (2), participation in type-1 communities increases monotonically until λ_t converge to its steady-state level. Returning to the participation expression, there are two reasons for this increase over time. First, $\underline{\omega}_t^{p,NI}$ is increasing in λ_t from equation (6), which implies from equation (5) that $\underline{\omega}_t^{p,I}$ is increasing in λ_t as well. It follows that both $\underline{\omega}_t^{p,NI}$ and $\underline{\omega}_t^{p,I}$ will be increasing over time from part (a) of the Proposition. Second, the weight λ_t on $\underline{\omega}_t^{p,I}$ is increasing over time, whereas the weight $(1 - \lambda_t)$ on the smaller term $\underline{\omega}_t^{p,NI}$ is decreasing over time. The gap in participation between type-1 and type-2 communities will thus widen over time until the dynamical system reaches the steady-state to complete the proof of Proposition 2.

3.5 Persistence without Identity

Persistent differences across communities in the supply of skilled labor (Proposition 1) and institutional participation (Proposition 2) are driven by identity in our model. Could these differences be generated by differences in institutional quality alone? If assumption A2 is satisfied, then any change in institutional participation (quality) is inframarginal and will have no effect on the Central Planner's labor allocation decision. We thus proceed to relax assumption A2 by allowing the participation threshold to lie to the right of the job threshold. Without identity, institutional participation in each type of community, $c = \{1, 2\}$, will then satisfy the condition,

$$\theta + g\left[(1 - P_{NS})\underline{\omega}_t^{e,c} + (1 - P_S)(\underline{\omega}_t^{p,c} - \underline{\omega}_t^{e,c})\right] - c_p \underline{\omega}_t^{p,c} = 0.$$

Collecting terms and simplifying,

$$\underline{\omega}_t^{p,c} = \frac{\theta + g(P_S - P_{NS})\underline{\omega}_t^{e,c}}{c_p - g(1 - P_S)} \equiv \alpha + \beta \underline{\omega}_t^{e,c}.$$

The overall benefit to the community from participating in the collective institution is given by the expression,

$$B_c = \left(\theta + g\left[(1 - P_{NS})\underline{\omega}_t^{e,c} + (1 - P_S)(\underline{\omega}_t^{p,c} - \underline{\omega}_t^{e,c})\right] - c_p \frac{\underline{\omega}_t^{p,c}}{2}\right) \underline{\omega}_t^{p,c},$$

which from the first equation above, simplifies to

$$B_c = \frac{c_p}{2}\left(\underline{\omega}_t^{p,c}\right)^2.$$

B_c is a function of $\underline{\omega}_t^{p,c}$, which in turn is a function of $\underline{\omega}_t^{e,c}$ and so the Central Planner's labor allocation decision is no longer independent of institutional participation. Without identity, moving costs are the same in all communities. Normalizing so that moving costs are zero and accounting for individual participation decisions as derived above, the Central Planner will set $\underline{\omega}_t^{e,1}$, $\underline{\omega}_t^{e,2}$ to maximize

$$\pi\left[\frac{c_p}{2}\left(\underline{\omega}_t^{p,1}\right)^2 - \frac{c_e}{2}\left(1-\underline{\omega}_t^{e,1}\right)^2\right] + (1-\pi)\left[\frac{c_p}{2}\left(\underline{\omega}_t^{p,2}\right)^2 - \frac{c_e}{2}\left(1-\underline{\omega}_t^{e,2}\right)^2\right]$$

subject to the market-clearing condition $\pi\underline{\omega}_t^{e,1} + (1-\pi)\underline{\omega}_t^{e,2} = \omega_{NS}$. Making use of the linear relationship between $\underline{\omega}_t^{p,c}$ and $\underline{\omega}_t^{e,c}$ derived above, the solution to this problem is $\underline{\omega}_t^{e,1} = \underline{\omega}_t^{e,2}$, which implies in turn that $\underline{\omega}_t^{p,1} = \underline{\omega}_t^{p,2}$. Relaxing assumption A2 does not, by itself, generate differences in institutional participation and the labor allocation across communities. Once we depart from linearity and introduce increasing returns to scale in the benefits that the institution provides, however, the relationship between $\underline{\omega}_t^{p,c}$ and $\underline{\omega}_t^{e,c}$ will no longer be linear and differences across communities will emerge in equilibrium. The Central Planner can now increase the overall benefit from institutional participation by shifting the job threshold to the right in some communities, say the type-1 communities, with an accompanying shift in the threshold to the left in type-2 communities (to clear the market). This increase in benefits will be partially offset by the accompanying increase in overall training costs as the labor allocation moves away from $\underline{\omega}_t^{e,1} = \underline{\omega}_t^{e,2}$, but it is straightforward to verify that job thresholds and participation thresholds will not be the same across communities in equilibrium, providing an alternative explanation for Propositions 1 and 2.[7]

While increasing returns to scale can explain differences in the supply of skilled labor and institutional participation across communities in principle, this assumption is difficult to justify theoretically or empirically in the Midwest. Individual inputs in the church, such as volunteering and making charitable contributions, are substitutes rather than complements. With the recent exception of mega-churches, which are relatively rare in the Midwest, there is little evidence of the market concentration that would accompany increasing returns to scale in the church's production function either. By the same argument, we would not expect to observe increasing returns to scale in supporting institutions such as parochial schools or in

[7]Without accounting for participation in the collective institution and setting moving costs to be the same across communities, it is straightforward to verify that $\underline{\omega}_t^{e,1} = \underline{\omega}_t^{e,2}$ is the optimal allocation. Once we add institutional participation to the Central Planner's objective function, marginal shifts away from this allocation will have a first-order effect on the benefits from participation (with increasing returns to scale) but only a second-order effect on training costs.

the labor-market networks that were historically connected to the church.[8] Could persistent differences be generated by relaxing assumption A1 instead? Let the benefit provided by the collective institution to *all* residents be B_1 in type-1 communities and $B_2 < B_1$ in type-2 communities. It is straightforward to verify that these differences in institutional quality will have no effect on the solution to the Central Planner's problem. The labor allocation decision will only be affected if individuals place a premium on participating in the specific institution located in their hometown, which is the case with identity.

4 Empirical Analysis

Based on the discussion in Section 2, fractionalization in 1860 can be interpreted as being positively associated with the probability that a county is a type-1 community. This allows Propositions 1 and 2 and the assumptions that are needed to test the theory to be restated in terms of 1860 fractionalization. In particular, we will estimate the effect of 1860 fractionalization on (i) participation over time in specific socializing institutions associated with the transmission of identity, and (ii) occupational choice and mobility many years in the future. We will also provide support for the assumptions that 1860 fractionalization is uncorrelated with standard determinants of occupational choice today such as the availability of skilled jobs, the ability distribution in the population, and school quality. The analysis concludes by assessing the robustness of the main results.

4.1 Historical Fractionalization and Current County Characteristics

A suitable research setting for our analysis must satisfy two conditions: (i) there must be sufficient variation in fractionalization across spatial communities or counties in the initial period when the economy in this setting is first developing and institutions are being established, and (ii) historical fractionalization must be uncorrelated with standard determinants of occupational choice today. The preliminary results presented in Figure 3 indicate that our measure of fractionalization satisfies the first condition. We now proceed to provide support for the second condition. More stringent tests will be discussed in Sections 4.3 and 4.4.

Labor market networks organized around the European-origin countries in Table 2 are no longer salient in the American Midwest (Gans 1979, Alba 1990). However, fractionalization in 1860 could have been correlated with other features of the economy at the time of initial

[8] Returning to the linear technology, an alternative mechanism to generate persistent differences without identity would relax the assumption that θ is the same across communities. If historically fractionalized communities, which are type-1 communities in the model, invested more heavily in fixed institutional characteristics that have a permanent effect on participation, then this would imply that $\theta_1 > \theta_2$. It is straightforward to verify from the equations above that this implies $\underline{\omega}_t^{e,1} < \underline{\omega}_t^{e,2}$, which is inconsistent with the empirical results.

development that had persistent effects. To explore this possibility, we must understand what determined fractionalization in the first place. In a rapidly expanding Midwest economy, some of the variation in fractionalization across counties was no doubt a consequence of accidental initial settlement by ethnic groups in particular locations, which fueled the arrival of more migrants as networks crystallized. At the same time, fractionalization would have been determined by the demand for labor, with more ethnic groups attracted to rapidly growing areas. We have already discussed the importance of transportation links in the development of the Midwest and Table 3 consequently investigates the effect of railroads and distance to canals and a Great Lakes harbor on fractionalization in 1860.[9] Counties with a railroad running through them and counties that are closer to a canal or harbor have significantly higher fractionalization in Table 3, Column 1 as well as a larger population in Column 4.

The results in Table 3 indicate that counties with superior transportation infrastructure, which were more populated and presumably growing more rapidly, were more fractionalized in 1860. The population of the county in 1860 could have determined subsequent agglomeration in economic activity, with long-term implications for the growth of the local economy and the demand for skilled labor. Recall, however, that the effect of fractionalization on identity derived in Section 2.2 is conditional on the number of workers, N, and the total surplus, R. All the empirical tests that follow in the paper will consequently estimate the effect of 1860 fractionalization *conditional* on N and R, although we verify that the main results go through unconditionally as well. The total surplus, R, will in general depend on the size of the local economy and the types of jobs that are available. The total number of migrant workers competing for this surplus, N, will be positively correlated with the total population. Using the county population in 1860 and the share of manufacturing and agriculture in that year to jointly measure N and R, we see in Table 4, Columns 1-3 that 1860 fractionalization (conditional on county characteristics in that year) is uncorrelated with characteristics of the economy in 1990, such as the share of agriculture, share of manufacturing, and the total population (as well as population density, not reported) that are associated with the availability of skilled jobs.

Given the high levels of mobility in the United States, few individuals living in the Midwest today could trace an unbroken line of descent to European ancestors arriving in the *same* county at the time of initial settlement. Not surprisingly, 1860 fractionalization is uncorrelated with (white) ethnic fractionalization in 1990 in Column 4. Even if 1860 fractionalization was correlated with the ability distribution among the initial migrants, it is

[9]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.

thus unlikely to be correlated with the current distribution of ability in the population and we will later observe that test scores are indeed uncorrelated with 1860 fractionalization. However, we must also consider the possibility that 1860 fractionalization could be correlated with school quality today. Although conflict theory predicts that fractionalization will strengthen in-group identity, it also predicts that diversity will increase out-group prejudice. The negative effects of fractionalization on inter-group cooperation and, hence, public good provision have been extensively documented. Alesina, Baqir, and Easterly (1999), for example, document a negative and significant relationship between racial fractionalization and the allocation of resources to local public goods, including education, in the United States.

An important advantage of measuring fractionalization as early as 1860 is that public infrastructure was relatively rudimentary at that time, limiting the negative consequences of inter-group competition. Depending on subsequent migration, ethnic fractionalization in 1860 could nevertheless be correlated with ethnic or racial fractionalization later in the twentieth century when it was relevant for public good provision. Reassuringly, 1860 fractionalization is uncorrelated with racial fractionalization in 1990 in Column 5 and we will later observe that 1860 fractionalization is also uncorrelated with education expenditures per student (and total public expenditures per capita). Notice, in contrast, that more ethnically fractionalized counties in 1860 have significantly *lower* religious fractionalization in 1990 in Column 6.[10] This intriguing result, which stands conspicuously apart from the other results reported in Table 4 serves as a useful preview for the analysis of church participation that follows.

4.2 Church Participation

Proposition 2 states that 1860 fractionalization should have a positive effect on participation in the collective institution, with this effect growing stronger over time. Although the model assumes that a single institution is available to transmit identity from period-0 onwards, multiple socializing institutions co-exist in practice. We focus on the church in our analysis because it was among the first institutions to be established when immigrants arrived in an area (Hoerder 1991, Bodnar, Simon, and Weber 1982). "[Migrants] from varying regions formed a community based on common nationality and religion centered on the central

[10]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. 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.

cultural institution – the church" (Gjerde 1991: 176). This institution would thus have been well situated to instill identity that discouraged migration by its members, where required, when communities in the Midwest were first being established.

Historically, the church congregation provided many forms of mutual assistance including credit, insurance, job referrals, business information, and social support (Gjerde 1985, Alexander 1991). Indeed, it has been argued that early immigrants to the Midwest participated in church communities to benefit from the economic and social services they provided, instead of being drawn to the church by a particular belief or ideology (Bodnar 1985). The Midwestern church continues to provide important forms of social support for its members. Church activities include 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, which bring families and friends together on a regular basis (Elder and Conger 2000). Although other institutions, such as parochial schools, will also play a role in building an attachment to the hometown, we focus on the church in the analysis that follows because of its central position in community life and because church participation data are available at the county level, by denomination, from the Census of Religious Bodies (CRB) at roughly ten-year intervals from period 0 (1860) till the present (2000).[11] Additional supporting results based on enrollment in parochial schools will be provided in Section 4.3.

Based on the CRB data, 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 roughly 80 percent of church participants over the 1860-2000 period. Among these denominations, the Catholics and Lutherans grew substantially in popularity, 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.

Given our focus on historical competition between migrant groups, we expect churches that were dominated by migrants to have played an especially important role in creating identity, as we have defined it. Based on the country of origin of the incoming migrants, reported in Table 1, most of the migrant churches would have been Catholic or Lutheran.[12]

[11]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.

[12]Apart from the Germans and the Irish, the English were also an important migrant group in 1860.

Regressing the population-share of different denominations in 1860, obtained from the CRB, on the share of migrants in that year, obtained from the population census, counties with a greater share of migrants are indeed disproportionately Catholic and Lutheran. Although we have shown in Figure 3 that 1860 fractionalization has a positive effect on overall church participation, a stronger test of Proposition 2 is that its predictions should apply to the migrant denominations alone. While the pool of potential members in those denominations would have been restricted to particular ethnicities to begin with, this constraint would have been subsequently relaxed, allowing identity to propagate through the population. Higher 1860 fractionalization should thus be associated with a greater population share of Catholics and Lutherans and this effect should grow stronger over time. In contrast, the model has no prediction for the relationship between historical fractionalization and participation in other denominations over time, providing us with a useful falsification test.[13] We consequently proceed to estimate two separate regressions in each census year; the first regression has the share of Lutherans and Catholics in the population as the dependent variable and the second regression has the share of all other denominations as the dependent variable.

Using the same regression specification as in Figure 3, with county population, agriculture share, and manufacturing share in 1860 as additional regressors to account for the total surplus (R) and the relevant labor force (N) in the initial period, the 1860 fractionalization coefficient, with the corresponding 95 percent confidence band, is reported in each census year in Figure 4. As predicted, the effect of 1860 fractionalization on the share of Catholics and Lutherans in the population is positive and significant, and gets steadily larger over time.[14] Although we do not report results separately by denomination, this pattern is obtained for both the Catholics and the Lutherans. A one standard deviation increase in 1860 fractionalization would increase the population share of Catholics and Lutherans in the county by four percentage points (22 percent) in 2000. In contrast, the fractionalization coefficient is much smaller and *negative*, but remains significant and stable over time in the companion regressions for the other denominations.[15] These results, taken together, provide

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.

[13]The direct effect of stronger Catholic and Lutheran churches is to lower participation in other denominations. However, the endogenous response by other denominations and the accompanying increase in religious competition could at the same time increase overall church participation (Finke and Stark 1992, Gruber 2005). The net effect on participation in other denominations is consequently ambiguous.

[14]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 fractionalization 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.

[15]We do not report the fractionalization coefficient in 1860 in Figures 3 and 4 for expositional convenience.

strong support for the model and an explanation for the negative correlation between ethnic fractionalization in 1860 and religious fractionalization in 1990 that we reported in Table 4.

The CRB data only allow us to measure church participation among the current *residents* of the county, whereas the theory is based on the assumption that individuals *born* in type-1 communities are more likely to be exposed to socializing institutions in childhood. We consequently turn to the National Longitudinal Survey of Youth (NLSY), which consists of a nationally representative sample of American high school seniors in 1979 who were subsequently interviewed annually from 1979 to 1994 and biennially thereafter. The NLSY collects information on the respondents' county of birth and the religious denomination they were raised in.

The dependent variable in Table 5, Column 1 is the share of Catholics and Lutherans in the county, while the dependent variable in Column 2 is the share of all other denominations, using CRB data in 2000. The dependent variable in Column 3 is the fraction of individuals *born* in a given county who were raised as Catholics or Lutherans, using NLSY data. And the dependent variable in Column 4 is the corresponding fraction raised in all other denominations. 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 1860 fractionalization, manufacturing share, agriculture share, and population in that county. Matching the patterns in Columns 1-2, and in Figure 4 over many census years, we see that individuals born in historically fractionalized counties are significantly more likely to be raised as Lutherans or Catholics in Column 3, whereas the coefficient on 1860 fractionalization in Column 4 is negative (but insignificant).[16]

4.3 Occupational Choice and Mobility

Proposition 1 states that individuals born in historically fractionalized counties should be more likely to select into low-skill jobs associated with low mobility. The NLSY is uniquely suited to test this prediction because the county of birth is available for each individual in the nationally representative sample of high school seniors, as well as residential location and the type of job (where relevant) from 1979 onwards. We will study the (conditional) effect of 1860 fractionalization on migration and the type of job in 2000, when respondents

This coefficient for the non-migrant denominations is substantially more negative than what we see for the other years in Figure 4 and its inclusion would obscure changes in the fractionalization coefficient over time that we observe in the Figures. The coefficient in 1860 for the migrant denominations is negative, but is small in magnitude and statistically insignificant.

[16]Our explanation for the pattern observed in Figure 4 and Table 5 is that identity is propagating through the population over time in historically fractionalized counties, increasing participation in socializing institutions. An alternative explanation is that Catholics and Lutherans have been migrating into historically fractionalized counties over time. Without a role for identity, such selective migration even if it was a response to superior local institutions, would not explain the mismatch between the supply and the availability of skilled jobs in historically fractionalized counties, as discussed in Section 3 and shown below.

were old enough to be settled in their careers and to have made job related moves.

Including 1860 manufacturing share, agriculture share, and population to account for the total surplus (R) and the labor force (N) in the initial period, as well as the individual's race, gender, and age as regressors, we see in Table 6, Column 1 that individuals born in historically fractionalized counties are significantly less likely to have migrated from their county of birth by 2000. On average, 58 percent of the individuals in the sample migrate from the county of birth. The 1860 fractionalization coefficient in Column 1 indicates that a one standard deviation increase in fractionalization would reduce migration by 8 percentage points (14 percent).

To test the accompanying prediction that individuals born in historically fractionalized counties will select into less mobile low-skill jobs, we compute the average migration rate within each occupation in the NLSY. The rate corresponding to the individual's occupation in 2000 (conditional on being employed) is then specified as the dependent variable in Column 2. Using the same set of regressors as in Column 1, we see that the 1860 fractionalization coefficient is indeed negative and significant, as predicted by the theory. In a related exercise, the dependent variable is measured by average college completion in the individual's occupation in Column 3. Migration and college completion are strongly correlated across occupations (the correlation is 0.6) as assumed in the model. Not surprisingly, the 1860 fractionalization coefficient is negative and significant in Column 3 as well.

The results in Table 6, Columns 1-3 are consistent with the proposition that identity, as we have defined it, reduces the individual's propensity to migrate, with implications for occupational choice. We complete the analysis in this section by testing the assumption that the availability of skilled jobs does not vary across types of communities.[17] If this assumption is satisfied, then individuals residing in historically fractionalized counties should hold the same type of jobs as individuals residing in historically less fractionalized counties once the market clears. The dependent variable in Table 6, Column 4 is the average migration rate in the individual's occupation, as in Column 2. However, the 1860 variables are now measured in the individual's county of *residence* rather than in the county of *birth*. Reassuringly, the 1860 fractionalization coefficient in Column 4 is much smaller in magnitude than the corresponding coefficient in Column 2 and statistically indistinguishable from zero. When the individual's occupation is measured by average college completion, the 1860 fractionalization coefficient in Column 5 actually changes sign and is positive and significant (at the 10 percent level). Individuals residing (and working) in historically fractionalized counties are engaged in occupations that are at least as skilled and as mobile as the occupations available to

[17] If occupation-specific human capital is transmitted from the parent to the child, then training costs would be effectively lower for individuals born in counties with a greater availability of skilled jobs. The supply of skilled labor would then be greater in those counties, without a role for identity.

individuals in historically less fractionalized counties, consistent with the preliminary results reported in Table 4.

4.4 Human Capital

We have focussed on the church in the analysis of institutional participation because this institution has occupied a central position 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 from the theory is that individuals born in historically fractionalized counties are more likely to have been exposed to these socializing institutions in childhood as well. This section thus begins by estimating the (conditional) relationship between enrollment in parochial school and 1860 fractionalization.

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 (which are mostly Catholic and Lutheran in the Midwest). The dependent variable in Table 7, Column 1 is the share of students enrolled in parochial schools in 1970, while the dependent variable in Column 2 is the corresponding share in 1980. The regressors in both specifications are 1860 fractionalization, manufacturing share, agriculture share, and population. The 1860 fractionalization coefficient is positive and significant and nearly identical in Columns 1 and 2. The share of students in parochial school is 0.08 on average across our Midwestern counties. The estimates in Columns 1-2 indicate that a one standard deviation increase in 1860 fractionalization would increase that share by 0.014 (20 percent).

The 1990 census did not collect information on parochial schooling, but total private school enrollment is available in all three census rounds. Parochial schools account for the bulk of private schooling in the United States and, not surprisingly, the 1860 fractionalization coefficient continues to be positive and significant with the share of students in private school as the dependent variable (not reported). This coefficient is just slightly larger in magnitude than the point estimates in Columns 1-2 and remains very stable across the 1970-1990 census rounds. The results on parochial schooling, together with the results presented on church participation and denominational affiliation, indicate that individuals growing up in historically fractionalized counties are more likely to come in contact with socializing institutions. The consistency of the results on church participation, mobility, and parochial school enrollment, obtained with three independent data sets, also increases our confidence in the model and the interpretation of the results.

We complete this section by providing support for the important assumption in the model

that the cost of investing in human capital, reflecting ability in the population or school quality, should not be systematically higher in historically fractionalized counties.[18] 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 fractionalized counties are significantly more likely to be enrolled in parochial schools, this would suggest that they actually have access to schools of superior quality. One concern, however, is that public schools in those counties may consequently receive fewer resources, lowering overall school quality. To explore this possibility, we match information on education expenditures from the Annual Survey of Governments with total enrollment in public schools from the population census. The dependent variable in Table 7, Columns 3-4 is constructed as the ratio of education expenditures to public school students in each county in 1970 and 1980, respectively. Using the same set of regressors as in Columns 1-2, we see that 1860 fractionalization has no effect on public school expenditures. Results not reported indicate that 1860 fractionalization has no effect on total public expenditures per capita either.[19] Once again, we see that 1860 fractionalization has a significant effect on specific (socializing) institutions alone.

The results presented in Table 7 provide no evidence that school quality is lower in historically fractionalized counties. However, school quality is notoriously difficult to measure and we must also allow for the possibility that the ability distribution varies across counties. To provide direct support for the claim that individuals born in historically fractionalized counties are no less capable and no less prepared to invest in the human capital that is necessary to secure skilled jobs, we estimate the (conditional) relationship between measures of educational attainment obtained from the NLSY and 1860 fractionalization in Table 8. Conditional on 1860 manufacturing share, agriculture share, and population, as well as the individual's age, gender, and race, we see in Table 8, Columns 1-3 that 1860 fractionalization has no effect on AFQT scores, high school completion, and college completion.

Why are individuals born in historically fractionalized counties as likely to complete college, even though they end up in less mobile occupations associated with lower college completion? Apart from providing access to particular occupations, there are nonpecuniary benefits to college education. Given the results on parochial schooling in Table 7, one reason why individuals born in historically fractionalized counties may be overqualified is that school quality is higher in those counties, lowering the cost of investing further in human capital. Regardless of the explanation, going beyond the model and allowing for

[18] It is straightforward to verify from the model that type-1 communities would supply more unskilled workers, without a role for identity, if c_e was higher or ability was lower in those communities.

[19] Using the same data, Alesina, Baqir, and Easterly (1999) document a negative and significant relationship between contemporaneous racial fractionalization and the allocation of resources to certain public goods, particularly education, roads, and welfare. Our findings are not inconsistent with their results since we saw in Table 4 that 1860 fractionalization was uncorrelated with 1990 racial fractionalization.

multiple skilled jobs, we still expect college students from historically fractionalized counties to make choices that will restrict their future mobility. To test this prediction, we compute average migration within each college major listed in the NLSY. The average migration level corresponding to the individual's major is then specified as the dependent variable in Table 8, Column 4. Using the same set of regressors as in Columns 1-3, we see that individuals born in historically fractionalized counties select majors associated with significantly lower migration, as predicted, conditional on completing college.

4.5 Robustness Tests

We complete the empirical analysis by discussing a number of tests that were conducted to assess the robustness of the results. Each 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) as the dependent variable. To preserve space, we only report the 1860 fractionalization coefficient for each test in Table 9.

We begin in Panel A with alternative construction of the fractionalization variable. Columns 1-6 report estimates with fractionalization measured at other points in time, 1850 and 1870, during the period of initial development. Subsequently, we increase the accuracy of the fractionalization measure by excluding all counties with less than 20 observations in the 1860 IPUMS (approximately 10 percent of the sample) in Columns 7-9. Panel A concludes by computing 1860 fractionalization with men only in Columns 10-12.[20]

Panel B continues with alternative construction of the fractionalization variable. Farmers account for 50 percent of the 1860 workforce in Table 2. Ethnic competition may have been less relevant in this occupational category and so we compute the fractionalization statistic without farmers in Columns 1-3. The fractionalization statistic that we use in the regressions reported in the paper is computed as the weighted average of ethnic fractionalization within each of the 11 occupational categories listed in Table 2. Competition is assumed to occur within a coarser set of four occupational categories in Columns 4-6 and fractionalization is computed without regard to occupation in Columns 7-9.[21] The implicit assumption when computing the fractionalization statistic is that migrants compete with each other within

[20]Both men and women participated in the workforce in the nineteenth and early twentieth centuries, with ethnic networks channeling women into jobs as well (Bodnar 1980). We might nevertheless expect labor networks to have been organized along gender lines within ethnic groups, and this robustness test allows for the possibility that male networks occupied a dominant position in the labor market and the communities they were drawn from.

[21]The four aggregate categories are white collar, agriculture, manufacturing, and service and laborers. These categories correspond to the broad headings in Table 2 except that Blue collar, nonfarm is divided into manufacturing (craftsman, operative) and service and laborers (household service, service, laborer nonfarm).

a distinct segment of the labor market. The specification in Columns 10-12 relaxes this assumption by including native workers as a separate group. Reassuringly, the 1860 fractionalization coefficient declines substantially in (absolute) magnitude with each of the three outcomes and is no longer statistically significant with migration as the dependent variable.

Panel C, Columns 1-6 report results with alternative construction of the 1860 population variable. We include total population in 1860 in all the regressions reported in the paper to partially account for both the migrant workforce (N) as well as the total surplus that was available historically (R). The robustness tests replace total population with the migrant population and the total workforce in 1860, respectively. Table 9 concludes with specifications that include historical characteristics that could potentially determine the outcomes of interest and are correlated with 1860 fractionalization. The population share of each ethnicity in the 1860 census (computed at the county level) is included in Columns 7-9 to allow for the possibility that individual-group share rather than inter-group fractionalization determines our outcomes. To complete the panel, we allow for the possibility that the results are driven by historical church competition rather than historical labor market competition by including 1860 religious fractionalization as an additional regressor in Columns 10-12.

The results in Table 9 indicate that the 1860 fractionalization coefficient remains significant at the 5 percent level and is very stable across the alternative specifications, almost without exception. As a final robustness test, we report conditional and unconditional estimates of the nonparametric relationship between the three outcomes and 1860 fractionalization in Figure 6.[22] The individuals in the NLSY 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 fractionalization 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 we have uncovered.

5 Conclusion

This paper draws a connection between competing migrant networks in the Midwest when it was first developing, the in-group identity that emerged endogenously to support these networks in particular locations (counties), and institutional participation and occupational choice today. Individuals born in counties with greater ethnic fractionalization in 1860 are today – 150 years later – (i) significantly more likely to participate in socializing institutions

[22]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 5.

such as churches and parochial schools that are associated with the transmission of identity from one generation to the next, and (ii) significantly less likely to select into mobile skilled occupations.

Instead of focussing on differences in occupational choice and mobility across denominations or between church participants and non-participants, we study how historical circumstances, measured by ethnic fractionalization, shaped the orientation of churches *within* particular denominations dominated by the early migrants. The effect of historical fractionalization on participation in these denominations actually grows stronger over time, consistent with a model in which identity is propagating through the population in select counties over the generations and emphasizing the fact that small differences in initial conditions can have large long-term effects on institutions and individual choices. This last observation reinforces Greif's (2006) view that institutions cannot be simply described by a set of formal rules, but are in fact more complex arrangements whose formation and evolution are determined by history and context.

References

[1] Abramitzky, Ran, Liran Einav, and Oren Rigbi. Forthcoming. Is Hanukkah Responsive to Christmas? *Economic Journal*.

[2] Akerlof, George A. and Rachel E. Kranton. 2000. Economics and Identity, *Quarterly Journal of Economics* 105(3): 715-53.

[3] Akerlof, George A. and Rachel E. Kranton. 2005. Identity and the Economics of Organizations, *Journal of Economic Perspectives* 19(1): 9-32.

[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] Atack, Jeremy, Fred Bateman, Michael Haines, and Robert A. Margo. Forthcoming. Did Railraods Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850-60, *Social Science History*.

[10] Bénabou, Roland. 1996. Equity and Efficiency in Human Capital Investment: The Local Connection, *Review of Economic Studies* 63(2): 237-264.

[11] Bisin, Alberto and Thierry Verdier. 2000. "Beyond the Melting Pot": Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits, *Quarterly Journal of Economics* 115(August): 955-988.

[12] Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, and Yves Zenou. 2008. *Bend It Like Beckham*: Ethnic Identity and Integration. New York University typescript.

[13] Bisin, Alberto, Georgio Topa, and Thierry Verdier. 2004. Religious Intermarriage and Socialization in the United States, *Journal of Political Economy* 112(3): 615-664.

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

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

[16] 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.

[17] Brewer, M.B. and R.J. Brown. 1998. Intergroup Relations, in *Handbook of Social Psychology*, eds. D.T. Gilbert, S.T. Fiske, and G. Lindzey, 4th edn. New York: Oxford University Press.

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

[19] Durlauf, Steven. 1996. A Theory of Persistent Income Inequality, *Journal of Economic Growth* 1:75-93.

[20] 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.

[21] 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.

[22] Fernandez, Raquel and Richard Rogerson. 1996. Income Distribution, Communities, and the Quality of Public Education. *Quarterly Journal of Economics*, 111(1): 135-164.

[23] Finke, Roger and Rodney Stark. 1992. *The Churching of America, 1776-1990*. New Brunswick, NJ: Rutgers University Press.

[24] 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.

[25] Fryer, Roland G., Jr. Forthcoming. An Empirical Analysis of 'Acting White,' *Journal of Public Economics*.

[26] Galor, Oded and Joseph Zeira. 1993. Income Distribution and Macroeconomics, *Review of Economic Studies*, 60: 35-52.

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

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

[29] Giles, Michael W. and Arthur Evans. 1986. The Power Approach to Intergroup Hostility, *Journal of Conflict Resolution* 30(3): 469-186.

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

[31] 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.

[32] 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.

[33] Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge England: Cambridge University Press.

[34] Gruber, Jonathan. 2005. Religious Market Structure, Religious Participation, and Outcomes: Is Religion Good for You? NBER Working Paper 11377.

[35] 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.

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

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

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

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

[40] Putnam, Robert D. 2007 *E Pluribus Unum*: Diversity and Community in the Twenty-First Century (The 2006 Johann Skytte Prize Lecture), *Scandinavian Political Studies* 30(2): 137-174.

[41] Ragin, C.C. 1979. Ethnic Political Mobilization: The Welsh Case, *American Sociological Review* 44(August): 619-635.

[42] Sanders, Scott Russell. 2007. To Stay or Move On: The Artist's Dilemma. *Bloom*, February/March: 65-68.

[43] Sumner, William G. 1906. *Folkways*. New York: Ginn.

[44] Tabellini, Guido. 2008. The Scope of Cooperation: Values and Incentives, *Quarterly Journal of Economics* 123(3): 905-950.

[45] 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.

[46] Tullock, Gordon. 1980. Efficient Rent Seeking, in *Toward a Theory of the Rent-Seeking Society*, eds. James M. Buchanan, Robert D. Tollison, and Gordon Tullock, College Station: Texas A&M Press.

[47] Weber, Max. 1930. *The Protestant Ethic and the Spirit of Capitalism*. London, England: Allen and Unwin.

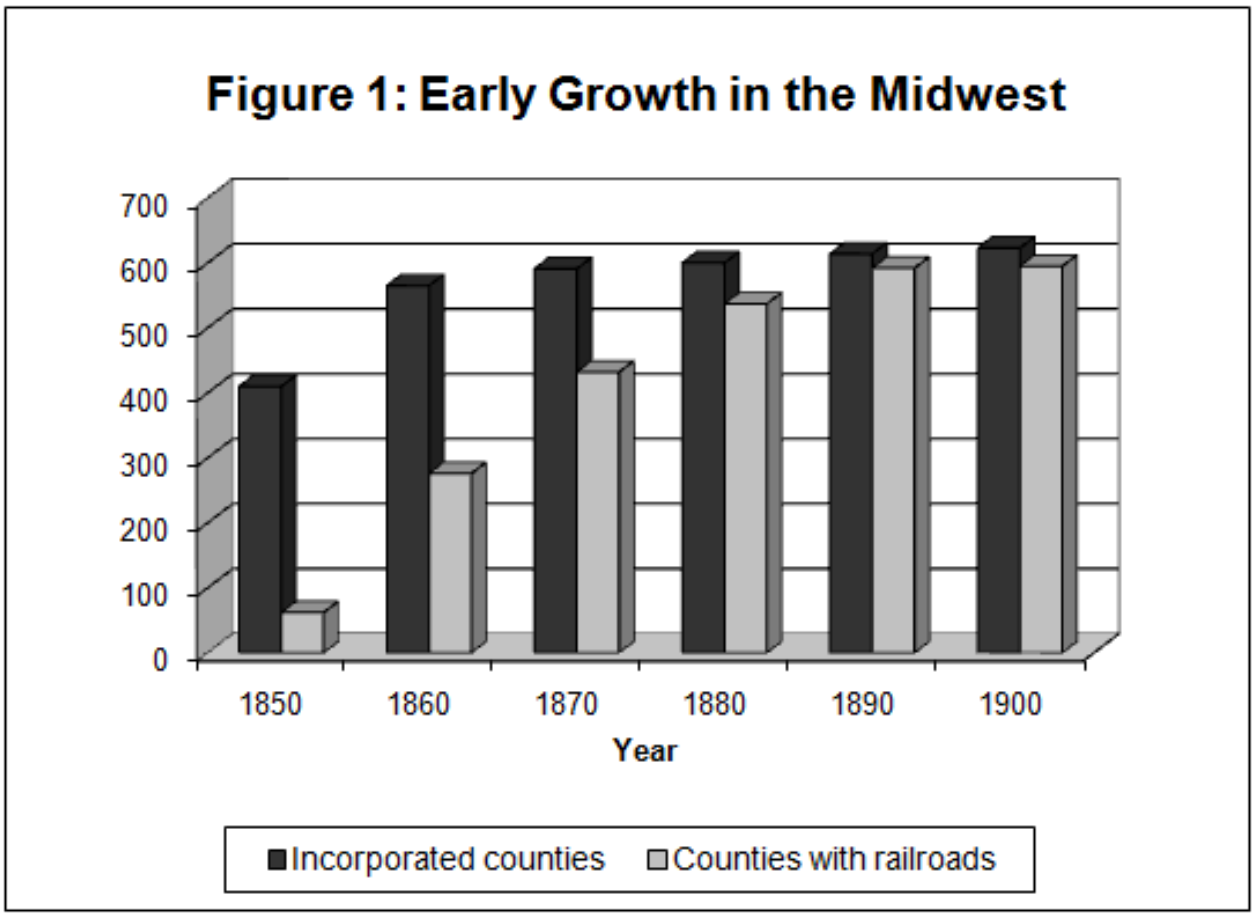
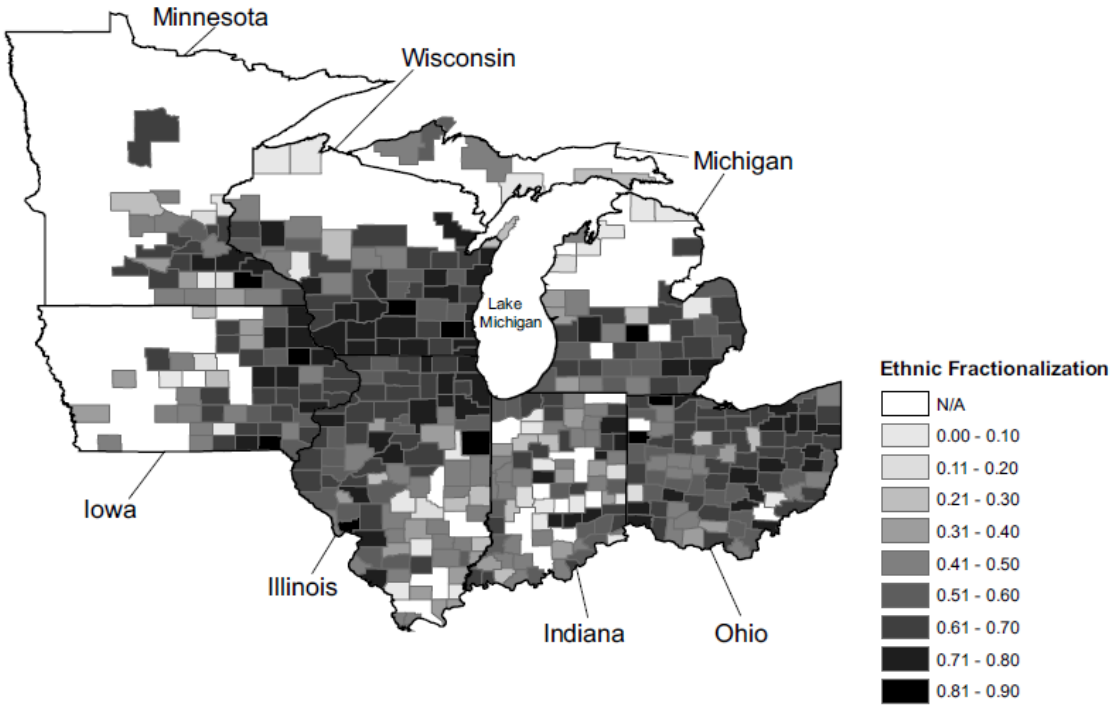


Figure 1: Early Growth in the Midwest
700
600
500
400
300
200
100
0
1850
1860
1870
1880
1890
1900
Year
Incorporated counties
Counties with railroads

Figure 2: Ethnic Fractionalization in 1860



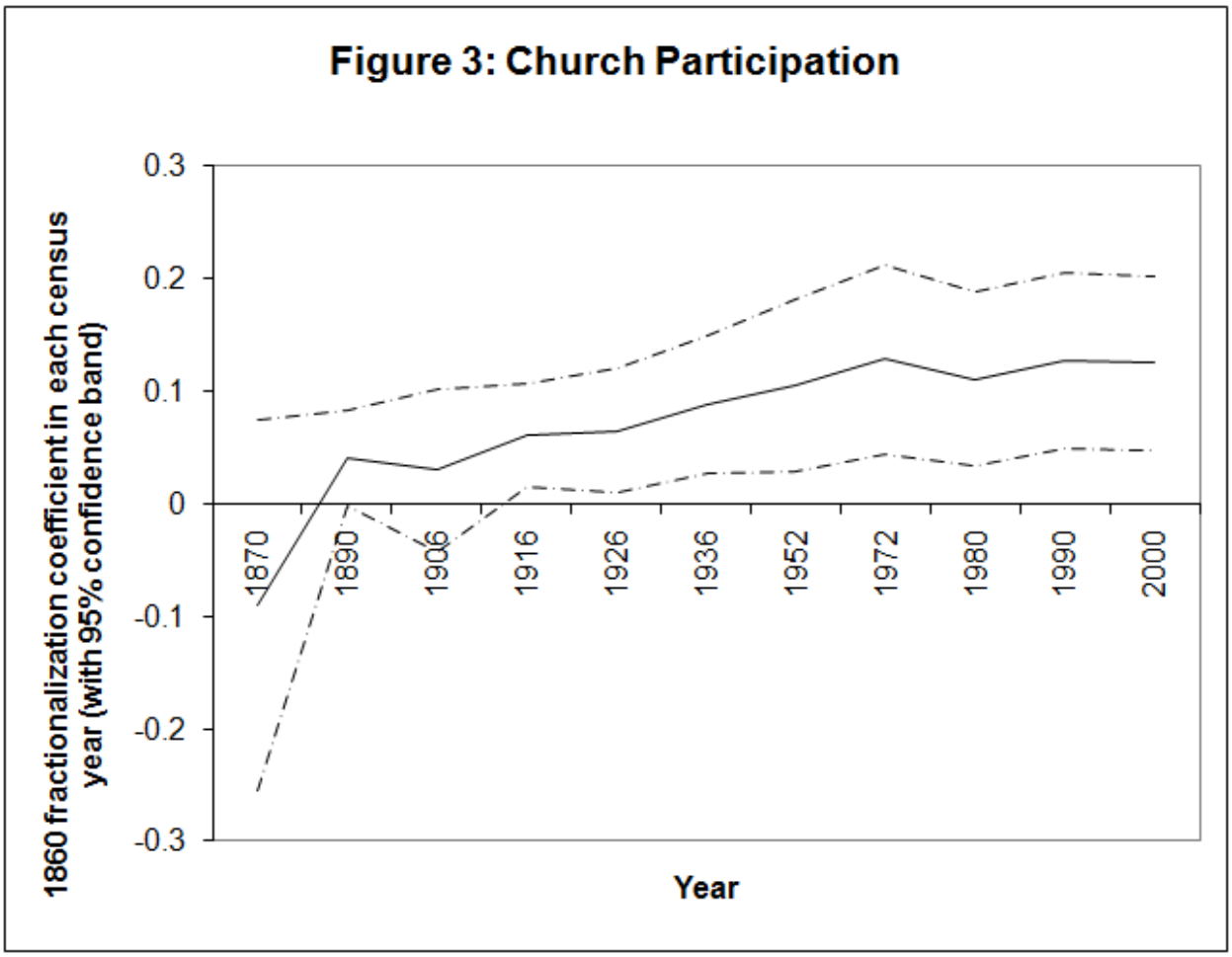


Figure 3: Church Participation
1860 fractionalization coefficient in each census year (with 95% confidence band)
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
1870
1890
1906
1916
1926
1936
1952
1972
1980
1990
2000
Year

Figure 4: Church Participation by Denomination

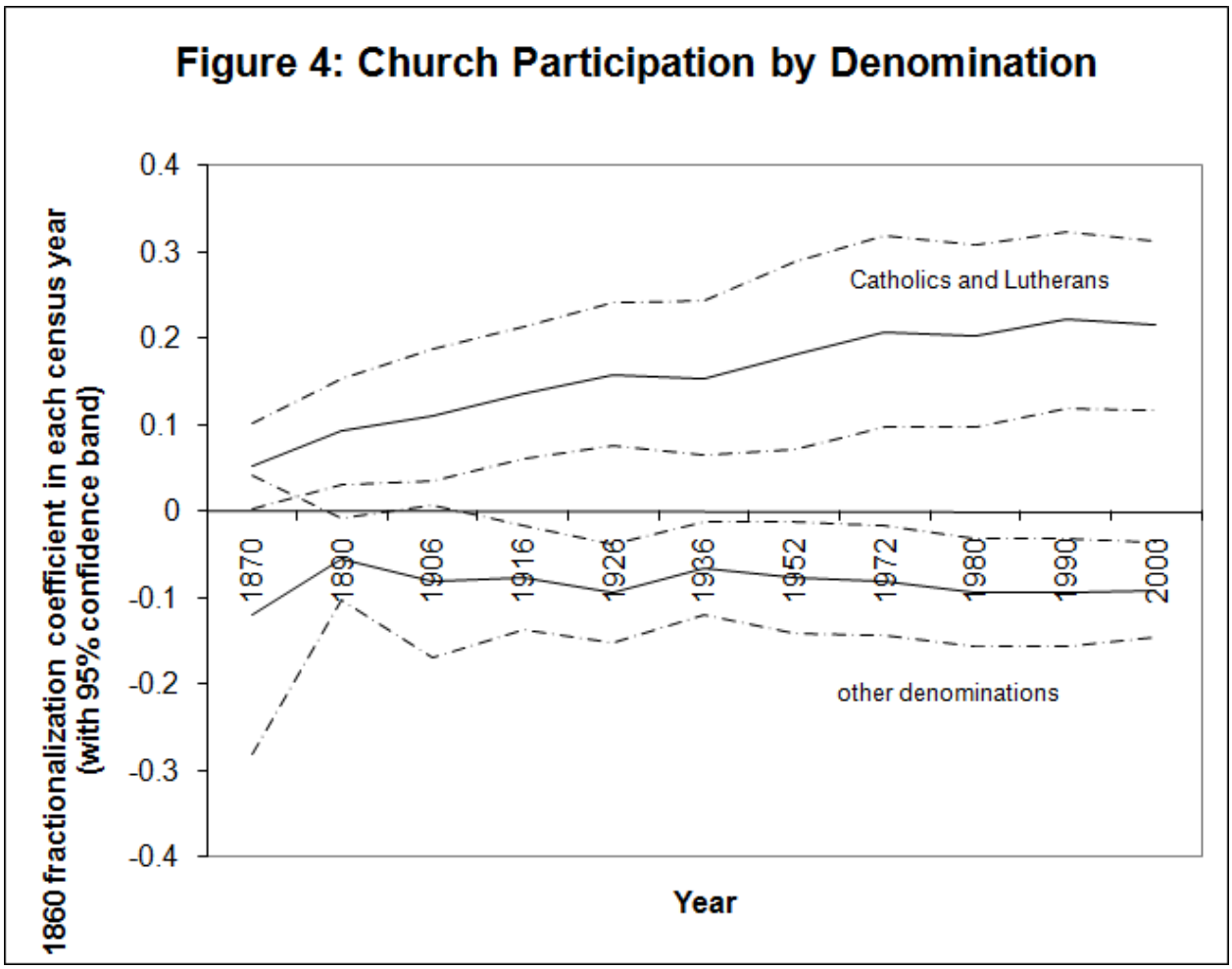


Figure 5: Historical Fractionalization and Current Outcomes

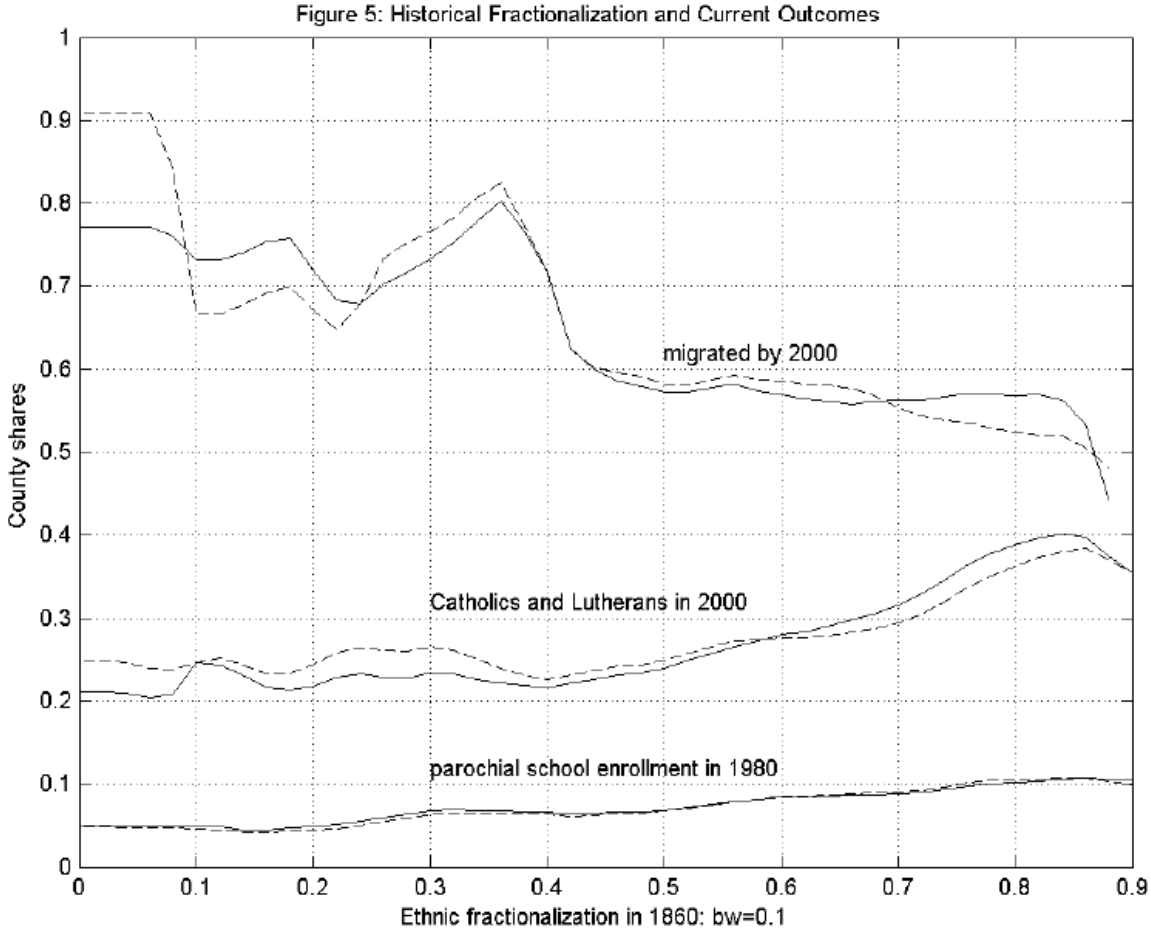


Table 1: Ethnic Distribution, 1860-1900

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: Transportation Infrastructure and County Characteristics, 1860

Dependent variable:	ethnic fractionalization	manufacturing share	agriculture share	population
	(1)	(2)	(3)	(4)
Railroad through county, 1860	0.049	0.001	-0.013	0.102
	(0.022)	(0.008)	(0.013)	(0.012)
Distance to canal, 1890	-0.668	0.169	-0.034	-0.469
	(0.191)	(0.097)	(0.125)	(0.102)
Distance to Great Lakes harbor	-0.252	-0.066	0.100	-0.136
	(0.073)	(0.029)	(0.048)	(0.056)
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.
Fractionalization is one minus the the Herfindahl index of ethnic concentration, averaged across occupational categories.
Manufacturing share and agriculture share in 1860 computed using IPUMS.
Population divided by 100,000.

Table 4: Fractionalization in 1860 and County Characteristics in 1990

Year:	1990					
Dependent variable:	agric share	manufac share	pop	racial frac	ethnic frac	religious frac
	(1)	(2)	(3)	(4)	(5)	(6)
Fractionalization, 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.
Fractionalization is one minus the Herfindahl index of ethnic concentration, averaged across occupational categories.
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: Fractionalization in 1860 and 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)
Fractionalization, 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.
Fractionalization is one minus the Herfindahl index of ethnic concentration, averaged across occupational categories.
Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.
Population divided by 100,000.

Table 6: Fractionalization in 1860 and 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	avg migration in occupation	avg college completion in occupation
	(1)	(2)	(3)	(4)	(5)
Fractionalization, 1860	-0.414	-0.108	-0.168	-0.021	0.127
	(0.138)	(0.058)	(0.084)	(0.058)	(0.076)
Manufacturing share, 1860	-0.242	-0.159	-0.378	0.154	0.246
	(0.260)	(0.114)	(0.189)	(0.094)	(0.215)
Agriculture share, 1860	-0.317	-0.133	-0.349	0.159	0.206
	(0.222)	(0.079)	(0.133)	(0.075)	(0.123)
Population, 1860	0.0648	0.026	0.029	0.000	0.035
	(0.040)	(0.013)	(0.023)	(0.013)	(0.017)
White	0.284	0.077	0.085	0.013	0.096
	(0.052)	(0.018)	(0.014)	(0.020)	(0.020)
Female	-0.014	0.000	0.008	-0.004	0.009
	(0.024)	(0.013)	(0.014)	(0.012)	(0.013)
Age	0.015	0.001	0.003	0.002	0.005
Observations	1437	1216	1216	1237	1237

Note: Standard errors in parentheses are clustered at the county level.
Fractionalization is one minus the Herfindahl index of ethnic concentration, averaged across occupational categories.
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.

Table 7: Fractionalization in 1860 and Education (Population Census)

Dependent variable:	parochial share		education expenditure per student	
Year:	1970	1980	1970	1980
	(1)	(2)	(3)	(4)
Fractionalization, 1860	0.071	0.070	-0.072	0.274
	(0.016)	(0.012)	(0.119)	(0.187)
Manufacturing share, 1860	0.126	0.072	0.027	1.058
	(0.049)	(0.044)	(0.443)	(0.594)
Agriculture share, 1860	-0.014	-0.016	0.317	0.044
	(0.039)	(0.032)	(0.343)	(0.425)
Population, 1860	0.025	0.050	0.013	0.781
	(0.021)	(0.015)	(0.236)	(0.388)
Observations	437	437	437	437

Note: Robust standard errors in parentheses.
Fractionalization is one minus the Herfindahl index of ethnic concentration, averaged across occupational categories.
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.
Parochial share is parochial school enrollment grades K-12 divided by school enrollment grades K-12.

Table 8: Fractionalization in 1860 and Education (NLSY)

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

Note: Standard errors in parentheses are clustered at the county level.
Fractionalization is one minus the Herfindahl index of ethnic concentration, averaged across occupational categories.
Manufacturing share in 1860 and agriculture share in 1860 are computed using IPUMS.
Population is divided by 100,000.
Fractionalization, 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 9: Robustness Tests

Dependent variable:	a	b	c	a	b	c	a	b	c	a	b	c
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Alternative construction of ethnic fractionalization variable												
Fractionalization	0.356	-0.368	0.102	0.361	-0.292	0.075	0.288	-0.410	0.081	0.203	-0.154	0.057
	(0.053)	(0.121)	(0.015)	(0.050)	(0.195)	(0.012)	(0.051)	(0.169)	(0.012)	(0.044)	(0.095)	(0.013)
Observations	290	1157	290	517	1471	517	415	1417	415	434	1436	434
Robustness test	fractionalization, 1850			fractionalization, 1870			exclude counties w/ less than 20 obs, 1860			men only, 1860		
Panel B: Alternative construction of ethnic fractionalization variable, continued												
Fractionalization, 1860	0.255	-0.327	0.078	0.225	-0.452	0.066	0.120	-0.422	0.055	0.006	-0.002	0.001
	(0.046)	(0.168)	(0.012)	(0.049)	(0.156)	(0.012)	(0.047)	(0.136)	(0.012)	(0.001)	(0.005)	(0.0002)
Observations	385	1369	385	437	1437	437	437	1437	437	437	1437	437
Robustness test	exclude farmers			4 occupational categories			no occupation categories			include native born		
Panel C: Alternative construction of population variable and additional regressors												
Fractionalization, 1860	0.131	-0.396	0.060	0.210	-0.411	0.069	0.095	-0.482	0.034	0.211	-0.443	0.067
	(0.049)	(0.136)	(0.012)	(0.049)	(0.138)	(0.012)	(0.049)	(0.178)	(0.013)	(0.049)	(0.139)	(0.125)
Observations	437	1437	437	437	1437	437	437	1437	437	437	1437	437
Robustness test	migrant population			working population			pop share of each ethnicity			religious fractionalization		

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 and religious fractionalization in 1860 are included separately as additional regressors in Panel C, Columns 7-12.