Internal Migration and the Diffusion of Schooling in the United States*

Vasily Rusanov†

New York University

Download most recent version here.

Abstract

This paper investigates the influence of internal migration on education, and specifically, the uneven rise of school enrollments in the nineteenth-century United States. Despite having the highest average enrollment in the world, Eastern states varied greatly in terms of how many children went to school, with states in the South lagging behind those in the North. I track the locations of millions of white adults moving west using the Population Census in 1850–1880. Focusing on a sample of seven destination states, I estimate a discrete choice model to show that parents from more educated origins had a higher willingness to pay for education. I then compare the destination counties that received different mixes of migrants from the Eastern origins. Internal migrants created spillovers for local populations: parents born in the destination states were more likely to send children to school if their migrant neighbors originated from states with higher school enrollments. Detailed county-level data on school finances, which I assemble by digitizing the Census of Social Statistics, show that counties with migrants increased public spending on schools. Taking the estimated choice model and the spillover effects together, I find that tripling the distance costs of migration would have reduced the enrollment rates of natives by 1.5 percentage points. These findings suggest that internal migration mattered beyond the reallocation of labor and preference sorting.

Keywords: internal migration, schools, diffusion of technology, school funding
JEL Classification: I22, J61, N31, N91, R51

---

*I am grateful to my advisors Elena Manresa, Petra Moser, and Martin Rotemberg for their guidance and feedback. I am grateful for Paul Rhode for helping me access Social Statistics schedules for the state of Michigan. I thank James Fenske, Andreas Ferrara, Michael Gilraine, Walker Hanlon, Zahin Haque, Trevon Logan, Margaux Luflade, Robert Margo, Paul Rhode, Dmitry Sorokin, Sharon Traiberman, and Dan Waldinger for discussions and comments. Participants of seminars at NYU, UW-Madison, NBER SI, IPUMS Data-Intensive Research Conference, UEA Meetings, and EHA Meetings provided useful feedback. I gratefully acknowledge financial support from the Economic History Association's Exploratory Data and Travel Grant, the CV Starr Center for Applied Economics Data Grant, and the NSF Doctoral Dissertation Improvement Grant #2049699. Any errors are my own.

†Department of Economics, New York University. Email: vasily.rusanov@nyu.edu; Website: vasilyrusanov.com

Introduction

High rates of internal migration are a distinctive characteristic of economic development in the United States (Ferrie, 2003). In 1850, for example, 38% of the US-born whites did not live in the state where they were born, and this share even increased slightly by the end of the century. At the time, American rates of internal migration exceeded those of other countries by a factor of three (Appendix A).

How did such high rates of internal migration affect the economic development of the United States? This paper begins to answer this question by examining the effects of internal migration on education. As the population gradually moved west, receiving locations were exposed to internal migrants with vastly divergent experiences with education — dependent on conditions in their origin locations. Native parents interacted with migrant parents who had different views on whether their children should go to school and how schools should get funded. My research investigates the influence of such exposure on natives' school enrollments and schools' funding.

Education relied on simple "common schools" that had little support from the federal government, with the costs paid initially by private spending and later by locally set property taxes. Local control and funding of schools, coupled with a lack of compulsory schooling laws, led to significant variation in whites' enrollments between and within US states. This paper seeks to explain this inequality through patterns of internal migration. As one historian of public education in Wisconsin put it, "Perhaps the chief explanation of the general acceptance of the free-school idea in early Wisconsin is the fact that many of the settlers had witnessed its operation in the East and were convinced of its soundness" (Jorgenson, 1956, p. 93).

I establish two main results. First, I find that school enrollments of migrants' children varied significantly by their state of origin. I estimate the revealed preferences for schools in a discrete choice framework where people choose a location and whether to enroll their child, taking into account the price of schools. I then compare the estimated willingness to pay of the parents by the education levels in their origin states. I find that people from more educated states (measured by enrollments) were willing to pay more for enrolling their children. Connecticut, Maine, and Massachusetts, had over 90% of native children enrolled in schools as early as 1850, while the national average for white children was 66%. Migrants from these states had the highest willingness to pay for enrollment.

Second, I look at the spillover effects from receiving migrants coming from more educated origins. Defining mothers as "native" if they were born in the state of residence, I find that counties that received a larger share of migrants from more educated states had higher levels of enrollment for the the children of native mothers. For instance, receiving migrants exclusively from Massachusetts as opposed to Louisiana—respectively, the most and the least educated states in 1880—increased natives' enrollment by 28.43 percentage points (off the mean enrollment across all years of 59.1%).

I exploit the significant variation in the school enrollments in the states in the East. Some people who moved from the eastern states came from Massachusetts, the first US state with a robust public school system following Horace Mann's promotion of free schools. Others came from the Southern states, where the slave-holding planters controlled the legislature, and the schools mostly comprised expensive academies and other private schools (Stoddard, 2009; Go and Lindert, 2010). The enrollment gap between the North and the South widened after the Civil War due to the collapse of the Southern enrollments (Collins, 2007).

The changing composition of US-born migrants explains a significant part of the inequality in school enrollments of the local children, and internal migration was an important channel in the spread of schools. There are two sources of variation: first, as explained above, the origin states differed in terms of their school systems and the resulting enrollments. Second, migration was not uniform, resulting in a different mix of migrants arriving in the destination states.

Location choices may be related to the expected quality or price of schools in the destinations. Migrants choose where to live strategically, which could bias the estimates of the spillover effect from migrants onto natives. For example, if people from more educated places preferred destinations with better expected schooling amenities, naïve estimates of the spillover effect would have an upward bias.

To identify the causal effects of the composition of migrants, I propose a strategy that exploits the similarity of counties in the West to the Eastern states in terms of suitability for growing different crops. The idea that migrants follow the crops that they grow because they develop relevant skills dates back to at least Steckel (1983).[1] I construct an instrumental variable based on the numbers of migrants predicted by what crop a farmer should choose in the origins and the destination. With the stocks of internal migrants that are predicted by how similar the origin and the destination are, I compute the level of schooling of *predicted* migrants, where the prediction is not based on the level of schooling in the destination.

This new instrument is an alternative to the "early settlement" instrument, commonly used in the migration literature.[2] Compared to the "early settlement" predictor, similarity in agricultural crop yields is explicit on how the initial settlement occurred. I conduct the tests proposed in Goldsmith-Pinkham et al. (2020) and find that the crops prediction leads to estimates that have a more diverse set of states as sources of variation driving the results.[3] Despite the conceptual dif-

[1] In a study of internal migration with lottery assignment of destination, Bazzi et al. (2016) finds that farmer's productivity is significantly higher when allocated to a lot that is more similar to the farmer's origins.

[2] The "early settlement" (also known as "enclave") instrument was first introduced in Altonji and Card (1991). This traditional instrument constructs the predicted number of migrants by interacting the past shares of movers from each state with the current-period outmigrants from that state. Jaeger et al. (2018) provide an overview of its use in the migration literature. Recent papers that use "early settlement" IV to study the local effects of foreign migration to the US include Hunt (2017) and Tabellini (2020).

[3] Specifically, I calculate the Rotemberg weights that show the weight in the overall estimated spillover effect of the just-identified coefficients when using each origin-year separately. For the past settlement, the greatest weight is

ference, the crop suitability instrument and the past migration instrument yield similar estimates in magnitude.

After establishing the causal impact of the migrants on the natives, I study the mechanism behind spillovers. To that end, I digitize the Census of Social Statistics to build a dataset with the number of schools, teachers, and pupils as well as the sources of school funding and general taxes in each county. I calculate the per-student public and private spending on education. Counties that received migrants from more educated places saw significant growth in public spending per student. People coming from more educated origin states in the East promoted the idea of taxation, public funding, and free common schools, which reduced the costs of attendance and expanded access to education. Thus, the composition of internal migrants played an essential role in shaping local institutions.

To better understand the results, I dismiss some potential mechanisms for the spillover effects that could be alternatives to the public spending channel. First, I check whether the impact went through marriages between local mothers and fathers coming from educated states. I change the definition of nativity to include only children, both of whose parents were native, not just the mother — and find that the effect on the fully native children is the same. Children of mothers who married a native father were affected similarly to the children of mothers who married an internal migrant. Second, I split the sample of native children into farmers and non-farmers, estimating the effect separately for two groups. The effect is the same for both groups, suggesting that the increased enrollments were not due to the changes in occupational structure and returns to education induced by migrants from educated states.

Mass schooling was one of the most important social and economic innovations of the 19$^{\text{th}}$ century. The fact that the US was a leader in its adoption may be surprising, given how decen-

from New York in 1870, and it is greater than the next four weights combined. See Appendix F for details.

tralized decisions on education were. Goldin and Katz (2009) emphasize the "grassroots nature" of American education, with its local decision-making, usually at the county level. The local population would independently vote on the property taxes to finance a public school for their children, without any regulation prescribing they do so. The social perceptions of the return to schooling, the value of the taxable property, and various political economy issues resulted in different county-level outcomes.

This contrasts with the experience of other countries, where universal elementary schooling usually followed compulsory schooling laws. Rather than compulsory schooling laws, I show that migration from places with high educational enrollment was crucial for the diffusion of education. The only two states with compulsory schooling laws in my sample are Michigan (1871) and Wisconsin (1879).[4] I find that not only are the results robust to omitting them from the sample, but also that restricting the sample to the cross-section of counties in just those two states in 1880 yields a very similar estimated spillover effect. This suggests that the composition of migrants likely mattered after the introduction of compulsory schooling.

Economic historians often view the "grassroots" rise of universal formal learning for white students as a key driver of the US economic success (Goldin and Katz, 2009). I build on the existing research on the rise of schools in the US in the 19th century (Goldin and Katz, 2009; Go and Lindert, 2010; Parman, 2018) that emphasizes returns to education and favorable political institutions to explain the geographic inequalities in school enrollment. Focusing on a later period and the rise of high schools, Goldin and Katz (2000) estimate the returns to schooling in Iowa in 1915. Their work suggests that the relatively early onset of universal education (compared to the other countries) came from a relatively higher income premium from a marginal year of school rela-

[4] Importantly, compliance with compulsory schooling laws was low until the states added regulations regarding child labor that occurred even later. In fact, in 1880, both Wisconsin and Michigan reported enrollments below 80%. See Bandiera et al. (2018) for the timeline of compulsory schooling and child labor laws adoption in the US and across the world.

tive to the cost. Some rural and mostly agricultural areas had remarkably high shares of children enrolled, whereas others lagged in adopting elementary and middle education.

This paper shows how differential migration can explain the variation in enrollments within states, and not just the general North-South divide in school enrollments.[5] I contribute to some recent literature on migration and peer effects in educational attainment. Internal migration is less studied than migration from other countries, and this paper is the first to link the geographic mobility of 19$^{\text{th}}$-century US and the educational outcomes for the locals. Hunt (2017) uses state-level panel data to discuss how foreign migration increased the educational attainment in the US in 1940–2010. Parman (2012) finds significant returns to education in a sample of farmers in 1915 Iowa and documents spillover effects from the education of a farmer's neighbors. Bandiera et al. (2018) discuss how the rise of compulsory schooling (which most states implemented after the period studied in my paper) was a response to low enrollment rates by some of the foreign migrants, in order to teach them American civic norms.

Go and Lindert (2010), discussing the role of public spending and voters' preferences for schools, emphasize the decentralization of power as an important feature behind the public spending on schools. Crucially for this paper, all males above certain age could typically vote in local elections, although no records are preserved on the margins in the votes for school taxes, which is why I rely on revealed preferences of migrants for schools rather than on preferences reported in popular voting. Paulsen et al. (2021) suggest that initial high public investments in common schools were self-sustaining. Using random allocation of lands following the Revolutionary War, they show that initial higher investments in schools increased earnings, reduced inequality, and increased democratic participation.

The contribution of this paper is also to the broader literature on the migration-induced

[5]See Appendix C for the map of school enrollments in 1880, the final year in my sample. The North-South divide is evident, which warrants the use of county fixed effects in all specifications.

changes in the preferences of natives, including preferences for public spending and welfare state (Dahlberg et al. (2012), Alesina et al. (2016), Tabellini (2020), Giuliano and Tabellini (2020)). Some other studies look at how preferences are shaped by the place of origin and then are passed on to children (Grosjean (2014), Fernández (2011)).

To quantify the changes in school enrollments due to internal migration, I combine the estimates from the discrete-choice location model and the spillover effects. I estimate a counterfactual where the costs of migration are increased by a factor of three, which reduces the share of people who leave origin states from 40% to 27%. This reduces the share of migrants in the destination states and changes their composition because people from different origins respond differently to the increased cost of migrating. I find that this reduction in the migration rate would have lowered the enrollments by 1.5 percentage points (or 2 percent of the baseline enrollment of 72.1% in 1870). If the internal migration had not existed, the enrollments would have reduced by 9 percentage points.

Taken together, my findings suggest that internal migration mattered beyond the reallocation of labor and preference sorting. Crucially, the migrants both differed in their enrollment choices and caused spillovers onto the native population in their destination.

Data and Sample

Starting from 1850, the Population schedules of the US Census recorded the state of birth of every respondent and whether the respondent had attended school over the last year. This information allows for tracking the migration of white adults[6] and the school enrollments of their children in 1850, 1860, 1870, and 1880. I use the full-count Population Census data from

[6] Slavery and subsequent segregation created substantial barriers for migration and education of the Black population in the period that I study, often excluding them entirely from the school system. This is why I only study white migrants and students.

IPUMS (Ruggles et al., 2019).

The main outcome of interest is the school enrollment for children aged 7–14 in a given county. Decisions about school funding, taxes, curriculum, and policy were usually made at the county or more local level. This is why this paper focuses on counties as a unit of observation. The Population schedules included a question "Did respondent attend any educational institution within the last year", and the census marshals were explicitly instructed to exclude Sunday schools. The Census Day was always June 1 in the period studied, making it a consistent (albeit lax, as even a short period of attendance counted) measure of schooling which I refer to as "enrollment", as is common in the literature (Goldin and Katz, 2009).

The age cutoffs of 7–14 correspond to the enrollment in primary school, mainly the *common schools* which usually had only one teacher and lacked grades. The schools taught reading and writing[7], and, for older children, provided instruction in geography, history, natural sciences and sometimes classics.

Migrants are defined as adults aged 23 or older who do not live in their state of birth.[8] For children, I define the origin as the state of birth of his or her mother. In 1850–1870, this information is imputed from the reported state of birth of adults who live in the same household as the child. IPUMS created the family relation variables based on the characteristics of family members, making the mother's place of birth available whenever the mother lives with the child. In case of large households, mothers relation is assigned based on age, surname match, and the records order in the Census rolls. I use mothers rather than fathers because they are slightly more likely to be available in the data.

[7]Literacy could potentially be an outcome of interest, but the Census recorded literacy rates above 90% for native white children in all states in my sample as early as 1860, giving little variation in the outcome.

[8]This is a crude measure and my biggest data limitation. I focus on state of birth because mothers are less likely to be missing than fathers. I do not use migration rates based on the linked censuses (Abramitzky et al., 2020) because currently these methods only exist for males.

Twenty-two Eastern states sent migrants to the counties in the seven destination states, resulting in a diverse composition of migrants in the destinations. I focus on a sample of seven "younger" destination states that were receiving a large influx of settlers: Arkansas, Illinois, Iowa, Michigan, Minnesota, Missouri, and Wisconsin. These states had significant inflows from the older states, but few people moved back from them to the East, which alleviates the concerns related to the classic reflection problem when estimating spillover effects[9] as described by Manski (1993). For the 22 origin states in the East, I do not have the county-level school data, and so for them, school prices and enrollments are aggregated to the state level.

Table 8 in Appendix B shows state-level migration and the level of schooling in the 22 origin states and the seven destination states. The actual estimation relies on county-level data, but the county-level summary of migrant composition in 403 counties would be too big to report here. The substantial variation in the origins of migrants is evident from those stocks of state-to-state migrants, revealing that locations differed in the composition of migrants they received.

For the estimation of the spillover effects, I calculate the school enrollments of native children to use as the outcome variable. I define a child as native if his or her mother is born in the state of residence. As a separate robustness check, I change the definition to both parents being native, finding that the results are robust to this change in the definition of nativity. Since only the state of birth is recorded in the Census, I have to ignore the within-state migration.[10] The nativity indicator assumes that mothers born in a given state were born in the county of their current residence.

[9]Suppose that migrants from more educated places increase the education levels of the locals. Suppose also that we observed significant numbers of migrants going both from origins to destination and from destinations to origins. Then regressing the education of the locals on the average education in the home of migrants would greatly overestimate the effect of the composition of migrants, because it will capture the multiplier effect from the migrants affecting the locals in destination, who in turn affect the locals in the origin through the migration in the opposite direction etc. It is impossible to distinguish the marginal effect from the overall effect that includes the multiplier.

[10]See (Rosenbloom and Sundstrom, 2004) for a discussion of the state of birth as a measure of migration. They compare the flows based on place of birth to other measures, such as the "place of residence 5 years ago", which has been a census question since 1940.

The states in the sample of destination had a high variation in school enrollments between their counties. Table 1 shows the county-level enrollment variation of native children in the sample in 1880. Every state had some counties where two-thirds had attended some school over the preceding year. At the same time, in some counties less than 20% of children had attended any school at all, even for a day.

Table 1: County school enrollments in 1880, native white children aged 7–14

State	mean	st.dev	min	max	N	Native Children
Arkansas	38%	0.137	0%	67%	61	24,986
Illinois	74%	0.082	41%	86%	101	120,841
Iowa	75%	0.157	0%	100%	99	17,627
Michigan	74%	0.142	19%	94%	75	48,041
Minnesota	55%	0.297	0%	100%	66	860
Missouri	66%	0.142	14%	85%	113	111,469
Wisconsin	69%	0.128	20%	84%	57	16,013

Summary statistics for native (mother born in state) children aged 7-14 in 572 counties in 7 destination states that constitute the sample. Major cities (Chicago, Minneapolis, St. Louis, Detroit, Milwaukee) are excluded. Mean and st. deviation are at the county level, not weighted by the number of children.

I use county borders from NHGIS (Manson et al., 2020). For county-level regressions, I normalize county borders in 1850–1880 to the borders in 1870 using the procedure outlined in Hornbeck (2010).[11] Although the Population Census data is available for 1850–1880, the CSS only covers the years 1850–1870, hence the choice of 1870 as the year of "standard" county borders. As a robustness check, I use 1880 borders and do not find any substantial changes in the estimates. For the final estimation, I only use the counties that had non-zero native children population in all years between 1850–1880.

[11]This procedure assumes a uniform distribution of population across space. If, for example, a county split into two counties between 1860 and 1870, then the population in 1860 will be assigned to the two "new" counties proportionally to their areas.

Census of Social Statistics

I augment the Population schedules with the schedules of Social Statistics (which I refer to as the Census of Social Statistics, or CSS). The Social Statistics schedules were a part of the same Decennial Census as the Population schedules. The forms recorded the information on local schools, newspapers, taxes, wages, property valuations, and other social and economic conditions, and were usually filled by the same Marshals who performed the population enumeration. The marshals were instructed to obtain the information either from local official publications or inquiring with the county officials (Census Office, Department of the Interior, 1860). CSS is a unique source that describes schools in antebellum US consistently across all the states covered by the Decennial Census, before the first federal institution (the Office of Education) was formed to centrally compile information on schools across the country.

Census Bureau publications did not report county-level information in its publications that followed the census. Instead, it only reported state-level aggregates. This paper is the first one to use county-level data on schools from CSS across multiple states. I collected and digitized the original hand-filled schedules of CSS for the seven Western states in the sample. For the 22 origin states, I use the state aggregates published by the Census Office (Haines, 2010).

In the section on schools and academies—the key section of CSS for this paper—it was often the case the public schools were reported to charge tuition ("rate-bills"), and private schools received public funds. The line between private and public schools was blurred. Hence, to proxy the importance of public and private role in education, I use measures of funding, rather than the designations of private and public schools.

Similarity in potential yields between origins and destinations

The empirical analysis focuses on 1) the migration and enrollment preferences of internal migrants and 2) the impact of these preferences on the natives in migrants' destinations. For both parts, I rely on the importance of crop choice for the internal migrants. This section describes the *dissimilarity of comparative advantage*, used throughout this paper as the measure of how attractive different locations were for internal migrants.

Most of the internal migrants in the nineteenth-century US were farmers and farm laborers. In 1880, among the internal migrants in the sample states, 64% of men and 58% of women who reported occupation reported being farmers and farmer laborers or managers. Of the remainder, many worked in occupations that were directly related to agriculture. General for-hire laborers, millers, carpenters, and blacksmiths most likely also had a preference for moving to places that grew similar crops because the tools they operated or maintained were familiar. Migrating workers chose their destination to maximize the return on human capital they acquired before moving, choosing a place with similar soil and climate. Steckel (1983) discusses how East-West patterns of migration in the second half of the nineteenth century can be explained by similarity in agro-climatic characteristics. Skills related to cultivation and harvesting of specific crops were essential for farmers who chose their destinations.

The numerous contemporaneous guides confirm the importance of crops for location choice for settlers, published by states and independent publishers. For example, the guide for future migrants published by Iowa Board of of Immigration in 1870, circulation of 35,000 copies, gave location suggestions to migrants based on what crops they intended to grow (Iowa Board of Immigration, 1870):

> [Corn] is successfully raised in all parts of the state, but the southern portion is best

> adapted to it. ... [Oats] rarely fail to yield abundantly in all sections of the state, though perhaps the northern portion is best adapted to it. ... Potatoes usually yield well throughout the state; but the new counties in the northwest bear off the palm in this crop.

Migrants could obtain more detailed information than this crude description from the land surveys conducted prior to settlement of new states. Before the new territories were settled, the federal government sent land surveyors to carefully study and delineate the land[12]. Thus, it was possible to learn about the type of soil, landscape, water access, and other features through General Land Offices that were established to distribute land to the new waves of settlers.

I use similarity in comparative advantage in growing different cultures to measure the distance between origins and destinations. I start by using four major cultures: Corn, Oats, Wheat, and Potatoes. Agro-Ecological Zones (GAEZ) project of the Food and Agriculture Organization (Fischer et al., 2012), a specialized agency of the United Nations, provides data on potential crop yields of those cultures (which I refer to as "suitability"). The data is at the level of cells of 5 arc minutes by 5 arc minutes, which is approximately 10 km $\times$ 10 km (this is not exact since the size of arc minute varies with latitude). FAO uses different natural characteristics, such as chemical composition of soil, precipitation, and elevation to calculate potential yield of different crops given an assumed level of crop management and input use. I assume intermediate level inputs with rain-fed water supply, which is consistent with the agricultural practices of that time. I use this grid-level data to calculate, for each county, the median level of potential yield in tonnes/hectare. Since the origins of migrants are recorded at the state level, I aggregate the suitability measure by taking a weighted average of the county-level suitability, using the areas of farm lands in each

[12]See (Atack et al., 2000) for the discussion of land surveys and the operation of General Land Offices.

county as the weights.[13]

For example, denote $corn_\ell$ and $potato_\ell$ to be the corn and potato suitability in tonnes/hectare in location ℓ. The dissimilarity $z_{od}^{corn,\ potato}$ between o and d is calculated as:

$$z_{od}^{corn,\ potato} = \Big| \underbrace{\frac{corn_d}{potato_d}}_{\text{dest comp adv}} - \underbrace{\frac{corn_o}{potato_o}}_{\text{origin comp adv}} \Big| \tag{1}$$

Higher levels of z_{od} reflect that o and d are more different in terms of their comparative advantage in growing crops. Comparative advantage determines specialization, and the idea behind this dissimilarity measure is that migrants will want to specialize in the same crop in their destination. One advantage of using this particular measure is that I am not taking a stance on converting potatoes to corn or wheat. Taking a Euclidean (or any other) distance between the 5-dimensional vectors without normalization would assume that a tonne of potatoes is the same as a tonne of corn.[14] Appendix D shows the maps of for corn/potato and wheat/potato measures, at the county level for destination states and at the stat level for origin states in the East.

Another advantage of using comparative advantage is that it ignores the absolute levels of potential yields, converting everything to relative units. Thus, I am not picking migration between places that have better agricultural conditions overall. Rather, migration flows are predicted to be higher between places that are similar in their *relative* suitability.

Comparing preferences by origin

I estimate a discrete choice model where parents of each child make two decisions: (1) where to live, including the choice to stay in their home state or to leave and (2) whether to enroll their

[13] In a related paper, Fiszbein (2021) also used FAO data to study the effect of historical agricultural diversity in the US on present-day incomes and educational attainment. In another project, Raz (2021) shows that more heterogeneous counties in terms of soil composition (with the data from Digital General Soil Map of the United States) have historically had weaker communal ties and higher individualism, as measured by surname uniqueness.

[14] One alternative to normalization would be to use average prices to convert volume measures into values. However, market access of Midwestern states was very low before the 1890s, suggesting that no centralized market existed (Donaldson and Hornbeck, 2016) and prices from commodities exchanges could be a misleading measure for crop choice.

child in school. The goal is to measure the willingness to pay (WTP) for school for parents coming from different origins. Consider a household with the mother from origin $o(i)$, deciding whether to move to the destination d and whether to enroll their child in school.

$$U(loc = d, school)_i = \beta_{home}\mathbb{1}_{stay} + \underbrace{\beta_c z_{od}^{corn} + \beta_w z_{od}^{wheat} + \beta_d \log(z_{od}^{km})}_{\text{cost of moving to } d} \mathbb{1}_{\text{move to } d} + \tag{2}$$

$$\underbrace{M_i\Gamma_{sch} + \beta_{sch} + \beta_{price} \cdot price_d}_{\text{utility from school}} \mathbb{1}_{\text{enroll in school}} + \nu_{i,d,school}$$

If the household decides to stay in the home state, it receives the utility from the home bias, β_{home}. The cost of moving to a different location varies by the dissimilarity in crop suitability (corn/potatoes and wheat/potatoes ratios) and the geographic distance. Destination counties and states[15] d vary by the price of schools (which can be zero if school are free), and the coefficient β reflects the price sensitivity. If the child is enrolled in school, the utility increases by β_{sch}. I include many family characteristics relevant for school decisions in M_i: family real estate per child, gender and age of the child, literacy of the household head, and also whether the household lives on a farm or the household head is a farmer/laborer. M_i includes time fixed effects. I omit the time index t from all variables to make the notation lighter. Each household simply chooses from the options in a given year: 1850, 1860, or 1870.

The households are assumed to choose $\{d, school\}$ to maximize their utility. The error term $\nu_{i,d,school}$ is the unobserved part of the utility, assumed to have Type 1 Extreme Value distribution. It rationalizes the fact that households with the same characteristics choose different options, and is assumed to be independent across all children.

All coefficients are allowed to vary by the state of origin $o(i)$. Thus, people from different origins can have a different home bias, price sensitivity and crop dissimilarity, and also put different

[15] I specify a full set of choices. For example, a mother i with origin $o(i)$ being Massachusetts can choose to stay in her home state, move to any of the $22 - 1 = 21$ other states or to one of the counties in the 7 states for which county-level school data is available.

weights on demographic characteristics for school enrollment decisions. Subscripts $o(i)$ for each estimated coefficient in Equation 2 are omitted.

The coefficients in (2) do not have a direct interpretation, given that only the differences in utility are identified (Train, 2003). To compare the preferences of migrants from different origin, I use Wilingness to Pay for schools, calculated as WTP $= (M_i\Gamma_{sch} + \beta_{sch})/\beta_{price}$. It is the price at which a household will be indifferent between enrolling and not enrolling a child with characteristics M_i (since the utility for enrolling varies by the characteristics, so does the WTP). Estimated WTP is in dollars, and it is comparable across origins.

Results: differences in preferences

I estimate the choice model with the utility described in (2) using maximum likelihood. I sample 12,000 households in each year[16], with some households having multiple children.

Figure 1 summarizes the estimation results. It shows, for every origin state, the willingness to pay for enrolling a child in school, $W\hat{T}P = (M_i\hat{\Gamma}_{sch} + \hat{\beta}_{sch})/\hat{\beta}_{price}$. Since $W\hat{T}P$ varies by the child characteristics M_i, I choose to compare it across boys in a household with a literate head, with family wealth of \$100 and one sibling. Since the wealth is fixed, the differences WTP have the interpretation of being conditional on wealth. Standard errors are calculated using the delta method. For some states, the willingness to pay is estimated to be negative, which can be interpreted as average parents valuing their children working and receiving informal education as home so much that they would prefer not to have their children in school.

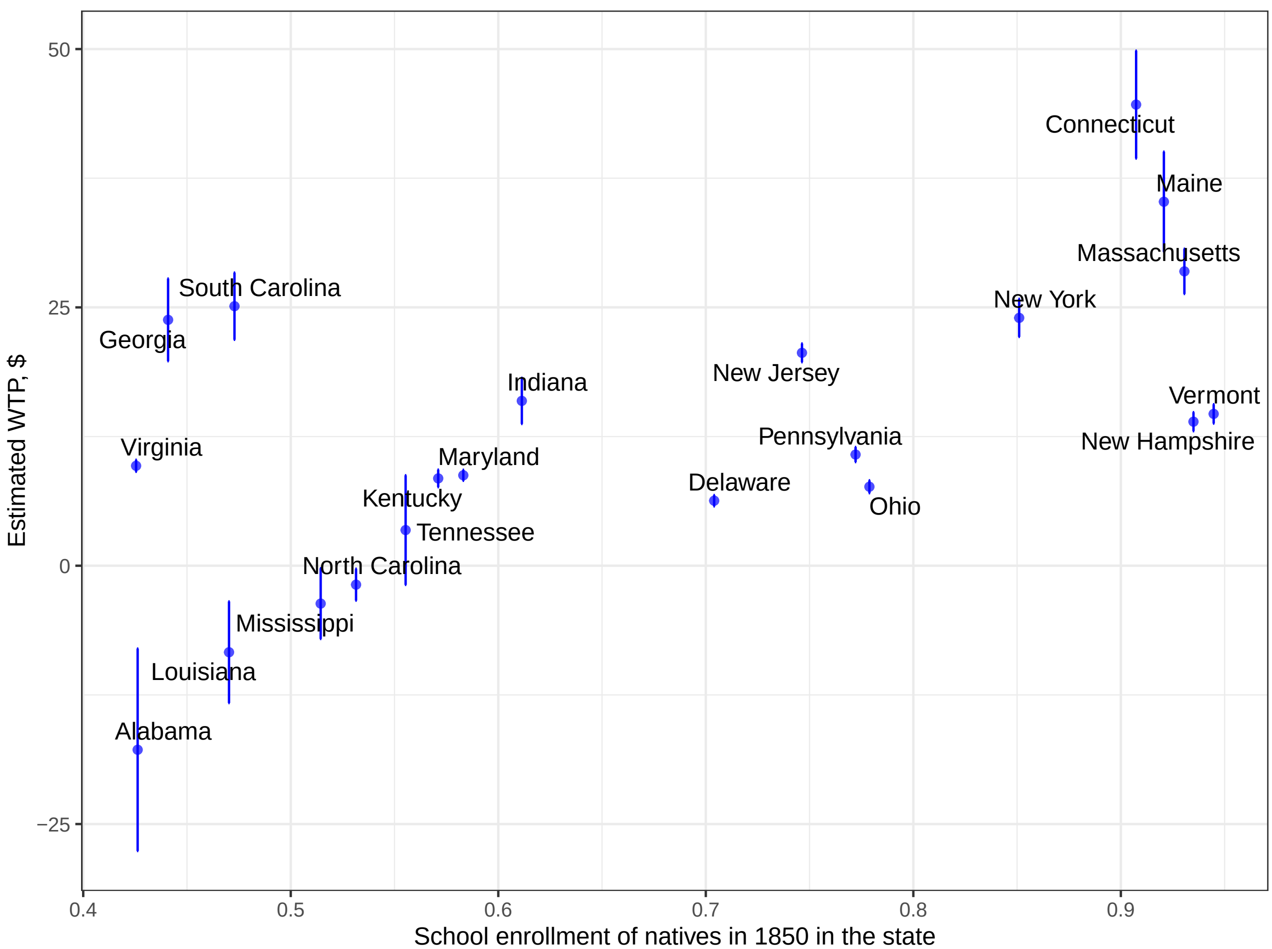
The estimated WTP correlates strongly with the school enrollments in the origins of migrants. The estimates show that mothers from Connecticut, Maine, and Massachusetts had the highest

[16]There are 420 counties + 22 origin states available as destinations, and in each destination there are two options, enrolling or not enrolling a child in school. With the number of alternatives nearing 1000, the estimation requires sampling.

willingness to pay. These states had well-development school public systems, and so, on average, parents who remained in their home state did not have to pay much to enroll. The high WTP is identified, essentially, from the migrants from those states who moved to states with less developed public school systems where one had to pay tuition, and still enrolled their children in school.

These results suggest that people coming from more educated states (as measured by school enrollments in those states in 1850, the first year in the sample) had additional preference for education. Since the model allows location choice, it allows for selection into locations based on differential preferences for education.

Figure 1: Estimated Willingness to Pay and the origins' school enrollment in 1850

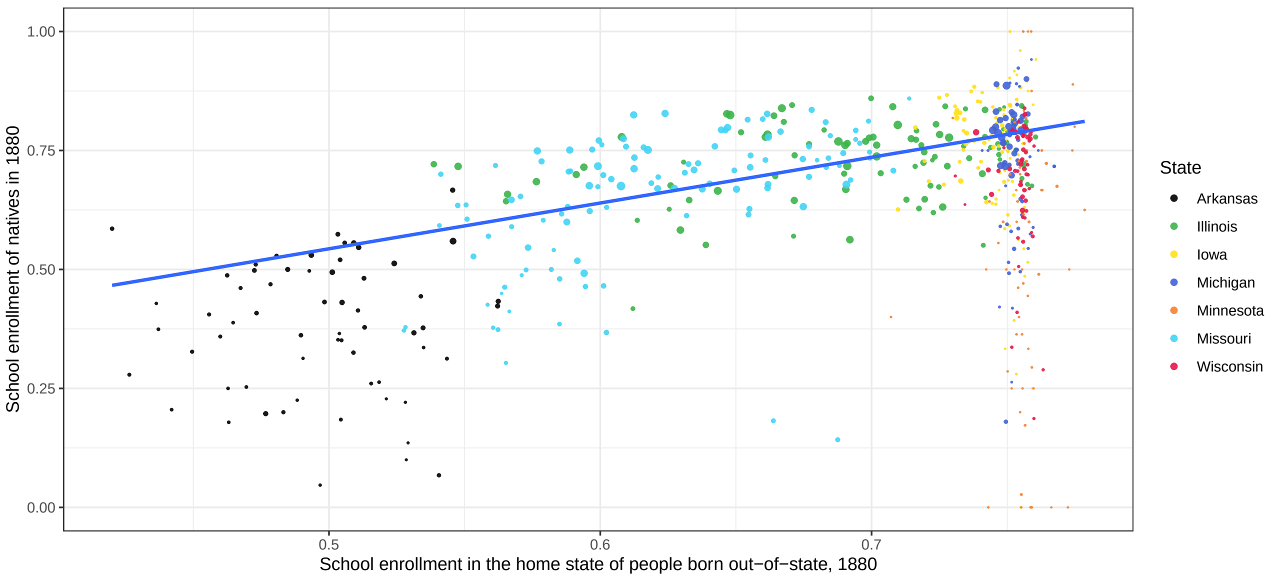


Estimated Willingness to Pay for enrolling child against the school enrollment in mother's origin. WTP is based on the estimates of Equation (2), calculated as $\hat{WTP} = (M_i \hat{\Gamma}_{sch} + \hat{\beta}_{sch})/\hat{\beta}_{price}$. WTP is for a 10-year old boy with one sibling in a family with the total household wealth of \$100. Standard errors are calculated using the delta-method.

Empirical strategy: spillover effects

Figure 2 displays the key relationship showing the spillover effects. It plots the county-level school enrollment rates of the natives against the average enrollment rates in the origins of internal migrants in 1880.

Figure 2: Schooling in origins of internal migrants and natives' school enrollments in 1880



I test the relationship suggested in Figure 2 in a regression framework that includes county and year fixed effects. Consider migrants who move from locations o (22 states) to d (422 counties in the 7 states in my sample), observed every 10 years between 1850–1880. I estimate the following regression:

$$h_{dt} = \alpha_d + \delta_t + \beta_M \sum_o s_{odt} h_{ot} + X_{dt}\Gamma + \varepsilon_{dt} \tag{3}$$

Here, h_{dt} is schooling for *native* children in either the origin or destination, and s_{odt} is share of adults in living in county d who are born in o among all US migrants in d. This denominator for this share is all migrants from the 22 origin states, so $s_{odt} = \frac{N_{od}}{\sum_d N_{od}}$ where N_{od} is the number of people from o to d. The coefficient of interest is β_M, which reflects the impact of the average level of schooling in the home states of migrants on the level of schooling of locals in d. I include county fixed effects α_d control for the initial levels of enrollments and the unobserved fixed characteristics of counties. Year fixed effects δ_t pick the overall increasing trend in school enrollments. As an alternative, in some specifications I include year fixed effects interacted with third-degree polynomials of latitude and longitude to control flexibly for geographic unobservables. X_{dt} includes control variables: shares of immigrants immigrants from several foreign countries.

I flexibly control for the shares of migrants from the top foreign countries in terms of the number of immigrants in the US. Controlling for the share of the share of Germans follows the existing work (Ager and Cinnirella, 2020) on the role of German immigrants in the diffusion of education and the kindergarten movement (note, though, that my sample only includes children aged 7–14, while the kindergarten age is 5–6). I include, separately, the shares of German, Canadian, Irish, British, and Scandinavian migrants. Together, these origins represent 94% of all foreigners in the destination states in 1880.

Selective migration

Selective migration based on quality, price, or other unobserved characteristics of schools would introduce bias in OLS estimates of equation (3). Such selection could mean that the flows of migrants s_{odt} are correlated with ε_{dt}. For example, if migrants from more educated locations o formed expectations about future levels of schooling in d, they would make strategic choices about where to go given the education in their origin. They could exhibit homophily and try to go to states that also had higher levels of education.Homophily would lead to a positive bias in the estimate of β_M. Expression (4) below formalizes this, assuming that ε_{dt} is not correlated with h_{ot}, the education in the origin.

$$\text{cov}(s_{odt}h_{ot};\varepsilon_{dt}) = \underbrace{\text{cov}(s_{odt}, h_{ot}\varepsilon_{dt})}_{>0 \text{ if there is homophily}} + \underbrace{\text{cov}(h_{ot}, \varepsilon_{dt})}_{\text{assumed to be } 0} \mathbb{E}(s_{od}) \tag{4}$$

To address endogeneity concerns, I use predicted shares of migrants $\hat{s}_{od}$ instead of the actual shares s_{od}, forming an instrument $z_{dt} = \sum_j \hat{s}_{odt} h_{ot}$. The predicted stocks of migrants $\hat{s}_{od}$ are constructed from the agricultural similarity between o and d in terms of how suitable locations are for growing corn, potatoes, and wheat . The similarity measures are interacted with the current-year number of out-migrants from a given state, which gives the variation across years. I complement this novel instrumental variable strategy with a more traditional approach of using past migration as a predictor of future migration.

The key assumption behind the crop instrumental variable is that agro-climatic suitability for different crops correlates with the migrants attracted to a particular location, but does not influence education through other channels. The assumption required for exclusion restriction is that the *level* of suitability (which affects how similar a location is to different origin states) does not determine the school enrollment through other channels, for example, through differential demand for child labor by farmers growing different cultures. This effect is unlilkely: using the exogenous shock of Boll Weevil in Georgia, Baker (2015) shows that the production of cotton did not have an effect on the school enrollment of white children. Despite this finding I do not use cotton as one of the predictors because of its association with slavery and related institutions.

I employ a three-stage procedure. First, I create different predicted stocks of migrants using the (dis)similary in suitability for growing different crops. Second, I construct $z_{dt} = \sum_o \hat{s}_{od} h_o$.[17] Finally, I perform a 2SLS estimation of 3 using those instruments: I use $\sum_o \hat{s}_{od} h_o$ as instruments for $\sum_o s_{od} h_o$.

At the prediction stage, I use Poisson Pseudo Maximum Likelihood (PPML) with the following conditional expectation of the number of migrants N_{odt} from origin o living in destination d

[17] An auxiliary regression that predicts a total/average value based on a combination of individual contributions is common in trade literature (Frankel and Romer, 1999). Deij et al. (2021) discuss the econometric properties of this family of estimators and emphasize that it's important to keep zero values. In my context, zero values are not very common, as the migration matrix is not sparse. I do keep all the zero predicted values.

in period t:

$$\mathbb{E}(N_{odt}|z,\ outmigration) = \exp\left(\beta_c z^{corn}_{od} + \beta_w z^{wheat}_{od} + \beta_o z^{oats}_{od} + \beta_d \log(z^{km}_{od}) + \sum_{d' \in \mathbf{D}} N_{od't}\right) \qquad (5)$$

The prediction is based on 1) the dissimilarity measures between origin and destination, 2) distance (in kilometers) and 3) $\sum_{d'} N_{od't}$, the number of total outmigrants from the origin states do all destinations $\mathbf{D}$, including to other origin states.[18]

The results of the prediction are reported in Table 2. Dissimilarity in agricultural conditions is a good predictor of migration flows. This yields four vectors of predicted stocks $\hat{s}_{odt} = \hat{N}_{odt} / \sum_{d'} \hat{N}_{od't}$, corresponding to columns (1), (2), (3) and (4) in Table 2.

Table 2: Predicting migration flows through dissimilarity in potential yields

	(1)	(2)	(3)	(4)	(5)
Dissimilarity in corn/potato	−4.224 ***				−2.886 ***
	(0.362)				(0.303)
Dissimilarity in wheat/potato		−14.967 ***			−1.042
		(1.312)			(1.233)
Dissimilarity in oats/potato			−4.352 ***		−15.641 ***
			(0.772)		(0.957)
Distance, ’000s km				−1.882 ***	−1.621 ***
				(0.069)	(0.076)
Outmigrants from origin, ’000s	0.004 ***	0.004 ***	0.004 ***	0.004 ***	0.004 ***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Destin. county FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
N	49324	49324	49324	49324	49324

*** p < 0.01; ** p < 0.05; * p < 0.1

Estimates of Poisson regression model. Each observation is an {origin state; destination county} dyad in one of years 1850, 1860, 1870, or 1880. Dependent variable is Stock of migrants born in state o who live in county d.

I obtain different predicted stocks $\hat{s}_{odt}$. Then, for every county d, I construct three different[19] predicted average levels of schooling in origins of migrants: $\sum_o \hat{s}_{od} h_o$. Each of those predicted levels is based on one pair of crops.

[18]For example, it includes the number of outmigrants from Massachusetts to New York in a given year, even though New York is not in the sample of destination states. I also use the leave-one-out predictor $\sum_{d' \neq d} N_{od't}$ and find that the predictions and the second-stage results are almost the same, which is not surprising given that there are 403 small destination counties, and so $\sum_{d' \in \mathbf{D} \setminus d} N_{od't} \approx \sum_{d' \in \mathbf{D}} N_{od't}$.

[19]I use three different predictions because it allows me to conduct overidentification tests. The Sargan overidentification test does not reject the null that model is correctly specified (p-value = 0.9). Pooling all the crops together and only using one prediction (column (5) in Table 2) does not alter the results significantly.

Results: spillover effects

Table 3 shows the results for both the OLS and 2SLS estimation of equation (3). The composition of migrants was important for the school enrollment of natives in their destinations. Column (3), which is my preferred specification, shows that an increase in school enrollments in the home states of migrants by 1 percentage point is associated with an increase in the schooling rates for the white native children by around 0.584 percentage points. This is a sizable effect: one st. deviation increase in school enrollment in the origins of internal migrants increases the school enrollment of the locals by 0.356 st. deviations.

Table 3: Estimation results: spillover effects

	OLS		IV (crops)		IV (past migr.)	
	(1)	(2)	(3)	(4)	(5)	(6)
Education in origins of internal migrants	0.584 *** (0.142)	1.015 *** (0.338)	0.584 *** (0.139)	1.592 * (0.824)	0.639 *** (0.148)	1.201 ** (0.566)
Shares of foreign migr.	✓	✓	✓	✓	✓	✓
Year FE	✓		✓		✓	
County FE	✓	✓	✓	✓	✓	✓
Year FE × Lat-long polyn.		✓		✓		✓
N	1612	1612	1612	1612	1209	1209
R^2	0.716	0.725	0.716	0.722	0.729	0.736
First stage F-stat			758.14	35.76	3448.93	263.82

*** p < 0.01; ** p < 0.05; * p < 0.1

A unit of observation is a county in 1850, 1860, 1870, 1880. Outcome variable is the share of native children (aged 7-14, mother born in the state of residence) enrolled in school. Mean of outcome: 0.592; mean of Education in origins of migrants: 0.651. Standard errors in parentheses are adjusted for spatial autocorrelation per Conley (1999) with 100 mile cutoff. Observations are weighted by the number of native children.

The estimates from the 2SLS and OLS estimation are not different in magnitude. I interpret this as a result of two forces that cause bias in the opposite directions. There is a measurement error in the regressor $\sum_o s_{odt} h_{ot}$, since o only indexes origins at the state level. Some states are large and diverse, and the composition measure does not capture within-state migration, leading to a downward, attenuation bias in $\hat{\beta}_M^{OLS}$. On the other hand, as shown in the first part discussing differences in migrants' preferences, it is more likely that migrants from more educated states preferred to go to places with better schools (if preference for enrollment is a good proxy for a preference for quality and other school characteristics). This would lead to an upward bias in $\hat{\beta}_M^{OLS}$ is OLS estimation.

Following Goldsmith-Pinkham et al. (2020), I decompose the 2SLS estimate of β_M into a weighted average of estimates when using each of the 22 origins separately. Examining these weights, I find that identifying variation comes from both the Northern and the Southern origin states, and the results would be similar if β_M was estimated only off one of the top-weight origins.[20]

I conduct several additional robustness checks, reported in Appendix G, adding and removing control variables and fixed effects. I find that the estimates of β_M are not sensitive to the specification. Also, as suggested in Jaeger et al. (2018), I show that there is sufficient variation in $\hat{s}_{odt}$ across time to ensure that $\hat{\beta}_M$ does not conflate short- and long-term responses to the changing composition of internal migrants (see Appendix H).

Share of internal migrants in population

So far, the estimation focused on the diversity of the migrants' origins. In this section, I account for the share of internal migrants in the population in addition to their composition. Some counties population had fewer than 5% of internal migrants, while in some the internal migrants accounted for 85% of the population.[21] The counties in the sample saw a steady decrease in the average share of internal migrants during the period studied, from 75.4% in 1850 to 49.5% in 1880. This decrease not due to the reduced share of people living the origin states, which remained rather stable at 40% of adults in the east leaving their state of birth. Rather, it was caused

[20]See Appendix F for the results of Rotemberg weight decompostion.
[21]See Appendix I for the distribution of the share of internal migrants.

by the increase in the foreign migration and the births of the previous cohorts of migrants.

The share of internal migrants represents the intensity of treatment of the natives. Counties were treated by a different composition of migrants, and if there were more of them in the population, it was easier for them to promote education and public funding of schools. I replace the $\sum_o s_{odt} h_{ot}$ with $\sum_o s_{odt} h_{ot} \cdot s_{\text{migrants},dt}$ and estimate the modified version of Equation (3), where I rescale the treatment by the share of internal migrants:

$$h_{dt} = \alpha_d + \delta_t + \beta'_M \sum_o s_{odt} h_{ot} \cdot s_{\text{migrants},dt} + X_{dt}\Gamma + \varepsilon_{dt} \tag{6}$$

Here, $s_{\text{migrants},dt} = \sum_o N_{odt} / (\sum_o N_{odt} + N_{\text{foreign},\, t} + N_{\text{native},\, t})$ is the share of internal migrants in the population of county d. Note that this normalization is equivalent to using the entire adult population in the denominator of the weighting shares:

$$\sum_o s_{odt} h_{ot} \cdot s_{\text{migrants},dt} = \sum_o \frac{N_{odt}}{\sum_o N_{odt} + N_{\text{foreign, t}} + N_{\text{native, t}}} h_{ot} \tag{7}$$

Table 4: Estimation results: share of internal migrants normalized

	OLS	IV (crops)	IV (past migr.)
	(1)	(2)	(3)
Education in origins of internal migrants, normalized	0.630 ***	0.415 **	1.002 ***
	(0.134)	(0.178)	(0.293)
Shares of foreign migr.	✓	✓	✓
Year FE	✓	✓	✓
County FE	✓	✓	✓
N	1612	1612	1209
R^2	0.714	0.713	0.726
First stage F-stat		715.39	266.53

*** p < 0.01; ** p < 0.05; * p < 0.1

A unit of observation is a county in 1850, 1860, 1870, 1880. Outcome variable is the share of native children (aged 7-14, mother born in the state of residence) enrolled in school. Mean of outcome: 0.592; mean of Education in origins of migrants: 0.352. Explanatory variable is $\sum_o s_{odt} h_{ot} \cdot s_{\text{migrants},dt}$, where $s_{\text{migrants},dt}$ is the share of internal migrants in the population of county d. Standard errors in parentheses are adjusted for spatial autocorrelation per Conley (1999) with 100 mile cutoff. Observations are weighted by the number of native children.

Table 4 shows the estimates of Equation (6). The estimates confirm that the spillover effect is robust to adjusting for the share of migrants in the population.

Internal migrants and school funding

Table 5 shows the impact of migrants coming from more educated places on school funding. I use the same procedure as before, constructing predicted flows of migrants from agricultural dissimilarity to tackle selection of destinations. Restricting the sample years to 1850–1870, I find that migrants from states with high school enrollments increased the public, but not private spending on schools. An increase of school enrollments in the origins of internal migrants by 1pp corresponds to a 28.5¢ increase in public expenditure per one student (Column 2). No effects are found for private spending (the "rate-bills" in common schools and the tuition payments in private schools).

Table 5: Composition of Migrants and School Spending

	Public $/student		Priv $/student		Pub $/(Pub $ + Priv $)
	(1)	(2)	(3)	(4)	(5)
Educ. in origins of migr.	28.778 ***	28.500 ***	2.994	0.580	0.131
	(4.942)	(5.115)	(2.452)	(2.584)	(0.142)
log(Wealth/Children)		0.451		1.572 ***	-0.042 ***
		(0.445)		(0.342)	(0.011)
Shares of Europ. migr.	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
N	1087	1087	1087	1087	1087
R^2	0.184	0.185	0.254	0.290	0.436
First stage F-stat	488.12	479.69	488.12	479.69	479.69

*** p < 0.01; ** p < 0.05; * p < 0.1

A unit of observation is a county in years 1850, 1860, 1870. Outcome variable public (taxes + public funds) or private (other sources) spending per enrolled student. Robust standard errors in parentheses. Following (Go and Lindert, 2010), funds from "endowment" are not included in either category.

In Appendix G, I reject some of the alternative channels through which internal migrants could affect the locals. First, I test whether the spillovers went through marriage (eg, native mothers being more likely to marry men from an educated state who wanted to invest more in their child's education). I make the definition of nativity stricter by calculating the school enrollments for children of native mothers whose father was also native. With this alternative definition of nativity, the estimate of β_M does not change. This implies that the composition of migrants affected the native-migrant and native-native couples in similar way, and the spillovers did not go through marriage.

Second, I check whether the spillover went through the migrants from more educated states changing the occupational structure of the counties they chose as destinations and increasing the returns to schooling. If this was the case, then the education effect would have been higher for the farming families, and so I restrict the sample to the non-farmers.[22] Again, this was not the case: the effect on non-farmers was similar to that on the general population of natives.

Counterfactual: increasing the costs of internal migration

Combining the estimated preferences of migrants and the spillover effects, what would the school enrollments be if the internal migration in the US was lower? In this section, I look at the counterfactual levels of school enrollments under a hypothetical reduction in the rate of internal migration. I combine the estimates of migrants' location and schooling choices and the estimates of the spillover effects.

The first row of Table 6 reports the school enrollments from combining the predicted stocks of internal migrants from the choice model with the estimated spillover effects. I plug the number of migrants from each origin in Equation (6), using the estimated parameters from Column (2) in Table 4, and calculate $\hat{h}_{dt}$, the predicted enrollments of natives under given stocks of migrants.

As a first step, I increase migration costs, and use the choice model to calculate the predicted stocks of internal migrants in the counties of the seven destination states in my sample. In the second step, I use the composition and the number of these counterfactual migrants to calculate the school enrollments of the locals in the destination states, keeping the number of foreigners

[22] Specifically, I only include families where the head of household does not report occupation of a farm owner, farm manager, farm laborer, or farm foreman and also the family is not living on a farm.

and natives fixed.

As a first counterfactual scenario, I triple the cost of distance, replacing $\log(z_{od}^{km})$ with $\log(3z_{od}^{km})$ in equation 2. I chose the factor for three because that's close to the ratio in Fogel (1969) for the cost of transporting goods by railroad as opposed to the pricier wagon transportation.

This increase in the distance cost of migration reduced the average share of leavers from the origin states from 40% to 27%, as more people prefer the home option. The share of internal migrants in the destination states reduces, and also there are changes in the relative composition.[23] The costs of migration with distance are different for different groups, and thus the tripling of the distance has differential impact on the likelihood of moving of people from different states.

Table 6 shows the results of this scenarios. Compared to the baseline, the school enrollments in the destination states decrease in each year by about 1.5 percentage points. In a more extreme scenario, I completely ban the internal migration, which decreases the school enrollments of natives by 9–13 pp.

Table 6: Average school enrollments predicted by the model and under counterfactual scenarios

	1850	1860	1870
predicted by the choice model	61.6%	75.0%	72.1%
Counterfactual: distance x3	60.6%	73.5%	70.6%
Counterfactual: no internal migration	48.3%	62.4%	62.3%

This table compares the baseline predictions of the model with two counterfactual scenarios. I use estimates of Equation (2) to construct $\hat{N}_{odt}$, the predicted number of migrants from each origin state to each destination county. I then plug this prediction in estimated Equation (6) and obtain $\hat{h}_{dt}$, averaging them for each year.

[23] I find that these are small. Negative effect really comes from the reduced number of migrants. The increase in the distance cost is uniform, and, although the parameters are different for different origins, there is not enough heterogeneity to significantly change the composition of migrants.

Conclusion

This paper discusses how high internal migration rate contributed to the adoption of universal schooling in the US in the 19th century. I combine the data from the Population census with the newly-digitized description of schools and education finding. Using this data, I estimate the preferences of internal migrants for locations and enrolling their children in schools, finding that they differed significantly in terms of their willingness to pay for education. I then turn to the spillover effects from the migrants, calculating the average education in the origins of migrants that natives were exposed to. I find that the native parents who had their neighbors arrive from more educated states were more likely to enroll their children in schools.

Taking the estimated preferences and the spillover effects together, I estimate a counterfactual change in the school enrollments following the increase in the cost of migrating. I find that if the cost of migration had tripled with distance, the ensuing decrease in internal migration would have lowered the school enrollment of natives by 1.5 percentage points.

References

Abramitzky, Ran, Leah Boustan, and Myera Rashid. 2020. "Census Linking Project: Version 1.0 [Dataset]."

Ager, Philipp, and Francesco Cinnirella. 2020. *Froebel's Gifts: How the Kindergarten Movement Changed the American Family.* Technical report. Social Science Research Network, August.

Alesina, Alberto, Johann Harnoss, and Hillel Rapoport. 2016. "Birthplace Diversity and Economic Prosperity." *Journal of Economic Growth* 21, no. 2 (June): 101–138. ISSN: 1381-4338, 1573-7020.

Altonji, Joseph G, and David Card. 1991. *The Effects of Immigration on the Labor Market Outcomes of Less-Skilled Natives.* Edited by John Abowd and Richard Freeman. Immigration, Trade and the Labor Market. University of Chicago Press.

Atack, Jeremy, Fred Bateman, and William Parker. 2000. "Northern Agriculture and the Westward Movement." In *The Cambridge Economic History of the United States,* edited by Stanley L. Engerman and Robert E.Editors Gallman, 2:285–328. Cambridge Economic History of the United States. Cambridge: Cambridge University Press.

Baker, Richard B. 2015. "From the Field to the Classroom: The Boll Weevil's Impact on Education in Rural Georgia." *The Journal of Economic History* 75 (4): 1128–1160.

Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo. 2018. "Nation-Building through Compulsory Schooling during the Age of Mass Migration." *The Economic Journal* 129 (617): 62–109.

Bazzi, Samuel, Arya Gaduh, Alexander D. Rothenberg, and Maisy Wong. 2016. "Skill Transferability, Migration, and Development:Evidence from Population Resettlement in Indonesia." *American Economic Review* 106, no. 9 (September): 2658–2698.

Census Office, Department of the Interior. 1860. *Instructions to US Marshals. Instructions to Assistants.* Washington: Geo. W. Bowman, Public Printer.

Collins, William J. 2007. "Education, Migration, and Regional Wage Convergence in American History." In *The New Comparative Economic History: Essays in Honor of Jeffrey G. Williamson,* edited by Timothy J. Hatton, Kevin H. O'Rourke, and Alan M. Taylor. MIT Press.

Dahlberg, Matz, Karin Edmark, and Heléne Lundqvist. 2012. "Ethnic Diversity and Preferences for Redistribution." *Journal of Political Economy:* 36.

Deij, Sabine, Jakob B. Madsen, and Laura Puzzello. 2021. "When Are Instruments Generated from Geographic Characteristics in Bilateral Relationships Invalid?" *Journal of Applied Econometrics* (May): jae.2815. doi:10.1002/jae.2815.

Die Volkszählung Im Deutschen Reich Am 1. Dezember 1880. 1969. Reprint of 1883 edition. Vol. Band 57. Statistik Des Deutschen Reichs. Otto Zeller.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *The Quarterly Journal of Economics* 131 (2): 799–858.

Fernández, Raquel. 2011. "Does Culture Matter?" *Handbook of social economics* 1:481–510. ISSN: 1570-6435.

Ferrie, Joseph P. 2003. "Internal Migration." In *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition,* edited by S.B. Carter, S.S Gartner, M.R. Haines, A.L. Olmstead, R. Sutch, and G. Wright. Cambridge University Press.

Fischer, Günther, Freddy O Nachtergaele, Sylvia Prieler, Edmar Teixeira, Géza Tóth, Harrij van Velthuizen, Luc Verelst, and David Wiberg. 2012. "Global Agro-Ecological Zones (GAEZ v3. 0)-Model Documentation."

Fiszbein, Martin. 2021. "Agricultural Diversity, Structural Change, and Long-Run Development: Evidence from the US." *American Economic Journal: Macroeconomics.*

Frankel, Jeffrey A, and David Romer. 1999. "Does Trade Cause Growth?" *American Economic Review* 89 (3): 130.

Giuliano, Paola, and Marco Tabellini. 2020. "The Seeds of Ideology: Historical Immigration and Political Preferences in the United States," Working Paper 27238 (May).

Go, Sun, and Peter Lindert. 2010. "The Uneven Rise of American Public Schools to 1850." *The Journal of Economic History* 70 (1): 1–26.

Goldin, Claudia. 1998. "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century." *The Journal of Economic History* 58 (2): 345–374.

Goldin, Claudia, and Lawrence Katz. 2000. "Education and Income in the Early Twentieth Century: Evidence from the Prairies." *The Journal of Economic History* 60 (3): 782–818.

———. 2003. "Mass Secondary Schooling and the State," Working Paper Series, no. 10075 (November). doi:`10.3386/w10075`.

———. 2009. *The Race between Education and Technology.* Harvard University Press.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review* 110 (8): 2586–2624. ISSN: 0002-8282.

Grosjean, Pauline. 2014. "A History of Violence: The Culture of Honor and Homicide in the US South." *Journal of the European Economic Association* 12 (5): 1285–1316.

Haines, Michael R. 2010. "ICPSR, "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002" ICPSR Study 02896-V3." *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributer].*

Hornbeck, Richard. 2010. "Barbed Wire: Property Rights and Agricultural Development." *The Quarterly Journal of Economics* 125 (2): 767–810.

Hunt, Jennifer. 2017. "The Impact of Immigration on the Educational Attainment of Natives." *Journal of Human Resources* 52 (4): 1060–1118.

Inter-university Consortium for Political and Social Research. 1992. "Social, Demographic, and Educational Data for France, 1801-1897." doi:`10.3886/ICPSR00048.v1`.

Iowa Board of Immigration. 1870. *Iowa: The Home for Immigrants, Being a Treatise on the Resources of Iowa, and Giving Useful Information with Regard to the State, for the Benefit of Immigrants and Others.* Mills & Company.

Jaeger, David, Joakim Ruist, and Jan Stuhler. 2018. *Shift-Share Instruments and the Impact of Immigration.* Technical report w24285. Cambridge, MA: National Bureau of Economic Research, February. doi:10.3386/w24285.

Jorgenson, Lloyd P. 1956. *The Founding of Public Education in Wisconsin.* State Historical Society of Wisconsin.

Leasure, J. William, and Robert A. Lewis. 1968. "Internal Migration in Russia in the Late Nineteenth Century." *Slavic Review* 27 (3): 375–394. doi:10.2307/2493340.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60 (3): 531–542.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. 2020. "IPUMS National Historical Geographic Information System: Version 15.0 [Dataset]." *Minneapolis, MN: IPUMS.* doi:10.18128/D050.V15.0.

Minnesota Population Center. 2020. "Integrated Public Use Microdata Series, International: Version 7.3." doi:10.18128/D020.V7.2.

Parman, John. 2012. "Good Schools Make Good Neighbors: Human Capital Spillovers in Early 20th Century Agriculture." *Explorations in Economic History* 49 (3): 316–334.

———. 2018. "Education and Human Capital in American Economic History." In *The Oxford Handbook of American Economic History, Vol. 1,* edited by Louis P. Cain, Price V. Fishback, and Paul W. Rhode, 400–424. Oxford University Press.

Paulsen, Tine, Kenneth Scheve, and David Stasavage. 2021. "Foundations of a New Democracy: Schooling, Inequality, and Voting in the Early Republic": 46.

Raz, Itzchak Tzachi. 2021. "Learning Is Caring: Soil Heterogeneity, Social Learning and the Formation of Close-knit Communities," Working Paper.

Rosenbloom, Joshua L., and William A. Sundstrom. 2004. "The Decline and Rise of Interstate Migration in the United States: Evidence from the IPUMS, 1850–1990." *Research in economic history:* 289–325.

Ruggles, Steven, Sarah Flood, Grover Goeken Ronald, Meyer Josiah, Pacas Jose Erin, and Matthew Sobek. 2019. "IPUMS USA: Version 9.0." doi:10.18128/D010.V9.0.

Steckel, Richard H. 1983. "The Economic Foundations of East-West Migration during the 19th Century." *Explorations in Economic History* 20 (1): 14–36.

Stoddard, Christiana. 2009. "Why Did Education Become Publicly Funded? Evidence from the Nineteenth-Century Growth of Public Primary Schooling in the United States." *The Journal of Economic History* 69, no. 1 (March): 172–201.

Tabellini, Marco. 2020. "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration." *The Review of Economic Studies* 87 (1): 454–486.

Train, Kenneth E. 2003. *Discrete Choice Methods with Simulation.* Cambridge: Cambridge University Press. doi:10.1017/CBO9780511753930.

Appendices

A Internal migration across countries

Table 7 compares the rates of internal migration across several countries in various years, showing that the US clearly stood out in terms of its rate of internal migration. It reports the share of native-born people of all ages who left their first-level administrative divisions of birth (states in the US, departments, counties, and provinces in other countries). People born abroad are excluded from these calculations. I do not adjust for the age structure because for some countries internal migration statistics are not broken down by age. Note that the sizes of first-level administrative divisions varied greatly: for example, French departments are greater in number and smaller in population and area than US states.

Table 7: Internal migration rates in select countries

	Year	Units	Natives not living in state/province/dept of birth (%)
USA (whites)	1850	states	37.94
USA (whites)	1880	states	41.00
USA (Blacks)	1960	states	29.33
France	1872	departments	12.90
Sweden	1880	counties	12.72
Russian Empire *(excl. Poland and Finland)*	1897	provinces	9.72
Argentina	1869	provinces	9.03
Prussia	1880	provinces	6.25

Share of native-born people who do not live in their first-level administrative division (state/department/county/province) of birth; by country. Foreign-born people are excluded from this calculation. *Sources:* USA: Population censuses from IPUMS; Sweden, Argentina: IPUMS-International (Minnesota Population Center, 2020); Russia: Russian Census volumes edited by Troynitsky and compiled in Leasure and Lewis (1968), Table 1 and Table 2 (to exclude Poland); Prussia: *Die Volkszählung Im Deutschen Reich Am 1. Dezember 1880* (1969); France: Inter-university Consortium for Political and Social Research (1992)

B Summary of migration flows

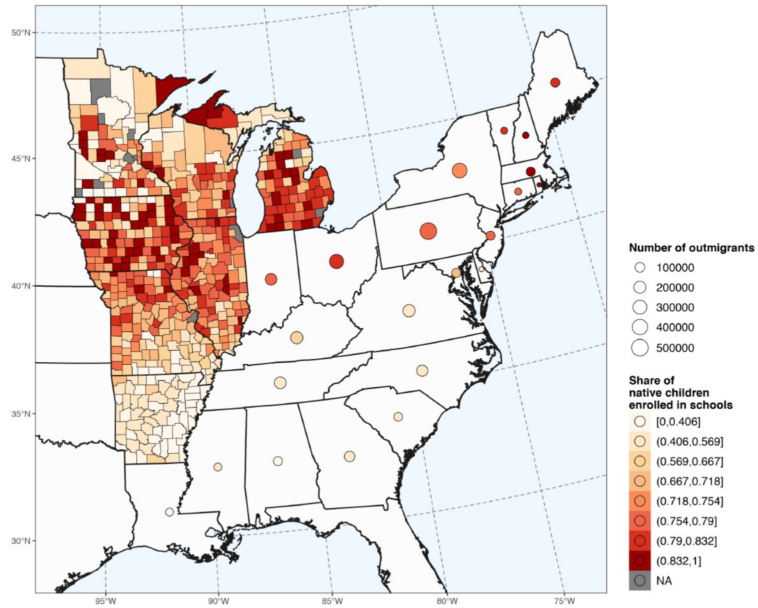
Table 8: States of birth of adults in 1880, with levels of schooling of the natives in those states

Origin State	Arkansas	Illinois	Iowa	Michigan	Minnesota	Missouri	Wisconsin	School Enrollment
Massachusetts	0.12	1.01	1.08	1.12	1.35	0.34	1.25	86.8%
New Hampshire	0.04	0.43	0.60	0.40	0.85	0.12	0.63	84.5%
Rhode Island	0.02	0.12	0.11	0.11	0.13	0.04	0.14	83.4%
Maine	0.07	0.43	0.71	0.57	2.40	0.19	1.28	80.7%
Ohio	1.50	10.03	14.20	7.02	3.38	7.54	3.08	79.3%
Vermont	0.07	1.01	1.55	1.61	2.04	0.27	2.20	79.2%
Pennsylvania	0.75	6.20	8.81	3.63	3.11	3.35	2.84	78.4%
Connecticut	0.06	0.56	0.62	0.78	0.59	0.19	0.91	77.6%
Indiana	1.98	5.20	6.57	1.04	1.56	5.21	0.75	75.7%
New Jersey	0.09	0.99	0.73	0.91	0.37	0.28	0.40	75.7%
New York	0.81	6.69	9.70	26.23	11.24	2.61	13.73	75.1%
Maryland	0.19	0.80	0.68	0.14	0.17	0.67	0.11	67.4%
Kentucky	4.39	4.18	1.56	0.08	0.37	10.52	0.19	59.9%
Virginia	2.15	2.22	2.16	0.23	0.47	5.81	0.25	56.3%
Tennessee	22.10	2.64	0.64	0.03	0.08	7.75	0.07	52.2%
Mississippi	5.78	0.09	0.04	0.01	0.02	0.18	0.01	51.7%
Delaware	0.02	0.12	0.11	0.06	0.04	0.10	0.04	47.7%
North Carolina	5.29	0.75	0.51	0.03	0.07	1.92	0.04	45.3%
South Carolina	3.59	0.14	0.06	0.01	0.02	0.27	0.01	44.1%
Georgia	7.88	0.10	0.03	0.01	0.02	0.34	0.02	41.6%
Alabama	8.67	0.17	0.03	0.01	0.02	0.47	0.01	36.6%
Louisiana	0.92	0.08	0.04	0.01	0.03	0.10	0.02	35.3%
Arkansas	21.54	0.08	0.03	0.00	0.01	0.62	0.00	43.9%
Illinois	2.28	25.97	5.44	0.43	2.13	5.27	1.19	73.8%
Iowa	0.23	0.36	9.67	0.08	0.67	1.25	0.17	77.8%
Michigan	0.15	0.36	0.80	20.45	0.81	0.34	0.53	78.4%
Minnesota	0.01	0.02	0.07	0.02	2.35	0.03	0.08	59.2%
Missouri	4.58	1.24	0.92	0.05	0.23	30.80	0.13	69.1%
Wisconsin	0.08	0.39	1.91	0.56	4.51	0.36	14.70	72.2%
Other US	0.89	0.15	0.12	0.05	0.11	0.34	0.06	
Foreign	3.73	27.48	30.52	34.34	60.87	12.74	55.12	
White Adults	229,687	1,064,431	701,902	706,074	286,142	685,194	522,723	

This table shows, for every of the 7 destination states, the distribution of adult places of birth in 1880 . Adults are defined as 23 year old and older, which is old enough to have children of schooling age (7–14). Composition percentages sum up to 100%. The last column shows education levels for the *native* (mother born in-state) white children in the origin and the destination states. "Other US" includes D.C., the Western states, and the adults with unknown state of birth. Major cities (Chicago, Milwaukee, Minneapolis, St. Louis) are excluded from the destination states.

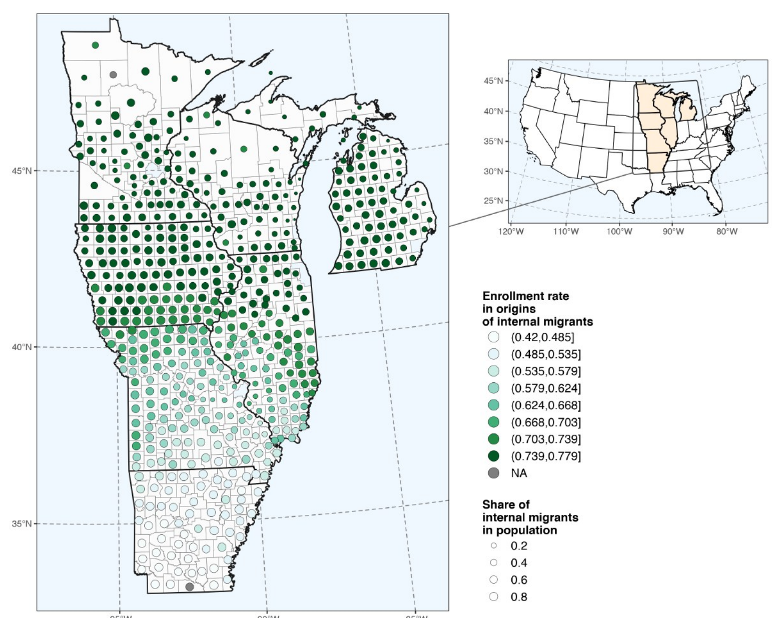
C Maps of school enrollments

Figure 3: School enrollments of native (mother born in state) children in 1880



Source: Population census (Ruggles et al., 2019). For origin states, the size of the circle represents the number of outmigrants from those states to the destination states in the West. Big cities (Minneapolis, Chicago, St.Louis, Milwaukee) were dropped from the sample.

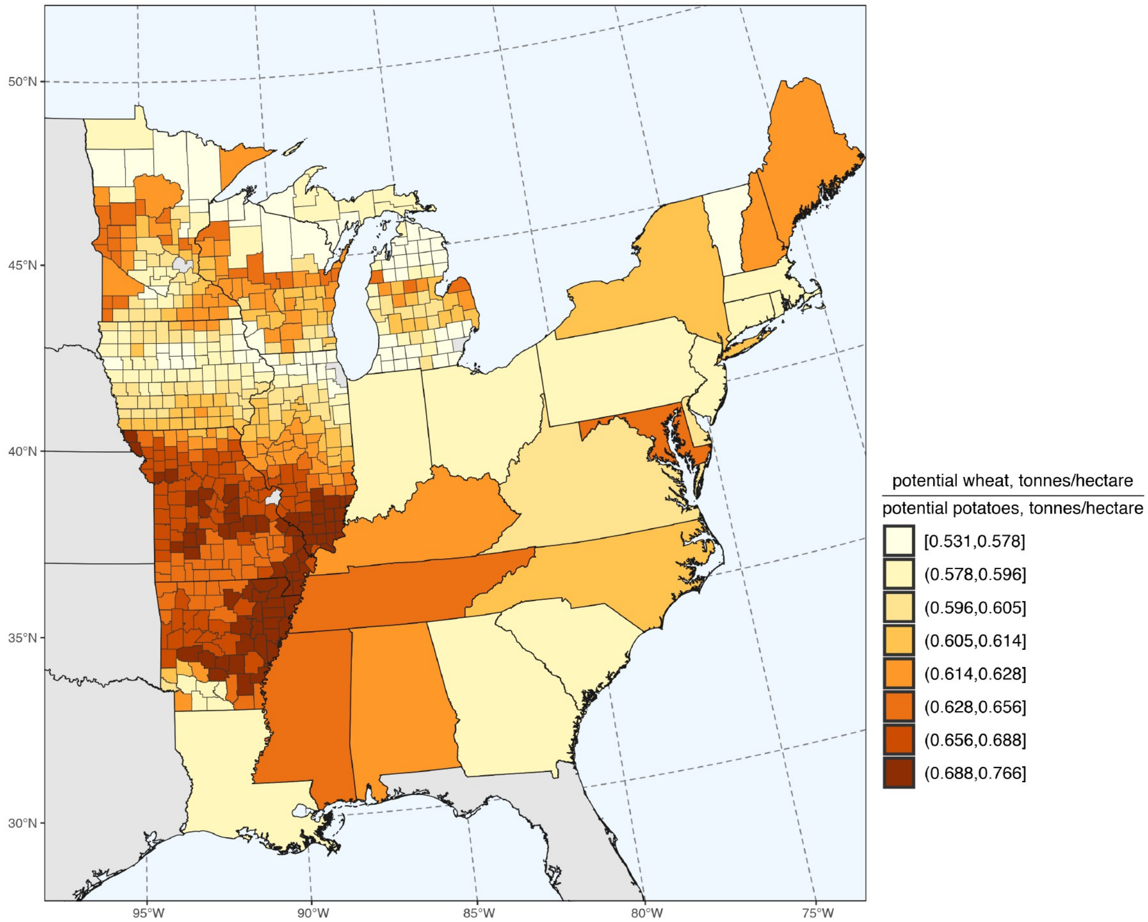
Figure 4: Average school enrollments in the origins of migrants in 1880



Source: Population census (Ruggles et al., 2019). This map shows, for every destination county, the average school enrollment in the origins of internal migrants who moved to that county.

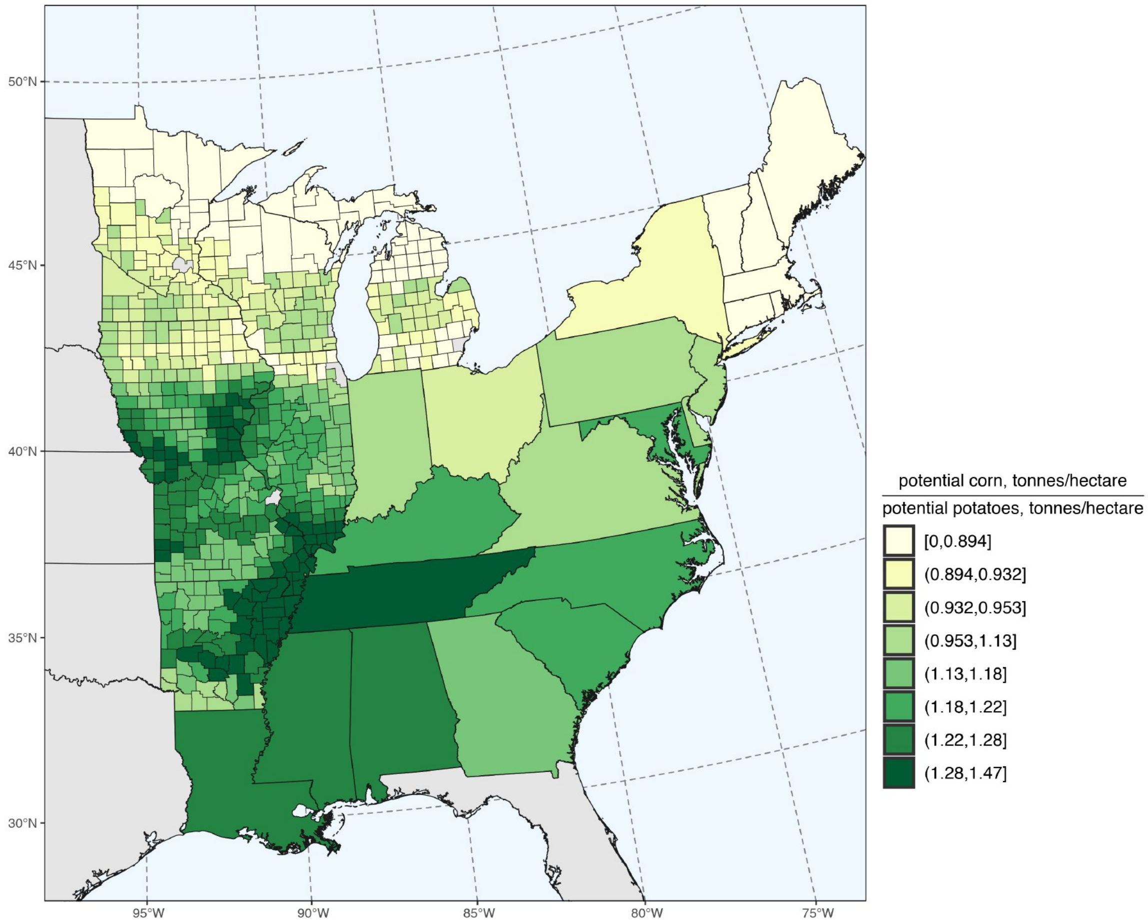
D Maps of relative crop suitability

Figure 5: Potential yields of wheat relative to potatoes, $\frac{wheat}{potato}$



Source: NHGIS county and state shapefiles (Manson et al., 2020), FAO GAEZ potential crop yields (Fischer et al., 2012). For 22 origin states, suitabilities are aggregated by weighting county-level suitabilities by the area of farmland.

Figure 6: Potential yields of corn relative to potatoes, $\frac{corn}{potato}$



Source: NHGIS county and state shapefiles (Manson et al., 2020), FAO GAEZ potential crop yields (Fischer et al., 2012). For 22 origin states, suitabilities are aggregated by weighting county-level suitabilities by the area of farmland.

E Prediction of migration flows

This section shows the first stage for the 2SLS regressions in Table 3. Both figures have the actual average levels of school enrollment in the origins of migrants, $\sum_o s_{odt}h_{ot}$ on the vertical axes, and the average levels of school enrollments of predicted migrants, $\sum_o \hat{s}_{odt}h_{ot}$, on the horizontal axes. Figure 7 uses all the crops (column (5) in Table 2) to construct s_{odt}. Figure 8 uses "past migration predicts future migration" instrument: $\hat{s}_{odt} = \frac{s_{odt-1}}{\sum_o s_{odt-1}} \sum_o s_{odt}$ (1850, the first year, has to be omitted because of the use of a lagged value).

Figure 7: First stage: schooling in origins of actual vs predicted migrants (crops)

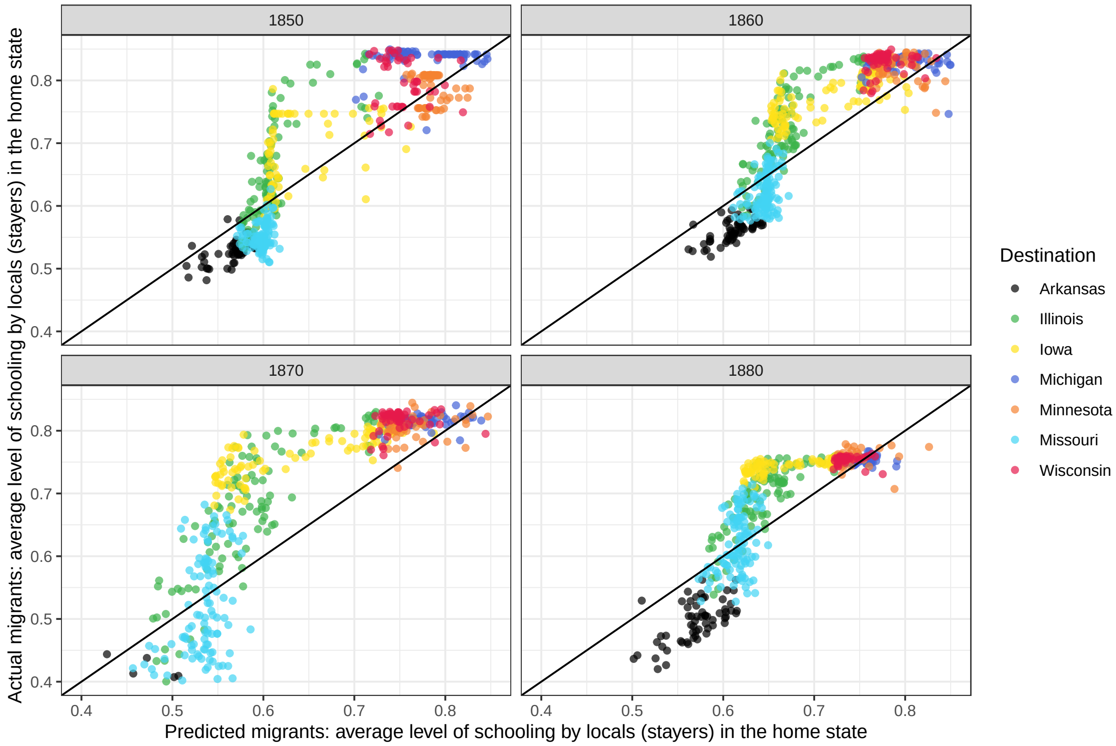
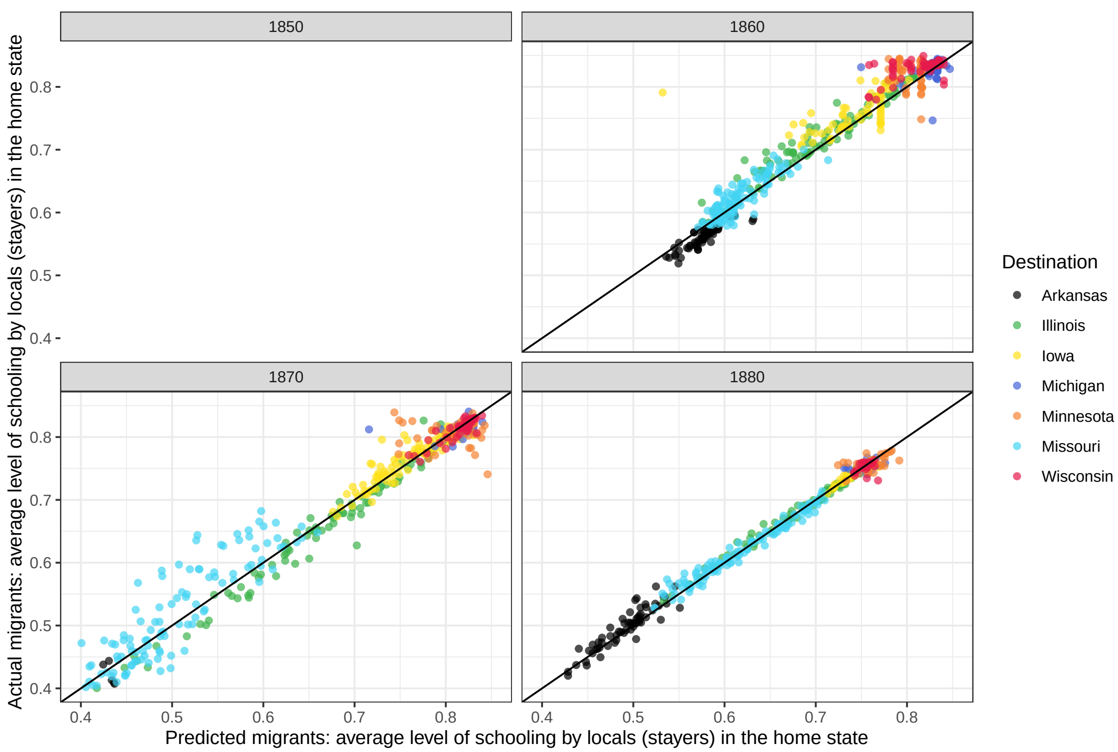


Figure 8: First stage: schooling in origins of actual vs predicted migrants (past migration)



F Rotemberg weights

I follow Goldsmith-Pinkham et al. (2020) to decompose the 2SLS estimate of $\hat{\beta}_M$, the spillover effect of internal migrants on the natives in Equation 3. The decomposition $\hat{\beta}_M = \hat{\alpha}_{ot}\hat{\beta}_{M,ot}$ presents the estimate as a Rotemberg-weighted average of individual estimates $\hat{\beta}_{M,ot}$, where each estimate relies on migrants from one of the 22 origin states in one of the 4 time periods.

The weights $\hat{\alpha}_{ot}$ sum up to one; however, they can be negative. They are higher for origin-years with a stronger correlation of the predicted stocks of migrants $\hat{s}_{odt}$ and the actual stocks of migrants s_{odt} (i.e., stronger first-stage in the just-identified regressions). Examining them characterizes the influence of each origin in each year on the estimated IV coefficient. Goldsmith-Pinkham et al. interpret them as measures of sensitivity to misspecification (i.e., endogeneity of stock of migrants from a particular origin in a particular year). I calculate and examine the Rotemberg weights separately for estimation that predicts stocks of migrants with the dissimilarity in crops and with the early settlement. I report the results for the two types of prediction in Tables 9 and 10 respectively.

For both types of prediction, the five origins with the highest absolute weights include a diverse mix of North and South states. The estimated coefficients for these states $\hat{\beta}_{M,ot}$ are close to the overall estimate $\hat{\beta}_M$. Thus, the sensitivity to reducing the influence of high-weight origins is low. Omitting the origins with negative estimated weights gives a mean estimate that is very close to the baseline (Panels D). Taken together, this evidence shows that the estimation does not rely heavily on including particular states in the estimation, and the results are not sensitive to misspecification of the impact of migrants from any individual origin state. As a robustness check, I in Panel D I present the estimated β_M when excluding the most influential industries.

Panels B show that for both types of instruments, the weights $\hat{\alpha}_{ot}$ are not strongly correlated with neither the levels of enrollment in the origin h_{ot} or with the variance of predicted stocks of migrants $\text{Var}(s_{odt})$. Thus, the estimates are not driven by emigration from either the leading or the backward (in terms of school enrollments) origin states.

The reported weights in Tables 9 and 10 reveal an advantage of prediction based on similarity in crop suitability rather than early settlement. The weights in the crops prediction have lower variance, meaning that different origins play a more uniform role in estimation rather than having extremely high weights on New York (the state with the greatest number of out-migrants and thus the strongest first-stage in the just-identified regressions). In the early settlement prediction, New York's positive weight in 1870 is greater than the combined weights of the next four states. The likely reason is that the prediction relies heavily on the states that happened to be important sources of internal migrants in 1850, when New York was already an important origin. The crops prediction is relatively better at capturing the average of the changing migration patterns in 1850–1880.

Table 9: Summary of Rotemberg weights: crops IV

Panel A: Negative and positive weights α_o

Weights	Sum	Mean	Share
positive	5.412	0.123	0.500
negative	−4.412	−0.100	0.500

Panel B: Correlations

	$\hat{\alpha}_{ot}^{crops}$	$\hat{\beta}_{M,ot}^{crops}$	h_{ot}	$\text{Var}(\hat{s}_{odt})$
$\hat{\alpha}_{ot}^{crops}$	1.000			
$\hat{\beta}_{M,ot}^{crops}$	0.006	1.000		
h_{ot}	0.090	0.020	1.000	
$\text{Var}(\hat{s}_{odt})$	0.021	−0.185	0.055	1.000

Panel C: Top and bottom 5 Rotemberg weight origins

Origin State, o	Year, t	Enrollment, h_{ot}	$\hat{\alpha}_{ot}^{crops}$	$\hat{\beta}_{M,ot}^{crops}$
Top 5 positive weights				
New York	1870	0.830	1.654	0.632
Ohio	1870	0.841	0.797	0.571
Tennessee	1880	0.522	0.586	0.475
Kentucky	1880	0.599	0.426	0.379
Pennsylvania	1870	0.806	0.421	0.614
Bottom 5 negative weights				
Pennsylvania	1880	0.784	−0.280	0.287
Kentucky	1870	0.487	−0.414	0.674
Tennessee	1870	0.364	−0.547	0.595
Ohio	1880	0.793	−0.677	0.247
New York	1880	0.751	−1.294	0.544

Panel D: Estimates $\hat{\beta}_{M,ot}$ for positive and negative weights

Weights	$\hat{\alpha}$-weighted Sum	Share in $\hat{\beta}_M$	$\hat{\alpha}$-weighted Mean
positive	2.571	5.254	0.475
negative	−2.081	−4.254	0.472
Excluding top and bottom five			
positive	0.372	0.761	0.244
negative	−0.526	−1.076	0.438

This table reports Rotemberg weights for the estimation of Equation (3), following Goldsmith-Pinkham et al. (2020). For simplicity, I use the prediction from all crops combined together.

Table 10: Summary of Rotemberg weights: early settlement IV

Panel A: Negative and positive weights α_o

Weights	Sum	Mean	Share
positive	6.073	0.202	0.455
negative	−5.073	−0.141	0.545

Panel B: Correlations

	$\hat{\alpha}_{ot}^{past}$	$\hat{\beta}_{M,ot}^{past}$	h_{ot}	$\mathrm{Var}(\hat{s}_{odt})$
$\hat{\alpha}_{ot}^{past}$	1.000			
$\hat{\beta}_{M,ot}^{past}$	0.004	1.000		
h_{ot}	0.080	0.216	1.000	
$\mathrm{Var}(\hat{s}_{odt})$	0.040	0.048	0.102	1.000

Panel C: Top and bottom 5 Rotemberg weight origins

Origin State, o	Year, t	Enrollment, h_{ot}	$\hat{\alpha}_{ot}^{past}$	$\hat{\beta}_{M,ot}^{past}$
Top 5 positive weights				
New York	1870	0.830	2.312	0.674
Tennessee	1880	0.522	0.651	0.620
Kentucky	1880	0.599	0.632	0.796
Ohio	1870	0.841	0.563	0.670
Virginia	1880	0.563	0.337	0.882
Bottom 5 negative weights				
Virginia	1870	0.360	−0.203	1.097
Ohio	1880	0.793	−0.375	0.935
Kentucky	1870	0.487	−0.516	1.166
Tennessee	1870	0.364	−0.586	0.587
New York	1880	0.751	−1.968	0.497

Panel D: Estimates $\hat{\beta}_{M,ot}$ for positive and negative weights

	$\hat{\alpha}$-weighted		$\hat{\alpha}$-weighted
Weights	Sum	Share in $\hat{\beta}_M$	Mean
positive	3.722	5.822	0.613
negative	−3.083	−4.822	0.608
Excluding top and bottom five			
positive	0.847	1.325	0.555
negative	−0.850	−1.330	0.619

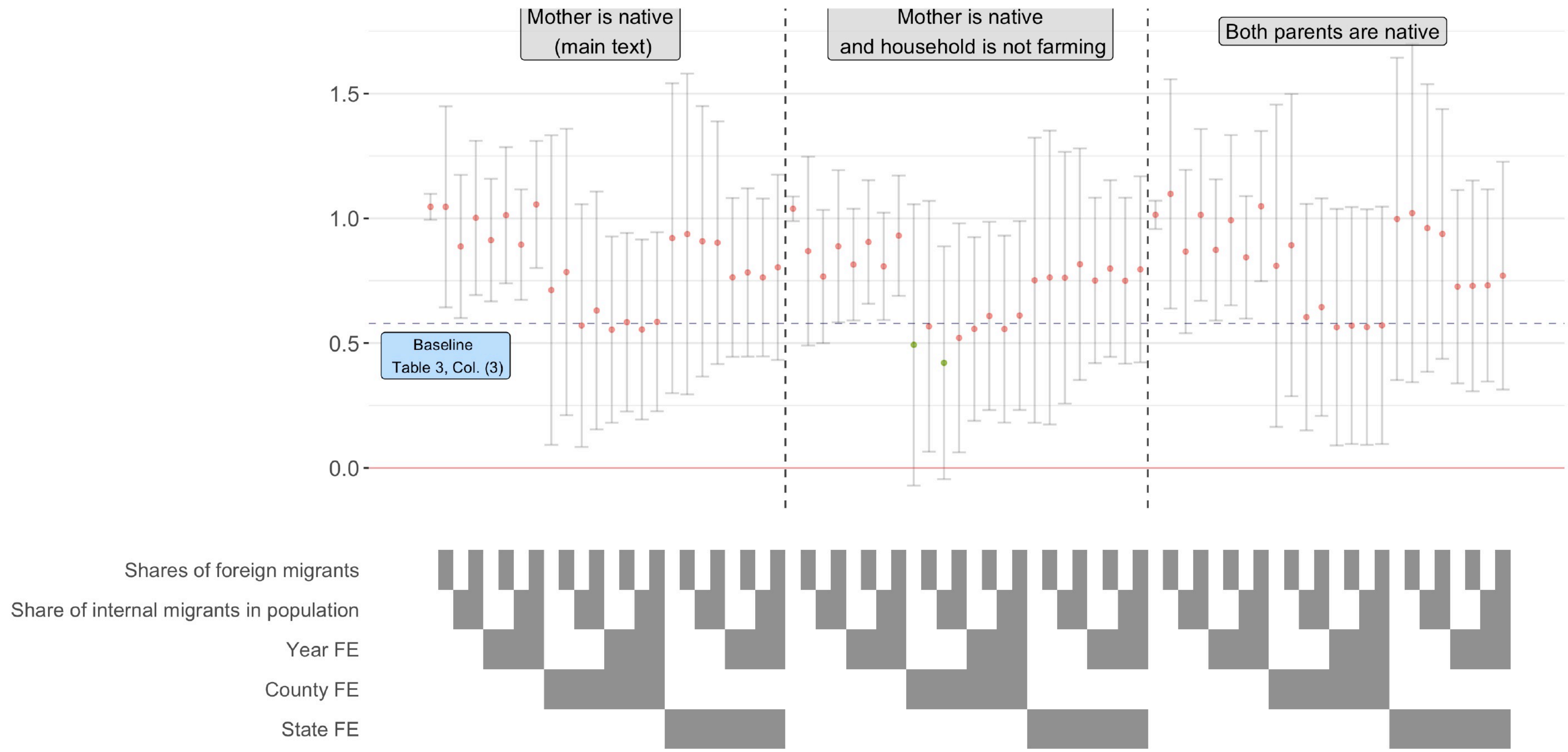
This table reports Rotemberg weights for the estimation of Equation (3), following Goldsmith-Pinkham et al. (2020).

G Alternative specifications and robustness checks

I conduct robustness checks for the estimates β_M in Equation (3), reported in Table 3. Figure 9 below shows the estimates $\hat{\beta}_M$ across different specifications. The right-hand-side variable is the county-level school enrollment of children of native mothers (as in the main text), the children of non-farming parents and the children both of whose parents are native. The effects are similar across all three outcome variables. In specifications with county fixed effects, estimated coefficients are closed to $\hat{\beta}_M = 0.579$ (the preferred estimate from Table 3).

The results from alternative specifications suggest that the effect did come from farmers introducing more education-heavy practices to other farmers (the spillover effects on non-farmers would have been lower). They also suggest that the spillovers from internal migrants onto the natives were not driven by the marriage between internal migrants and locals, as the sample that includes mixed couples with only mother being native experience a similar effect to the fully native couples.

Figure 9: Estimates of $\hat{\beta}_M$ in Equation (3): Robustness checks and alternative outcome variables



Estimates of $\hat{\beta}_M$ from Equation (3) with three alternative outcome variables: the school enrollment of children of native mothers (used in the main text), the school enrollment of children of native mothers whose families are not in farming, and the school enrollments of children of native mothers and native fathers (a stricter definition of nativity). Native parents are defined as those who are born in the state of residency. For the non-farmers households, I exclude families where either (1) the head of the household reported a farming-related occupation (farmer, laborer, farm manager) or (2) the family is living on a farm. The bottom part shows various control variables and fixed effects that are included or excluded from estimation. Coefficients are estimated using 2SLS, where the average school enrollment in the origins of internal migrants $\sum_j s_{odt} h_{ot}$ is instrumented with $\sum_j \hat{s}_{odt} h_{ot}$. The predicted stocks of migrants $\hat{s}_{odt}$ are based on crop dissimilarities between o and d. The control variable share of internal migrants is calculated as $\sum_o \hat{s}_{odt}$. Conley st. errors with a 100 mile cutoff are used to construct the 95% confidence intervals around the estimates.

H Testing for conflating short- and long-term responses

In this section, I follow the double-instrumentation approach proposed by Jaeger et al. (2018) to examine whether the crop similarity IV and the past migration IV conflate the short-term and the long-term responses.[24] This approach tackles the issue of dynamic adjustments following the shocks to the composition of internal migrants.

Equation (3) assumes that the impact of the stock of migrants is contemporaneous: the current composition of migrants affects the current enrollments of the locals. In this section, I test the robustness of the results to misspecification with respect to the potential dynamic responses. The tests suggested by Jaeger et al. (2018) checks to what extent the assumption of contemporaneous responses in Equation (3) matters for the estimation.

Specifically, the estimated effect $\hat{\beta}_M$ would pick up the effect of both the contemporaneous and the past migration if the instrument $\hat{s}_{odt}h_{ot}$ is not substantially better at predicting $s_{odt}h_{ot}$ than the $\hat{s}_{odt-1}h_{ot-1}$, the lagged instrument. The "double-instrumentation test" proposed by Jaeger et al. consists of running a first-stage regression of $s_{odt}h_{ot}$ on both $\hat{s}_{odt}h_{ot}$ and $\hat{s}_{odt-1}h_{ot-1}$.

Unlike (Jaeger et al., 2018), my model is estimated in levels with county-fixed effects rather than in changes. Thus, instead of performing the double-instrumentation test individually for 1870 and 1880, I pull them together and perform the test in a two-period panel with fixed effects, like in the original estimation reported in Table 3.

The results are in Table 11. For both the crops prediction and the past migration prediction, the lagged instrument does not predict the contemporaneous regressor (the average enrollments in the origins of internal migrants). For the prediction that relies on crops, only the coefficient

[24](Jaeger et al., 2018) discuss the use of the past migration IV in the context of wage adjustment to the labor supply from increased number of immigrants. They show that the past migration IV does not tackle the endogeneity that arises because current-period response to shocks happen while the wage market adjusts to the shocks from the previous periods. Thus, the estimation that relies on the past migration IV will yield biased estimates of the responses in the current period because it will pick wage responses from the migration shocks in the past period.

on the contemporaneous value of migration is statistically significant (column (1)). However, for the past migration, there is strong negative effect of lagged migration. Taken together, the results suggest that the concerns raised by Jaeger et al. (2018) are unlikely to corroborate the estimation results in my context.

Table 11: First Stage: test from Jaeger et al. (2018)

	Dependent Variable			
	Educ. in origins of migrants	Lagged Educ. in origins of migrants	Educ. in origins of migrants	Lagged Educ. in origins of migrants
	(1)	(2)	(3)	(4)
Educ. in origins of migrants predicted by crops	1.648 ***	0.357 **		
	(0.202)	(0.177)		
Educ. in origins of migrants predicted by crops (lagged)	0.141	2.156 ***		
	(0.248)	(0.397)		
Educ. in origins of migrants predicted by past			0.733 ***	0.268 **
			(0.061)	(0.115)
Educ. in origins of migrants predicted by past (lagged)			−0.248 ***	1.169 ***
			(0.080)	(0.143)
Shares of foreign migr.	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
County FE	✓	✓	✓	✓
N	1209	1209	806	806
R^2	0.974	0.971	0.993	0.988
Kleibergen-Paap statistic	125.100	38.829	361.205	51.597
Kleibergen-Paap p-value	0.000	0.000	0.000	0.000

*** p < 0.01; ** p < 0.05; * p < 0.1

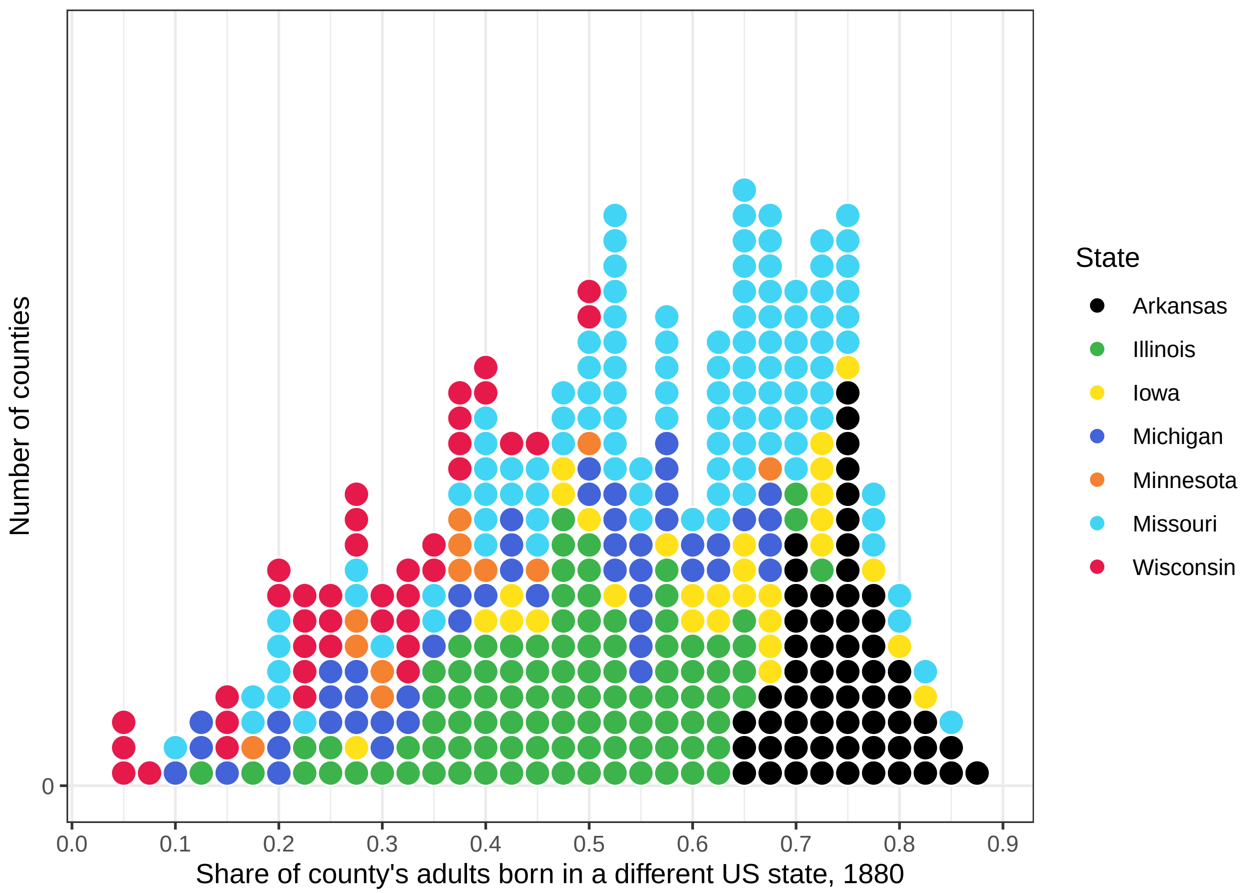
This table exmaines the dynamics of the explanatory variable and the IV following the procedure proposed by Jaeger et al. (2018). The average enrollment in the origins of internal migrants and its lag is regressed on the IV (the average enrollment in the origins of predcited migrants) and the IV's lag.

I Share of Internal Migrants across counties

Figure 10 shows the distribution of the share of internal migrants in the counties of the seven destination states in the sample. The number of adults migrants from 22 origin states in the East is divided by the total adult population, which also includes foreign-born and native adults. Thus, the share is calculated as:

$$s_{\text{internal migrants},\, dt} = \frac{\sum_o N_{odt}}{\sum_o N_{odt} + N_{\text{foreign},\, t} + N_{\text{native},\, t}}$$

Figure 10: Distribution of the share of internal migrants across counties in the sample



Distribution of the share of internal migrants in the population in 1880, by county. Each dot represents one of the the 403 counties, colored by the state where it is located. Only adults aged 23 are above are included.