Migration, Social Change, and the Early Decline in U.S. Fertility*

Alessandra Fogli†

Stefania Marcassa‡

Abstract

We study the impact of internal migration on the U.S. fertility transition in the Nineteenth century. We show that fertility declined faster in counties characterized by a higher outward migration, especially towards the Western frontier. We exploit the number of acres granted to veterans of the Civil War to estimate the causal effect of migration on fertility decline. Our theory is based on the diffusion of new family values governing intergenerational behavior with respect to saving and fertility. Migration and the lack of remittance technology lowered expected transfers from children, and incentivized precautionary savings of parents. Results are robusts to several measures of fertility and internal migration.

JEL classification: C33, J11, J13, N32, O15, R23

Keywords: Fertility rate, fertility transition, counties, outward migration.

*We are grateful to seminar participants at Meetings of the SOHE 2019 and 2018, DIAL Conference 2017, Université de Louvain, Università di Trento, CY Cergy Paris Université, SABE Conference 2016, COSME 2015, Universitat de Barcelona, Econometric Society Meeting 2014, EEA Toulouse 2014. An earlier version of this paper has been circulated under the title "The Geography of Social Change". This research has been conducted as part of the project Labex MME-DII (ANR11-LBX-0023-01).

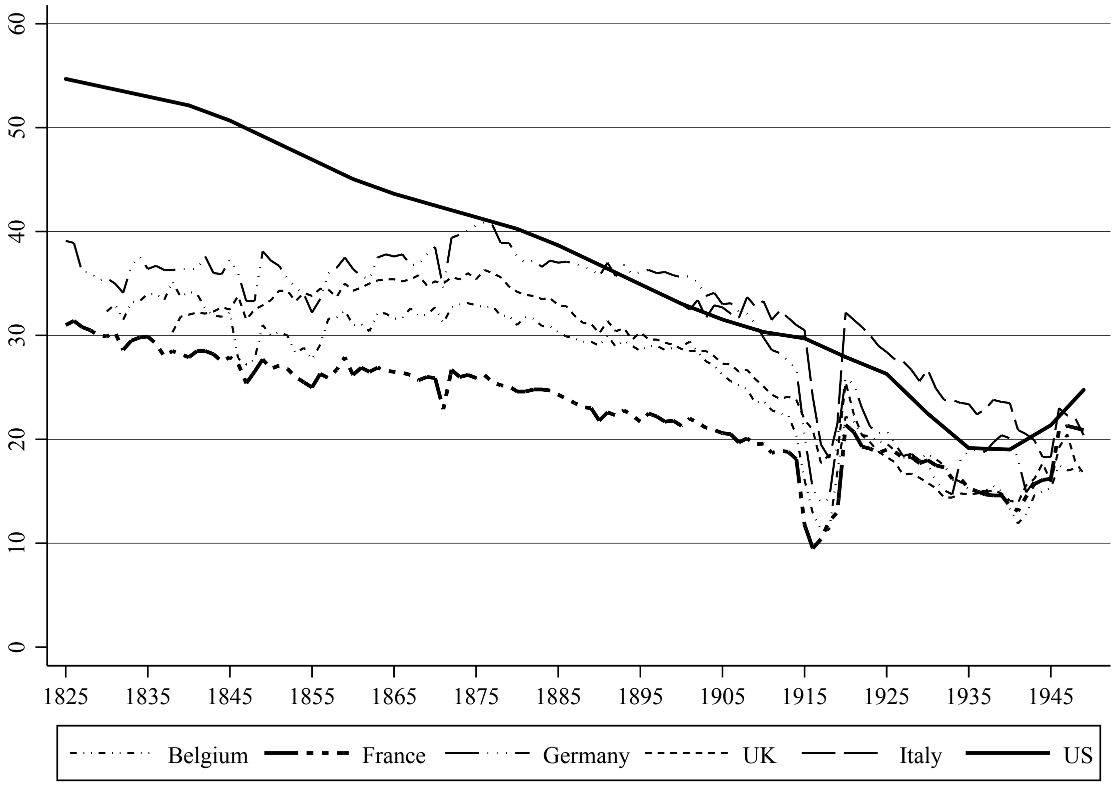
†Federal Reserve Bank of Minneapolis, Minneapolis, MN, US. Email: afogli00@gmail.com.

‡Corresponding author. CY Cergy Paris Université (THEMA), France. Email: stefania.marcassa@cyu.fr.

1. INTRODUCTION

In the U.S., fertility transition was underway from the beginning of the 19th century while all other Western countries, with the exception of France (Daudin et al., 2019), began their sustained decline in birth rates much later, in the late 19th or early 20th century. As shown in Figure 1, the U.S. crude birth rate went from about 55 in 1825 to about 33 in 1900. This timing is difficult to reconcile with conventional theories that place great reliance on structural changes in child costs and benefits associated with modern economic development such as urbanization, industrialization, the rise in literacy, and the increase in female labor force participation. As we argue below, the American fertility decline started well before many of these changes became important.

Figure 1 – Births per 1000 population



Sources: European data are from Mitchell (2003). US data are from Vital Statistics.

We provide an alternative explanation that is consistent with the timing and geographical pattern of the U.S. fertility decline. Our theory relies on a change in family social norms: patriarchal family, based on intergenerational transfers, dissolved as young Eastern sons migrated towards the West. As described by Ferrie (1997), migrants were attracted by natural resources at locations distant from the narrow band of initial settlement on the Atlantic coast. Farmers moved to more productive land in the Great Plains by the middle of the 19th century, and mineral and timber resources were used to good advantage by migrants to the

West and the Northwest. During this process, the lack of remittance technology lowered expected transfers from migrant children, and incentivized precautionary savings of parents especially when the distance from the home county was high. Lewis (1983) documents that savings rate rose from 16 to 22 percent between 1830 and 1900. Moreover, he also estimates that, between 1830 and 1900, about one-quarter of the 6-percentage-point rise in the savings rate can be attributed to a decline in the dependency rate (the ratio of dependent children to adults). By the end of the 19th century, the rates of population growth in each region converged, and the geographic distribution of population became stationary (see Vandenbroucke, 2008b, Figure 2).

We test whether the characteristics of the Westward migrants, and the distance of migration have contributed to the drop in fertility.[1] Our identification strategy relies on the heterogenous geographical patterns of fertility decline. The decrease in fertility started in the East and, by the end of the century, the differences in the child-woman ratio across counties had disappeared, and its variance had halved, as shown in Figure 11.[2] In our baseline regression, we find that a one percent increase in the percentage of migrants contributes about 4.4 percentage points to the fertility decline in the home county from 1850 to 1880. Most importantly, the size of the correlation increases with the distance of migration. That is, a one percent increase in the percentage of migrants moving to the *frontier*, i.e. towards counties located in states that do not belong to the original settlements, contributes 4.7 percentage points to the fertility decline. Interestingly, in the last two cases, the effect of within state or non-frontier movers on fertility is positive but not significant, indicating that distance is an important determinant of the correlation. We control throughout the analysis for state-level fixed effects, which absorb any heterogeneity in migration patterns across U.S. states.

To address endogeneity concerns, we take advantage of the historical institutional backgroud and study how child-woman ratio and migration have been conditioned by land policy set by the U.S. government. Between 1847 and 1855, the Congress granted acres of land to veterans of the Civil War and their heirs through the Homestead Acts. Our source of exogenous variation is the average number of acres granted as a percentage of the total improved acres available in the home state. The idea is that families whose members were involved in the Civil War received federal land for private ownership which provided them with incentives to migrate in the states where the acres were located. This generates as-good-as-random assignment of migrants. We also estimate the IV model for a sample of states where fertility norms should not be affected by migration because they are located on the West coast. As expected, our estimates are close to zero for these counties. This suggests that our empirical model captures the effect of Westward migration from the settlements and not other confounding factors.

[1]The *Westward expansion* denotes the 19th-century movement of settlers into the American West, which began with the Louisiana Purchase and was fueled by the Gold Rush, and the Oregon Trail.

[2]The child-woman ratio is defined as the ratio of the number of white children of age 0 to 9, and the number of white women of age 15 to 44, in a given year.

We shed light on the mechanism by providing evidence of the change in the composition of the American family, and its impact on the saving rate. Relying on Census data and the historical literature, we show that the percentage of multigenerational families declined by more than 10 percent by the end of the 1800s. Moreover, the improvement of the economic conditions outside of the farm provided incentives to move out of the family and diminished the role of the patriarch. The arising of new economic opportunities in urban areas and in the Western part of the country coincided with the revolution of the family structure, and with the rising of the saving rate. In order to measure the correlation between financial development and fertility, we compute the velocity of money using 1850 county level data on nominal manufacturing output value and the amount of money in circulation. The idea being that a high (low) money velocity implies a low (high) saving rate. We show that the 1850 child-woman ratio is positively correlated with the velocity of money. Hence, counties with a lower level of fertility were associated with a lower money velocity, or a higher saving rate. This finding supports our idea that the old-age support guaranteed by the work of children in the farm was replaced by banking deposits.

Our empirical analysis relies on historical data drawn from several sources. To quantify the number of internal migrants at the county level, we use an innovative dataset: the IPUMS Linked Representative Samples (Ruggles et al., 2015) which link records from the 1880 complete-count database to the 1850 complete-count of the U.S. census of the population. For each individual we have information about her *county* of residence at the time of the 1850 Census, her destination *county* of residence in the linked sample of 1880, and other demographic characteristics. We believe that this measure of outward migration is more accurate than those proposed by Carter et al. (2001) and Rosenbloom and Sundstrom (2003) which use *state* level data to *infer* the number of migrants from either the change population composition over Census decennials (Carter et al., 2001), or from the difference between state of birth and state of residence declared at the time of the Census interview (Rosenbloom and Sundstrom, 2003). As county-level data of internal migration are not available for the period 1800-1840, our analysis will focus on the second half of the Nineteenth century. For fertility and county characteristics, we follow Haines and Hacker (2011), and use the ICPSR data set "Historical, Demographic, Economic and Social Data: The United States, 1790-2002" (Haines et al., 2010). In particular, fertility is approximated by the child-woman ratio, i.e., the ratio between the number of white children of age 0 to 9, and the number of white women of age 15 to 44 or 49, depending on sample availability.

We perform several robustness checks on both the fertility measure and the sample of migrants. The child-woman ratio (CWR henceforth) does not account for age structure or marriage patterns, both of which changed significantly during this period. A shift in age structure in the absence of a fertility reduction would increase or decrease the CWR. By failing to take into account the changing population age structure, the CWR may understate the change in behavior. To prove the robustness of our results, we use two alternative measures of fertility. First, we compute the number of children ever born as in Jones and Tertilt (2006), available at the state level. Second, we consider children of age one to four and five to nine years old to take into account the high child mortality rate of that period of time. In both cases, results are confirmed and reported in Section 3.2. In Appendix D, we show estimates of a sample where we exclude counties that belong to the bottom and top 5% of the distributions of CWR and percentage of migrants, to verify that results are not led by outliers in the dependent or independent variable. We also run the regressions on a sample of migrants that includes women. All these robustness checks reduce unobserved heterogeneity between counties but leave our results unchanged. Finally, we show that the percentage of migrants from 1850 to 1880 has some persistent effects on the decline of the CWR from 1850 to 1900, and, as we would expect, this effect is decreasing over time.

As a further test of validity of our mechanism, we measure the correlation between fertility decline and migration in Europe. We plot the number of emigrants from several European countries to the U.S. from 1820 to 1920, and applied the same strategy of Knodel and van de Walle (1979) to identify the date at which marital fertility declined by 10 percent. At first sight, we cannot rule out the link between the year associated to the highest migration flow and the fertility decline. Few exceptions are present, as France and Ireland. To get a measure of the correlation, we use micro data from Mitchell (2003) and compute both the percentage of migrants from Europe by decade, and the change in the child-woman ratio (as in our analysis for the U.S.). We find that the correlation is significantly positive, but smaller than the ones estimated for the U.S.

1.1. Other Theories of Fertility Transition

In this section we go through a review of the existing theories of fertility decline, and discuss our contribution. First, we know from demographers that fertility transition usually starts after or at the same time as the reduction in mortality. Notestein (1945) argued that couples in high-mortality societies have a lot of births to ensure a surviving brood of the desired size. An exogenous mortality decline induces couples to have fewer children because they do not need so many "spares". As discussed in Guinnane (2011), Notestein's account was motivated by the experience of developing countries after World War II, where public-health interventions first reduced infant and child mortality, and in some places those developments were followed by declines in fertility. But, the Notestein's argument does not fit the timing of the historical declines in fertility and mortality. As shown in Figure 10 using data from Haines (2001), there is no sustained fall in the infant mortality rate until the 1890s.

Second, Yasuba (1962) and Easterlin (1976) argued that rising population density or diminishing land availability was the prime mover in the downward secular course of fertility. Yasuba (1962) proposed that East-West differences in population density could account for the geographical pattern of fertility. Acquisition of new land in the settled areas became increasingly difficult and costlier and the average distance from the settled to the new areas where land was plentiful became farther. Consequently, fertility in the older communities may have been reduced directly in response to the decreased demand for children or indirectly as a result of the rise in the age at marriage and the fall in the incidence of marriage. Easterlin (1976) recasted Yasuba (1962)'s argument: He suggested that parents had an altruistic motive to preserve and augment the family's wealth and to pass those assets on to their children when they died. Over time fertility would decline because, in any given community, land would become increasingly scarce, more expensive, and more difficult to acquire.

This theory has few limitations. First, improvements in transportation and communication, the continuing release of the public domain at land auctions, and rising agricultural incomes should have made it easier to purchase a farm. Second, the land-scarcity model has difficulty explaining why fertility was so high at the beginning of the 19th century and why the onset of the fertility decline occurred at the time it did. Fertility began to fall at precisely the time American land policy changed, opening up vast expanses of public domain to settlement. Beginning with the Congressional Act of 16 September 1776 and the Land Ordinance of 1785, a wide variety of Congressional acts governed the distribution of federal land in the thirty public land states. Various acts opened up new territories, established the practice of offering land as compensation for military service, and extended preemption rights to squatters (e.g., the Indian Removal Act in 1830, the Preemption Act in 1841, and the Homestead Acts in 1862). These acts each resulted in the first transfer of land from the federal government to individuals. Relatively speaking, the threat of land scarcity must have appeared much greater in 1800 than at any time during the period between 1815 and 1840.

We show that the impact of these variables is different whether we consider their effect on the level of the CWR in 1850 or on the decline of the CWR from 1850 to 1880. We show that, in a state fixed effect regression, the coefficients on land availability and farm value are statistically significant and have signs consistent with those of Easterlin (1976) and Yasuba (1962). When we consider the effect of land availability and value of farms in 1850 on the decline of the CWR from 1850 to 1880, their coefficients are not statistically significant. Density (or population per square mile) is significantly negatively correlated to both the level of the CWR in 1850, and the decline of the CWR. A possible reason is that density reflects the extent of the urbanization of a population. Since urban fertility was lower than rural fertility throughout the country in 1850, this may produce a strong correlation between density and fertility.

An alternative theory that tried to overcome the aforementioned issues is the one by Sundstrom and David (1988) based on state level data. They suggested that the high demand for children in the early years was motivated by parents' desire to provide for their own old-age security.[3] By having a large number of children and by offering these children a portion

[3]Note that before the Social Security Act of 1935, the United States had no social insurance system. See Caucutt et al. (2013) for an analysis of the association between urbanization, industrialization and the rise of social insurance.

of the farm family's wealth as a potential inheritance in exchange for their continuing support, parents could ensure for themselves a flow of goods and services even after their own ability to support themselves was diminished by old age. According to their argument, the old-age security motive for having many children would have weakened substantially when opportunities outside of agriculture began to improve some time in the early-nineteenth century. The importance of inheriting farmland would be diminished. Testing their model using state-level data for 1840, they concluded that nonagricultural labor market opportunities had a large, negative effect on rural white fertility.

We make three contributions with respect to Sundstrom and David (1988). First, we are able to exploit the geographic variation of migration at the county level and obtain a more precise estimate of its effect on CWR from 1850 to 1880. This analysis would not be as robust with state level data, especially at a time when the number of politically organized states were limited to 26.[4] Second, we show that distance of migration plays a role, weakening the theory of Sundstrom and David (1988), which does not allow to differentiate between nearby or far nonagricultural opportunities. It is not enough to move out of its own family, but the distance has to be such that, given the geographical constraints of that time, interactions and remittances become very costly. Hence, we do not rule out the importance of industrial development, but we claim that it was not the most relevant determinant of fertility decline in nearby rural areas. In fact, our data show that the CWR level of 1850 depend negatively on the fraction of men manufacturer workers, and the decline of the CWR from 1850 to 1880 is positively correlated to the increase of the fraction of manufacturer workers. Nonetheless, the fraction of migrants has a higher correlation with the decline of the CWR than the percentage of manufacturing workers.

Our paper is also related to the macroeconomic literature on the fertility transition. In particular, Vandenbroucke (2008a) and Vandenbroucke (2008b) calibrate an endogenous growth model and show that the decrease in transportation costs induced the westward migration, and population growth was key for productive investments in land. He also shows that the role of international immigration is negligible. Our analysis adds the role of internal migrants, and show that the presence of the railroad and navigable rivers or canals had a negligible role in the decision to depart from the East. The major incentive to migration has instead been provided by the availability of land in the Midwest.

The paper proceeds as follows. In Section 2, we review the historical features of the CWR and westward migration. Section 3 presents the empirical strategy, results, and robustness checks. The mechanism is discussed in section 4. In section 5 we speculate on the fertility transition in Europe. Section 6 concludes.

[4]See Figure 3.

2. HISTORICAL DATA

In this section we describe the variables used in the empirical analysis: the dependent variable (i.e. the child-woman ratio), and the independent variables (i.e. the westward migration and its characteristics, and the economic and social characteristics of the counties or states).

2.1. Dependent Variable: The Child-Woman Ratio

The aggregate census data from Haines et al. (2010) do not allow us to compute fertility rates, as the total number of children born per woman is not available. We therefore apply the child-woman ratio (CWR), an indirect fertility measure traditionally defined as the number of white children aged 0-9 per white woman aged 15-49 (Carter et al., 2006), at the county level.[5]

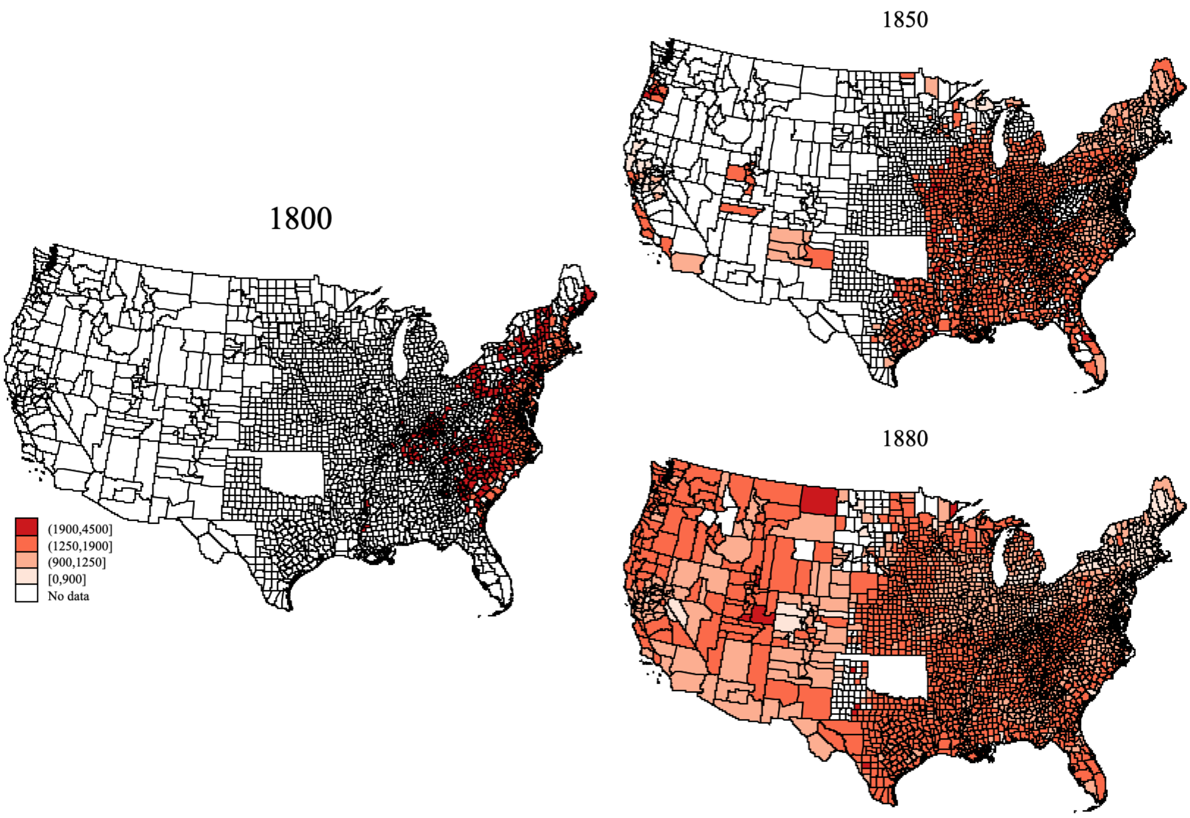
Figure 2 depicts the decline of the child–woman ratios by county from 1800 to 1880.[6] The maps clearly show that CWRs decreased remarkably from 1800 to 1880, and the decline was not homogenous across the country. On average, the CWR decreased from 1,955 (1,406) in 1800 (1850) to 1,267 children per 1,000 women in 1880. Dispersion varied also from 1800 (1850) to 1880: the standard deviation dropped from 401 (281) to 80 in 80 (30) years. In 1800, the county with the lowest CWR (859) is in Virginia (Norfolk City), while the highest CWR county (Muhlenberg) is in Kentucky with 4,370 children per 1,000 women. In 1850, the county with the lowest CWR (182) is in California (Colusa), while the highest CWR county (Manistee) is in Michigan with 5,000 children per 1,000 women. By 1880, the CWR was homogenous across the territory.

Haines and Hacker (2011) showed that the east-west differences are apparent throughout and New England was the area where CWRs was the lowest at the beginning of the century. There is a less apparent suggestion of a north-south gradient, with the South having had higher CWRs. Urban places had lower CWRs than did rural areas, but the decline took place in both rural and urban areas between 1800 and 1840. Rural CWR remained above urban CWR, but absolute differences diminished as both types of residents limited their family size. Variation across space narrowed from 1800 onwards. In 1810, the South had fertility ratios over 30 percent higher than in New England (the lowest fertility region). This differential had increased to about 60 percent in 1860, and the relative difference was nearly the same at the end of the century. The Midwest moved from being a region of quite large families to, by 1900, one with fertility close to the leaders in the transition, i.e. New England and Middle Atlantic regions.

[5] Fertility rates as commonly defined by demographers can be computed starting from 1933 on as vital statistics are not available earlier than that.

[6] See Figure 11 in Appendix A to observe the complete transition from 1800 to 1880. The historical boundary files have been downloaded from the IPUMS and the NHGIS websites.

Figure 2 – Map of CWR, 1800 to 1880



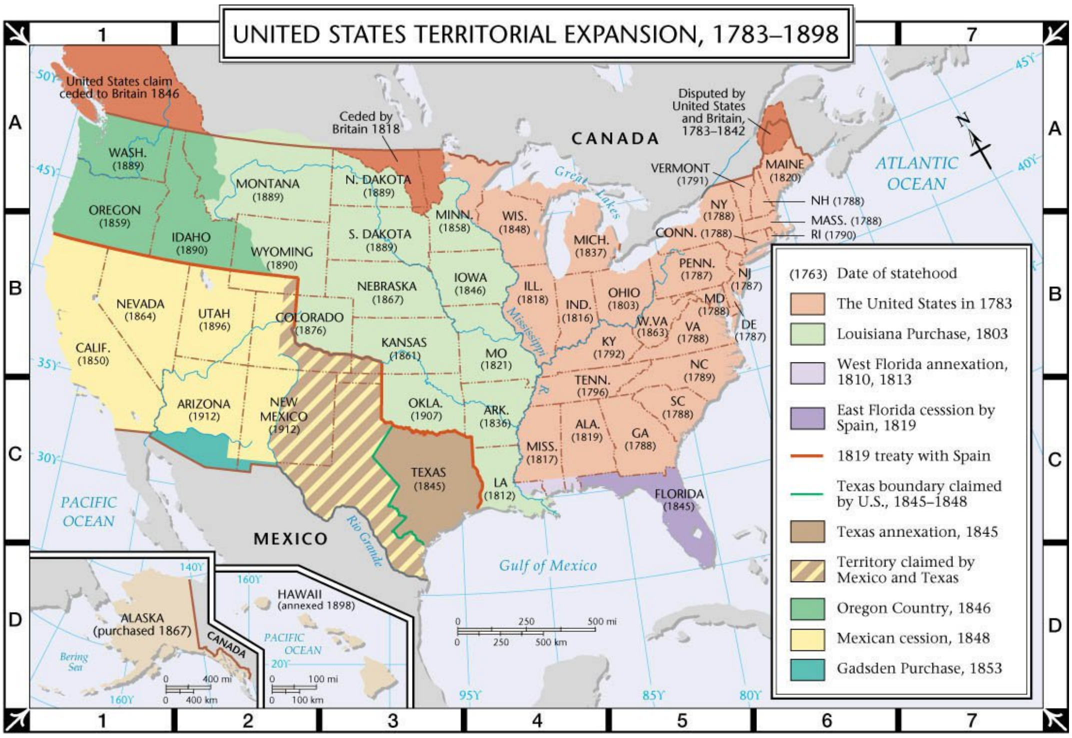
Source: Authors' computations using data from Haines et al. (2010) and Ruggles et al. (2015).

Even though this is the commonly used measure of historical fertility, we are aware of the fact that it does not account for age structure or marriage patterns, both of which changed significantly during this period. A shift in age structure in the absence of a fertility reduction would increase or decrease the CWR. By failing to take into account the changing population age structure, the CWR may understate the change in behavior. It clearly suffers from the differential mortality and under enumeration of women and children and conflates the separate effects of marriage and marital fertility. Nonetheless, it is a useful indicator of differential fertility across space and time for geographic areas. As a first robustness check and to take into account of infant mortality, we compute the CWR using the number of children of age 1 and older, as infant mortality is defined as the death of young children under the age of 1. Second, we also estimate the empirical model using the measure of children ever born employed by Jones and Tertilt (2006) as dependent variable. All results are in line with our benchmark findings.

2.2. Westward Migration

In 1803, President Thomas Jefferson purchased the territory of Louisiana from the French government, and set the beginning of the westward expansion. In Figure 3, we can see that the Louisiana Purchase stretched from the Mississippi River to the Rocky Mountains and from Canada to New Orleans, and it doubled the size of the United States.

Figure 3 – Territorial Expansion of the United States



Source: Pearson Education.

By 1840, nearly 7 million Americans, i.e. 40 percent of the nation's population, lived in the trans-Appalachian West. Following a trail blazed by Lewis and Clark, most of these people had left their homes in the East in search of economic opportunity. The western frontier offered the possibility of independence and upward mobility.

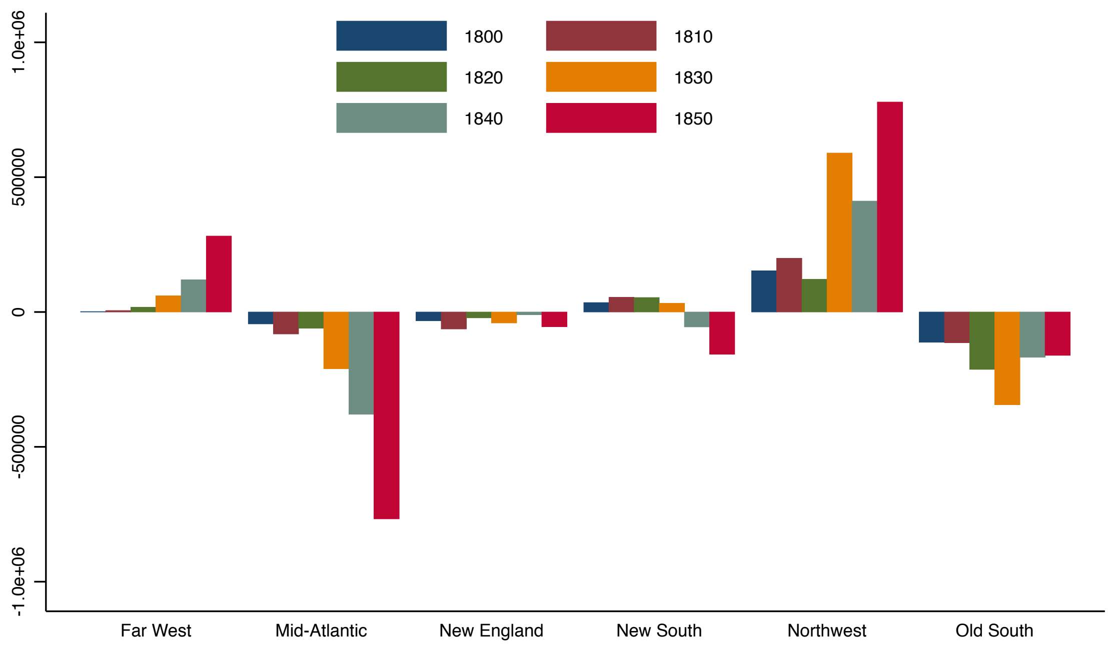
In 1843, during the "Great Emigration", one thousand pioneers took the Oregon Trail to cross the Rockies to the Oregon Territory, which belonged to Great Britain, and thousands more moved into the Mexican territories of California, New Mexico and Texas. In 1837, American settlers in Texas obtained independence from Mexico, and joined the union as a slave state in February 1846. In June of the same year, Great Britain allowed Oregon to join as a free state. During the 1830s and the 1840s, pioneers streamed westward towards Michigan, Arkansas, Wisconsin, and Iowa. Even more adventurous families took the Oregon Trail to reach the Pacific coast.[7] In 1849, while California got populated by gold searchers,

[7] Sequeira et al. (2020) show that a county's connection to the railway network affected the number of im-

the Mormons settled down in Utah. Before the beginning of the Civil War, migrants filled the Mississippi River valley, Texas, the southwest territories, and the new states of Kansas and Nebraska. During the war, gold and silver mines called later settlers into Oregon, Colorado, Nevada, Idaho, and Montana. But by the late 1880s, with the decline of the range cattle industry, families took advantage of the unsettled Great Plains to build their own farms, and signed the end of the westward movement. By the early 1890s a frontier waas no longer in place within the 48 continental states.

In order to provide evidence of the trend in migration from 1800, in Figure 4 we plot estimates of net interregional migration from 1800 to 1850, by decade, a rare and precious measure collected by McClelland and Zechauser (1982).

Figure 4 – Net Interregional Migration of White Men 1800-1860



Source: McClelland and Zechauser (1982).

The largest net loser of population throughout this period was the Mid-Atlantic region (New York, New Jersey, and Pennsylvania), east of the Appalachians. McClelland and Zechauser (1982) reports that the total exodus from this region for the entire period exceeded the total net loss of the Old South by more than 75 percent. This region was the the most important supplier of population to the west. The out-migration accelerated between 1800 and 1820, and then doubled from 1830 to 1860. The total loss amounted to approximately two thirds of the total net in-migration into the Northwest region in the same sixty years. The region that gained the most population was the Northwest, west of the Appalachians and north of Tennessee. The net influx between 1800 and 1860 was almost four

migrants that settled in the county.

million. New England and Old South were net losers of population but total net loss for the latter was almost four times that of the former. From 1800 to 1840, the region experienced a net influx of population. But population migrated from the South after 1840, as the nation became progressively divided.

For our empirical analysis, we make use of Census data on internal migrants from the IPUMS linked representative sample 1850-1880.[8] The sample is created by linking all men of 15 years and older who were in the U.S. both in 1850 and in 1880 census years.[9] This measure of migration is the best measure at the county level available nowadays despite the fact that it fails to indicate the timing of an individual's move between birth and the census and fails to count moves in between censuses. These data are different to those in Figure 4: the former accounts for the outflow of migrants; the latter shows the difference between outflow and inflow. Summary statistics are in Table A.3.1, and the characterization of the migrants is in Table 1.

The sample includes white men born in the US of age 10 and older. Geographic controls and region fixed effects are included in every regression. Columns (1)-(2), (3)-(4), and (5)-(6) examine the effect of individual characteristics on the probability of moving counties, states, and frontier, respectively. We see that age and is negatively associated with migrating, as is being married. Moreover, this correlation is higher for state and frontier migrants. The second part of the table shows that, compared to New England, the other regions predict a significantly lower probability of migrating. Interestingly, the occupational score (proxy of income), the literacy level, and living in a farm are not predictors of the fraction of internal movers. Hence these characteristics are not significant determinant of the migration choice, and do not mark a difference between movers and stayers.

These results indicate several important things. The long-distance movers that are the focus of this study consist mainly of unmarried men; this may be due to the lack of family constraints these men would have faced. The finding that occupational score and farming do not predict aggregate migration flows is also interesting. Adding destination county characteristics do not change these predictions. The fact that aggregate migration is not associated with economic characteristics of the home and destination counties supports the hypothesis that a heterogeneous migrant population made heterogeneous locational choices.

[8] The data are available from the website https://usa.ipums.org/usa/linked_data_samples.shtml.

[9] The samples of all women present in both censuses years are also available, but they do not include women who got married and changed their last name in between censuses. In Appendix D, we show that results hold true when considering both men and women.

Table 1 – Characterizing Migration Patterns

Dependent Variable:	Moved counties			Moved states			Moved frontier		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Age	0.008***	0.008***	0.007***	0.006***	0.006***	0.005***	0.006***	0.006***	0.005***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Occupational Score	0.002	0.001	0.002	0.000	0.000	0.001	0.000	-0.000	0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Illiterate	0.028	-0.031	-0.041	0.042	-0.002	0.011	0.038	-0.007	0.009
	(0.043)	(0.045)	(0.051)	(0.041)	(0.044)	(0.048)	(0.041)	(0.044)	(0.048)
Farmer	-0.019	-0.035	-0.020	-0.038	-0.035	-0.046	-0.028	-0.041	-0.033
	(0.034)	(0.035)	(0.035)	(0.030)	(0.031)	(0.029)	(0.029)	(0.030)	(0.028)
Married	-0.083**	-0.084**	-0.068**	-0.100***	-0.100***	-0.079**	-0.098***	-0.099***	-0.081**
	(0.027)	(0.028)	(0.028)	(0.024)	(0.025)	(0.023)	(0.024)	(0.025)	(0.023)
Home region fixed effects (New England excluded):									
Middle Atlantic	-0.066***	-0.018	-0.018	-0.070***	-0.030	-0.030	-0.069***	-0.031	-0.029
	(0.017)	(0.019)	(0.019)	(0.018)	(0.020)	(0.020)	(0.018)	(0.020)	(0.019)
East North Central	-0.268***	-0.229***	-0.195***	-0.256***	-0.221***	-0.190***	-0.268***	-0.231***	-0.194***
	(0.016)	(0.020)	(0.019)	(0.016)	(0.019)	(0.018)	(0.016)	(0.018)	(0.018)
West North Central	0.435***	0.378***	0.378***	0.024	-0.018	-0.018	0.031	-0.014	-0.014
	(0.034)	(0.044)	(0.044)	(0.016)	(0.086)	(0.086)	(0.096)	(0.087)	(0.087)
South Atlantic	-0.279***	-0.229***	-0.204***	-0.261***	-0.219***	-0.168***	-0.313***	-0.267***	-0.209***
	(0.018)	(0.027)	(0.025)	(0.018)	(0.024)	(0.024)	(0.016)	(0.022)	(0.021)
East South Central	0.046	0.044	0.039	0.060	0.046	0.041	0.052	0.042	0.035
	(0.038)	(0.041)	(0.038)	(0.041)	(0.043)	(0.039)	(0.040)	(0.044)	(0.039)
West South Central	-0.224***	-0.142***	-0.124***	-0.245***	-0.182***	-0.155***	-0.253***	-0.188***	-0.161***
	(0.019)	(0.024)	(0.023)	(0.017)	(0.022)	(0.022)	(0.017)	(0.022)	(0.021)
Mountain	0.023	0.036	0.006	0.051	0.066	0.030	0.058	0.068	0.028
	(0.045)	(0.044)	(0.041)	(0.050)	(0.050)	(0.048)	(0.050)	(0.050)	(0.048)
Pacific	0.058	0.070	0.009	0.013	0.023	-0.019	0.015	0.026	-0.019
	(0.054)	(0.051)	(0.057)	(0.046)	(0.044)	(0.050)	(0.046)	(0.044)	(0.050)
1850 Home county controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
1850 Destination county controls	No	No	Yes	No	No	Yes	No	No	Yes
Number of clusters (home county)	147	147	143	147	147	143	147	147	143
Observations	1,113	1,109	1,047	1,113	1,109	1,047	1,113	1,109	1,047
R-squared	0.221	0.299	0.306	0.280	0.314	0.340	0.306	0.340	0.374

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. OLS regressions. Each column presents the estimates from a separate regression. The unit of observation is a county. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. *1850 county controls*: percentage of US born pop., age, percentage of illiterate pop., percentage of white pop., sex-ratio, railroad access, navigable river access, percentage pop. living in urban areas, population density, occupational score, land availability, farm value per acre, percentage of workers in manufacturing.

3. EMPIRICAL STRATEGY

In this section we describe the empirical strategy and the equations to be estimated in order to measure the correlation between migration and fertility decline. Before going into the details, we study how fertility is related to the characteristics of our county level data in comparison with the state level data used in the previous literature. Table 2 shows the correlations between the CWR in 1850 and several variables that have been at the center of previous theories on fertility transition.

Table 2 – OLS estimates of historical literature variables

Dependent Variable:	*Log of 1850 CWR*									*Decline of CWR from 1850 to 1880*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Yasuba (1962)		Easterlin (1976)				Sundstrom and David (1988)									
1850 Population density	-0.000**	-0.000**		-0.000		0.000**			0.000***	-0.000***			-0.000***			
	(0.000)	(0.000)		(0.000)		(0.000)			(0.000)	(0.000)			(0.000)			
1850 Land availability			0.487***	0.487***	0.353***	0.355***			0.345***		0.006		-0.010			
			(0.039)	(0.039)	(0.034)	(0.034)			(0.010)		(0.026)		(0.025)			
1850 Log farm value			-0.094***	-0.094***	-0.089***	-0.093***			-0.077***		-0.001		0.022**			
			(0.011)	(0.011)	(0.010)	(0.010)			(0.010)		(0.009)		(0.009)			
% 1850 men in manufacturing							-3.741***	-1.617***	-1.017***			-0.807***	-0.946***			
							(0.262)	(0.239)	(0.195)			(0.170)	(0.171)			
% change men in manufacturing 1850-1880														0.018***		0.018***
														(0.004)		(0.004)
% change land availability 1850-1880															-0.035**	-0.037**
															(0.011)	(0.013)
% farm value 1850-1880															0.000	-0.009
															(0.011)	(0.012)
State fixed effects	No	Yes	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,604	1,604	1,173	1,173	1,173	1,173	1,177	1,177	1,173	1,566	1,173	1,177	1,173	1,085	1,169	1,081
R-squared	0.015	0.562	0.379	0.379	0.739	0.739	0.307	0.658	0.750	0.421	0.530	0.506	0.551	0.549	0.538	0.559

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. OLS regressions. Each column presents the estimates from a separate regression. The unit of observation is a county. Robust standard errors in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables.

In columns (1) and (2), Yasuba (1962)'s theory about population density is tested, and the coefficient is always significantly equal to zero, which rule out population density as the main determinant of county CWR decline. In columns (3) to (6), we measure the correlation between some economic characteristics (land availability and farm value) of the county and the 1850 CWR. Results are consistent with Easterlin (1976)'s model of fertility decline, as counties where available acres are high are also counties where the CWR is high. On the contrary, farm value seems to disincentive fertility. Columns (7) to (9) show the correlation between the industrial structure of the counties and the CWR. Once again, results are in line with Sundstrom and David (1988)'s model, and the presence of manufacturing firms does not promote fertility. When the dependent variable is the change of CWR from 1850 to 1880 (columns (10)-(16)), available acres and farm value do not affect the CWR decline. The 1850 percentage in manufacturing has a negative correlation with the decline of the CWR, but change in manufacturing from 1850 to 1880, has the expected effect on the CWR decline. We can conclude that level of industrial development in 1850 does not necessarily predict the decline in the CWR, but only the growth rate does. In terms of land availability and firm value, the former does not affect the transition while the sign of the coefficient of the latter confirms the Easterlin's hypothesis. We will show that migrants increase the explanatory power of our model, and does not impact the above discussed correlations.

3.1. Estimating Equations

Our baseline empirical strategy exploits the fact that both the CWR and the percentage of migrants fluctuated across counties, and test the predictions of Azariadis and Drazen

(1993)'s model.[10]

We begin our analysis by showing that migrating is not sufficient to impact fertility decisions, but it is necessary to move far enough to end participation in the original family's production activity. Hence, distance and origin of migration play a role. We start our analysis by estimating the following model by OLS:

$$\log\left(\frac{CWR_{ij}^{1850}}{CWR_{ij}^{1880}}\right) = \alpha_j + \gamma Migr_{ij} + \mathbf{X}_{ij}\Gamma + \epsilon_{ij}, \tag{3.1}$$

where the dependent variable is the percentage change of CWR from 1850 to 1880 in the home county i, state j, in similar fashion of Goldstein and Klüsener (2014). The independent variable of interest is $Migr$, i.e. the fraction of migrants moving from the "home" county i in state j in 1850 to the "destination" county in 1880. The structure of our historical data enables us to introduce state fixed effects (α_j) in order to account for common time-shocks state characteristics. Since the dependent variable is computed as a rate of change, each county appears in the dataset only once, hence we cannot control for county and year fixed effects. ϵ_{ij} is the error term. γ should be interpreted as the percentage change in the geometric mean of CWR change. It should be that $\gamma > 0$ because being a migrants has an increasing effect on the fertility decline of the home county.

To assess the extent of migration distance, we estimate the following model by OLS:

$$\log\left(\frac{CWR_{ij}^{1850}}{CWR_{ij}^{1880}}\right) = \alpha_j + \gamma_1 MigrCounty_{ij} + \gamma_2 MigrState_{ij} + \mathbf{X}_{ij}\Gamma + \epsilon_{ij}. \tag{3.2}$$

In this regression, the percentage of migrants is divided into migrants that moved to another county in the same state ($MigrCounty$), and the fraction of migrants who moved to another state ($MigrState$) departing from county i in state j. Our hypothesis is that only migrants who moved far from their home state have contributed to the decline of CWR. Hence, we expect $\gamma_2 > 0$ and γ_1 to be non significant or $\gamma_1 < \gamma_2$.

We also consider a third measure, i.e. the *frontier*, which includes all destination states different than Maine, New Hampshire, Vermont, New York, Massachusetts, Rhode Island, Connecticut, New Jersey, Delaware, Maryland, West Virginia, Pennsylvania, Virginia, North Carolina, and South Carolina. The empirical model becomes

$$\log\left(\frac{CWR_{ij}^{1850}}{CWR_{ij}^{1880}}\right) = \alpha_j + \beta_1 MigrNoFront_{ij} + \beta_2 MigrFront_{ij} + \mathbf{X}_{ij}\Gamma + \epsilon_{ij}, \tag{3.3}$$

[10] They propose an overlapping generation model where children bargain with their father over the farm's output and have the outside option to leave and earn a salary that they would not remit to their original family. The authors show that there exists a unique steady state equilibrium in which higher wages would induce sons to migrate: the value of their outside options increases as well as their bargaining power. Therefore, as income share of fathers is reduced for any number of offspring, families would find it optimal to bear fewer children today in expectation of higher future wages.

where $MigrNoFront$ is the fraction of non-frontier migrants; and, $MigrFront$ is the fraction of *frontier* migrants. Once again, we expect $\beta_2 > 0$, but β_1 could be positive or not significant, as we combined migrants of any distance.

X is a matrix of migrant characteristics, and home and destination county-level variables in 1850 and 1880 that may have influenced individuals' decisions to have children. The former group of variables includes: age, age square, occupational score, literacy, farm status, and marital status. The latter group of variables includes: percentage of US born population, age, percentage of illiterate population, percentage of white population, sex-ratio, railroad access, navigable river access, percentage of population living in urban areas, population density, occupational score, land availability, farm value per acre, percentage of workers in manufacturing, percentage of churches per population. These variables are described in Appendix A. Results are in Table 3.

Table 3 – OLS estimates of the effects of migration on the decline of the CWR

Dependent Variable: ***Decline of CWR from 1850 to 1880***	(1)	(2)	(3)	(4)	(5)	(6)
Migrants	0.041**	0.044**				
	(0.020)	(0.020)				
County migrants			0.017	0.033		
			(0.029)	(0.026)		
State migrants			0.048**	0.047**		
			(0.023)	(0.023)		
Other than frontier migrants					0.012	0.035
					(0.024)	(0.026)
Frontier migrants					0.050**	0.047*
					(0.024)	(0.024)
Migrant Controls	No	Yes	No	Yes	No	Yes
Home and destination county controls 1850	No	Yes	No	Yes	No	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of clusters (home county)	147	143	147	143	147	143
Observations	1,176	1,087	1,176	1,087	1,176	1,087
R-squared	0.496	0.631	0.497	0.631	0.497	0.631

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. OLS regressions. Each column presents the estimates from a separate regression. The unit of observation is a county. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables.

We estimate equations (3.1)-(3.3) using ordinary least squares, and results are in columns (1) to (6) when the dependent variable is the log change in CWR from 1800 to 1850. Note that for ease of interpretation, the decline of CWR is expressed as $-\log\Delta CWR$. In general, coefficients are stable across specifications. In terms of percentage change, migrants contribute from about 4 to 5 percent to the geometric mean of the CWR decline from 1850 to 1880. Controlling for all the aforementioned background variables (column 2) does not alter the size of the coefficients.

In columns (3)-(4), we report the results of the estimation of equation (3.2), where we distinguish between county and state movers. We see that there is no effect if migrants changed county but do not cross the state border. On the contrary, state movers increase the mean of the CWR decline up to 4.7 percent (column 4). Hence, in our estimations, distance plays a role that has not been considered by the previous literature, as Sundstrom and David (1988). Columns (5)-(6) show the results of the estimation of equation (3.3). Westward migration increases the average decline of CWR by about 5 percent from 1850 to 1880, while migration in a different county or state not belonging to the frontier is not statistically significant. This result underlines that the direction of the move (from the East to the West) is as important as the distance alone, and the decrease of CWR is explained by this category of migrants only.

Full regression results are available in Tables B.1 and B.2, Appendix B. Table B.1 shows a summary of the results, focusing on the coefficients of migrants. Tables B.2 shows all of the specifications. In Table B.2, we can observe that the coefficients of the covariates (when significant) are robust across specifications. Being born in the US increases the likelihood of living in a county where the CWR decreases over time. It must be recalled that US fertility transition started earlier than in other countries. Hence, it could have taken longer to international migrants to change their fertility norms. As predicted by the fertility literature, we observe that literacy is negatively correlated to CWR. Similarly, a high male to female ratio in the county of residence has a negative effect on the decline of CWR. Urban density as well as labor force participation and income (occupational score) are positively correlated with the decline of the CWR. In column (3), as well as in columns (6) and (9), the percentage of workers in manufacturing in the home county is negatively correlated with the decline of CWR, while the same percentage in the destination county is positively correlated with it. This result may suggest that income effect prevails when family members remain in the proximity, but the presence of manufacturing firms far from the home county, contributed to the decrease of the CWR in the home county, as transfers are not feasible. The percentage of unimproved acres contributed to the decrease of the CWR, regardless of the location.

In Table 4, we show the correlation between the level of the CWR in 1880 and the percentage of migrants between 1850-1880. The coefficients on migrants are less stable than in the previous regressions, and decrease when including all control variables, but remain statistically significant across specifications. State and frontier migrants account for about 3 percent of the level of CWR in 1880. In columns (2), (5) and (8), the level of the CWR in 1850 accounts for about 30 percent of the CWR level in 1880. Hence, the initial level of CWR is an important determinant of the contemporaneous level of CWR. Yet, migration explains a significant percentage of it. Another interesting observation is that, introducing migrant and county characteristics makes county migrants and non-frontier migrants powerless in explaining the level of CWR in 1880, while increasing the fitting of the model.

Table 4 – OLS estimates of the effects of migration on the 1880 CWR

Dependent Variable: ***Log of 1880 CWR***	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Migrants	-0.063**	-0.050**	-0.020**						
	(0.020)	(0.017)	(0.015)						
County migrants				-0.083**	-0.043*	-0.016			
				(0.031)	(0.026)	(0.021)			
State migrants				-0.058**	-0.052**	-0.021*			
				(0.022)	(0.019)	(0.017)			
Other than frontier migrants							-0.089**	-0.042*	-0.020
							(0.030)	(0.025)	(0.020)
Frontier migrants							-0.056**	-0.052**	-0.020*
							(0.022)	(0.020)	(0.018)
Log of 1850 CWR		0.600***	0.331***		0.600***	0.331***		0.600***	0.331***
		(0.049)	(0.034)		(0.050)	(0.034)		(0.050)	(0.034)
Migrant Controls	No	No	Yes	No	No	Yes	No	No	Yes
Home and destination county controls 1850	No	No	Yes	No	No	Yes	No	No	Yes
Home and destination county controls 1880	No	No	Yes	No	No	Yes	No	No	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of clusters (home county)	147	147	143	147	147	143	147	147	143
Observations	1,176	1,176	1,086	1,176	1,176	1,086	1,176	1,176	1,086
R-squared	0.633	0.767	0.893	0.633	0.767	0.893	0.633	0.767	0.893

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. OLS regressions. Each column presents the estimates from a separate regression. The unit of observation is a county. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables.

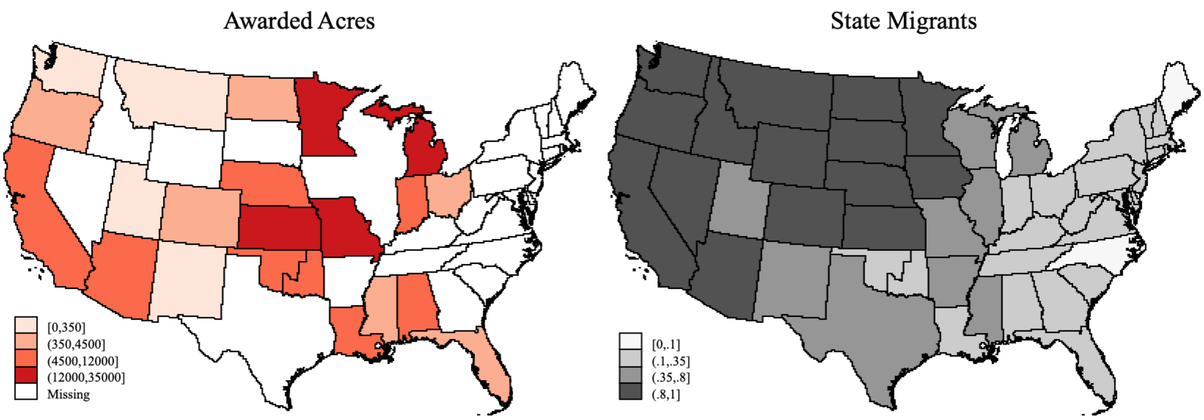
3.2. Instrumental Variable Estimates

To corroborate our theory and test the prediction of our explanation, we propose an instrumental variable analysis. First, as unobserved characteristics of the counties other than the flow of westward migrants may have contributed to the decrease in CWR, the IV analysis allows us to estimate the coefficient of interest consistently, and free of bias caused by the omitted variable issue. In particular, the positive correlation should not be driven by, say, a decrease in the sex-ratio (decreasing the probability of marriage, and hence the number of children). If this were the main driver, then there would be no dynamic implication to a shock that caused couples to have less children than they did previously. Moreover, to attach a causal interpretation to the relationship between the Westward migrant share and the fertility rate, the location of migrants should be orthogonal to factors that might have independently contributed to shape the fertility preferences of the residents. One specific concern is that migrants might come from counties that were booming at the time of their departure, and that such thriving economic conditions persisted over time, directly influencing Americans' preferences for smaller families. Similarly, migrants might have been coming from areas with more liberal views towards family changes, which may in turn be correlated with fertility preferences later in time.

To overcome these and similar concerns, in addition to controlling for historical demographic and economic county characteristics and for state fixed effects, we construct a measure of freely granted acres that would have provided incentives to the Eastern population to migrate to the West. Between 1847 and 1855 the Congress passed four land warrant acts which granted 60 million acres of land to veterans and their heirs. Approximately one in nine U.S. families received a land warrant for earlier military service. We use the ICPSR Military Bounty Land Warrants in the United States (1847-1900) dataset (Oberly, 1991) to compute the average number of acres awarded by the state of residence of veterans when applying for the land warrant. As Lindert (1986) describes, the ICPSR sample represents a randomly drawn half percent of those who received bounty land warrants for earlier wartime service. This includes widows and minor heirs, but over three-quarters of the warrant recipients were the original veterans. The 320,000 veterans of the War of 1812 who received warrants qualified because they had served a wartime minimum of 14 days. The sample represents about 30 percent of the 1810 white male population between the ages of 16 and 45 and appears to be broadly representative of U.S. white males at the time of the War of 1812.

In the left panel of Figure 5 we show the location of the acres awarded, while in the right panel we show the states chosen by migrants as their destination.

Figure 5 – Military Bounty Land Warrants and Migrants by Destination States



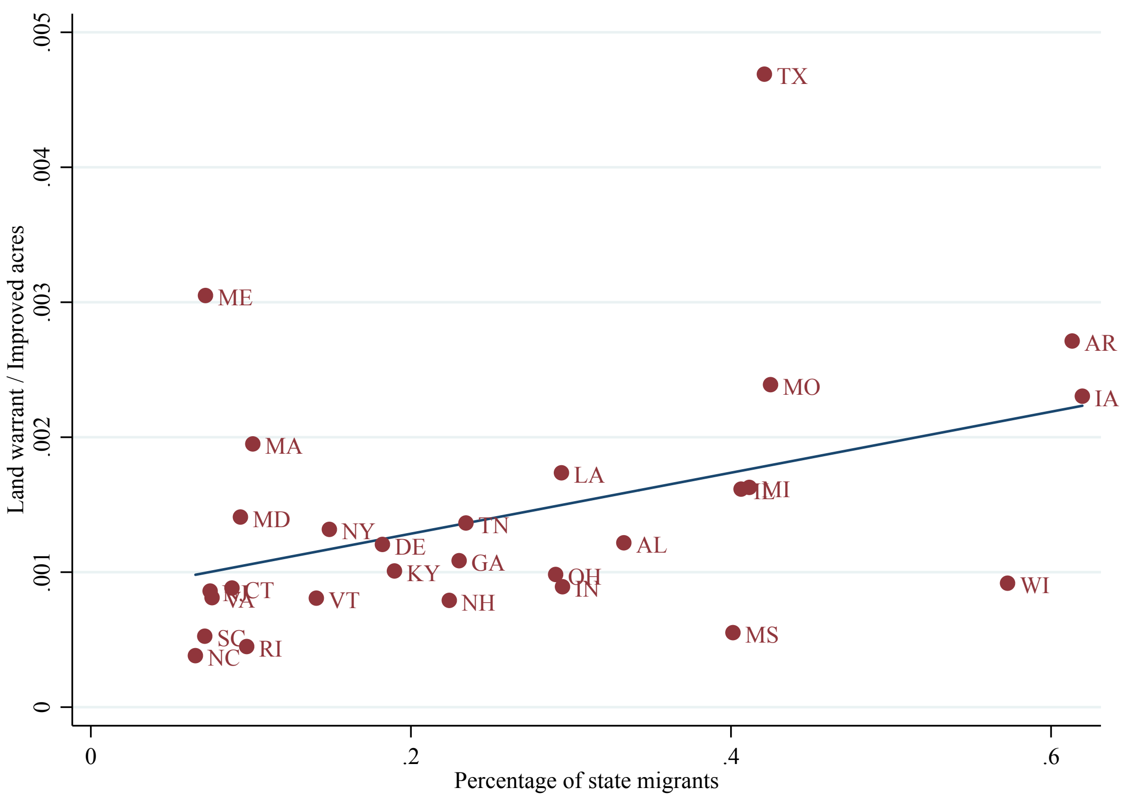
Source: Left panel: Authors' computations using data from Oberly (1991); Right panel: Authors' computations using data from Ruggles et al. (2015)

Our instrumental variable is defined as follows:

$$Instrument_i = \frac{Land\ Warrant_i}{Improved\ Land_i} \tag{3.4}$$

which corresponds to the number of acres obtained through the warrant per total of improved acres in the state of residence i of the recipients. We believe that the number of awarded acres as a percentage of local available improved acres is a better indicator of the incentives to migrate from the home state. The absolute level of awarded acres is only informative in relation to the improved acres available at the home county. The (positive and significant) correlation between the IV and the percentage of migrants who moved to a different state is scattered in Figure 6.

Figure 6 – Migrants and Granted Acres



Source: Granted acres are computed using data from Oberly (1991); Migrants are computed using data from Ruggles et al. (2015).

Table 5 shows the results of the first-stage, IV-2SLS estimates, and the reduced-form estimates.

Table 5 – First-stage and IV-2SLS estimates of the effects of migration on the decline of CWR

	First-stage estimates			**IV-2SLS estimates**			**Reduced-Form**
Dependent Variable:	Migrants	State migrants	Frontier migrants	Decline of CWR from 1850 to 1880			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(Instrument)	0.100***	0.078***	0.140***				0.048***
	(0.016)	(0.014)	(0.013)				(0.008)
Migrants				0.479***			
				(0.092)			
State migrants					0.617***		
					(0.117)		
Frontier migrants						0.585***	
						(0.108)	
Migrant Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Home and destination county controls 1850	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic	39.565	30.206	32.452				
Observations	1,067	1,067	1,067	1,067	1,067	1,067	1,067

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. Each column presents the estimates from a separate regression. The unit of observation is a state. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables.

The first-stage relationship between the instrument and the share of migrants is strongly positive: granted acres are significantly related to migrants at over 95 percent confidence (columns (1) to (3)), and this relationship is robust to the inclusion of migrants, and county controls.[11] Granted acres provided incentives to migrate and exploited unimproved lands. Higher share of granted acres are associated with significantly higher decrease of CWR in the reduced-form regression (column 7), with a point estimate of 0.043 which suggests that a 1 percentage point increase in the number of acres received by warrant is associated with a 0.0004 percentage point decrease in the CWR. This is the first indication that higher share of granted acres makes fertility to decrease more in the residing counties of the benefiaciaries. We then proceed to the second-stage estimate to measure the impact of migration on the CWR. We find that the instrument is strongly correlated with the migrant share, with a Kleibergen-Paap F-statistic of approximately 30. Taken literally, the magnitude of the point estimate suggests that a 1 percentage point increase in the number of acres received by warrant is associated with a 0.001 percentage point increase in the average migrant share. According to the IV-2SLS estimates, states with a greater share of migrants from 1850 to 1880 have a significantly higher decrease of CWR.

Comparing the estimates in Table 3 and 5, it is clear that the OLS correlation in Table 3 between migrant share and the decline of the CWR is smaller than the 2SLS estimate. One explanation for this differences is that the 2SLS estimates are causal while the OLS estimates are not, with the difference between the two arising due to the negative selection by mi-

[11] Note that, as one observation is a state, fixed effects cannot be included.

grants, which results in OLS estimates that understate the effect on fertility. Negative selection would occur if migrants tended to move from counties that counterfactually would have experienced lower CWR decline. This is consistent with the historical evidence (Ransom and Sutch, 1986) that migrants were moving from farms, and high fertility was sustained by the rural population. Counterfactually, these counties would not have been likely candidates for a decrease in fertility a high fertility of the rural population and the self- sufficiency of agriculture were mutually reinforcing. Self-sufficiency meant an absence of well developed markets that, in turn, required a reliance upon family-based mechanisms of reciprocity to provide farm labor and old-age security. Large families supplied the needed labor during the seasonal peaks of agricultural work but they also supplied a surplus of labor in the off-season.

We next perform additional robustness checks. First, in Table 6, we conduct the IV analysis on Western and Eastern states, separately. Our assumption is that migration has affected fertility norms in Eastern counties and states. Hence, we verify that migration had a significant impact on families residing in the Eastern part of the country only.

Table 6 – IV-2SLS estimates of the effects of migration on the decline of CWR: Western and Eastern States

Dependent Variable: **Decline of CWR from 1850 to 1880**	Western states			Eastern states		
	(1)	(2)	(3)	(4)	(5)	(6)
Migrants	-0.463			0.505***		
	(0.516)			(0.085)		
State migrants		-2.708			0.620***	
		(11.580)			(0.101)	
Frontier migrants			-4.030			0.596***
			(25.433)			(0.096)
Migrant Controls	Yes	Yes	Yes	Yes	Yes	Yes
Home and destination county controls 1850	Yes	Yes	Yes	Yes	Yes	Yes
Observations	255	255	255	812	812	812

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. Each column presents the estimates from a separate regression. The unit of observation is a state. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables. Western states include: Tennessee, Mississippi, Kentucky, Alabama, Texas, Oklahoma, Arkansas, Louisiana, Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming, California, Oregon, Washington.

We show that IV estimates provide coefficients which are not significant for Western states. This suggests that our empirical model captures the effect of Westward migration from the settlements and not other confounding factors.

Second, we compute the CWR including children from 1 to 4 and 5 to 9 years of age. Data on children mortality in the 1800s are not available, but we know from CDC (1999) that in 1900, 30 percent of all deaths in the United States occurred in children less than 5 years of age compared to just 1.4 percent in 1999. The CWR is computed using Census complete count data from Ruggles et al. (2015). Table 7, column (1)-(6) shows results for the change of CWR between 1850-1880. We do an additional check, and compute the number of children ever born (CEB hereafter) using several 1% public samples of the U.S. Census data (Ruggles et al., 2015). We follow the same methodology employed by Jones and Tertilt (2006) of fertility by cohorts of women. We define a cohort to be five years of birth years. The decline of CEB is given by the difference (of the logarithm) between CEaB of the 1893 and 1818 cohort. The drawback of this analysis is that the CEB is computed at the state level, and not at the county level, as the number of observations per county available for the oldest cohorts in 1818 is extremely small. Results are in columns (7)-(9) of Table 7. Results are similarly large and statistically significant as in Table 5.

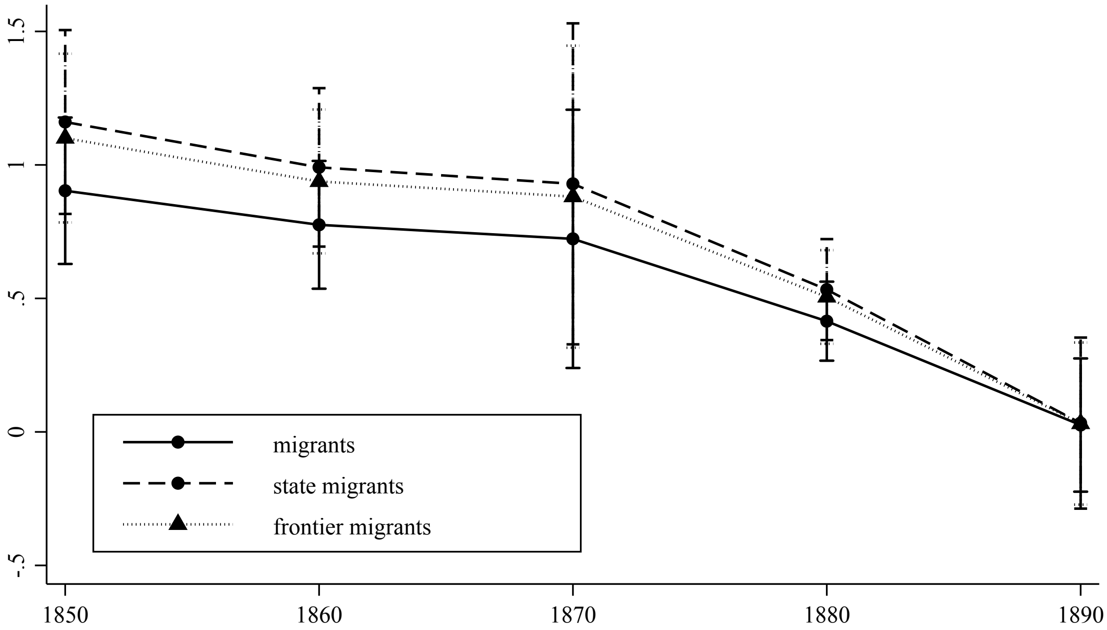
Table 7 – IV-2SLS estimates of the effects of migration on the decline of (alternative measures of) CWR

Dependent Variable:	Decline of CWR 1-4			Decline of CWR 5-9			Decline of Children Ever Born		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Migrants	0.552***			0.425***			0.594***		
	(0.100)			(0.084)			(0.099)		
State migrants		0.713***			0.548***			0.737***	
		(0.126)			(0.103)			(0.133)	
Frontier migrants			0.675***			0.519***			0.698***
			(0.116)			(0.096)			(0.123)
Migrant Controls	Yes	Yes	Yes	Yes	Yes	Yes			
Home and destination county controls 1850	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen Paap *F*-statistic	39.565	30.206	25.562	31.432	30.206	32.452	58.561	29.576	48.898
Observations	1,067	1,067	1,067	1,067	1,067	1,067	1,056	1,056	1,056

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. Each column presents the estimates from a separate regression. The unit of observation is a state. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables.

Finally, in Figure 7 we plot the coefficients of the IV-2SLS estimates where the dependent variable is the CWR decline from 1850, 1860, 1870, 1880, 1890 to 1900, and the independent variable is the fraction of migrants from 1850 to 1880. We can see that the trend is decreasing but statistically significant until 1880.

Figure 7 – IV-2SLS estimates of the effects of 1850-1880 migration on the decline of CWR over time



Sources: Each coefficient is the result of an IV-2SLS estimate where the dependent variable is the log of the decline of the CWR from each year and 1900. The independent variables include migrant, home and destination county controls.

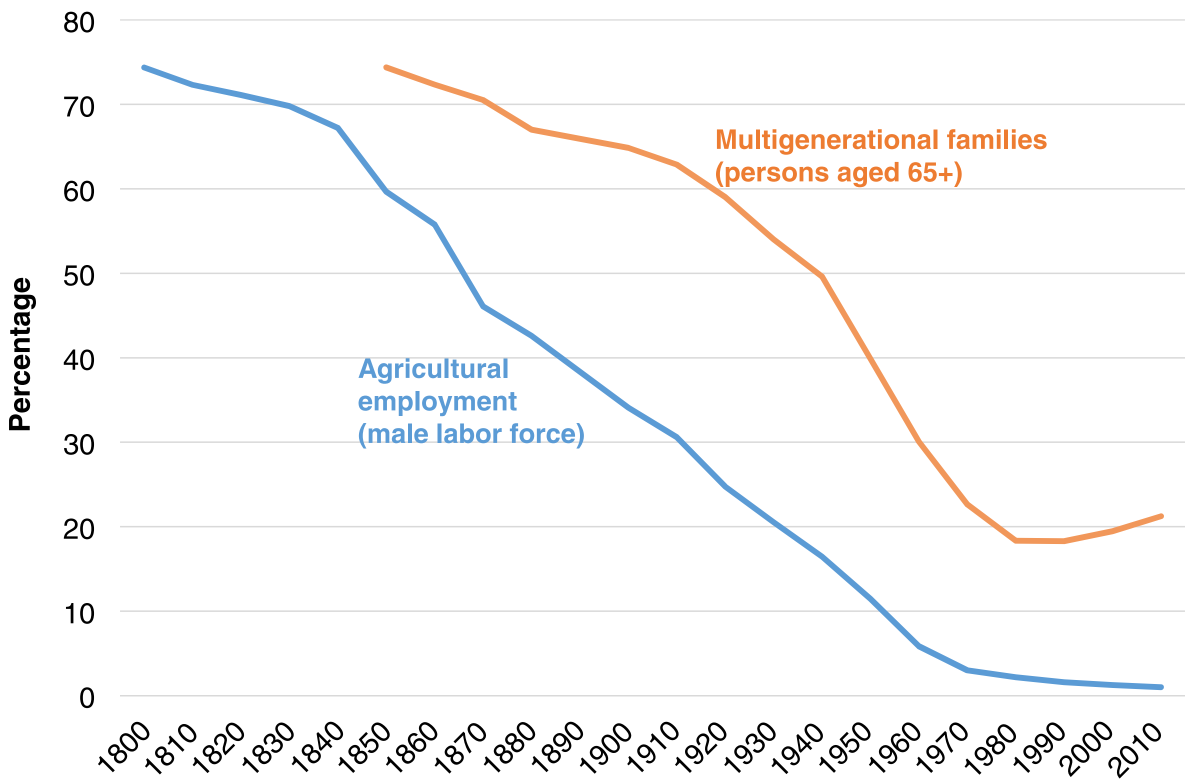
4. DISCUSSION

4.1. Changes in Family Composition

The decline in fertility is one of the demographic changes that American families experienced in the Nineteenth century. The size of the families got smaller, and patriarcal families were being replaced by simpler structures. Moreover, Ruggles (2015) documents that coresidence of persons older than 65 declined from 74% in 1850 to about 60% at the end of the century. For most of the nineteenth century, production was carried out by families. In 1800, three-quarters of the workforce was engaged in agricultural work, and a majority of the population lived in farms until 1850 (Ruggles, 2015; Weiss, 1992). Farms could not operate without family labor; all family members who were old enough contributed to farm production. Among the one-quarter of the population who did not work on farms at the beginning of the nineteenth century, most still made their living through the family economy. As the century progressed, new high-paying opportunities arose in factories. The number of factory jobs grew 600% between 1850 and 1900, and there were rapidly expanding opportunities in clerical, sales, and professional occupations (Lebergott, 1984). The growth of well-paying wage labor jobs for men undermined the economic underpinnings of patriarchal authority. As young men took jobs off the farm, they moved away from home and out of the control of the patriarch. Figure 8 compares the percentage of men in agriculture with the percentage of elderly residing in multigenerational families.

A generation after agricultural employment began to decline, multigenerational coresidence followed suit. As discussed by Ferrie (1997) and Ransom and Sutch (1986), the possibility of migration to cheaper western lands may be behind the overturn of long-standing family support patterns, as children looked for new fortunes far from their parents who were forced to invest in alternative resources to provide for their old-age support.

Figure 8 – Agricultural employment and multigenerational families: United States, 1800–2010



Sources: Ruggles et al. (2015), Weiss (1992), and Ruggles (2015) (p.1802).

In Europe, the process had similar characteristics: Ransom and Sutch (1986) documents that in rural England, the out-migration in the mid and late nineteenth century consisted almost entirely of young men and women who almost never returned and rarely sent remittances of money to family members left behind.

4.2. Saving Rate

In order to induce children to remain in the proximity of parents and provide care for them and work the farm or family business (if the parents were landowning farmers or business proprietors), inducements in the form of real property via inheritance or inter-vivos transfers were necessary. But as areas became more developed, with more urban populations and non-agricultural activity as well as transport connections, it became easier to rely on life-cycle saving in other forms, for example savings in banks and annuities. In fact, U.S. savings rate rose from 16 to 22 percent of GNP between 1830 and 1900. Lewis (1983) examined the hypothesis that a falling dependency rate accounts for part of the rise in nineteenth-century savings rates. He found that between 1830 and 1900 about one-quarter of the 6-percentage-point rise in the savings rate can be attributed to a decline in the dependency rate.

The spread of banks and financial alternatives was hence relevant in this context, since financial institutions provided an alternative to saving in the form of real property or children. For example, in 1800 there were 28 state banks (and the First Bank of the United States), located almost entirely in larger cities. By 1860, there were 1,562 state banks, much more widely spread across the country (Carter et al., 2006). We could ask whether CWR decline boosted the opening of banks or banking accounts. If the emergency of savings for old age was relatively important in boosting banks and other financial institutions, then financial development should be viewed as an indicator of the extent of the transition.

We measure the correlation between financial development and 1850 CWR, and estimate the following equation:

$$MoneyVelocity_i = \beta_0 + \beta_1 \log CWR_i + \mathbf{X}_i' \beta_2 + \epsilon_i \tag{4.1}$$

where the dependent variable $MoneyVelocity_i$ is computed as the ratio of 1850 nominal manufacturing output value and the amount of 1850 money in circulation. Money velocity is an approximation of saving rate, and the two variables are inversely correlated.[12]

Table 8 – Money velocity and CWR

Dependent Variable:	Money Velocity
$\log(CWR)$	1.826***
	(0.462)
County controls	Yes
Observations	939
R-squared	0.417

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. OLS regressions. The unit of observation is a state. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term.

The 1850 velocity of money is 2.7 on average which is consistent with the aggregate value of the velocity of money.[13] Results in Table 8 show that if the CWR increases (decreases)

[12] Data on state or county saving rate are not available. We chose not to use the relative number of banks or bank deposits as a percentage of the state population, to avoid the correlation with the CWR.
[13] See https://fred.stlouisfed.org/series/A14187USA163NNBR.

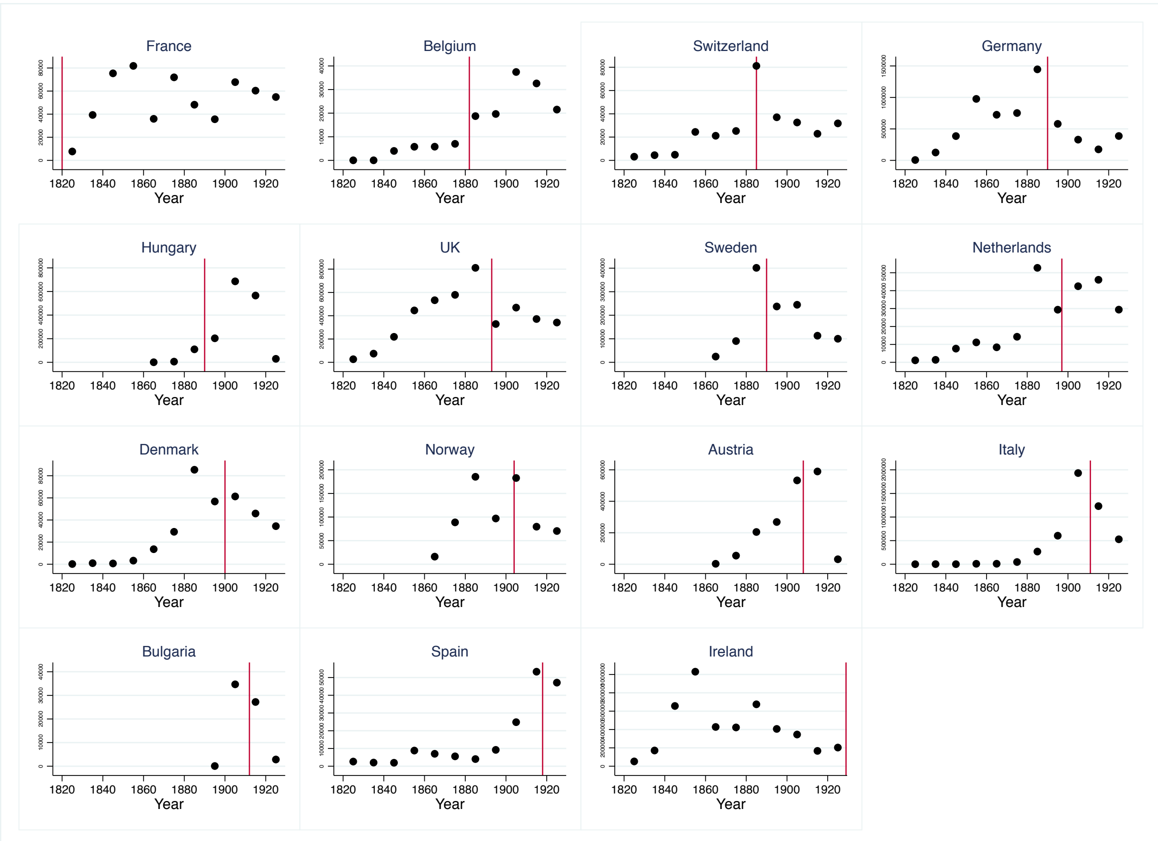
by one percent, money velocity increases (decreases) by about 0.014 units, and savings decrease (increase). We can also interpret the results by saying that counties with a lower CWR are counties with a lower money velocity and hence a higher saving rate, and viceversa.

5. RELATED EVIDENCE

In this section, we extend our analysis to the European fertility transition. A summary of the existing theories about European fertility decline can be found in Guinnane (2011), who also lists detailed data on birth crude rate and cohort fertility in the period 1800-1970 for five major countries: France, England and Wales, Germany, the United States, and Italy. He also shows that fertility declining started in the eighteenth or nineteenth century, and this decline accelerated in the second half of the nineteenth century.

We show the results of two exercices. In Figure 9 we use data from the U.S. Office of Immigration Statistics and show the trend of emigrants from several European countries to the U.S. from 1820 to 1920.

Figure 9 – Migration and Fertility Transition in Europe



Sources: Dots are immigrants from https://www.dhs.gov/xlibrary/assets/statistics/yearbook/2008/ois_yb_2008.pdf, Table 2 "Persons obtaining legal permanent resident status by region and selected country of last residence: fiscal years 1820 to 2008"; vertical red lines are placed in correspondence of the "Date of decline in marital fertility by 10 percent" from Knodel and van de Walle (1979), Table 1.

The vertical line is placed in correspondence of the date at which marital fertility declined by 10 percent, as reported in Knodel and van de Walle (1979). In general, we remark a strong correlation between the date of the highest emigration flow and the fertility decline. Few exceptions are present: France, where the decline of fertility started earlier than 1820 (as described by Daudin et al., 2019 and Murphy, 2015); in Ireland, on the contrary, fertility transition is estimated to happen later than 1920, while emigration peaked around 1850.

We also rely on data from Mitchell (2003) to compute the percentage of migrants from Europe by decade (Table A8, page 129), and the change in the child (age 5 to 9 years old)-woman (age 15 to 49 years old) ratio (Table A2, pages 12-44). Descriptive statistics are in Table C.1, Appendix C. In Table 9, we show the results of an OLS regression where the dependent variable is the decline of CWR from the earliest to the latest of the years available for each country, and the independent variable is the percentage of migrants (with respect to total population) from a specific country. The correlation is positive, meaning that the higher the fraction of migrants, the higher the decline of the CWR, and the impact is of about 1.65 percentage points.[14]

Table 9 – Decline of CWR and Migration in Europe

Dependent Variable: ***Decline of CWR***												
Percentage of Migrants from:	France	Belgium	Switzerland	Germany	UK	Sweden	Netherlands	Denmark	Norway	Austria	Italy	Spain
	0.443*	-0.363	0.058*	0.117*	0.022**	0.032	0.025	0.052*	0.023	0.060*	0.001	0.010
	(0.253)	(0.283)	(0.031)	(0.092)	(0.009)	(0.016)	(0.030)	(0.030)	(0.014)	(0.049)	(0.010)	(0.022)
Observations	18	9	11	8	5	9	12	11	10	11	8	11
R-squared	0.160	0.190	0.278	0.215	0.489	0.278	0.072	0.269	0.223	0.200	0.003	0.024

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. OLS regressions. The unit of observation is a pair country-year. Robust standard errors in parentheses. All regressions include a constant term.

6. CONCLUSION

We study the relationship between internal migration and the CWR decline in the nineteenth century in the U.S. We show that better economic opportunities, in the form of land availability, attracted individuals to the West, and contributed to changing fertility norms in the Eastern counties. In particular, with respect to previous studies, we underline that distance of migration matters as tranfers from far away children become limited.

We also show that the effect is persistent over time, even though it decreases and is not significant by 1880. Our results hold if we use different measures of fertility, such as the number of children ever born, or consider other age groups of children.

In order to support our mechanism, we refer to the historical literature, and discuss the change in family composition that occurred in the Nineteenth century. Moreover, we use data from output manufacturing to compute money velocity and link it to the saving rate.

[14] As the number of observations for each country is limited (from 8 to 17 at most), it is not possible to include country fixed effects in the estimation.

Results suggest that fertility decline may have provided incentives to the development of financial system, as social security was still inexistent in the nineteen century. Finally, we attempt to extend our analysis to Europe, and provide evidence of the relationship between European emigration towards the U.S. and the dates marking the beginning of fertility transition.

REFERENCES

Azariadis, C. and Drazen, A. (1993). Endogenous Fertility in Models of Growth. *Revista de Analisis Economico*, 8:3–18.

Carter, S. B., Gartner, S. S., Haines, M. R., Olmstead, A. L., Sutch, R., and Wright, G. (2006). *Historical Statistics of the United States: Earliest Times to the Present. Millennial Edition (Vol. 5).* New York: Cambridge University Press.

Carter, S. B., Ransom, R. L., and Sutch, R. (2001). Family Matters: The Life-Cycle Transition and the Unparalleled Antebellum American Fertility Decline. Technical report, University of California Project on the History of Saving.

Caucutt, E. M., Cooley, T. F., and Guner, N. (2013). The Farm, the City, and the Emergence of Social Security. *Journal of Economic Growth*, 18(1):1–32.

CDC (1999). Control of Infectious Diseases, 1900-1999. *Journal of the American Medical Association*, 282(11):1029–1032.

Daudin, G., Franck, R., and Rapoport, H. (2019). Can Internal Migration Foster The Convergence In Regional Fertility Rates? Evidence From 19th Century France. *Economic Journal*, 620:1618–1692.

Easterlin, R. A. (1976). Factors in the Decline of Farm Family Fertility in the United States: Some Preliminary Research Results. *Journal of American History*, 63:600–614.

Ferrie, J. P. (1997). Migration to the Frontier in Mid-Nineteenth Century America: A Re-Examination of Turner's "Safety Valve".

Goldstein, J. R. and Klüsener, S. (2014). Spatial Analysis of the Causes of Fertility Decline in Prussia. *Population and Development Review*, 40(3):497–525.

Guinnane, T. W. (2011). The Historical Fertility Transition: A Guide for Economists. *Journal of Economic Literature*, 49(3):589–614.

Haines, M. R. (2001). The urban mortality transition in the united states, 1800-1940. Working Paper 134, National Bureau of Economic Research.

Haines, M. R. and Hacker, J. D. (2011). Spatial Aspects of the American Fertility Transition in the Nineteenth Century. In Merchant, E. R., Deane, G. D., Gutmann, M. P., and Sylvester, K. M., editors, *Navigating Time and Space in Population Studies*, pages 37–63. Springer, Dordrecht.

Haines, M. R., university Consortium for Political, I., and Research, S. (2010). Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. https://doi.org/10.3886/ICPSR02896.v3.

Jones, L. E. and Tertilt, M. (2006). An Economic History of Fertility in the U.S.: 1826-1960. NBER Working Papers 12796, National Bureau of Economic Research, Inc.

Knodel, J. and van de Walle, E. (1979). Lessons from the Past: Policy Implications of Historical Fertility Studies. *Population and Development Review*, 5(2):217–245.

Lebergott, S. (1984). *The Americans: An economic record.* New York, NY: W. W. Norton & Company.

Lewis, F. D. (1983). Fertility and savings in the united states: 1830-1900. *Journal of Political Economy*, 91(5):825–840.

Lindert, P. H. (1986). Westward Who? Estimates of Native White Interstate Migration After the War of 1812. *The Journal of Economic History*, 46(2):431–440.

McClelland, P. D. and Zechauser, R. J. (1982). *Demographic Dimensions of the New Republic. American Interregional Migration, Vital Statistics, and Manumissions, 1800-1860.* Cambridge University Press.

Mitchell, B. R. (2003). *International historical statistics: Europe, 1750-2000 / B.R. Mitchell.* Palgrave Macmillan Basingstoke, England, 5th ed. edition.

Murphy, T. (2015). Old Habits Die Hard (Sometimes). *Journal of Economic Growth*, 20(2):177–222.

Notestein, F. (1945). Population—The Long View. In Schultz, T. W., editor, *In Food for the World*, pages 36–57. University of Chicago Press, Chicago.

Oberly, J. W. (1991). Military Bounty Land Warrants in the United States, 1847-1900. https://doi.org/10.3886/ICPSR09514.v1.

Ransom, R. and Sutch, R. (1986). Did rising out-migration cause fertility to decline in antebellum new england? a life-cycle perspective on old-age security motives, child default, and farm-family fertility. *Working Papers on the History of Saving*, 5(April).

Rosenbloom, J. L. and Sundstrom, W. A. (2003). The Decline and Rise of Interstate Migration in the United States: Evidence from the Ipums, 1850-1990. Technical Report 9857, NBER Working Paper Series.

Ruggles, S. (2015). Patriarchy, power, and pay: The transformation of american families, 1800–2015. *Demography*, 52:1797–1823.

Ruggles, S., Genadek, K., Goeken, R., Grover, J., and Sobek, M. (2015). Integrated Public Use Microdata Series: Version 6.0 [dataset]. http://doi.org/10.18128/D010.V6.0.

Sequeira, S., Nunn, N., and Qian, N. (2020). Immigrants and the making of america. *Review of Economic Studies*, 87(1):382–419.

Sundstrom, W. A. and David, P. A. (1988). Old-Age Security Motives, Labor Markets, and Farm Family Fertility in Antebellum America. *Explorations in Economic History*, (25):164–197.

Vandenbroucke, G. (2008a). The american frontier: Technology versus immigration. *Review of Economic Dynamics*, 11(2):283–301.

Vandenbroucke, G. (2008b). The U.S. westward expansion. *International Economic Review*, 49(1):81–110.

Weiss, T. (1992). U.s. labor force estimates and economic growth, 1800–1860. In Gallman, R. E. and (Eds.), J. J. W., editors, *American economic growth and standards of living before the Civil War*, page pp. 19–75. Chicago, IL: University of Chicago Press.

Yasuba, Y. (1962). *Birth Rates of the White Population in the United States, 1800-1860: An Economic Study*. The Johns Hopkins Press, Baltimore.

A. DATA APPENDIX

A.1. Data Source

Data for this paper are compiled from multiple sources. Ruggles et al. (2015) provide a nationally representative sample of census data linked between 1850 and 1880, which we use to infer the migration status of individuals. Haines et al. (2010) provide county-level data on population and several county and/or state characteristics for all decennials from 1790 to 2002. We make use of the data of 1850 and 1880 to compute CWR and its change. Moreover, we exploit economic characteristics for counties and states in 1850. Oberly (1991) provides data on military bounty land warrants from 1847 to 1900 at the state level.

A.2. Key Variables

Child-woman ratio: From Haines et al. (2010) for 1800 and from Ruggles et al. (2015) for 1880. It is defined as the ratio between the number of white children 0-9 (or 1-4, 5-9, and 1-9) to the number of white women 15-49 in 1880.

Children ever born: From Ruggles et al. (2015). We compute average fertility by cohorts of women. We define a cohort to be five years of birth years. The decline of CEB is given by the difference (of the logarithm) between CEB of the 1893 and 1818 cohort.

Moved county: From Ruggles et al. (2015), 1850-1880 linked samples. It is equal to 1 if the person reports living in a different county in 1880 than 1850, and recoded to equal zero if county borders changed during that time.

Moved state: From Ruggles et al. (2015), 1850-1880 linked samples. It is equal to 1 if the person reports living in a different state in 1880 than 1850, and recoded to equal zero if county borders changed during that time. This variable is also equal to 1 if the individual reports being born in a different state.

Frontier: From authors' computations based on Ruggles et al. (2015) 1850-1880 linked samples. It is equal to 1 if the person moved to a state that is not in this list: Maine, New Hampshire, Vermont, New York, Massachusetts, Rhode Island, Connecticut, New Jersey, Delaware, Maryland, West Virginia, Pennsylvania, Virginia, North Carolina, South Carolina.

Age: From Ruggles et al. (2015). Age reported in years.

Literate: From Ruggles et al. (2015). Equal to one if the individual reports being literate.

Married: From Ruggles et al. (2015). Equal to one if the individual reports being married in early period or in both periods.

Sex ratio: From Haines et al. (2010). Computed as the ratio between total number of men to total number of women ($mtot/ftot$).

Percentange of population living in urban areas: From Haines et al. (2010). Defined as the ratio between urban and total population ($urb850/totpop$).

Population density: From Haines et al. (2010). Defined as the ratio between total population and the 1880 county area ($totpop/area$).

Percentange of churches per population: From Haines et al. (2010). Defined as the ratio between total number of churches and total population ($churches/totpop$, where $churches$ is the sum of all churches in the county).

Percentange of white population: From Ruggles et al. (2015). Defined as the ratio between white and total population.

Percentange of unimproved acres: From Haines et al. (2010). Defined as the fraction of agricultural land that is unimproved. Calculated as the number of acres of unimproved agricultural land divided by total agricultural land ($acunimp/(acunimp + acimp)$).

Railroad access: From Haines et al. (2010). Defined as the presence of railroads ($rail$).

Navigable river access: From Haines et al. (2010). Defined as the presence of navigable rivers ($water$).

Farm value per acre: From Haines et al. (2010). It is equal to the log of the total value of farm property divided by the acres of improved agricultural land in a county ($farmval/acimp$).

Workers in manufacturing: From Haines et al. (2010). It is computed as the ratio between the total number of men working in manufacturing and the total population ($mfglabor/totpop$)

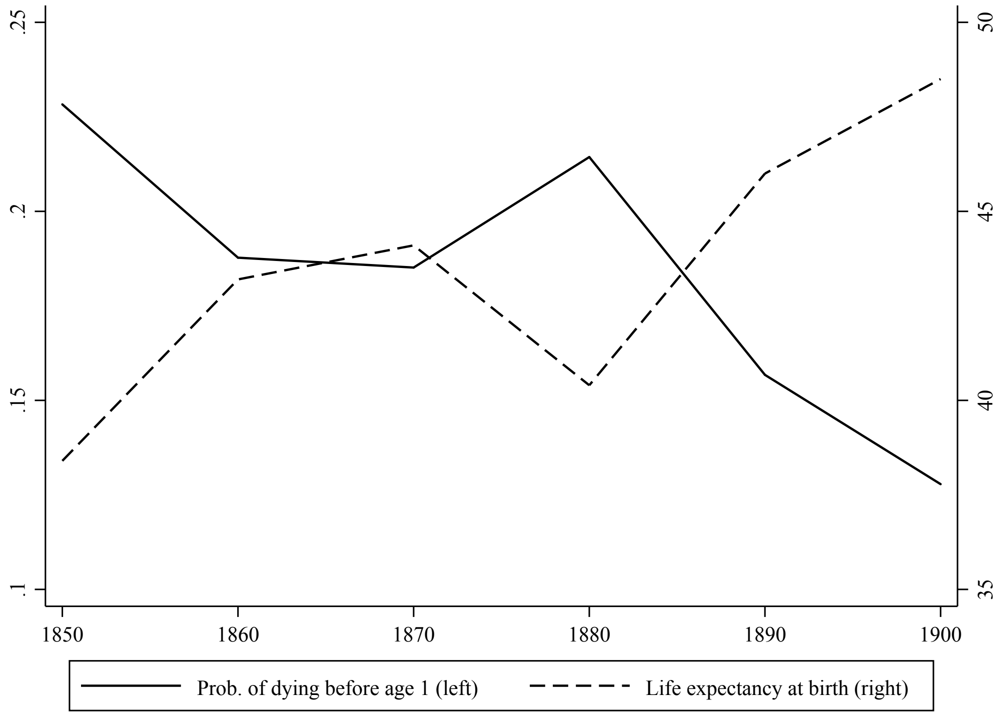
Acres: From Oberly (1991). It is the average acres granted to militaries or their families.

Money in circulation: From Haines et al. (2010) ($money$) at the state level.

Value of manufacturing output: From Haines et al. (2010) ($mfgrms2$) at the state level.

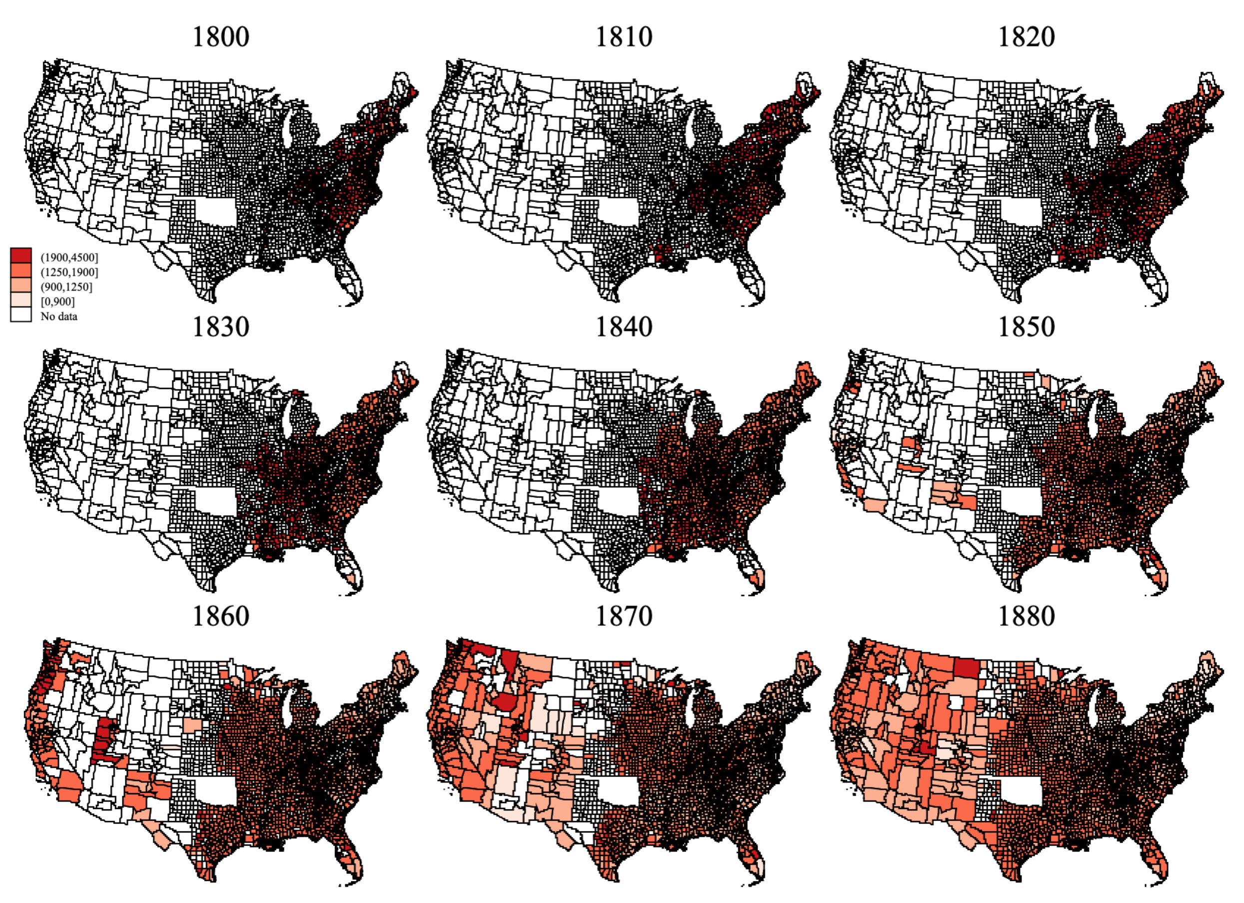
A.3. Descriptive Statistics

Figure 10 – Infant Mortality and Life Expectancy at Birth



Sources: Authors' computations using data from Haines (2001).

Figure 11 – Map of CWR, 1800-1880



Source: Authors' computations using data from Haines et al. (2010) and Ruggles et al. (2015).

Table A.3.1 – Descriptive Statistics

Variables:	Obs.	Mean	Std. Dev.	Min.	Max.
White Children 0-9 / White Women 15-44:					
1850	1,657	1,403.457	281.449	181.818	5,000
1880	2,505	1,267.215	127.604	234.375	1,500
White Children 1-4 / White Women 15-44:					
1850	1,615	667.832	138.179	156.863	2,285.714
1880	2,500	578.611	127.604	234.375	1,500
White Children 5-9 / White Women 15-44:					
1850	1,611	749.948	138.686	152.824	2,285.714
1880	2,500	657.585	119.942	205.128	1,777.778
Children ever born (state level):					
1818	43	5.483	1.754	2	11.055
1893	43	2.465	0.490	1.255	3.687
Percentage of men migrants who:					
Moved counties	1,176	0.336	0.265	0	1
Moved states	1,176	0.261	0.251	0	1
Frontier	1,176	0.250	0.253	0	1
Demographic characteristics of men migrants 1850-1880:					
Age	1,177	19.411	5.146	10.25	47.5
Occupational score	1,166	18.589	7.848	9	80
Percentage of literate	1,174	0.910	0.208	0	1
Percentage of farmers	1,177	0.719	0.341	0	1
Percentage of married	1,177	0.293	0.291	0	1
1850 County characteristics:					
Demographics:					
Age (home)	1,177	21.111	3.010	11.5	63
Age (dest.)	1,108	21.144	2.872	11.5	63
Sex ratio (home)	1,177	1.186	2.338	0.832	62.236
Sex ratio (dest.)	1,108	1.137	0.792	0.832	16.463
Percentage of US born (home)	1,177	0.943	0.099	0.289	1
Percentage of US born (dest.)	1,108	0.941	0.088	0.357	1
Percentage of white pop. (home)	1,177	0.981	0.051	0.508	1
Percentage of white pop. (dest.)	1,108	0.981	0.045	0.508	1
Percentage of illiterate (home)	1,177	0.067	0.056	0	0.298
Percentage of illiterate (dest.)	1892	0.578	0.190	0.001	0.992
Churches per pop. (home)	1,173	0.002	2.338	0.832	62.235
Churches per pop. (dest.)	1,104	0.002	0.001	0	0.012
Pop. density (home)	1,607	53.619	692.881	0.008	25,777.350
Pop. density (dest.)	1,104	0.002	0.001	0	0.012
Economics:					
Presence of navigable rivers (home)	1,176	0.468	0.499	0	1
Presence of navigable rivers (dest.)	1,107	0.450	0.428	0	1
Railroad access (home)	1,176	0.275	0.446	0	1
Railroad acces (dest.)	1,107	0.281	0.388	0	1
Percentage of urban pop. (home)	1,177	0.045	0.138	0	1
Percentage of urban pop. (dest.)	1,108	0.058	0.128	0	0.873
Percentage of unimproved acres (home)	1,174	0.594	0.185	0	1
Percentage of unimproved acres (dest.)	1,107	0.597	0.168	0	0.996
Log farm value per acre (home)	1,173	28.832	63.100	3.044	2,033.361
Log farm value per acre (dest.)	1,107	33.747	59.206	4.836	1,153.638
Percentage of workers in manufacturing (home)	1,177	0.020	0.030	0	0.268
Percentage of workers in manufacturing (dest.)	1,108	0.021	0.029	0	0.320
Labor force participation (home)	1,177	0.895	0.111	0	1
Labor force participation (dest.)	1,108	0.897	0.097	0	1
Acres of land warrants	1,432	11,787.710	6,926.097	160	32,720
Acres of land warrants as pct. of unimp. acres	1,117	0.002	0.004	0.000	0.087
Money in circulation (home)	1,433	7,777,089	7,709,370	833,960	1.55e+08
Value manufacturing output (home)	2,409	1.36e+07	3.01e+07	24,300	5.55e+08

B. MAIN RESULTS

Table B.1 – Evidence on Migration and Decline of CWR - Migrants

Dependent Variable: **Decline of CWR from 1850 to 1880**	**Children of age 0-9**									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PANEL (I)										
Migrants	0.041**	0.039**	0.036**	0.043**	0.047**	0.044**	0.040**	0.036*	0.033	0.031
	(0.020)	(0.017)	(0.016)	(0.017)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.021)
Number of clusters (home county)	147	147	147	147	143	143	143	143	143	143
Observations	1,176	1,172	1,166	1,154	1,089	1,087	1,087	1,086	1,086	1,086
R-squared	0.496	0.594	0.618	0.624	0.630	0.631	0.652	0.668	0.668	0.671
PANEL (II)										
County migrants	0.017	0.016	0.007	0.023	0.033	0.028	0.028	0.023	0.022	0.022
	(0.029)	(0.025)	(0.025)	(0.026)	(0.026)	(0.026)	(0.026)	(0.025)	(0.025)	(0.026)
State migrants	0.048**	0.046**	0.044**	0.051**	0.047**	0.044**	0.045**	0.040**	0.036	0.034
	(0.023)	(0.019)	(0.019)	(0.026)	(0.023)	(0.023)	(0.019)	(0.022)	(0.018)	(0.024)
Number of clusters (home county)	147	147	147	147	143	143	143	143	143	143
Observations	1,176	1,172	1,166	1,154	1,089	1,087	1,087	1,086	1,086	1,086
R-squared	0.497	0.594	0.619	0.624	0.631	0.652	0.652	0.668	0.668	0.671
PANEL (III)										
Other than frontier migrants	0.012	0.019	0.009	0.025	0.034	0.035	0.029	0.024	0.023	0.022
	(0.028)	(0.025)	(0.025)	(0.024)	(0.026)	(0.026)	(0.025)	(0.025)	(0.024)	(0.025)
Frontier migrants	0.050**	0.046**	0.044**	0.049**	0.052**	0.047**	0.045**	0.040**	0.036**	0.034
	(0.024)	(0.020)	(0.019)	(0.020)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)
Number of clusters (home county)	147	147	147	147	143	143	143	143	143	143
Observations	1,176	1,172	1,166	1,154	1,089	1,087	1,087	1,086	1,086	1,086
R-squared	0.497	0.594	0.619	0.624	0.631	0.631	0.652	0.668	0.668	0.671
1850 Home county controls:										
Demographics	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economics	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Migrant Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
1850 Destination county controls:										
Demographics	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Economics	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
1880 Home county controls:										
Demographics	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Economics	No	No	No	No	No	No	No	Yes	Yes	Yes
1880 Destination county controls:										
Demographics	No	No	No	No	No	No	No	No	Yes	Yes
Economics	No	No	No	No	No	No	No	No	No	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table B.2 – Evidence on Migration and Decline of CWR - Migrants

Dep. Var.:	***Decline of CWR from 1850 to 1880, Children of age 0-9***					
	(1)	(2)	(3)	(4)	(5)	(6)
Migrants	0.044**	0.031				
	(0.020)	(0.021)				
County migrants			0.028	0.022		
			(0.026)	(0.026)		
State migrants			0.044**	0.034		
			(0.023)	(0.024)		
Non-frontier migrants					0.035	0.022
					(0.026)	(0.025)
Frontier migrants					0.047**	0.034
					(0.024)	(0.025)
	1850 Home county controls:					
	Demographics					
% US born	0.508***	0.331***	0.503**	0.327***	0.502***	0.326***
	(0.065)	(0.068)	(0.065)	(0.068)	(0.065)	(0.068)
Age	-0.001	-0.002	-0.001	-0.002	-0.001	-0.002
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Illiteracy	-0.567***	-0.348***	-0.565***	-0.347***	-0.566***	-0.347***
	(0.089)	(0.097)	(0.089)	(0.097)	(0.089)	(0.097)
% white pop.	0.141	0.185*	0.141	0.184*	0.137	0.181
	(0.111)	(0.109)	(0.111)	(0.109)	(0.111)	(0.109)
Sex ratio	-0.058***	-0.043***	-0.058***	-0.044***	-0.058***	-0.044***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
% churches	-5.276	-4.893	-5.254	-4.809	-5.287	-4.838
	(3.382)	(3.319)	(3.368)	(3.320)	(3.361)	(3.307)
	Economics					
Rivers	-0.014	-0.019**	-0.013	-0.018**	-0.013	-0.018**
	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)	(0.009)
Railroads	0.005	-0.009	0.006	-0.008	0.005	-0.008
	(0.015)	(0.013)	(0.015)	(0.013)	(0.015)	(0.013)
% urban pop.	0.113**	-0.063	0.108**	-0.067	0.109**	-0.067
	(0.048)	(0.051)	(0.047)	(0.050)	(0.048)	(0.050)
Farm value	-0.000	0.000	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Labor force part.	0.088**	0.095**	0.088**	0.095**	0.088**	0.095**
	(0.029)	(0.028)	(0.030)	(0.029)	(0.030)	(0.029)
% unimp. acres	-0.013	0.063**	-0.012	0.063*	-0.011	0.064**
	(0.026)	(0.030)	(0.026)	(0.030)	(0.026)	(0.030)
Occup. score	0.002*	0.001	0.002*	0.001	0.002*	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
% men in manuf.	-0.478**	-0.411*	-0.483**	-0.411*	-0.485**	-0.415*
	(0.206)	(0.212)	(0.205)	(0.211)	(0.205)	(0.210)

			Migrant Controls:			
Age	0.000	-0.001	0.000	-0.002	0.000	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age square	-0.000	0.000	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Illiteracy	-0.010	-0.004	-0.011	-0.004	-0.010	-0.004
	(0.016)	(0.015)	(0.016)	(0.015)	(0.016)	(0.015)
% married	0.008	0.012	0.009	0.013	0.009	0.013
	(0.011)	(0.010)	(0.011)	(0.010)	(0.011)	(0.010)
Occup. score	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Farmer	0.009	0.004	0.009	0.004	0.009	0.004
	(0.012)	(0.011)	(0.012)	(0.011)	(0.012)	(0.011)
			1850 Destination county controls:			
			Demographics			
% US born	-0.124	-0.076	-0.115	-0.070	-0.115	-0.070
	(0.077)	(0.083)	(0.078)	(0.083)	(0.078)	(0.083)
Age	0.001	0.000	0.001	0.000	0.001	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Illiteracy	0.109	0.132	0.108	0.132	0.110	0.133
	(0.097)	(0.104)	(0.097)	(0.105)	(0.098)	(0.105)
% white pop.	0.027	-0.022	0.026	-0.021	0.028	-0.019
	(0.132)	(0.126)	(0.132)	(0.126)	(0.132)	(0.126)
Sex ratio	0.002	-0.006	0.002	-0.006	0.002	-0.006
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)
% churches	4.940	5.188	5.031	5.258	5.067	5.300
	(3.239)	(3.302)	(3.206)	(3.282)	(3.208)	(3.284)
			Economics			
Rivers	-0.011	-0.005	-0.012	-0.006	-0.012	-0.006
	(0.012)	(0.012)	(0.011)	(0.013)	(0.012)	(0.013)
Railroads	-0.015	-0.008	-0.015	-0.008	-0.015	-0.008
	(0.017)	(0.016)	(0.017)	(0.016)	(0.017)	(0.016)
% urban pop.	0.023	0.033	0.030	0.038	0.029	0.038
	(0.053)	(0.062)	(0.051)	(0.060)	(0.051)	(0.060)
Farm value	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Labor force part.	0.001	-0.010	0.001	-0.010	0.001	-0.010
	(0.037)	(0.036)	(0.037)	(0.036)	(0.037)	(0.036)
% unimp. acres	0.055*	0.066*	0.052	0.064*	0.051	0.063*
	(0.032)	(0.034)	(0.032)	(0.034)	(0.032)	(0.034)
Occup. score	-0.000	-0.001	-0.000	-0.001	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
% men in manuf.	0.092	0.460**	0.099	0.464**	0.105	0.470**
	(0.220)	(0.226)	(0.220)	(0.224)	(0.220)	(0.224)
1880 Home county controls	No	Yes	No	Yes	No	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of clusters	143	143	143	143	143	143
Obs.	1,087	1,086	1,087	1,086	1,087	1,086

R-squared	0.631	0.671	0.652	0.671	0.631	0.671

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. OLS regressions. Each column presents the estimates from a separate regression. The unit of observation is a county. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables.

C. EUROPE

Table C.1 – Descriptive Statistics

Variables:	Obs.	Mean	Std. Dev.	Min.	Max.
Austria (1869-1951):					
Children 5-9 / Women 15-49	8	329.389	51.697	240.149	390.757
Migrants by decade / Total population × 100	8	1.112	1.240	0.163	3.888
Belgium (1856-1947):					
Children 5-9 / Women 15-49	9	331.042	41.152	242.452	384.712
Migrants by decade / Total population × 100	9	0.236	0.161	0.022	0.408
Denmark (1870-1960):					
Children 5-9 / Women 15-49	10	357.513	33.338	281.905	397.163
Migrants by decade / Total population × 100	10	1.840	0.995	0.448	3.775
Finland (1890-1960):					
Children 5-9 / Women 15-49	8	404.604	57.326	310.480	476.764
Migrants by decade / Total population × 100	8	1.715	1.715	0.081	5.403
France (1851-1936):					
Children 5-9 / Women 15-49	17	293.951	26.620	208.238	327.719
Migrants by decade / Total population × 100	17	0.107	0.079	0.010	0.318
Germany (1871-1939):					
Children 5-9 / Women 15-49	8	328.653	64.499	219.954	381.370
Migrants by decade / Total population × 100	8	1.029	0.850	0.174	2.715
Italy (1861-1951):					
Children 5-9 / Women 15-49	9	371.492	34.168	310.118	417.715
Migrants by decade / Total population × 100	9	3.153	3.625	0.108	11.131
Netherlands (1859-1960):					
Children 5-9 / Women 15-49	11	384.419	27.984	343.205	432.810
Migrants by decade / Total population × 100	11	0.733	0.786	0.045	2.951
Norway (1855-1960):					
Children 5-9 / Women 15-49	11	377.516	51.473	257.075	453.333
Migrants by decade / Total population × 100	11	3.731	3.040	0.190	9.345
Spain (1855-1960):					
Children 5-9 / Women 15-49	11	372.290	55.997	285.329	478.900
Migrants by decade / Total population × 100	11	2.227	2.262	0.019	6.130
Sweden (1857-1960):					
Children 5-9 / Women 15-49	12	341.836	41.759	234.116	384.564
Migrants by decade / Total population × 100	12	2.252	2.254	0.125	6.834
Switzerland (1870-1960):					
Children 5-9 / Women 15-49	11	320.731	26.167	266.102	346.466
Migrants by decade / Total population × 100	11	0.976	0.764	0.239	2.898
United Kingdom (1851-1951):					
Children 5-9 / Women 15-49	9	352.131	35.073	296.935	393.894
Migrants by decade / Total population × 100	9	6.977	3.262	0.656	12.547

D. ROBUSTNESS CHECKS

Outliers. In Table D.1, we exclude counties that belong to the bottom and top 5% of the distributions of CWR or/and share of migrants. This allows us to sequentially remove potential outliers in terms of the dependent and the independent variables. We combine both approaches by excluding counties that meet any of these two criteria in columns (5) and (6). The upper panel of Table D.1 shows estimate results for state migrants, and the lower panel shows results for frontier migrants. Our estimates of interest are not substantially affected.

Table D.1 – OLS and IV-2SLS estimates of the effects of migration on the decline of the CWR: Outliers

Dependent Variable: ***Decline of CWR from 1850 to 1880***	Excluding 5th and 95th percentiles of CWR		Excluding 5th and 95th percentiles of migrants		Excluding 5th and 95th percentiles of CWR and migrants	
	OLS (1)	IV-2SLS (2)	OLS (3)	IV-2SLS (4)	OLS (5)	IV-2SLS (6)
County migrants	0.038*		0.029		0.034	
	(0.023)		(0.027)		(0.023)	
State migrants	0.059**	0.431***	0.042*	0.577***	0.047**	0.376***
	(0.019)	(0.100)	(0.024)	(0.123)	(0.019)	(0.103)
Migrant Controls	Yes	Yes	Yes	Yes	Yes	Yes
Home and destination county controls 1850	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	No	Yes	No	Yes	No
Kleibergen Paap F-statistic		21.317		25.578		17.866
R-squared	0.627		0.627		0.624	
Observations	1,009	989	1,035	1,102	958	938
Dependent Variable: ***Decline of CWR from 1850 to 1880***	Excluding 5th and 95th percentiles of CWR		Excluding 5th and 95th percentiles of migrants		Excluding 5th and 95th percentiles of CWR and migrants	
	OLS (1)	IV-2SLS (2)	OLS (3)	IV-2SLS (4)	OLS (5)	IV-2SLS (6)
Other than frontier migrants	0.039*		0.031		0.036*	
	(0.021)		(0.026)		(0.022)	
Frontier migrants	0.059**	0.411***	0.042*	0.544***	0.054**	0.358***
	(0.020)	(0.094)	(0.025)	(0.113)	(0.020)	(0.096)
Migrant Controls	Yes	Yes	Yes	Yes	Yes	Yes
Home and destination county controls 1850	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	No	Yes	No	Yes	No
R-squared	0.627	0.718	0.671	0.671		
Kleibergen Paap F-statistic		23.093		22.302		19.316
R-squared	0.615		0.626		0.624	
Observations	1,009	989	1,035	1,102	958	938

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. Each column presents the estimates from a separate regression. The unit of observation is a state. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables.

Table D.2 displays different approaches that allow us to ensure that our estimates of interest are not driven by outlying counties. The Table conveys results of estimations where we exclude counties for which the overall model performs poorly and produces residuals that exceed 2- and 3- standard deviations.

Table D.2 – OLS estimates of the effects of migration on the decline of the CWR: Outliers, contd.

Dependent Variable: ***Decline of CWR from 1850 to 1880***	Excluding 2-sigma outliers		Excluding 3-sigma outliers	
	(1)	(2)	(3)	(4)
County migrants	0.035		0.035	
	(0.023)		(0.025)	
State migrants	0.057***		0.061***	
	(0.016)		(0.018)	
Other than frontier migrants		0.037		0.037
		(0.022)		(0.024)
Frontier migrants		0.057***		0.061***
		(0.017)		(0.019)
Migrant Controls	Yes	Yes	Yes	Yes
Home and destination county controls 1850	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
R-squared	0.716	0.716	0.667	0.667
Observations	1,038	1,038	1,075	1,075

$***p<0.01$, $**p<0.05$, $*p<0.1$. Each column presents the estimates from a separate regression. The unit of observation is a state. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables.

Migrants. We append the linked samples of migrant women to the one of migrant men used in the main analysis. The drawback of this dataset is that it only contains women that preserved their last name from one wave to the other. It is this characteristics that allowed IPUMS to link individuals in two census years. Hence, women who got married and changed their last name in between census, are not present in the second census year. We estimate equations 3.1 to 3.3, and the IV-2SLS estimate for the new sample, and confirm the previous results. Coefficients are in Table D.3. Note that neither the number of observations nor the size of the coefficients differ importantly from those estimated in the previous sections. We also run the all the robustness checks: numbers are line with main findings and available under request.

Table D.3 – OLS and IV-2SLS estimates of the effects of (man and woman) migration on the decline of CWR

	OLS			IV-2SLS estimates		
Dependent Variable: ***Decline of CWR from 1850 to 1880***	(1)	(2)	(3)	(4)	(5)	(6)
Migrants	0.054**			0.219**		
	(0.017)			(0.064)		
County migrants		0.022				
		(0.026)				
State migrants		0.065**			0.272***	
		(0.019)			(0.078)	
Other than frontier migrants			0.025			
			(0.025)			
Frontier migrants			0.066**			0.266***
			(0.020)			(0.076)
Migrant Controls	Yes	Yes	Yes	Yes	Yes	Yes
Home and destination county controls 1850	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.636	0.637	0.637			
Kleibergen Paap F-statistic				86.17	58.25	41.36
Observations	1,082	1,082	1,082	1,023	1,023	1,023

$***p < 0.01$, $**p < 0.05$, $*p < 0.1$. Each column presents the estimates from a separate regression. The unit of observation is a state. White heteroskedastic standard errors adjusted for clustering at the county level in parentheses. All regressions include a constant term. See Appendix A.2 for sources and definitions of variables.