EFABG
7/17/07
4:40 pm

There's No Place Like Home: Local Identity and Occupational Choice in the Midwest *

Kaivan Munshi† Nicholas Wilson‡

July 2007

Abstract

This paper draws a connection between ethnic labor market networks in the American Midwest when it was first being settled, the local identity or attachment to place that emerged endogenously to support these networks, and occupational mobility today. Individuals born in counties with greater ethnic fractionalization in 1860, which we expect to be associated with stronger local identity, are significantly less likely to hold professional jobs, which come with greater geographical mobility, in 2000. A further connection is made between local identity and a particular social institution – the church – to explain the persistence of identity over multiple generations. We expect local identity to be positively correlated with the performance of the local church, which supports and is supported by this cultural trait, and as predicted counties with greater ethnic fractionalization in 1860 are associated with greater religious participation over many years in the future.

Keywords. Culture. Identity. Religion. Networks. Occupational choice. Mobility. Growth.
JEL. D85. J62. L14. L22.

*We are grateful to Andrew Foster, Chris Foster, Vernon Henderson, Dennis Hogan, Ashley Lester, Glenn Loury, Nancy Luke,David Meyer and Michael White for helpful discussions and suggestions. Ben Feigenberg 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

‡Brown University

1 Introduction

In a well functioning economy, individuals are born with an ability endowment that determines their investment in human capital and the jobs that they ultimately occupy. When markets function imperfectly, however, individuals with high ability can under-invest in human capital creating a mismatch between ability and occupational choice that has negative consequences for long-run growth (Banerjee and Newman 1993, Galor and Zeira 1993). The general implication of this argument is that the quality of *economic* institutions, measured by the structure of property rights and the performance of markets, can have important effects on growth. A recent series of empirical papers by Acemoglu, Johnson and Robinson (2001, 2002) have highlighted this "institutional" view, which currently serves as the dominant paradigm in economics.

Outside of economics, however, social scientists have long argued that culture, measured by shared beliefs and preferences within social groups, must be accounted for when explaining differences in the propensity to save and to invest in human capital across regions or countries. Weber's (1930) writings on the Protestant work ethic and its relation to capitalism are the classic references in this literature, but contemporary scholars such as Huntington (1996) have taken this position as well. Within economics, there have recently been some ingenious attempts to provide support for this "cultural" view by quantifying the relationship between culture, individual decisions that are associated with growth, and aggregate growth rates themselves (Barro and McCleary 2003, Tabellini 2005, Fernandez and Fogli 2007). To provide a credible alternative to the "institutional" view, however, it is not enough to demonstrate that culture matters. It is also important to explain why patterns of behavior should vary systematically across groups and why such behaviors should persist long after the economic circumstances that gave rise to them have ceased to be salient.

As in Bowles (1998), we take the view that preferences, and culture more generally, form endogenously in response to the economic environment but can subsequently persist independently of the economic forces that gave rise to them. The point of departure for the analysis in this paper is the observation that community-based networks often seek to restrict the occupational and spatial mobility of their members; such restrictions can be welfare enhancing when there is a negative externality associated with exit. Restrictions on mobility can take the form of explicit sanctions when a member exits or a less direct socialization process through which individuals come to identify strongly with their community. A strong sense of identity is not a cultural trait like thrift or hard work that is commonly

associated with economic growth. However, we will show that identification with a home community can reduce individual mobility and shape occupational choices in this paper, with accompanying implications for growth.

Recent evidence from urban India indicates that labor market networks can indeed restrict the mobility of their members (Munshi and Rosenzweig 2006). To isolate a role for identity in shaping mobility, however, we need to identify a research setting where networks were once active but are no longer relevant. Migrant workers had to bear most of the fluctuations in labor demand in the early stages of industrialization in the United States (Hoerder 1991). This is precisely the situation in which labor market networks are most useful and there is a wealth of historical evidence documenting the role played by ethnic migrant networks in finding jobs for their members at this time. With the arrival of the railroads around 1850, the Midwest in particular witnessed a tremendous influx of European migrants in response to the new occupational opportunities that became available. Manufacturing was still organized in relatively small workshops and the heavy industrialization and the homogenization of production that is characteristic of this region only began around 1880. Thus, between 1850 and 1880 there was a window of time during which the Midwest was rapidly settled and during which there was substantial fluctuation in labor market conditions and, hence, network characteristics across local areas. We exploit this variation *within* the Midwest to identify the effect of historical networks and the identity that they engendered on occupational choice and geographical mobility long after the networks themselves had ceased to be salient.[1]

Labor market networks and civic institutions based on a common European ancestry are largely irrelevant in the Midwest today (Gans 1979, Alba 1990). However, it would be incorrect to assume that these ethnic institutions disappeared without a trace. The migrant church was historically central to the maintenance of the ethnic network, providing a domain in which information could flow, commitments could be enforced, and identity could be instilled. Outside of the labor market, the church provided other services to its members as well, including social activity and support when they were sick or infirm. While the labor market networks might have disappeared, the church continues to provide important forms of mutual assistance, requiring a relatively high level of commitment in terms of time and effort from the congregation. If church inputs are complements, as suggested by Iannaccone (1998), then this *social* institution might continue to instill a strong sense of identity among

[1] Akerlof and Kranton (2000) show how identity can affect various economic decisions and outcomes in their pioneering paper. They do not, however, provide an explanation for why identity might differ across groups.

its members to restrict their geographical mobility. Professional occupations are associated with relatively high geographical mobility, which implies that individuals in tight-knit communities with a strong sense of local identity could tend to shade their career choices away from such occupations.

Attachment to place, rather than to an ethnicity or social group, has been identified as a distinct form of identity in the sociological literature (Hunter 1975, Guerson, Stueve and Fischer 1977, Hummon 1990) and the career trade-offs associated with this sense of local identity are well described in the following quote from a resident of Bloomington, Indiana who moved to that city to live with his wife who was born nearby, "... my chief ambition, I discovered during our early years in Bloomington, was not to make a good career but to make a good life. And such a life as I came to understand it, meant being a husband and a father first, and an employee second; it meant belonging to a place rather than to a profession ... So as I came to recognize 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. Most of what I valued in Bloomington was the result of efforts by people who loved the place, either because they grew up here and chose to stay, or because they landed here and chose to remain" (*Bloom Magazine*, February/March 2007: 67-68). We argue in this paper that what started as an ethnic identity to support labor market networks could have been transformed over time into a local identity or an attachment to place, to support a healthy church and broader community, connecting historical labor market networks to occupational choice over many future generations.

In which local areas do we expect this attachment to place to have been strongest? Historians describe the efforts made by migrant communities to establish "occupational beachheads" and "toeholds" in new locations, and to subsequently work hard to maintain these coveted positions once they were established. We will later show that the cost of a member's exit at the margin would have been most severe in labor markets with many small networks vying for scarce jobs and it follows that migrant communities in those locations would have invested most in instilling a sense of identity or group-loyalty in their members. Sumner's (1906) theory of ethnocentrism and group conflict generates the same qualitative prediction and more recently Bisin and Verdier (2000) argue, along similar lines, that individuals from minority groups at a greater risk of being acculturated will exert more effort to socialize their children in order to preserve their cultural identity. Later we will show that communities in locations with greater ethnic rivalries would have invested more in their local churches as well, as a way of instilling a strong sense of identity in their members, which over time would have translated into a stronger sense of local identity.

The framework outlined above generates two important testable predictions. The first prediction is that areas with greater ethnic rivalry at the time of settlement should be associated with stronger local identity, which implies that individuals born (and socialized) in those areas should be less likely to enter a professional career if this identity persisted over time. The second prediction is that greater ethnic rivalry should have given rise to better functioning churches and, hence, greater religious participation over many generations in the future.

To test the first prediction we combine data from the U.S. census and the National Longitudinal Survey of Youth (NLSY). Individual level data on occupations and ethnicity are available from the U.S. census going back to 1850. Following the discussion above, we focussed on the 1860 census when rapid growth in the Midwest was just commencing and constructed a measure of ethnic fractionalization, negatively related to the Herfindahl concentration index, within each broad occupational category in each county in that census year. Averaging over all occupational categories in each county we arrived at a measure of ethnic fractionalization in 1860 that we expect is positively correlated with network competition in the labor market at that time and, hence, with the strength of local identity in the county as it subsequently emerged. Matching the county-level measure of ethnic fractionalization with individual data from the NLSY, we find that individuals born in high fractionalization counties are significantly *less* likely to be employed in professional occupations and are significantly *less* likely to have migrated out of their birth county in 2000.

Our interpretation of this striking result is that local identity, determined endogenously by the labor market when the Midwest was first being settled in 1860, has subsequently persisted over time and continues to shape occupational choices in the county one hundred and forty years later. This interpretation is based on the assumption that the ethnic networks and the local economic conditions that gave rise to this identity are no longer salient and so any effect on current behavior must be due to this persistent cultural trait.[2] Although ethnic networks and European ancestry are clearly no longer directly relevant in the Midwest economy, ethnic fractionalization in 1860 could potentially be correlated with particular features of the economy at that time, which are correlated with the *demand* for professional labor today. For example, a county that was largely agricultural in 1860 is more likely to be agricultural today. Similarly, 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

[2]We argue later that cultural traits cannot persist over long periods of time without the continued support of social arrangements such as the family, the kin group, or the church. We consequently do not attempt to disentangle local identity from the *social* institution – the church – that supports this particular cultural trait in the current paper.

economy. Controlling for population, the share of agriculture, and the share of manufacturing in the county in 1860, we find that fractionalization in that year has no effect on a number of measures of current economic activity. Although these results are reassuring, they still leave open the possibility that some unobserved component of the local economy is lowering the demand for professional labor in the high fractionalization counties.

A stronger test of the identity hypothesis takes advantage of the observation that if the current demand for professional labor is uncorrelated with historical fractionalization, as we claim, then there must be a mismatch in the labor market. We show more formally that as long as there is some uncertainty in the demand for different types of labor, the supply of non-professional labor will exceed the expected demand in high fractionalization counties, with an accompanying shortfall in the supply of professional labor. To clear the labor market, professional labor must move into high fractionalization counties and non-professional labor must move out of those counties, on average, in competitive market equilibrium. It then follows that individuals *residing* in high fractionalization counties should be just as likely to hold professional jobs as individuals residing in low fractionalization counties, once the labor market clears. It is only individuals *born* in high fractionalization counties who are less likely to be professionals. The alternative hypothesis, based on differences in the demand for professional labor across counties, predicts that individuals born and residing in high fractionalization counties should be less likely to be professionals. We will later verify that labor flows systematically with the level of historical fractionalization as described above and, more importantly, that there is no correlation between occupational choice and the historical level of ethnic fractionalization in the county of residence.

The empirical strategy employed in this paper is similar to the strategy adopted by Fernandez and Fogli (2007) in a very interesting recent paper. Fernandez and Fogli establish that female labor force participation in the origin country affects labor force participation and fertility rates among second-generation female migrants in the United States. They rely on the assumption that culture will be transmitted across space and so cultural heterogeneity will continue to manifest itself in an economic environment – the U.S. labor market – that does not distinguish between social groups. The obvious alternative explanation, which they take care to rule out, is that human capital rather than culture is being transmitted across generations. In our case, human capital transmission is less of a concern since we are going back at least five generations; even if the ability distribution of the incoming migrants varied systematically with fractionalization in 1860, this heterogeneity would have

long since disappeared, given the high rates of inter-county migration (close to 60 percent on average) that we document below. With our strategy, the chief concern is that there is something about the *place*, other than identity, that has persisted over time, which the preceding test on the county of residence, rather than the county of birth, helps rule out.

If local identity did indeed persist over such a long period of time, what was the mechanism through which this cultural trait was sustained? Our proposed mechanism, based on the role of the local church, generates the prediction that high fractionalization counties should be associated with greater religious participation. A related implication is that denominations that were historically dominated by migrants – Lutherans and Catholics – should be especially popular in high fractionalization counties. Using data from the Census of Religious Establishments, available at roughly ten-year intervals from 1860 to 2000, we successfully verify each of these predictions. Recall that ethnic fractionalization has no effect on a number of current indicators of economic activity in the county once we control for local economic conditions in 1860. In contrast, the effect of ethnic fractionalization on religious participation actually grows stronger over the course of the twentieth century, emphasizing the important role played by the church in connecting historical settlement patterns to current individual choices.

Complementing recent studies that investigate the relationship between culture and growth, we find that local identity, measured by ethnic fractionalization in 1860, has an important effect on occupational choice today. A one standard deviation decline in fractionalization would increase the proportion of individuals holding professional jobs from 9 percent to 15 percent. Internal migration clears the labor market in our U.S. setting and so we do not expect identity to have a substantial effect on local growth rates. Nevertheless, the fact that the particular cultural trait that we study in this paper has such a large and persistent economic effect in a region as homogenous as the American Midwest suggests that cultural effects will be even stronger across countries with very different histories. Labor does not flow smoothly across countries, which implies that the effect of culture on growth could be quite large in practice.

2 The Institutional Setting

2.1 The Settling of the Midwest

The Midwest first began to be settled in the early nineteenth century with the expansion of the national canal system. The Erie Canal linking the Hudson to Lake Erie was completed in 1825 and numerous

inter-regional and intra-regional canals were built over the next two decades (Fishlow 2000). However, it was only with the arrival of the railroad that the Midwest took off on a steeper growth trajectory. Before 1850 the Midwest had less than one thousand miles of track, but almost ten thousand were added by 1857 (Meyer 1989).

Improved rail transportation stimulated industrialization and the Midwest'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. My 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).

In this paper, the Midwest is comprised of the states 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 increases sharply from 1850 to 1860 and then flattens out by 1880 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.[3] The rapid expansion of the railroad system and the industrialization 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, growing from 12 percent to 18 percent of the population, with the share of migrants in the total population remaining roughly constant thereafter.

Where did these migrants come from? 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. We use the 1-in-100 sample from the 1860-1900 IPUMS to study changes in the migrant population in these critical early decades in Table 1. The English (13%), the Irish (25%), and the Germans (32%), dominated the migrant population in the Midwest in 1860, just after the first wave of migration described above, with no other ethnic group accounting for more than a 3% share of the migrants in that year. Subsequently, the English and the Irish were displaced by the Germans

[3]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).

and the Scandinavians over the 1870-1900 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.

What jobs did the migrants occupy? Table 2 reports the occupational distribution of the migrants from the IPUMS sample in 1860, 1880, and 1900. 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. These trends are consistent with the growth of the manufacturing sector described above and they are similar for the foreign-born migrants and the native workers (not reported).

The apparent similarity in the occupational distribution for migrants and natives masks differences in the type of jobs they had access to within each occupational category. 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 employment (Gordon, Edwards, 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 (Hoerder 1991). We would expect these networks to have been particularly important during the rapid expansion phase around 1860, with the influx of migrants and the opening up of new labor markets. The initial conditions in the empirical analysis, measuring network characteristics at the time of settlement and the identity that they subsequently engendered, will consequently be measured using data at this point in time.

2.2 Ethnic Labor Networks in the Nineteenth Century

Accounts by contemporary observers and a rich 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 nineteenth century Midwest. Early historical studies used census data, which provides fairly detailed occupational and ethnic information, to identify ethnic clusters in particular locations and occupations. Gordon, Edwards, and Reich (1982) note that although foreign-born workers comprised just over 20 percent of the labor force in 1870, they accounted for 43 percent of the iron and steel operatives, 37 percent of the boot and shoe workers, 36 percent of the cotton mill operatives, 43 percent of the woolen mill workers, and 63 percent of the miners. Nearly a quarter of railroad

workers were Irish, a third of the miners were British, and about half the workers in the baking and confectionary business were German. While such clustering suggests that underlying ethnic networks were channelling their members into particular occupations, it could simply reflect the fact that migrants arrived with specific skills. Hutchinson's (1956) analysis of 1950 census data, however, indicates that clustering continued even among the migrants' children, with the concentration of particular ethnic groups in some industries actually increasing from the first to the second generation.

Although census data are a useful source of information, they do not provide details of the migration process and its connection to ethnic networks in the United States. Over the past four decades, however, 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 remarkable research effort has documented the formation of new settlements in the Midwest by pioneering migrants, the subsequent channelling 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. As with the census data, this research identifies occupational and geographic clustering, but at a dissagregate level. Over 45 percent of the Swedish emigrants from the parish of Rätvik eventually settled in Isanti County, Minnesota (Ostergren 1976). Two-thirds of the emigrants from Balestrand located in Norway Grove, Wisconsin in the first decade of migration from that Norwegian community (Gjerde 1985). And one-third to one-half of the German emigrants from Westerkappeln settled in Duden County, Missouri (Kamphoefner 1987). Although less detailed origin-country information is available for southern European migrants, similar ethnic clustering in particular neighborhoods of Midwestern cities has been documented for Polish, Italian, and Slovak immigrants from specific sending regions (Alexander 1991, Bodnar, Simon, and Weber 1982).

A possibly stronger indicator of the importance of migrant networks is the maintenance of ethnic ties over successive moves *within* the United States. Italians moved from Southern Illinois to the "Italian Hill" in St. Louis when coal mining operations were reduced in the 1920's and Slavs moved from mines in Western Pennsylvania to Detroit's growing automobile industry in the same decade (Bodnar 1985). Norwegians from Balestrand initially settled in Norway Grove, Wisconsin, but over time they established six satellite settlements in Wisconsin, Minnesota, Iowa, and Illinois (Gjerde 1985). A similar pattern has been documented for Norwegians immigrants from Fortun, who initially settled in Vernon and Crawford Counties, Missouri, but later established satellite communities throughout the

Midwest (Gjerde 1997).

While the preceding descriptions of ethnic clustering are informative, ethnic concentration within specific departments or firms in a local industry provides possibly the strongest evidence that labor networks were active. Nearly all three thousand employees of the Peninsular Gas Company in Detroit in 1900 were Polish, and Croatians held only three jobs in Indiana's oil refineries: stillman helper, fireman, and still cleaner (Bodnar 1985). Italians in Pittsburgh's steel industry dominated the carpentry, repair, and rail shops. And, relying on friends and relatives, Poles established occupational niches at the Jones and Laughlin and Oliver Mills on Pittsburgh's Southside, Heppenstalls and the Pennsylvania Railroad in Lawrenceville, and at the Armstrong Cork Company and the H.J. Heinz Plant. As John K. a Polish immigrant put it, "The only way you got a job [was] through somebody at work who got you in" (Bodnar, Simon, and Weber 1982: 56).

What kept ethnic networks in place so far from their origin locations? It has been argued that "[migrants] from varying regions [in the origin country] formed a community based on common nationality and religion centered on the central cultural institution – the church" (Gjerde 1991: 176). The building of a church was one of the first organized actions in the migrant community once it arrived in an area (Barton 1975, Bodnar, Simon, and Weber 1982). Churches provided both economic and social support to their members. Information about jobs and potential land transactions flowed within the congregation and the church also served as a public arena in which members who had reneged on their obligations could be sanctioned.

Given the variety of economic opportunities in the United States, individuals often had an incentive to move and seek employment elsewhere. Moreover, church congregations often consisted of groups from multiple regions in the origin country, giving rise to conflict and the potential for exit from the church. The stability of the church was essential for the viability of the labor market network. As discussed, one strategy to restrict mobility was to instill a sense of identity that transcended individual incentives and narrower familial and kin affiliation. Consistent with this restrictive role for the church, Gjerde (1985) finds that church members were significantly less likely to move than non-church members among the Norwegian migrants that he traces across space and over time. The discussion that follows will explore how such restrictions might have varied across Midwestern counties when the first wave of migrants arrived in the region.

2.3 Ethnic Networks Across the Midwest

"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." (Polish immigrant in Pittsburgh, quoted in Bodnar, Simon, and Weber 1982:62).

Numerous historians have described the efforts made by ethnic groups to establish a "toe-hold" (Thistlethwaite 1991) or a "beachhead" (Bodnar, Simon, and Weber 1982) in particular industries or establishments when they first settled in an area. The preceding quote suggests, in addition, that once a network had established a presence in a workshop, it was essential to *maintain* that presence. In the discussion that follows we will describe the labor market conditions under which migrant networks would have had the greatest incentive to restrict exit and, hence, instill a sense of identity among their members. The mechanism through which this sense of identity was instilled as well as an explanation for its persistence over multiple generations is postponed to Section 5 of the paper.

In general we would expect larger networks to be less vulnerable to negative labor demand shocks and to exit from the market by their members. To see why this should be the case, consider a market with M workshops and, to begin with, a single ethnic group with N members. M and N are both large numbers, with $M > N$. In the initial period when the ethnic group first arrives in the area, there is no network available to provide employment assistance and so we assume, for simplicity, that individuals are randomly assigned to workshops. It follows that the measure of workshops spanned by the network is $M\left[1-(1-1/M)^N\right]$, which is increasing in the size of the group N, but at a decreasing rate. Notice that we have placed no restriction on the number of individuals who can be assigned to a given department. A cap on this number would generate some unemployment, but would otherwise leave the span of the network unchanged.

Suppose that individuals work for a single period and that their positions are inherited by the cohort that succeeds them as long as they do not exit during their working lives. Exit from a given workshop will, however, erase the network's claim to future positions in that workshop if no member of the network remains. The network's positions remain intact as long as a single member remains, since the incumbent worker can use his influence to recruit members of his ethnic group, as in the quote

above. If we consider the effect of a single individual's exit at the margin, then the only vulnerable workshops are those occupied by a single member of the network. In our simple framework, the proportion of members occupying a position in such a workshop is $(1 - 1/M)^{N-1}$, which is decreasing in N. Thus, the (future) cost to the ethnic group of exit by a randomly selected individual is greater in smaller networks. More generally, once we allow for both demand shocks and exit we expect that there will be a threshold depth below which the network is vulnerable to the loss of its positions. The preceding discussion indicates that a greater proportion of the membership would be in such vulnerable workshops in small networks, which implies, in turn, that smaller networks would make greater efforts to restrict exit by their members. The same result is obtained in a related context by Bisin and Verdier (2000) who show that individuals from a minority group who are at greater risk of being acculturated into the general population will exert more effort to socialize their offspring to maintain their cultural identity.

Up to this point we have assumed that a single network is active in the market. Now suppose there are multiple networks and that when a network loses its position in a workshop, that position can be claimed by another network that is also active in that workshop. As above, there is little cost to exit at the margin when the network has sufficient depth in a workshop. There is little cost to exit in a workshop where the network is poorly represented as well if one or more competing networks has a sizeable presence there; even if exits can be controlled, demand shocks over time will shut down the network's few positions in that workshop. Exit costs will thus be greatest in workshops in which many evenly-sized networks are competing for positions. Assuming that positions are independently allocated across networks in the initial period, a randomly selected individual is more likely to belong to such a workshop in a market with many small networks. The cost of exit at the margin will consequently be greatest in such markets, with a commensurate increase in restrictions on mobility.[4]

How would we measure the distribution of network size in the market? We assume that networks were most active in the migrant population and that they were organized on the basis of the country of origin. Recall that the migrant church brought individuals from different regions in the origin country

[4]This prediction matches the empirical results in Bisin, Topa and Verdier (2004), who test the theoretical implications from Bisin and Verdier (2000) described above. While minorities do socialize their children more strictly, as a way of maintaining their cultural identity, socialization efforts across social groups in the population are found to be non-linear, with very small minorities and the majority group in an area exerting less effort. Bisin, Topa and Verdier attribute these differences in effort to underlying differences in the cost of socialization together with differences in the incentive to exert effort. In our framework, once we allow for competition between ethnic networks, reasonable assumptions on the consequences of exit imply that both the smallest and the largest networks will have the least incentive to restrict the mobility of their members.

together into a large and stable network that was capable of smoothing substantial fluctuations in the labor market. Numerous accounts of occupational mobility in the nineteenth century indicate that although networks supported the movement of their members across establishments and even across industries, migrants could rarely change the type or skill-level of their jobs. Competition between networks is consequently assumed to occur within the broad occupational categories specified in Table 2. The IPUMS provides the country of birth and the occupation (where relevant) for each sampled individual. Using these data, we computed ethnic fractionalization, defined as one minus the Herfindahl index of ethnic concentration, in each occupational category in each Midwest county in 1860, just around the time when migrants were flowing into the region.[5] The weighted average of the occupation-specific statistics, where the weight is measured by the share of migrants in a given occupation, then provides us with an overall measure of ethnic fractionalization in the county. Conditional on the number of migrants, greater fractionalization is associated with numerous small networks, which we expect to be associated with greater restrictions on mobility.

Figure 2 plots the fractionalization measure, which has a mean of 0.5 and a standard deviation of 0.2, across the seven Midwest states. Counties that were not incorporated and those without foreign-born migrants in 1860 are unshaded in the Figure. We cannot directly observe the role played by the networks or the restrictions that they placed on their members in 1860. Our strategy in this paper is to test the relationship between historical fractionalization and current occupational choice under the assumption that the restrictions on mobility, embodied in a local identity, persisted over time. This test is only valid if historical fractionalization is uncorrelated with the demand for professional labor in the county today. No particular spatial pattern is evident in the fractionalization measure in Figure 2. The discussion that follows will provide additional support for the claim that fractionalization in 1860 is uncorrelated with the nature of economic activity in the county today.

2.4 Historical Fractionalization and Current Economic Activity

To explore possible links between fractionalization in 1860 and the demand for professional labor today 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 fuelled the arrival of more

[5]The Herfindahl index of ethnic concentration is computed as the sum of the squared share of each ethnicity in the occupational category.

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.[6] 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 expected.

Transportation in 1860 clearly has no direct effect on the nature of the local economy today. The concern, however, is that transportation infrastructure in 1860 would have been associated with features of the local economy at that time, which did have long-term effects. Fractionalization could, in that case, be correlated with the demand for professional labor today, providing an alternative explanation for any observed relationship between historical fractionalization and current occupational choice.

Counties close to a harbor had a greater proportion of the workforce engaged in manufacturing and a smaller proportion in agriculture in 1860 in Table 3, Columns 2-3. Improved transportation, more generally, is associated with a larger population in 1860 in Column 4. Moreover, these transportation effects are persistent; superior transportation infrastructure in 1860 increases the share of manufacturing, reduces the share of agriculture, and increases population and population density in 1990 in Table 3, Columns 5-8.[7] This last set of results is very likely obtained because the occupational mix (manufacturing versus agriculture) and the population of the county in 1860, which are related to transportation infrastructure at that time, had a persistent effect on the local economy. Fractionalization in 1860 would also be correlated with these variables and not surprisingly we see in Table 4, Columns 3-4 that fractionalization has a positive and significant effect on population and population density in the county in 1990. However, we do not expect fractionalization in 1860 to have any effect on county characteristics in 1990 once important features of the local economy that would have had persistent effects, such as manufacturing share, agriculture share, and population in 1860, have been accounted for. This is indeed what we observe in Columns 5-8 of Table 4.

[6]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.

[7]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.

As expected, agricultural counties in 1860 continue to be agricultural, have smaller populations, and are less urban (with lower population density) in 1990. Larger counties in 1860 are larger, more urban, and less agricultural in 1990. Notice, however, that manufacturing in 1860 has no effect on current county characteristics. Although the factory system began to replace artisan shops by 1820, production continued to be largely organized in workshops managed by labor contractors who hired their own employees until 1870 (Gordon, Edwards, and Reich 1982). The heavy manufacturing that characterized the Midwest economy in the twentieth century, with its emphasis on the iron and steel industry, only came at the turn of the century (Meyer 1989). Recently it has been argued that the surge of foreign immigration in the second half of the nineteenth century provided the impetus for the factory system and the subsequent industrialization of the Midwest (Kim 2007). Whatever the explanation, it is clear that the pattern of manufacturing around 1860, spread throughout the Midwest in small towns, had little connection with the heavy manufacturing, concentrated in large cities, that followed in the twentieth century. This is presumably why 1860 characteristics have such little power in predicting the share of manufacturing in the county in 1990.

For our purpose, what is most important is that once we control for a few important features of the nineteenth century economy, fractionalization in 1860 has no effect on characteristics of the economy today, such as the share of manufacturing and urbanization (measured by population density), that are associated with the demand for professional labor. There is always the possibility that other unobserved demand characteristics are correlated with fractionalization. The simple model of occupational choice incorporating local identity that follows generates additional predictions for the *flow* of labor across counties that will allow us to rule out alternative demand-side explanations more generally.

3 Local Identity and Occupational Choice: Theoretical Framework

In the previous section we argued that counties with greater ethnic fractionalization at the time of initial settlement would be associated with greater restrictions on mobility, embodied in a stronger sense of local identity. Individuals in professional occupations are more likely to have to relocate during their careers. If attachment to place persists across many generations, then we would expect individuals born in high fractionalization counties with a stronger sense of identity to demonstrate a lower propensity to enter into professional occupations. We will see that this result is obtained quite

generally in competitive labor market equilibrium as long as there is some uncertainty in the demand for different types of labor across locations.

3.1 Population and Production Technology

Two types of jobs are available in this economy: professional and non-professional. Individuals who are *ex ante* identical live for two periods, working in the second period of their lives. Those individuals who choose to occupy professional jobs must invest in training in the first period of their lives, which costs C_e. Individuals who end up in non-professional jobs incur no such cost. The output obtained from a professional worker is $\overline{\theta}$ and the output obtained from an non-professional worker is $\underline{\theta} < \overline{\theta}$.

There are two locations in this economy with N individuals born in each location in each period. On average, sN professional jobs and $(1-s)N$ non-professional jobs are demanded in each location in each period. However, these locations also face demand shocks, with two states of the world occurring with equal probability:

State 1: $sN + \epsilon sN$ professional and $(1-s)N + \epsilon(1-s)N$ non-professional jobs in location 1.
$sN - \epsilon sN$ professional and $(1-s)N - \epsilon(1-s)N$ non-professional jobs in location 2.
State 2: $sN - \epsilon sN$ professional and $(1-s)N - \epsilon(1-s)N$ non-professional jobs in location 1.
$sN + \epsilon sN$ professional and $(1-s)N + \epsilon(1-s)N$ non-professional jobs in location 2.

Notice that these demand shocks are skill neutral, in the sense that the probability of receiving a shock is the same for professional and non-professional workers within each location. The demand shocks are also symmetric across locations.

Once demand shocks are introduced, labor must flow across locations at the beginning of each period to clear the market. In addition, we assume that professional workers are exogenously moved with probability P during the period in which they work. These workers are transferred outside the two locations and are immediately replaced by a professional worker drawn from outside as well. Individuals dislike moving, particularly those with a strong local identity. Let the cost of moving be C_{1t} for individuals born in location 1 and $C_{2t} < C_{1t}$ for individuals born in location 2 with lower historical ethnic fractionalization. This is the only difference between workers in the two locations. If productivity is sufficiently larger than moving costs, all workers will be employed in equilibrium and total output in this economy, in each state of the world, will be $2N\left[s\overline{\theta} + (1-s)\underline{\theta}\right]$. The competitive labor allocation will be obtained in that case as the solution to the Central Planner's problem of minimizing training and transportation costs across both locations.

3.2 Labor Market Equilibrium

Let the supply of professional labor in location 1 be x_1. From the structure of the demand shocks it then follows that the supply of professional labor in location 2 will be $2sN - x_1$ and that the supply of non-professional labor in location 1 will be $N - x_1$. We derive x_1 as the solution to the Central Planner's cost minimization problem in three regimes:

Regime 1: $x_1 \in [sN - \epsilon sN, sN + \epsilon sN]$

The supply of professional labor in each location is sufficient to satisfy the minimum demand in that location but does not exceed the maximum demand.

The labor flow in each state of the world can then be derived as:

Flow in state 1: $sN + \epsilon sN - x_1$ professional labor from location 2 to location 1.

$(1-s)N + \epsilon(1-s)N - (N - x_1)$ non-professional labor from location 2 to location 1.

Flow in state 2: $x_1 - [sN - \epsilon sN]$ professional labor from location 1 to location 2.

$(N - x_1) - [(1-s)N - \epsilon(1-s)N]$ non-professional labor from location 1 to location 2.

The Central Planner chooses x_1 to minimize expected cost

$$E(C) = 2sNC_e + \frac{1}{2}\epsilon N\left(C_{1t} + C_{2t}\right) + Px_1C_{1t} + P(2sN - x_1)C_{2t}, \tag{1}$$

where the second term on the right hand side is the cost associated with movement at the start of the period and the last two terms reflect movement of professional labor during the period. Because $C_{1t} > C_{2t}$ it is easy to verify that $E(C)$ will be minimized at $x_1^* = sN - \epsilon sN$. The equilibrium labor flows are then obtained as:

Equilibrium flow in state 1: $2\epsilon sN$ professional labor from location 2 to location 1.

$\epsilon N(1-2s)$ non-professional labor from location 2 to location 1.

Equilibrium flow in state 2: No flow of professional labor.

ϵN non-professional labor from location 1 to location 2.

Regime 2: $x_1 \in [sN - \epsilon sN - \epsilon N(1-2s), sN - \epsilon sN]$

We now reduce the supply of non-professional labor in location 2, but at most to the point where no non-professional labor flows to location 1 in state 1. In our set up, any reduction in non-professional labor supply in location 2 must lead to a reduction in professional labor in location 1 by the same amount. It then follows that the supply of professional labor in location 1 will no longer be sufficient

to meet the minimum demand in that location, while the supply of professional labor in location 2 will exceed the maximum demand in that location. Labor flows at the beginning of the period will necessarily increase, with an accompanying increase in moving costs, but we will see that this may be outweighed by the reduced cost of relocation for professional workers from location 1.

The labor flow in each state is derived as:

Flow in state 1: $sN + \epsilon sN - x_1$ professional labor from location 2 to location 1.

$(1-s)N + \epsilon(1-s)N - (N - x_1)$ non-professional labor from location 2 to location 1.

Flow in state 2: $sN - \epsilon sN - x_1$ professional labor from location 2 to location 1.

$(N - x_1) - [(1-s)N - \epsilon(1-s)N]$ non-professional labor from location 1 to location 2.

The expected cost can then be expressed as:

$$E(C) = 2sNC_e + \frac{1}{2}\left[\epsilon N + sN(1-\epsilon) - x_1\right](C_{1t} + C_{2t}) + Px_1C_{1t} + P(2sN - x_1)C_{2t}. \tag{2}$$

Collecting terms, $E(C)$ continues to be minimized at $x_1^* = sN - \epsilon sN$ if $P < 1/2(C_{1t} + C_{2t})/(C_{1t} - C_{2t})$. However, a new equilibrium allocation can be sustained, with $x_1^{**} = sN - \epsilon sN - \epsilon N(1-2s)$, if the sign of the inequality is reversed. This will be the case if P and $C_{1t} - C_{2t}$ are sufficiently large. Substituting x_1^{**} in equation (2) above, it is easy to verify that the term in square brackets is greater than ϵN, the term corresponding to it in equation (1), which implies that transportation costs at the beginning of the period have increased in this equilibrium. However, transportation costs during the period decrease with the reduction in x_1, and this effect dominates under the conditions on P and $C_{1t} - C_{2t}$ derived above. Equilibrium labor flows in each state can then be derived as:

Equilibrium flow in state 1: ϵN professional labor from location 2 to location 1.

No flow of non-professional labor.

Equilibrium flow in state 2: $\epsilon N(1-2s)$ professional labor from location 2 to location 1.

$2\epsilon N(1-s)$ non-professional labor from location 1 to location 2.

Regime 3: $x_1 \in [0, sN - \epsilon sN - \epsilon N(1-2s)]$

We now reduce the supply of professional labor in location 1, with an accompanying increase in non-professional labor, even further so that professional labor flows from location 2 to location 1 and non-professional labor flows in the opposite direction in both states of the world.

Labor flows are now derived as:

Flow in state 1: $sN + \epsilon sN - x_1$ professional labor from location 2 to location 1.

$(N - x_1) - [(1 - s)N + \epsilon(1 - s)N]$ non-professional labor from location 1 to location 2.

Flow in state 2: $sN - \epsilon sN - x_1$ professional labor from location 2 to location 1.

$(N - x_1) - [(1 - s)N - \epsilon(1 - s)N]$ non-professional labor from location 1 to location 2.

The corresponding expected cost expression is obtained as:

$$E(C) = 2sNC_e + \frac{1}{2}\left[2(sN - x_1)\right](C_{1t} + C_{2t}) + Px_1C_{1t} + P(2sN - x_1)C_{2t}. \tag{3}$$

It is straightforward to verify that this expression is minimized at $x_1 = sN - \epsilon sN - \epsilon N(1 - 2s)$, matching the equilibrium allocation x_1^{**} derived above.

3.3 Labor Allocation and Labor Flow Across Counties

The preceding discussion indicates that two equilibria can be sustained in the labor market. In each equilibrium, the supply of professional labor in location 1, with stronger local identity, falls short of the expected demand. This deficit is met by the flow of professional labor from location 2. The first empirical prediction from the model is consequently that the share of professionals among the workers born in location 1 should be lower than the corresponding share among workers born in location 2, despite the fact that labor demand does not vary across counties.[8] This result is obtained for any value of P as long as there is some uncertainty in the labor demand and $C_{1t} > C_{2t}$. Without uncertainty, the supply of professional labor in location 1 would only fall short of the expected demand if $P > (C_{1t} + C_{2t}) / (C_{1t} - C_{2t})$.[9]

In our set up, individuals *born* in location 1 are less likely to be professionals because they incur a greater cost when they move ($C_{1t} > C_{2t}$). However, this result could also be obtained if the demand for professional labor is lower in location 1. The second prediction of the model, which allows us to rule out this alternative demand-side explanation, is that once the labor market has cleared, the

[8] In equilibrium 1, the share of professionals is $s - \epsilon s$ in location 1 and $s + \epsilon s$ in location 2. In equilibrium 2, the corresponding shares are $s - \epsilon(1 - s)$ and $s + \epsilon(1 - s)$.

[9] Without uncertainty in labor demand, sN professional and $(1 - s)N$ non-professional jobs are available in each location in each period. Let $x_1 \in [0, sN]$ measure the supply of professional workers in location 1. It then follows that $sN - x_1$ professional workers would flow from location 2 to location 1 and $(N - x_1) - (1 - s)N$ non-professional workers would flow in the opposite direction at the beginning of each period. Using the same notation as above, the Central Planner chooses x_1 to minimize

$$E(C) = 2sNC_e + (sN - x_1)(C_{1t} + C_{2t}) + Px_1C_{1t} + P(2sN - x_1)C_{2t}.$$

It follows that $x_1 = sN$ if $P < (C_{1t} + C_{2t}) / (C_{1t} - C_{2t})$. $x_1 = 0$ and professional labor is under-supplied in location 1 if the sign of the inequality is reversed.

share of professionals *residing* in the two locations should on average be the same. In contrast, if the demand for professional labor is lower in location 1, then individuals born and subsequently residing in that location should be less likely to be professionals.

4 Local Identity and Occupational Choice: Empirical Analysis

4.1 Individual Data

To test the predictions from the previous section we need information on the individual's career choice, county of birth, and county of residence (post-employment). The National Longitudinal Survey of Youth 1979 (NLSY 1979) is the only large-scale data set that we are aware of that includes this information. The NLSY 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 survey collects basic information on the respondent's age, gender, race and, most importantly, county of birth. The Armed Forces Qualification Test (AFQT), which is designed to provide an unbiased measure of the individual's intelligence, was administered to all respondents in 1979. Subsequent survey rounds collected contemporaneous information on educational attainment, employment, occupation, income, and county of residence. We will study occupational choice and other outcomes related to that economic decision at two points in time – 1994 and 2000 – when the respondents were old enough to be settled in their careers and to have made some job-related moves. Occupational choices from the NLSY in these years will be matched to census data on historical fractionalization, both in the individual's county of birth and the contemporaneous county of residence, to test the predictions of the model.

Table 5 reports descriptive statistics for the individuals in our sample, who were aged 14-22 in 1979 and so around 33 years old in 1994 and 39 years old in 2000. Occupational categories in the NLSY (up to the 2000 round) are based on the 1970 codes from the census. Professional occupations are defined to include relevant codes listed under the Professional, Technical, and Kindred Workers category [1-196]. We exclude technical occupations listed in this category, which for the most part do not require a college degree and need to be filled on a permanent basis in most communities.[10] Such occupations are not particularly strongly associated with geographic mobility, which is the chief

[10]These occupations include Nurses, dieticians, and therapists [74-76], Religious workers [86,90], Social and recreation workers [100,101], Teachers, except college and university [141-145], Technicians [150-174], Writers, artists, and entertainers [175-196].

property that distinguishes professional and non-professional jobs in the model.

Based on this occupational classification, 9 percent of the respondents hold professional jobs, with little change from 1994 to 2000. 56 percent of all respondents had migrated out of their birth-county by 1994, with an increase to 59 percent by 2000. Consistent with the assumption that professional occupations are associated with greater mobility, 75 percent of the professionals and 53 percent of the non-professionals had migrated out of their birth-county by 1994 (these differences are significant at the 5 percent level and similar in 1994 and 2000).

Individuals in the sample are 33 years old by 1994 and should be established in the labor market. Nevertheless, employment levels continue to increase over time, from 81 percent in 1994 to 92 percent in 2000. Conditional on being employed, annual income (in 2000 dollars) also increases from 28 thousand in 1994 to 33 thousand in 2000. These changes in employment and income are presumably life-cycle effects, but they could, in principle, be due to selective attrition since this is a longitudinal survey. Notice, however, that racial composition and the proportion of women in the sample are very stable over the 1994-2000 period. Thus, we do not observe selective attrition from the sample, at least with respect to two important demographic characteristics that are associated with income and employment.

4.2 Historical Fractionalization and Current Occupational Choice

The first prediction from the model of occupational choice incorporating local identity is that individuals born in counties with greater ethnic fractionalization in 1860 should be less likely to hold professional jobs. Including race, gender, and age as regressors (although their omission would not affect the results) we see in Table 6, Columns 1-2 that individuals born in high fractionalization counties are indeed less likely to hold professional jobs in the 1994 and 2000 rounds of the NLSY. The individuals in our sample are drawn from 150 of the approximately 400 Midwestern counties that were incorporated and had attracted foreign migrants by 1860. While it thus seems unlikely that a few outlying counties are driving the results, we nevertheless report nonparametric estimates of the relationship between occupational choice and historical fractionalization in Figure 3. Less than 5 percent of the observations in Table 6, Columns 1-2 are drawn from counties with fractionalization below 0.4 and so to avoid smoothing the data too much we restrict the sample in Figure 3 to individuals from counties in the 0.4-0.8 range. We see that the probability that the individual is a professional declines

steadily with fractionalization, both in 1994 and in 2000, verifying the robustness of our results.[11]

Fractionalization had no effect on current economic activity in Table 4 once the 1860 population, manufacturing share, and agriculture share were included as controls. We consequently include these variables in all the regressions that follow, to control for important features of the 1860 economy that are correlated with 1860 fractionalization and could potentially affect the demand for professional labor today. Recall that the prediction for fractionalization was conditional on the size of the (ethnic) labor market in which networks were active. In principal, we would consequently want to include the number of migrant workers in the regression, but this variable is highly correlated with total population (the correlation is 0.9). Total county population in 1860 thus captures agglomeration effects as well as historical network effects.

We see in Table 6, Columns 3-4 that the coefficient on 1860 fractionalization becomes more negative and is more precisely estimated once the additional county-level controls are included in the occupational choice regressions. A one standard deviation decline in ethnic fractionalization would increase the probability of holding a professional job from 9 percent to as much as 15 percent. Individuals born in counties with a greater share of manufacturing and agricultural jobs in 1860 are also less likely to hold professional jobs, although this may simply reflect a greater demand for non-professional jobs in those counties today. Population in 1860, in contrast, has no effect on occupational choice in both the 1994 and the 2000 rounds. Finally, women and non-whites are significantly less likely to hold professional jobs in Columns 1-4.

Our interpretation of the negative fractionalization coefficient is that labor markets with many small ethnic networks were associated with greater exit costs, giving rise to endogenous restrictions on mobility. However, ethnic networks provided other forms of economic and social support in the nineteenth century. An immigrant group of sufficient size was needed to make widespread out-marriage unnecessary and to support a viable church congregation; two necessary conditions if networks of any sort were to function efficiently (Kamphoefner 1987). Consequently, what might have mattered most for the welfare of the community was not the size of the ethnic group within occupational categories but the size of the ethnic group as a whole. Table 6, Columns 5-6 consequently include ethnic fractionalization in the county, across all occupations, as an additional regressor. The coefficient

[11]The nonparametric regressions are estimated using the Epanechnikov kernel smoothing function. Although all the parametric regressions that follow will use the full set of counties, we verified that the fractionalization coefficient is unchanged when the sample is restricted to individuals drawn from counties in the 0.4-0.8 range as well as in the 0.45-0.8 range.

on this regressor is insignificant, with the remaining coefficients remaining largely unchanged. The coefficient on our labor market-based measure of ethnic fractionalization, in particular, continues to be precisely estimated and actually gets larger in (absolute) magnitude.

We complete the description of the occupational choice regressions by discussing a number of robustness tests (not reported). First, we computed ethnic fractionalization in 1860 with men only. Both men and women participated in the workforce in the nineteenth and early twentieth centuries, with ethnic networks channelling women into jobs as well. Bodnar (1980), for example, cites a 1930 study of two thousand foreign-born women, most of whom reported that they had secured their first job through social connections. He reports similar recruitment patterns among hosiery loopers and among Italian women in New York city and Buffalo around the same time. Nevertheless, we might expect labor networks to have been organized along gender lines within ethnic groups, with male networks occupying a dominant position in the labor market and in the communities they were drawn from. The coefficient on the alternative fractionalization measure (computed using men alone) continues to be negative and significant, both in 1994 and 2000, although it is smaller in size.[12]

Second, we replaced 1860 fractionalization with 1850 and 1870 fractionalization (separately). The coefficient continues to be negative (and significant at the 5 percent level in 1850), but it declines in size, emphasizing the importance of the time of initial settlement and rapid growth in determining future outcomes. Third, we replaced total population with the migrant population and, separately, with the workforce, in the occupational choice regression. These variables are all highly correlated, and not surprisingly the results were unchanged. Fourth, we included the share of migrants in the workforce, and the interaction of this variable with ethnic fractionalization, as additional regressors. This specification allows for the possibility that ethnic fractionalization has a larger effect in counties where migrants made up a larger share of the total population. However, both these variables have an insignificant effect on occupational choice, leaving the (uninteracted) fractionalization coefficient unchanged.

Fifth, we expanded the set of professional occupations by including individuals assigned to the

[12]We include all individuals who report being employed, regardless of their age, when computing the fractionalization statistic since there were no age restrictions on employment at that time. Apart from the 11 broad occupational categories in Table 2, some individuals in the census were also assigned to an undefined occupational category. Women were disproportionately represented in this category, which presumably covers home production and other informal activities (handicrafts played an important role in the economy at this time). We assume that networks were less relevant in this category and assign it zero weight when computing ethnic fractionalization in all the regressions that we report. Nevertheless, we verified that assigning a weight to this category based on its share of the migrant workforce in the county had no effect on the estimated fractionalization coefficient in the occupational choice regression.

"managers and administrators" code [245] within the broader Managers and Administrators, except Farm category [201-245]. Most of the specific occupations listed in this broad category do not conform to our definition of a professional occupation and the "managers and administrators" classification does not provide much information on the actual nature of jobs that it covers. However, 74.7 percent of the individuals assigned to that category migrated out of the county of their birth, which is close to the level of migration for the other occupations that we include in the professional category. We consequently included managers and administrators [245] in the expanded professional classification, increasing the share of professionals in the sample to 17.7 percent. Although these results need to to interpreted with caution since the additional professional occupations are selected on the basis of an outcome (migration) rather than their fundamental characteristics, we nevertheless verified that the fractionalization coefficient in the occupational choice regression with this expanded definition of professional activity remained negative and significant. Finally, we ran the occupational choice regression in other years – 1993, 1996, 1998 – generating a pattern of fractionalization coefficients very similar to the point estimates obtained in 1994 and 2000.

4.3 Historical Fractionalization and Outcomes Related to Occupational Choice

An individual born in a county with higher ethnic fractionalization in 1860 is less likely to select into a professional occupation in 1994 and 2000. Our interpretation of this result is that individuals born in high fractionalization counties identify strongly with their local communities and so wish to avoid the spatial mobility that comes with professional jobs. The regression results reported in Table 7, Columns 1-2 indicate that individuals from high fractionalization counties are indeed significantly less likely to migrate from the county of their birth. On average, around 58 percent of the individuals in the sample migrate from the county of birth. The point estimates indicate that a one standard deviation increase in fractionalization reduces migration by 8 percentage points (a 14 percent decline). Among the other regressors, none of the 1860 county-level controls significantly affect migration, but Whites are significantly more likely to move. Age is also (mechanically) positively associated with migration.

Does the effect of fractionalization on occupational choice and migration that we have just described have economic consequences? The results in Table 7, Columns 3-4 indicate that while individuals born in counties with a greater share of manufacturing and agriculture in 1860 are significantly more likely to hold a job by 2000, employment levels do not vary significantly with ethnic fractionalization in 1994

or 2000. Whites and males are, not surprisingly, significantly more likely to be employed, although the importance of these individual characteristics declines over time.

In contrast with the results for employment, the income regressions reported next in Table 7, Columns 5-6 indicate that high fractionalization is associated with significantly lower income (in 2000), conditional on being employed. Average annual income in 2000 was 33,000 dollars, and so our estimates indicate that a one standard deviation increase in fractionalization would have reduced income by 2,300 dollars (a 7 percent decrease). A greater share of manufacturing and agriculture in 1860 is associated with lower income in 2000, with Whites and males earning significantly more in 1994 and 2000.

4.4 Alternative Explanations for Variation in Occupational Choice

The results in Table 6 and Table 7 indicate that fractionalization affects occupational choice, with significant consequences for migration and income. Our view is that these results are driven by variation in local identity across counties, but could variation in innate ability or the demand for professional jobs generate these results instead? High fractionalization counties had superior transportation infrastructure in 1860 and so were most likely growing relatively rapidly at that time. It is entirely possible that particular types of migrants were attracted to those counties, although we would expect migrants drawn to areas with many competing ethnic groups to have been (positively) self-selected on ability. Given the high rates of inter-county migration, even in the Midwest, we would in any case expect few families to have maintained an unbroken line of descent to the present day in the same county. Consistent with this view, we see in Table 8, Column 1 that fractionalization, and all the 1860 variables for that matter, have no effect on AFQT scores. The same result is obtained with high school completion as the dependent variable. In contrast, individual characteristics have a strong effect on AFQT scores (but not on high school completion).

Even if innate ability does not vary systematically with fractionalization, we might still expect individuals from high fractionalization counties to invest less in the college education that is necessary to obtain a professional job. Individuals born in high fractionalization counties are less likely to have completed college, as expected, although the effect is not significant at conventional levels. They are also less likely to have attended college out of state, conditional on having attended, presumably because they prefer to remain close to home and because the quality of the education they receive is less relevant for their future careers. In general, there is no indication that individuals born in high

fractionalization counties are intrinsically less capable of securing professional jobs. The discussion that follows will consequently focus on differences in the demand for professional jobs across counties as the most credible alternative explanation for the fractionalization effect.

As discussed in the theoretical section, an important feature of the model of occupational choice with local identity is that the supply of professionals will fall short of the expected demand in high fractionalization counties. Non-professionals must then flow out of those counties, while professionals flow in, to clear the labor market. The second empirical prediction of the model is consequently that historical fractionalization in the county of residence should have no effect on occupational choice once the labor market has cleared. In contrast, if the current demand for professional jobs is lower in high fractionalization counties then individuals born *and* residing in those counties should be less likely to hold professional jobs.

Table 9, Columns 1-4 separate individuals born in a county into those who move and those who stay (by 1994 and 2000). Matching the results in Table 6, which are based on all individuals born in the county, the fractionalization coefficient in Columns 1-4 is negative for movers and stayers in 1994 and 2000. The coefficient for the two groups is comparable in 1994, but it is substantially larger – roughly twice as large – for the out-migrants in 2000. Some of the mismatch between the supply and the demand for labor in the high fractionalization counties is evidently reduced through the exit of non-professionals. Labor flows in the opposite direction reduce this mismatch further in Columns 5-6; although the fractionalization coefficient for the in-migrants is insignificant in 1994, it is positive and significant in 2000.

The labor flows that we have described all work towards reducing the mismatch between the supply and the demand for professional and non-professional labor across counties. Measuring historical fractionalization in the county of residence in Table 9, Columns 7-8 we see that the fractionalization coefficient is no longer significant and actually switches signs in 2000. It is only individuals born in high fractionalization counties who are less likely to be professionals. Once the market has cleared, individuals residing in those counties are just as likely to hold professional jobs.[13]

[13]The second prediction of the model that fractionalization in the county of residence should be uncorrelated with occupational choice is based on the assumption that the entire labor market can clear. The test that we implement, instead, is based on a restricted NLSY cohort. Labor supply and demand need no match within specific cohorts and it is reassuring that we nevertheless successfully test this prediction of the model.

5 Religion and the Persistence of Identity

In a recent paper, Akerlof and Kranton (2005) describe how firms and other economic organizations can instill a sense of identity or loyalty among their workers to solve agency problems. In our setting, the natural organization to instill such a sense of identity among workers in the network would have been the local church. The church was among the first institutions to be established when immigrants arrived in an area (Hoerder 1991, Barton 1975, Bodnar, Simon, and Weber 1982). The church congregation provided many forms of mutual assistance including credit, insurance, job referrals, business information, and social support (Gjerde 1985, 1997, Alexander 1991). Indeed, it has been argued that immigrants participated in church communities to benefit from these economic and social services, instead of being drawn to the church by a particular belief or ideology (Bodnar 1985).

Despite these material attractions, exit from the church and the local labor market was always a threat in nineteenth century America, especially in areas that had been recently settled. There was a tension between a "folk society" centered around the church parish and a competitive "individualistic society" with its many opportunities (Ostergren 1976). In addition, migration to the United States, and subsequent internal migration, was often organized under the auspices of relatively small communities from the origin country. These migrant networks were too small to maintain a viable church congregation and labor network when they first arrived in an area and so churches in the Midwest typically brought together many regional groups with the same national origin. Inter-regional conflict was consequently common, and there was always the threat that groups within the church would move again to take advantage of new opportunities that became available elsewhere (Gjerde 1997); as Kamphoefner (1987) notes, the more homogeneous was a congregation's origins, the less out-migration it was likely to experience. Under these circumstances, the church had to make a special (costly) effort to instill a sense of loyalty among its members that transcended narrower individual, kin, and regional affiliations.

If local identity played a role in reducing exit in the nineteenth century, how and why did it persist long after the ethnic labor market networks it supported ceased to be salient? Our explanation for this persistence is based on the observation that churches continue to provide important forms of *social* support to their members. Church activities include Sunday school service, pot-lucks, informal home parties, and food, visits, and other forms of support when members of the congregation are ailing or infirm. As Iannaccone (1998) notes, there are complementarities associated with these activities; if

an individual expects other members of the church to reciprocate in the future, he is more willing to contribute time and effort today. A strong sense of local identity in the community would increase the individual's confidence that the rest of the congregation would not renege on their social obligations in the future.

Although families in a county can rarely claim an unbroken line of descent from migrating European ancestors in 1860, church congregations are large enough to have maintained sufficient overlap from one generation to the next. Take the case of a church in a high fractionalization county. The first-generation members of that church in 1860 would have been instilled with a strong sense of identity primarily to maintain the ethnic labor networks that were active at that time. This strong identity would have resulted in a higher level of social support, and a better functioning church, in that first generation. Now suppose that labor networks and ethnic affiliation ceased to be relevant from the second generation onwards, but that social support continued to be useful. Given the complementarity in inputs, as long as a sufficiently large fraction of the first generation migrants overlapped with the next generation, they would have been particularly interested in ensuring that the newcomers into the congregation were also instilled with a strong sense of identity to maintain the high level of social support that was being sustained in equilibrium. A strong sense of local identity could have been sustained in this way over many overlapping generations, as the following description of life in rural Iowa would suggest:

"The farm of Robert Heinz, for example, has been in the family since the late 1800s, and his family has long played a role in the local community and in its Lutheran church. His wife, a bookkeeper for the local feed mill, currently serves as president of the PTA, just as his mother did when he attended the high school.

Heinz sits on the boards of a farmer cooperative and regional development council, supports the local FFA by attending functions with his two sons, and serves as a church trustee. His energetic role in the community is mirrored by the activity of his boys ... they are active in the church's lively youth group and take part in a school-sponsored organization of the Future Farmers of America." (Elder and Conger, 2000: 46).

The preceding discussion suggests that stronger identity is associated with a higher level of church inputs and better functioning local religious institutions in equilibrium. It then follows that counties with stronger identity should be associated with greater religious participation, which could extend over many generations. This prediction distinguished our model, with identity as a coordinating

device, from the model of religious cults proposed by Iannaccone (1992) in which self-sacrifice and strict norms of behavior are used to screen out free-riders and ensure higher levels of participation among those that remain. While this alternative mechanism may induce high levels of participation at the intensive margin, it results in small congregations and is typically short-lived.

The relationship between ethnic fractionalization, identity, and religious participation can be tested with data from the Census of Religious Bodies (CRB), which has been conducted at roughly ten-year intervals from 1860 to 2000. This census 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 1906, 1916, 1926, and 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.

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. Despite these changes in the management of the CRB and the measure of religious participation, we uncover clear changes in the mix of denominations as well as the effect of 1860 fractionalization on religious participation over time.

Table 10 reports changes in the mix of denominations and overall participation rates in our Midwestern counties over the 1860-2000 period. To take account of the fact that the measure of participation was changing over time, we report statistics in the first and the last census-year that each measure was used. Thus, participation is measured by the number of church seats from 1860 to 1890, by the number of members from 1890 to 1952, and by the number of adherents from 1972 to 2000.[14] The participation rate is then computed as the number of participants divided by the contemporaneous population in the county. 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 Lutherans grow rapidly in popularity over the 1860-1890 period and the 1890-1952 period, remaining stable thereafter. In contrast, the Methodists and the Presbyterians decline steadily over time. There is no clear trend among the Baptists and the Catholics.

The inability of the Baptist church to increase its share of church participants contrasts with the

[14] Although the number of members was also collected in the 1972-2000 census rounds, this statistic is not available for Catholics, a major denomination in our Midwestern counties, in the 1972-1990 rounds. Despite this change in the measure of participation, the mix of denominations is fairly similar in 1952 and 1972 in Table 10.

surge in popularity of this denomination elsewhere in the United States, as documented by Finke and Stark (1992). The Midwest stands apart from the rest of the country in that the "traditional" denominations, particularly the Catholics and the Lutherans, continue to dominate and we will connect this observation to the link between migrants churches, local identity, and religious participation below. Even without the Baptist surge, religious participation increased steadily over time in the Midwest, rising steadily over the 1860-1890 and 1890-1952 periods. Based on the statistics in the most recent 1972-2000 period, 55-60 percent of the population in our Midwestern counties report that they are church adherents.

The relationship between identity and the social institution – the church – that supports this cultural trait must be measured at the local level. We consequently proceed to estimate the relationship between ethnic fractionalization in 1860 and religious participation in the county in each round of the CRB. Everything else equal, we expect more fractionalized counties to have stronger local identities and better functioning churches, which would be reflected in greater religious participation. However, the type of migrant and the individual incentive to participate in the church would have varied with the nature of the local economy at the time of initial settlement. The relationship between fractionalization and religious participation would thus have been ambiguous in those early years. Once fractionalization ceased to be correlated with the nature of the local economy, we would expect its positive effect on local identity to show up unambiguously.

The religious participation regression includes the same set of county-level controls as the occupational choice regression. Because the regression is estimated over many census years, we simply report the fractionalization coefficient, together with the 95 percent confidence band, in Figure 4. The coefficient is negative in the early census years but not precisely estimated.[15] As discussed, the theoretical predictions must really be tested many years after the period of initial settlement and, indeed, we see that the fractionalization coefficient is positive and significant by 1916. Thereafter, it grows steadily larger and remains statistically significant all the way through till 2000. This result contrasts with the absence of any effect of 1860 fractionalization on current economic activity in Table 4, once the county controls were included, emphasizing the special link between historical fractionalization, local identity, and the church. The mutually reinforcing relationship between identity and the church, in

[15]The 1860 fractionalization coefficient with standard error in parentheses is -0.66(0.39). This outlying coefficient is omitted from Figure 4 to clarify changes in the fractionalization coefficient over time. As noted, church seats and the number of members are both available in the 1890 CRB. We use the first statistic to measure church participation in Figure 4, and in Figure 5 below, because it is more in line with trends in the fractionalization coefficient over time.

turn, allows historical fractionalization to affect economic decisions many generations in the future.[16]

Our interpretation of the relationship between ethnic fractionalization in 1860 and subsequent religious participation is based on the idea that migrant churches supported migrant networks at the time of settlement, with the local identity that they engendered maintaining the strength of their congregations over time. If this interpretation is accurate, then migrant denominations such as the Lutherans and the Catholics should be especially popular in high fractionalization counties where local identity is strongest.[17] We noted in Table 10 that these denominations dominate in our Midwestern counties. The additional prediction from our framework is that the share of Lutherans and Catholics among church participants should be higher in high fractionalization counties and that this relationship should be maintained over time, matching the persistent fractionalization effect reported in Figure 4.

We now regress the share of the migrant denominations among all church participants in each census year on ethnic fractionalization and the usual county controls, measured in 1860. As above, we report only the fractionalization coefficient, with the 95 percent confidence band, to preserve space. Matching the prediction from the theoretical framework, the fractionalization coefficient in Figure 5 is positive and significant in all CRB years, with a steady upward trend, particularly after 1906. Although we do not report results separately by denomination, Baptists fair particularly poorly in high fractionalization counties. As documented by Finke and Stark (1992), this decentralized denomination has grown rapidly in popularity over time throughout the United States. The migrant denominations are presumably strong enough to withstand the challenge of this upstart sect in the Midwest, especially in the high fractionalization counties, emphasizing the resilience of the local church in those counties.

6 Conclusion

This paper draws a connection between ethnic labor market networks in the American Midwest when it was first being settled, the local identity that emerged endogenously to support these networks and then persisted over many generations, and occupational mobility today. Individuals born in counties

[16] Although not reported, the coefficient on 1860 agriculture share is also positive and significant in the religious participation regressions and grows larger over time. Unlike ethnic fractionalization, which soon ceased to be directly relevant, recall from Table 4 that agricultural counties in 1860 remained disproportionately agricultural in 1990. It is well known that farming communities tend to be more religious and community-oriented and so the persistent effect of the 1860 agriculture share may simply reflect the persistence of agricultural activity in particular areas over time.

[17] Although the social history literature emphasizes the role of the Lutheran and Catholic churches in supporting migrant networks, this relationship between ethnic background and church denominations cannot be directly derived from the CRB. We did, however, verify that counties with a greater proportion of migrants in 1860 had a disproportionately large share of Catholics and Lutherans in that census year.

with greater ethnic fractionalization in 1860, which we expect to be associated with stronger local identity, are significantly less likely to hold professional jobs, which come with greater geographical mobility, in 2000.

The results in this paper speak directly to the debate on "institutions" versus "culture" as determinants of growth. We find that local culture generates significant variation in occupational choice across what appears to be a relatively homogenous region – the American Midwest – and we conjecture that these cultural effects might be even larger across countries with very different histories. At the same time, culture cannot be sustained without institutional support. Social institutions such as the church and the family help sustain cultural traits, which in turn keep these institutions alive. As long as these institutions continue to be useful, culture can persist long after the economic circumstances that gave rise to both the social institutions and the cultural traits have ceased to be relevant. The economics literature has focussed much of its attention on economic and political institutions. Our results from this research, as well as previous work on the role of social institutions in directly shaping economic decision and outcomes (Munshi and Rosenzweig 2005, 2006), indicate that these institutions might have important effects on growth as well.

References

[1] Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, 91(5): 1369-1401.

[2] Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2002. Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics*, 117(4): 1231-1294.

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

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

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

[6] 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.

[7] Banerjee, Abhijit and Andrew Newman. 1993. Occupational Choice and the Process of Development, *Journal of Political Economy* 101(2): 274-298.

[8] Barro, Robert J. and Rachel M. McCleary. 2003. Religion and Economic Growth Across Countries, *American Sociological Review* 68: 760-781.

[9] Barton, Josef J. 1975. *Peasants and Strangers: Italians, Rumanians, and Slovaks in an American City, 1890-1950*. Cambridge, Massachusetts: Harvard University Press.

[10] 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.

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

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

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

[14] 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.

[15] Bowles, Samuel. 1998. Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions, *Journal of Economic Literature* 36: 75-111.

[16] 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.

[17] Fernandez, Raquel and Alessandra Fogli. 2007. Culture: An Empirical Investigation of Beliefs, Work, and Fertility. New York University Department of Economics mimeograph.

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

[19] 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.

[20] Galor, O. and J. Zeira. 1993. Income Distribution and Macroeconomics, *Review of Economic Studies*, 60: 35-52.

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

[22] Gerson, Kathleen, C. Ann Stueve, and Claude S. Fischer. 1977. Attachment to Place, in *Networks and Places: Social Relations in the Urban Setting*, eds. Claude S. Fischer et al., New York: The Free Press.

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

[24] 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.

[25] Gjerde, Jon. 1997. *The Minds of the West: Ethnocultural Evolution in the Rural Middle West, 1830-1917*. Chapel Hill: University of North Carolina Press.

[26] 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.

[27] 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.

[28] Hummon, David M. 1990. *Commonplaces: Community Ideology and Identity in American Culture*. Albany: State University of New York Press.

[29] Hunter, Albert. 1975. The Loss of Community: An Empirical Test Through Replication, *American Sociological Review* 40(5): 537-552.

[30] Huntington, Samuel P. 1996. *The Clash of Civilizations and the Remaking of World Order*. New York: Simon and Schuster.

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

[32] Iannaccone, Laurence R. 1992. Sacrifice and Stigma: Reducing Free-Riding in Cults, Communes, and Other Collectives, *Journal of Political Economy* 100(2): 271-297.

[33] Iannaccone, Laurence R. 1998. Introduction to the Economics of Religion, *Journal of Economic Literature* 36: 1465-1496.

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

[35] Kim, Sukkoo. 2007. Immigration, Industrial Revolution and Urban Growth in the United States, 1820-1920: Factor Endowments, Technology and Geography, *NBER Working Paper* No. 12900.

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

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

[38] Munshi, Kaivan and Mark Rosenzweig. 2005. Why is Mobility in India so Low? Social Insurance, Inequality and Growth, *BREAD Working Paper* No. 092.

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

[40] Ostergren, Robert C. 1976. *Rätvik to Isanti: A Community Transplanted*. Ph.D. dissertation, University of Minnesota.

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

[42] Tabellini, G. 2005. Culture and Institutions: Economic Development in the Regions of Europe, CESifo Working Paper No. 1492.

[43] 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.

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

Figure 1: Early Growth in the Midwest

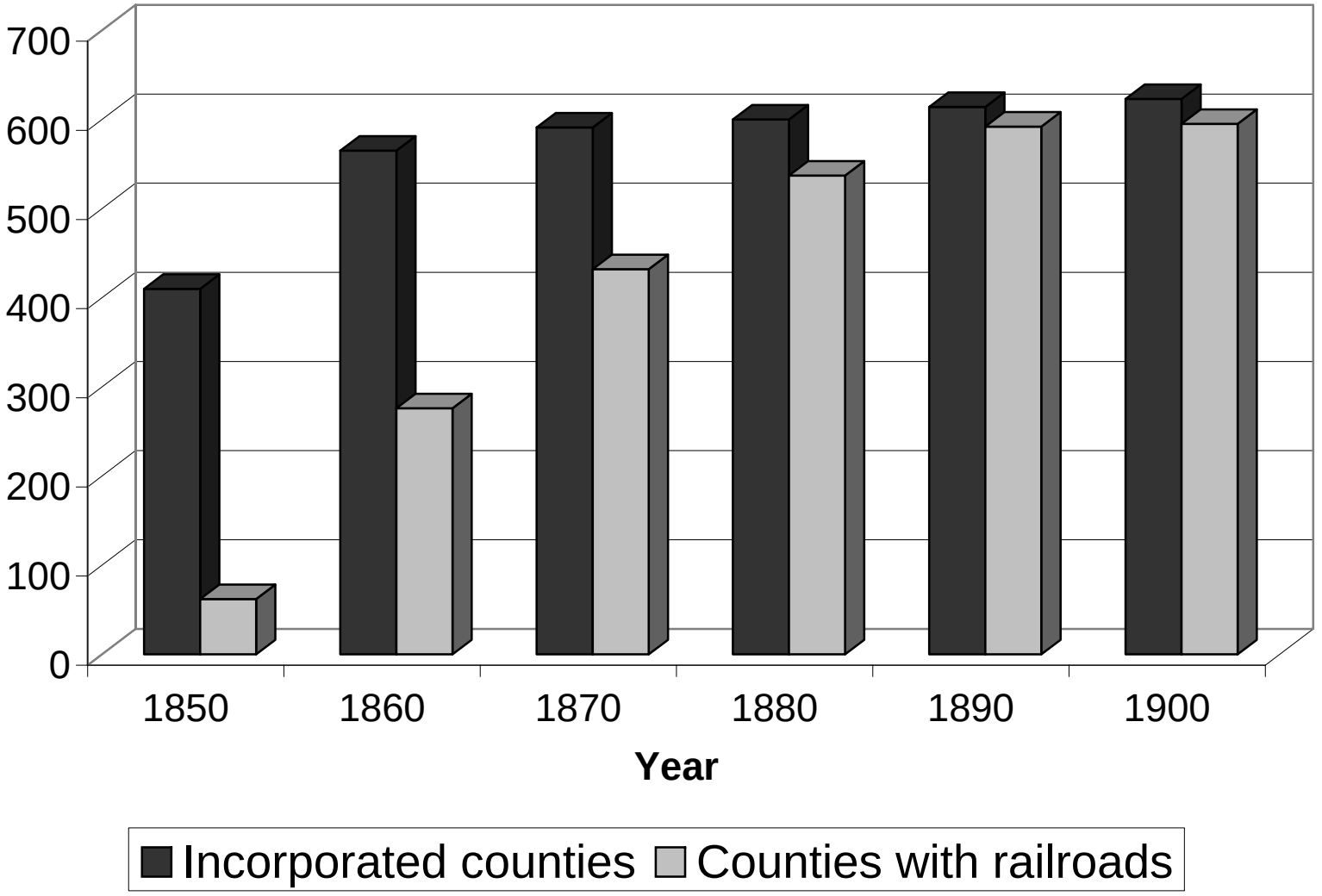

Figure 2: Ethnic Fractionalization in 1860

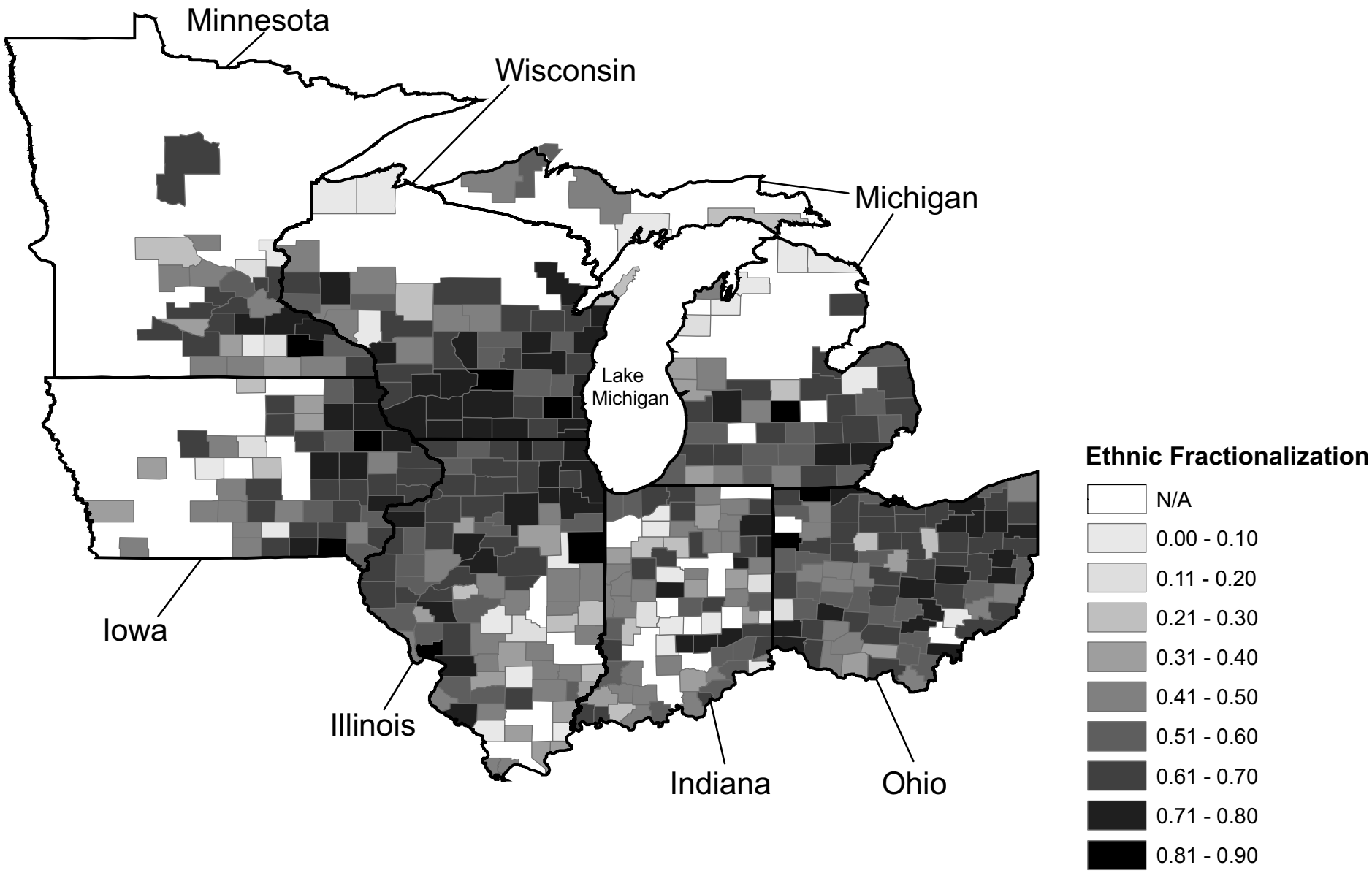

Figure 3: Historical Fractionalization and Current Occupational Choice

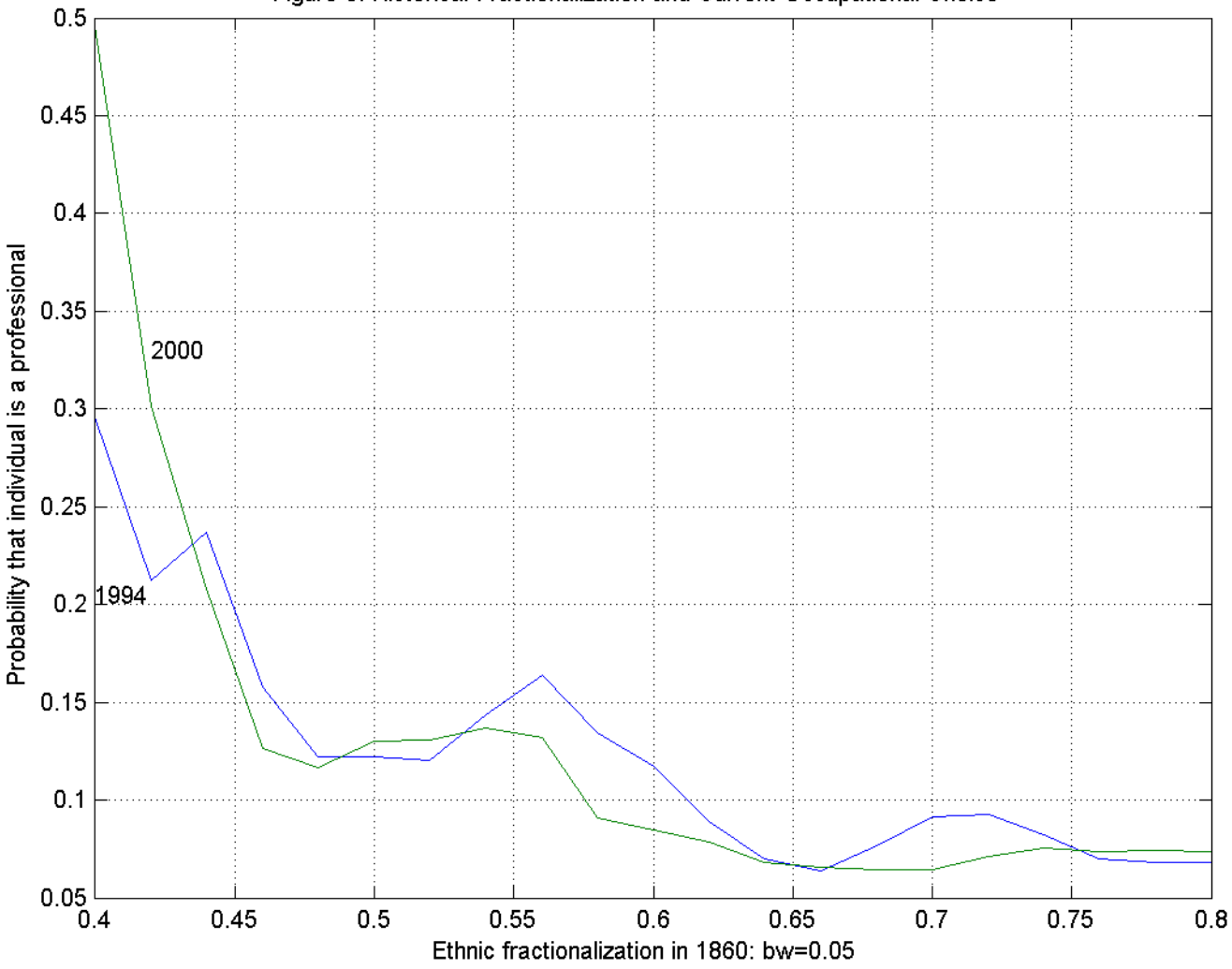

Figure 4: Religious Participation

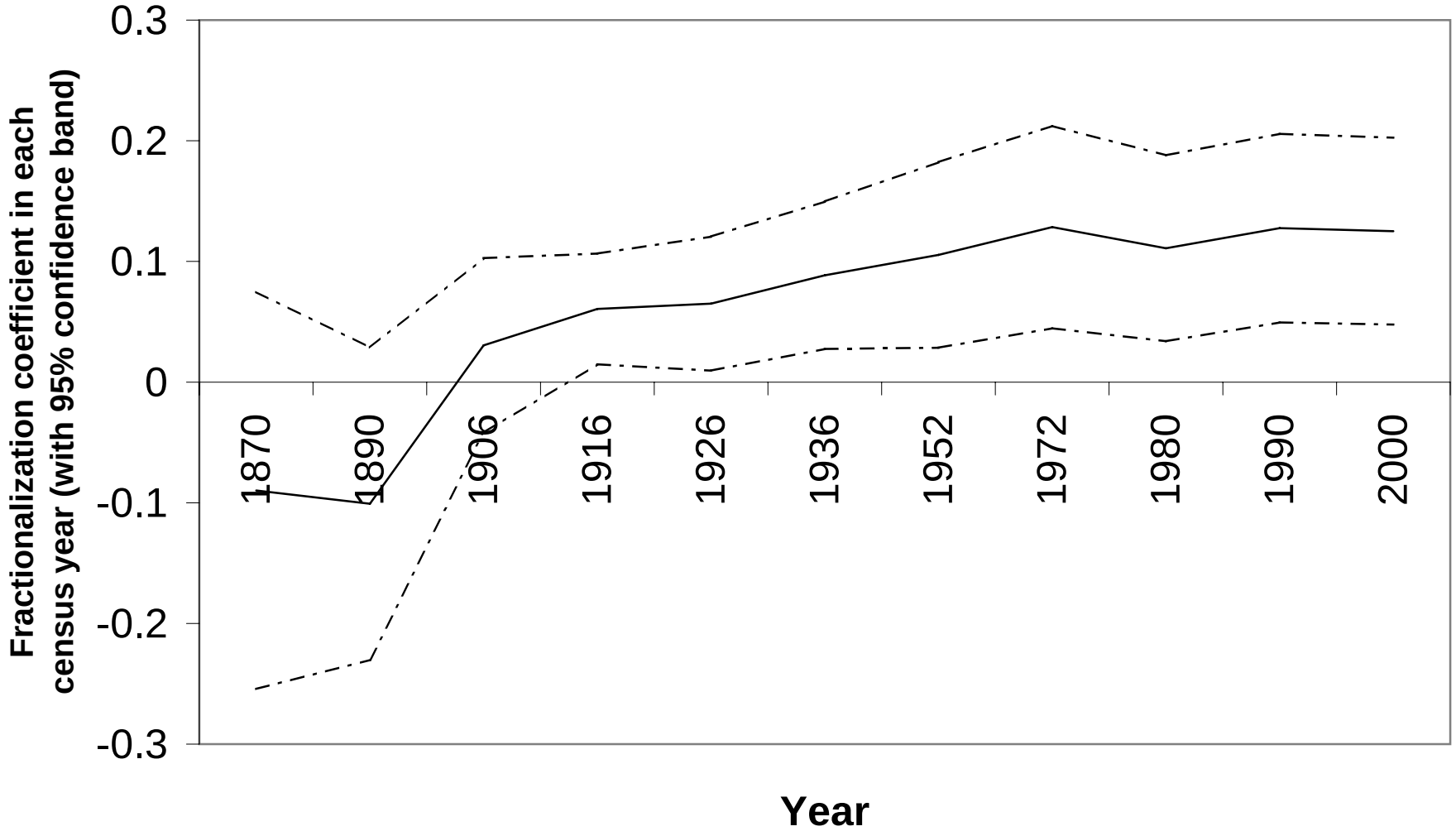

Figure 5: Share of Migrant Denominations

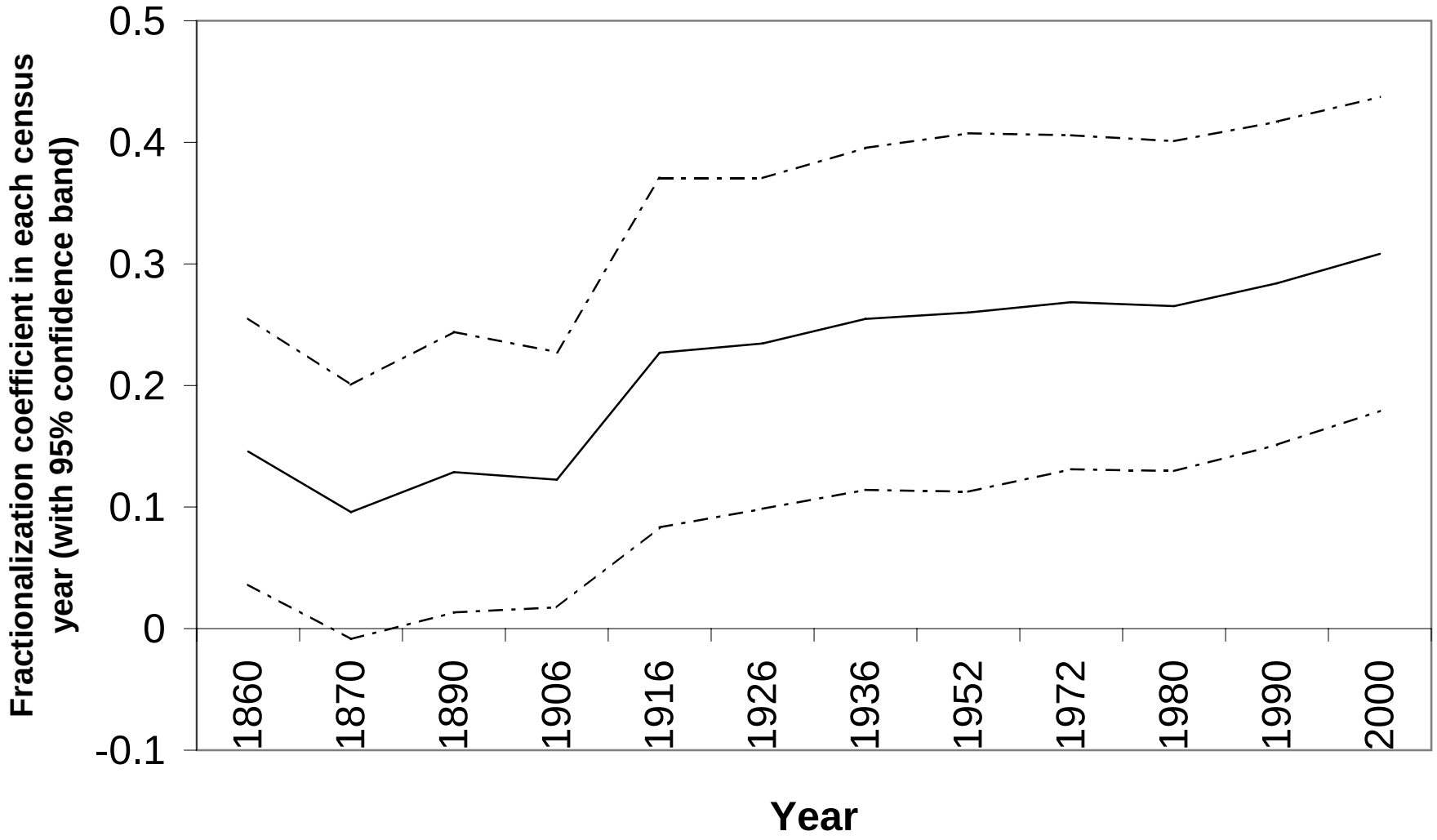

Table 1: Ethnic Distribution, 1860-1900

Census year:	1860	1880	1900
	(1)	(2)	(3)
Scandinavia			
Danish	0.01	0.02	0.02
Finish	0.00	0.00	0.01
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			
Belgian	0.00	0.01	0.01
Dutch	0.01	0.01	0.01
French	0.03	0.02	0.01
German	0.32	0.37	0.41
Italian	0.00	0.00	0.01
Swiss	0.02	0.02	0.02
Eastern Europe			
Czech	0.00	0.01	0.02
Hungarian	0.00	0.00	0.00
Polish	0.00	0.01	0.02
USSR	0.00	0.00	0.01
Other	0.14	0.05	0.04
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 and 1990

Year:	1860				1990			
Dependent variable:	fractionalization	manufacturing share	agriculture share	population	manufacturing share	agriculture share	population	population density
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Railroad through county, 1860	0.049	0.001	-0.013	0.102	0.013	-0.036	0.758	0.153
	(0.022)	(0.008)	(0.013)	(0.012)	(0.008)	(0.006)	(0.131)	(0.027)
Distance to canal, 1890	-0.668	0.169	-0.034	-0.469	-0.216	-0.097	-1.944	-0.385
	(0.191)	(0.097)	(0.125)	(0.102)	(0.066)	(0.039)	(2.694)	(0.380)
Distance to Great Lakes harbor	-0.252	-0.066	0.100	-0.136	-0.223	0.128	-4.211	-0.739
	(0.073)	(0.029)	(0.048)	(0.056)	(0.025)	(0.020)	(1.525)	(0.213)
Observations	401	401	401	401	401	401	401	401

Note: Robust standard errors are 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 and population density is measured in thousands per square mile.

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.

Table 4: Fractionalization in 1860 and County Characteristics in 1990

Year:	1990							
Dependent variable:	agriculture share	manufacturing share	population	population density	agriculture share	manufacturing share	population	population density
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fractionalization, 1860	-0.022	0.024	2.590	0.471	0.003	0.018	-0.269	0.031
	(0.014)	(0.019)	(0.792)	(0.113)	(0.015)	(0.020)	(0.854)	(0.094)
Manufacturing share, 1860	--	--	--	--	-0.090	-0.056	1.636	0.445
					(0.038)	(0.053)	(1.063)	(0.196)
Agriculture share, 1860	--	--	--	--	0.136	0.057	-0.962	-0.400
					(0.027)	(0.037)	(0.746)	(0.175)
Population, 1860	--	--	--	--	-0.087	0.030	11.716	1.778
					(0.031)	(0.026)	(5.372)	(0.518)
Observations	437	437	437	437	437	437	437	437

Note: Robust standard errors are 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 and population density is measured in thousands per square mile.

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.

Table 5: NLSY79 Descriptive Statistics

	year	
	1994	2000
	(1)	(2)
Professional	0.09	0.09
	(0.29)	(0.28)
Migrated out of county of birth	0.56	0.59
	(0.01)	(0.01)
Employed	0.81	0.92
	(0.01)	(0.01)
Income	27.80	33.06
	(0.51)	(0.58)
White	0.79	0.79
	(0.01)	(0.01)
Female	0.50	0.51
	(0.01)	(0.01)
Age	33.36	39.38
	(0.05)	(0.05)

Note: Standard errors are in parentheses.
Professional occupations are relevant codes in the Professional, Technical and Kindred Workers category.
All variables except income and age are binary.
Income is measured in thousands of dollars (2000).

Table 6: Occupational Choice

Dependent variable:	professional					
Year:	1994	2000	1994	2000	1994	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Fractionalization, 1860	-0.146	-0.163	-0.253	-0.316	-0.397	-0.367
	(0.083)	(0.088)	(0.096)	(0.109)	(0.163)	(0.126)
White	0.031	0.045	0.047	0.059	0.045	0.058
	(0.022)	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)
Female	-0.065	-0.023	-0.064	-0.021	-0.064	-0.020
	(0.020)	(0.017)	(0.020)	(0.016)	(0.020)	(0.016)
Age	0.0003	0.003	0.001	0.003	0.001	0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Manufacturing share, 1860	--	--	-0.213	-0.521	-0.119	-0.492
			(0.172)	(0.213)	(0.185)	(0.228)
Agriculture share, 1860	--	--	-0.272	-0.431	-0.218	-0.414
			(0.144)	(0.160)	(0.142)	(0.173)
Population, 1860	--	--	0.018	0.011	0.031	0.016
			(0.017)	(0.025)	(0.019)	(0.027)
Gross ethnic fractionalization, 1860	--	--	--	--	0.213	0.072
					(0.173)	(0.145)
Observations	1209	1122	1209	1122	1209	1122

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.

Gross ethnic fractionalization is measured without regard to occupations.

Occupational concentration is measured using the Herfindahl index multiplied by 1000.

Denominational concentration is measured using the Herfindal index.

Professional is a binary variable indicating whether the individual is employed in a professional occupation.

Professional occupations are relevant codes in the Professional, Technical, and Kindred Workers category.

Table 7: Outcomes Related to Occupational Choice

Dependent variable:	migrated		employed		income	
Year:	1994	2000	1994	2000	1994	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Fractionalization, 1860	-0.388	-0.414	0.114	0.037	1.659	-11.722
	(0.155)	(0.138)	(0.108)	(0.065)	(4.804)	(6.183)
Manufacturing share, 1860	-0.071	-0.242	0.078	0.362	-1.195	-26.379
	(0.305)	(0.260)	(0.230)	(0.136)	(10.268)	(12.231)
Agriculture share, 1860	-0.171	-0.317	0.031	0.253	-8.197	-21.016
	(0.311)	(0.222)	(0.164)	(0.107)	(7.808)	(9.359)
Population, 1860	0.021	0.065	-0.015	0.013	-0.210	0.449
	(0.039)	(0.040)	(0.026)	(0.015)	(1.041)	(1.691)
White	0.232	0.284	0.153	0.043	5.773	4.791
	(0.048)	(0.052)	(0.024)	(0.021)	(1.850)	(1.567)
Female	-0.014	-0.014	-0.155	-0.059	-11.986	-14.476
	(0.017)	(0.024)	(0.021)	(0.015)	(1.122)	(0.988)
Age	0.013	0.015	0.001	0.003	0.764	0.211
	(0.005)	(0.005)	(0.004)	(0.004)	(0.208)	(0.312)
Observations	1598	1437	1614	1332	1251	1122

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

Fractionalization is one minus the 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.

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

Migrated is a binary variable that indicates whether the individual resides outside the county of birth.

Employed is a binary variable that indicates whether the individual currently holds a job.

Income is measured in thousands of dollars (2000).

Table 8: Educational Performance

Dependent variable:	AFQT score	high school completion	college completion	college out of state
	(1)	(2)	(3)	(4)
Fractionalization, 1860	-0.535	0.016	-0.095	-0.228
	(7.457)	(0.100)	(0.104)	(0.145)
Manufacturing share, 1860	11.570	0.197	0.003	-0.420
	(17.828)	(0.160)	(0.241)	(0.311)
Agriculture share, 1860	-3.015	-0.017	-0.261	-0.535
	(14.672)	(0.143)	(0.193)	(0.308)
Population, 1860	-2.769	-0.052	-0.018	-0.029
	(1.981)	(0.023)	(0.023)	(0.030)
White	22.608	0.010	0.083	0.008
	(1.626)	(0.024)	(0.018)	(0.045)
Female	-2.728	0.000	-0.039	-0.002
	(1.248)	(0.017)	(0.020)	(0.024)
Age	2.880	0.006	0.005	0.025
	(0.332)	(0.003)	(0.004)	(0.006)
Observations	2187	2023	2023	1154

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.

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.

College out of state is a binary variable indicating whether the individual attended college out of state of birth.

Table 9: Labor Flows

Dependent variable:	professional							
Sample:	stayers		out-migrants		in-migrants		current residents	
Year:	1994	2000	1994	2000	1994	2000	1994	2000
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fractionalization, 1860	-0.182	-0.153	-0.213	-0.343	-0.016	0.205	-0.089	0.091
	(0.105)	(0.104)	(0.124)	(0.146)	(0.094)	(0.088)	(0.074)	(0.067)
Manufacturing share, 1860	-0.243	-0.078	-0.197	-0.674	-0.431	-0.098	-0.367	-0.046
	(0.194)	(0.225)	(0.221)	(0.260)	(0.167)	(0.237)	(0.131)	(0.172)
Agriculture share, 1860	-0.225	-0.159	-0.277	-0.480	-0.254	0.045	-0.248	-0.001
	(0.147)	(0.206)	(0.192)	(0.237)	(0.147)	(0.139)	(0.114)	(0.104)
Population, 1860	0.041	-0.021	-0.006	0.023	0.048	0.008	0.041	-0.012
	(0.023)	(0.018)	(0.028)	(0.046)	(0.034)	(0.024)	(0.017)	(0.016)
White	0.054	0.026	0.016	0.061	0.086	0.062	0.072	0.048
	(0.023)	(0.021)	(0.045)	(0.025)	(0.021)	(0.024)	(0.013)	(0.020)
Female	-0.026	0.028	-0.097	-0.057	-0.095	-0.076	-0.065	-0.031
	(0.020)	(0.020)	(0.026)	(0.022)	(0.020)	(0.021)	(0.015)	(0.016)
Age	0.004	-0.005	-0.005	0.006	-0.008	0.001	-0.002	-0.001
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)
Observations	557	479	648	637	648	651	1205	1130

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.

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

Professional is a binary variable indicating whether the individual is employed in a professional occupation.

Professional occupations are relevant codes in the Professional, Technical, and Kindred Workers category.

"Stayers" are individuals whose county of birth and county of residence are the same.

"Out-migrants" are individuals who reside in a county outside of their county of birth.

Independent variables for "out-migrants" are measured in county of birth.

"In-migrants" are individuals who reside in a county outside of their county of birth.

Independent variables for "in-migrants" are measured in county of residence.

Table 10: Distribution of Denominations and Religious Participation, 1860-2000

Measure of religious participation:	religious participation					
	church seats		church members		church adherents	
Census year:	1860	1890	1890	1952	1972	2000
	(1)	(2)	(3)	(4)	(5)	(6)
Distribution of denominations						
Baptist	0.14	0.11	0.08	0.06	0.06	0.09
Catholic	0.11	0.13	0.29	0.27	0.30	0.33
Lutheran	0.05	0.12	0.14	0.22	0.21	0.20
Methodist	0.38	0.27	0.21	0.18	0.19	0.13
Presbyterian	0.14	0.08	0.06	0.05	0.05	0.03
Other	0.18	0.28	0.22	0.22	0.19	0.22
Total	1.00	1.00	1.00	1.00	1.00	1.00
Proportion religious	0.59	0.67	0.30	0.50	0.58	0.54

Source: Census of Religious Establishments.

Proportion religious is computed as the number of church seats divided by the population 1860-1890, the number of church members divided by the population 1890-1952, and the number of church adherents divided by the population 1972-2000.