UNIVERSITY OF OTTAWA

DOCTORAL THESIS

Essays in Economic History

Author:
Joanne HADDAD

Supervisor:
Abel BRODEUR
Co-Supervisor:
Jason GARRED

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Department of Economics
Faculty of Social Sciences
University of Ottawa

Declaration of Authorship

I, Joanne Haddad, declare that this thesis titled, "Essays in Economic History" and the work presented in it are my own. Chapter one of this thesis is done by myself jointly with Dr. Abel Brodeur. My contribution is equal to his. The second chapter of this thesis is done jointly with Lamis El Kattan. My contribution is equal to hers. I, hereby acknowledge the contribution of Dr. Abel Brodeur, my thesis supervisor, for the research related to chapter one. I acknowledge the contribution of Lamis El Kattan, for the research related to chapter two.
I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Joanne HADDAD

Date: December 22, 2020

Acknowledgements

Special thanks go to my thesis supervisors Abel Brodeur and Jason Garred for the solid academic training and continuous support. I am indebted to them for reading and editing previous drafts of this dissertation. I would like to express my highest gratitude to Abel Brodeur for setting the bar high and therefore forcing me to surpass my limits. Throughout my time as his student, Abel never failed to give me guidance as well as moral and academic encouragement. He strongly endorsed my profile and helped me benefit the most out of my doctoral academic years. I also sincerely thank Jason Garred for all his advices, comments, suggestions and enlightenment as my thesis co-supervisor.

I thank my thesis committee composed of three internal examiners: Catherine Deri Armstrong, Louis-Philippe Morin and Matthew D. Webb for their suggestions, ideas and recommendations during my first and second thesis workshops and thesis defence as well as during departmental seminars. I also thank Laura Salisbury for agreeing to evaluate my thesis as an external examiner and for her comments and suggestions.

I am grateful to Samuel Bazzi and Martin Fiszbein from Boston University for offering me the opportunity to visit their campus and for their outstanding suggestions and comments, specifically related to the third chapter of this thesis. I also thank all faculty members at the University of Ottawa and all seminar participants at the AEA Economics of LGBTQ+ Individuals Virtual Seminar Series, CEA, Development Reading Group at Boston University, SCSE and departmental seminars at the University of Ottawa. I also thank Carole-Anne Gratton, Diane Ritchot and Irene Pare for their administrative help.

Lastly, I thank Farzaneh Davarzani, Fabien Forge, Lamis El Kattan, Taylor Wright and my colleagues at the University of Ottawa for their moral and academic support througout my years of study.

Dedicated to my father, mother, two sisters and brother for believing in me and supporting me througout my doctoral studies.

Contents

List of Figures

List of Tables

Abstract

The first chapter relates the size of the present-day LGBT population to the discovery of gold during the nineteenth century gold rushes. Comparing the surroundings of gold rush counties to other mining counties, we find that there are currently 15% more same-sex couples in former gold rush counties and that residents of these counties have more favorable attitudes toward homosexuality. Gold rush counties were initially isolated, mostly uninhabited and lacked strong formal institutions, which helped shaping pro-LGBT attitudes. Examining channels of persistence, we provide empirical evidence for selective migration and the lack of strong religious institutions.

The second chapter examines the impact of gender focused labor legislation on women's labor force participation and economic empowerment. We rely on historical legislative acts passed by state legislatures and exploit whether or not states passed regulatory laws regulating overall and industry specific employment and work conditions for women, night work laws and labor laws requiring provision of seats for working women. We exploit the fact that not all states enacted these laws as well as the variation in the timing of enactment of such laws. Our results show that women in comparison to men in treated states are more likely to be in the labor force post introduction of night work laws in comparison to control states. We also document the effect of industry-specific labor policies on women's likelihood to be employed in the affected industry and in higher-wage occupations within the industry of interest. Policy implications of our findings endorse the adoption of labor laws in favor of women to advocate their empowerment through a higher involvement in the labor market and financial independence.

The third chapter tests the doctrine of first effective settlement by relating early settlers' culture to within state variation in gender norms in the United States. In 1973, the cultural geographer Wilbur Zelinsky postulated that the distinctive traits of early settlers at initial stages of institutional development may be crucial for cultural formation. I capture settlers' culture using past female labor force participation, women's suffrage and financial rights at their place of origin. I document the distinctive characteristics of settlers' populations and provide suggestive evidence in support of the spatial (across locations) and vertical (over time) transmission of gender norms. My results show that women's labor supply is higher, in both the short and long run, in U.S. counties that historically hosted a larger settler population originating from places with favorable gender attitudes. My findings shed new light on the importance of immigrants' characteristics and their countries/states of origin for cultural formation in hosting societies.

General Introduction

This dissertation includes three essays on economic history. All three chapters consider historical economic shocks that happened in the United States and examine their consequences in the short and long run. This research contributes to several strands of the literature.

First, I contribute to the literature on the persistent effects of historical shocks by examining the impact of gold rush discoveries in the United States on the spatial distribution of a minority group, the LGBT population, as well as attitudes towards them more than 100 years later. Second, I contribute to the literature on women's formal involvement on the labor market and their empowerment by providing evidence in support of gender focused labor laws favoring them. I formally investigate the causal impacts of such laws and offer an important contribution to the debate on public policies targetting female empowerment. This is crucial in light of the rising demand for policies enhancing women's well-being, specifically in emerging countries, where work conditions for women are still detrimental and/or their involvement on the labor market is still low. Finally, I contribute to the literature on the roots and persistence of cultural traits by examining the impact of settlers' culture in explaining within state variation in gender norms in the United States.

The first chapter examines a potential historical determinant of the geographic distribution of minority groups and attitudes toward these groups given that these vary significantly across and within societies. Specifically, we examine the present-day size and spatial distribution of the LGBT population in the United States and relate it to the discovery of gold during the nineteenth century gold rushes. By comparing the surroundings of gold rush counties to other mining counties, we find that there are currently 15% more same-sex couples in former gold rush counties. We investigate the attitudes towards the LGBT population more than 150 years after the gold rushes and we find that the residents of these counties have more favorable attitudes toward homosexuality. Gold rush counties were initially isolated, mostly uninhabited and lacked strong formal institutions, which helped shaping pro-LGBT attitudes. Examining channels of persistence, we provide empirical evidence for selective migration and the lack of strong religious institutions.

In the second chapter, the impact of gender focused labor legislation on women's labor force participation and economic empowerment is examined. We rely on historical legislative acts passed by state legislatures and exploit whether or not states passed regulatory laws regulating overall and industry specific employment and work conditions for women, night work laws and labor laws requiring provision of seats for working women. We exploit the fact that not all states enacted these laws as well as the variation in the timing of enactment of such laws. Our results show that women in comparison to men in treated states are more likely to be in the labor force post introduction of night work laws in comparison to control states. We also document the effect of industry-specific labor policies on women's likelihood to be

employed in the affected industry and in higher-wage occupations within the industry of interest. Policy implications of our findings endorse the adoption of labor laws in favor of women to advocate their empowerment through a higher involvement in the labor market and financial independence.

The last chapter explores the question of cultural formation in the United States, specifically gender related norms. In 1973, the cultural geographer Wilbur Zelinsky postulated this in his doctrine of "first effective settlement". There is however little empirical evidence supporting the role of early settlers in shaping culture over the long run. This chapter tests this hypothesis by relating early settlers' culture to within state variation in gender norms in the United States. I capture settlers' culture using past female labor force participation, women's suffrage and financial rights at their place of origin. I document the distinctive characteristics of settlers' populations and provide suggestive evidence in support of the spatial (across locations) and vertical (over time) transmission of gender norms. My results show that women's labor supply is higher, in both the short and long run, in U.S. counties that historically hosted a larger settler population originating from places with favorable gender attitudes. My findings shed new light on the importance of immigrants' characteristics and their countries/states of origin for cultural formation in hosting societies.

Chapter 1

Institutions, Attitudes and LGBT: Evidence from the Gold Rush

1.1 Abstract

This chapter relates the size of the present-day LGBT population to the discovery of gold during the nineteenth century gold rushes. Comparing the surroundings of gold rush counties to other mining counties, we find that there are currently 15% more same-sex couples in former gold rush counties and that residents of these counties have more favorable attitudes toward homosexuality. Gold rush counties were initially isolated, mostly uninhabited and lacked strong formal institutions, which helped shaping pro-LGBT attitudes. Examining channels of persistence, we provide empirical evidence for selective migration and the lack of strong religious institutions.

1.2 Introduction

The geographic distribution of minority groups and attitudes toward these groups vary significantly across and within societies. Many contemporaneous socioeconomic determinants of these differences have been proposed, but few studies have analyzed their origins. In this chapter, we investigate the role of the California Gold Rush (start date 1848) and other gold rushes (Oregon (1851), Colorado (1858), Washington (1873) and Alaska (1896)) on the current place of residence of the U.S. LGBT population and attitudes toward this group.[1]

Identifying the causal effect of historical shocks such as natural resources discoveries on current outcomes is difficult for a number of reasons. For example, characteristics of mining counties and non-mining counties differ along many dimensions and economic shocks other than gold discoveries could have happened later. We address these issues by comparing gold rush counties to two control groups: (1) counties in which gold was not discovered during one of the gold rushes (see Fig. 1.1 and 1.2) and (2) counties with at least one discovery of either copper, iron, nickel or silver during the gold rushes and no gold discoveries (1848–1899).

This identification strategy is attractive because counties in the control groups are more comparable to gold rush mining counties (henceforth "gold rush counties") than non-mining counties, i.e., the sample of gold rush and non-gold rush mining counties (and other mineral counties) is balanced across a wide range of geographic variables. Nonetheless, they differ in few key dimensions. A first major difference between gold rush counties and the control groups is the fact that gold

[1] A mineral "rush" is often defined in popular culture as an onrush of miners seeking high-profit opportunities.

rush counties were mostly uninhabited at the moment of the first gold discovery, whereas the control counties were already settled. Second, we show that gold discoveries during one of the gold rushes led to a large increase in population ("rush") and temporary unbalanced sex ratios. This is due to the fact that the gold rushes led to one of the largest mass migration in the history of the U.S. (Table 1.1). A third dimension along which gold rush counties differ from control counties is the quality and existence of local institutions at the time of gold discovery. Gold rush counties were mostly in remote, isolated places, without strong legal institutions (Clay and Wright, 2005).[2] Historical records suggest that this contextual setting made it easier for LGBT individuals to avoid oppression (see Section 1.3).

In a first step, we test whether counties with gold rush discoveries currently have more unmarried partner households of the same sex. The estimates of our model with state fixed effects suggest that there are about 15 percent more same-sex couples in gold rush counties than in our main control group, i.e., non-gold rush counties. These findings are robust to several specification checks such as the inclusion in our model of a broad range of county variables to proxy for geography, isolation, demography, economic activity, education and urbanization, as well as the use of our second comparison group, i.e., other mineral discoveries. We also provide evidence that gold rush counties, in comparison to non-gold rush counties, have more same-sex couples who are married.

In a second step, we test whether individuals currently living in gold rush counties have pro-LGBT attitudes. We rely on data from the General Social Survey (GSS) and show that residents in gold rush counties are more likely to report that homosexuality is not wrong than respondents in control counties. This result suggests that pro-LGBT attitudes have persisted in gold rush counties.

These results are intriguing given that we are directly comparing gold rush and non-gold rush mining counties. Both set of counties are mostly located in the West, have similar geographical characteristics and have had balanced sex ratios for more than 75 years. It thus remains unclear why gold rush counties still have higher levels of pro-LGBT attitudes 150 years after the initial shock. We provide indirect evidence in support of selective migration to gold rush counties. We find that both migrants and non-migrants currently living in gold rush counties are more likely to report pro-LGBT attitudes. These results suggest that both individuals growing up and moving to gold rush counties have more favorable attitudes toward this group. Furthermore, we find that there are both more male and *female* same-sex couples in gold rush counties than in our control groups, suggesting that the gold rushes left an ethos of homosexual tolerance, rather than of gay tolerance.

We also provide suggestive evidence that the differential in attitudes is partly driven by the lack of (local) formal institutions at the moment of gold discovery and in the following decades, which may be an outcome of the gold rushes. We show that 9 percent of gold rush counties had a place of worship at the time of gold discovery and that 42 percent still do not have a notable place of worship. For the remaining 49 percent of gold rush counties, it took on average about 30 years from gold discovery to build a notable place of worship.

We make three contributions in our research. First, to our knowledge, we are the first study to empirically document the origins of the spatial distribution of the LGBT population in the United States. Second, we provide suggestive evidence on

[2]Lack of law enforcement is documented in many historical narratives of the American Frontier (Bazzi, Fiszbein, and Gebresilasse, 2017). For instance, there was no federal mining law at the onset of the California gold rush (Clay and Wright, 2005).

historical determinants of divergence in current attitudes regarding homosexuality.[3] Our research thus contributes to a small literature on the economics of LGBT (Black, Sanders, and Taylor, 2007).[4] Previous studies have also analyzed homosexual/heterosexual earnings and labor supply differences (Allegretto and Arthur, 2001; Hamermesh, 2019). Other papers have examined hiring and workplace discrimination (Aksoy et al., 2019; Sansone, 2019; Weichselbaumer, 2003) and antigay sentiment (e.g., Coffman, Coffman, and Ericson, 2016). Recent contributions also document the spatial distribution of LGBT population (e.g., Black et al., 2002; Gates, 2013; Smart and Whittemore, 2017) and revealed attitudes toward homosexuals using average home prices (Christafore, Leguizamon, and Leguizamon, 2013). A recent study by Fernández, Parsa, and Viarengo, 2019 document a sharp upward jump in the share of people who approve of same-sex relations around 1992, possibly because of the 1992 presidential election.

Finally, we contribute to a growing literature documenting that the initial institutional environment plays an important role in shaping social norms and attitudes in the short- and long-run (Fernandez and Fogli, 2009).[5] The most relevant study is perhaps Bazzi, Fiszbein, and Gebresilasse, 2017 who document the long run effects of the westward-moving frontier of settlement in the U.S. They provide evidence that the frontier had a persistent impact on individualism and preferences to redistribution partly through selective migration.

1.3 Background

Period of Settlement and Mass Migration California Gold Rush (start date 1848) resulted in one of the largest mass migration in American history producing a swift demographic change. Shortly after the discovery of gold in California, Oregon's Gold Rush happened (start date 1851). Pike's Peak region of the Kansas and Nebraska Territories also witnessed a gold rush which lasted between 1858 and 1861, the Colorado Gold Rush. In 1873, the only significant gold rush in the State of Washington took place. Montana's Gold Rush era lasted for a few years during the 1860–70s.

The last gold rushes were Alaska's Gold Rush dating back to 1896. Between 1898 and 1899 gold was discovered in Nome, Alaska, launching what is referred to as the Nome Gold Rush.

A number of characteristics make the gold rushes different than non-gold rush mineral discoveries. First, both tend to be in isolated regions, but the period of settlement and the extent of community development in the locations prior to the discovery of gold differ. Thus, what differentiates gold rush from non-gold rush mining counties is primarily the fact that non-gold rush counties were already settled (Table 1.1).

Second, the gold rushes attracted hundreds of thousands of young men to the West. The mass migration of men to areas where gold was discovered resulted in

[3] Asal, Sommer, and Harwood, 2013 relate the decriminalization of homosexual acts for a large sample of countries to variables such as legal origin, religion and economic development.

[4] A large literature documents positive attitudes towards the LGBT population in urban areas (e.g., Chauncey, 1995). We control for contemporaneous urbanization in our analysis.

[5] We contribute to a recent literature on the long-term effects of mineral discoveries (e.g., Buonanno et al., 2015; Couttenier, Grosjean, and Sangnier, 2017). Our findings also relate to papers analyzing the consequences of disequilibria in sex ratios. See, for instance, Anderson and Ray, 2010, Angrist, 2002, Grosjean and Khattar, Forthcoming, Hesketh and Xing, 2006 and Schacht and Kramer, 2016.

a large increase in population and temporarily unbalanced sex ratios in gold rush counties.

The most dramatic example may be the California Gold Rush. Clay and Jones, 2008 document that about 2–3 percent of all native-born men aged 20–40 in the U.S. were residing in California by the end of 1850. Men from Mexico, Canada, the United Kingdom and mainland Europe also migrated to gold rush counties. California's population increased from 8,000 Californios at the end of the 1840s to 120,000 in 1850 (U.S. California Census). This increase was mostly driven by the influx of about 110,000 male gold rushers. By 1852, California's population increased to 200,000. Women constituted only 8% of the Californian population in 1850. This percentage increased to 30 percent by 1860 and to approximately 50 percent by 1940 (Clay and Troesken, 2000; Holliday, 2015).[6]

The other gold rushes had a similar (temporary) effect on population growth. For instance, population in Alaska nearly doubled following the Alaska Gold Rush with about 64,000 inhabitants, out of which 46,000 were males and 18,000 females (Nugent, 2007).

Local Institutions The gold rushes were also different than non-gold rush mineral discoveries because of the lack of (local) formal economic, political and social institutions. McDowell, 2002 writes that "Not only were there no institutions to enforce the laws in California when gold was discovered, there were no laws." Laws governing property rights in mineral lands were absent and a new system of mining claims was later introduced during the California gold rush. At first, mining regions were treated as commons from which anyone could gather gold (McDowell, 2002).

The lack of places of worship was also evident. Goodrich, 1859 writes that "Away from law, away from public opinion, away from the restraints of home [...] puritan became a gambler; the boy taught to consider dancing a sin soon found his way to masked balls; *monte* became as familiar as the communion, and the catechism was forgotten, while the champagne popped sparked and excited."

Using historical data on Baptist, Congregational, Methodist, Presbyterian and Roman Catholic churches, Maffly-Kipp, 1994 documents that there were twice as many churches per capita in New York and Ohio than in California in 1860. In Iowa, a frontier region that was settled and evangelized during the same years as California, there were three times more churches per capita than in California. Maffly-Kipp, 1994 writes that "Gold rush California has consequently been depicted as a society without religion–'a prehistory'."

In this paper, we focus on religious institutions (and not other economic, social and political institutions) because of its documented importance in explaining attitudes toward the LGBT population (see Adamczyk and Pitt, 2009, for instance) and data availability.

1.3.1 Gold Rushes and the LGBT Population

Historical records indicate more favorable attitudes regarding homosexuality in gold rush mining areas. During the nineteenth century, the LGBT lived in hiding to avoid persecution. Sodomy was a felony in all states and considered a capital offense in some states. The lack of strong formal institutions in gold rush counties made it easier for LGBT individuals to avoid oppression and discrimination. With very weak

[6] This decrease in the male-to-female ratio resulted from the immigration of families and births in California, as well as the immigration of single females (Hurtado, 1999).

political, religious and social institutions at the time, sexual relations between men were not as socially opposed as in the rest of the country (Posner, 1994). Hurtado, 1999 argues that what was otherwise damned as immoral, was not completely removed from gold rush societies. This seemed designed to satisfy the appetites of a population of young men.

The lack of strong (formal) institutions and punishment may have favored pro-LGBT attitudes in gold rush counties. William, 1999 writes that miners "lived cheek-by-jowl with one another, even sharing beds and blankets." Men accepted to become dancing partners of other men (Boyd, 2003). See Appendix Figures 1.5 and 1.6. Lipsky writes that "The people of the gold rush understood sexuality differently than they do today. Concepts of homosexuality, bisexuality and heterosexuality had not yet been invented, so people did not define themselves by their sexual orientation. Having intimate relations with a person of the same gender did not automatically make anyone into 'a certain kind' of individual."

While it may have been a matter of convenience for some migrants to move to these male abundant gold rush areas, some of the migrant men may have been homosexuals who seized this economic opportunity to move West at a time where norms where almost absent. Lipsky, 2006 writes that "With few women in the city and even fewer in the mining camps, men undoubtedly turned to each other for comfort of all kinds. For some, in their physical and sexual primes, it may have been a matter of convenience only, but others certainly seized the opportunity to live exactly as they pleased, far from home and free of traditional social pressures."

1.3.2 Channels of Persistence and Transmission of Attitudes

There are many possible mechanisms explaining a long-run relationship between the gold rushes and homosexuality and attitudes. A first channel is selective migration to former gold rush counties. A rich literature on migrants point out that locational decisions are linked to the size of the minority group already in that location (e.g., Munshi, 2003). Locational preferences could be a function of LGBT attitudes and formal enforcement in the area (Yue, 2013). This channel is particularly important in our context given the very high mobility of the LGBT population (Black, Sanders, and Taylor, 2007).[7]

Another plausible channel is the development and persistence of pro-LGBT attitudes. Many studies provide evidence that historical shocks may persist via culture and norms (Couttenier, Grosjean, and Sangnier, 2017; Guiso, Sapienza, and Zingales, 2006; Tabellini, 2008). Attitudes, culture and norms may be transmitted horizontally and vertically (Giuliano and Nunn, 2017).[8] This is important in our context given that experimental findings and self-reported data suggest that increased contact with persons known to be homosexuals increases tolerance and acceptance toward the LGBT population (e.g., Altemeyer, 2002).

Bisin and Verdier, 2011 review models of cultural transmission and explain that beliefs and culture are partly transmitted through generations and acquired by different forms of learning. Importantly, the initial conditions in some of these models determine the long-run equilibrium. The lack of strong institutions potentially allowed individuals to transmit pro-LGBT values and attitudes to other generations (Tabellini, 2008). Of note, pro-LGBT attitudes in gold rush counties may be linked

[7] See Appendix 1.12 for a discussion of another possible mechanism, situational homosexuality.

[8] In 2011, about 15 percent of same-sex couple households had at least one biological, step or adopted child in their household.

to the likelihood for LGBT individuals to self-identify and self-disclose their sexual orientation.

We argue that the institutional context and pro-LGBT attitudes persisted and reinforced one another over time. Obviously, the lack of strong formal institutions might itself be an outcome of the gold rushes. Selective migration of population might have affected local institutions and local culture at a critical juncture of institutional formation, thus leaving a persistent impact (Bazzi, Fiszbein, and Gebresilasse, 2017).

Last, there may be a broader selective migration of more liberal people to gold rush areas who sought freedom from religious, social and political pressures. We find very weak evidence in Section 1.8 that residents of gold rush counties are currently more liberal in general.

1.4 Data Sources

1.4.1 LGBT Data

In order to construct the size of the lesbian and gay population at the county level, we use decennial census data. The 2010 U.S. census identifies same-sex cohabiting couples based on information about the sex of the members of the household and information from the relationship question. The Census Bureau designates as the head of household (or householder) "the member (or one of the members) in whose name the home is owned, being bought, or rented." Then data are collected on all household members, with each person being identified in terms of her or his relationship to the householder. Our measure of unmarried partnered households is thus based on households of the same sex reporting an unmarried partnership. In 2010, there were 2.2 unmarried partner households of the same sex per 1,000 inhabitants, and about 515,000 unmarried partner households of the same sex. Figure 1.3 displays the number of unmarried partner households of the same sex per capita in 2010. This map illustrates that San Francisco, Seattle and Oakland are some of the (large) cities with the highest number of same-sex couples per inhabitants. See Appendix Section 1.12 for more details on LGBT data and a discussion of measurement error.[9]

We also use data from the American Community Survey (ACS) for the years 2013–2017. The ACS data is available at the Public Use Microdata Areas (PUMA) level. PUMAs are nested within states and contain at least 100,000 people. In 2016, about 3.3 percent of respondents reported being in a same-sex married couple. Of note, though, our research focuses on a time period in which the number of same-sex married couples was growing. There were 2.3 percent of same-sex married couples in the U.S. in 2013 in comparison to 4 percent in 2016. This is likely due to the rapid expansion of states legalizing same-sex marriage and to the ruling of the Supreme Court in the civil rights case of *Obergefell v. Hodges*.

1.4.2 Mineral Resources Data

The Mineral Resources Data System (MRDS) provides information on all mineral discoveries in the United States. For the purpose of this paper, we keep only mineral discoveries for which the information on the exact geographic location and year

[9]Note that San Francisco county is not defined as a gold rush county. This is due to the fact that there were no gold discovery in this county. Nonetheless, San Francisco was greatly affected by the gold rush. Adding San Francisco to our treated group increases the size of our main estimates.

of discovery is available. Our sample comprises all different categories of development status defined by the MRDS: occurrence, prospect, producer, past producer and plant. The year and location of discovery enables us to categorize gold discoveries as part of a gold rush or not. Gold discoveries are deposits that include gold, exclusively or among other minerals, as a primary commodity. Our data set includes 3,803 gold deposits discovered between 1700 and 2000 across 25 states.

We construct a variable *GoldRushes* that is equal to one for a county if at least one gold discovery occurred during one of the gold rushes and zero otherwise. We define the gold rushes as discoveries that occurred between the years 1848–1899 in the Pacific and Mountain Census divisions, i.e., Western region. We also include the few discoveries in South Dakota during that time period.[10] There are 304 counties in the U.S. with at least one gold discovery during the period 1700–2000. Of these, 211 are codified as gold rush counties. Figure 1.1 illustrates all gold rush discoveries in the U.S. (See Appendix Figure 1.8 for discoveries in Alaska.)[11] These figures confirm that there were many discoveries in Colorado, Montana, Northern California, Sierra Nevada, etc. Figure 1.2 shows the other gold counties, i.e., gold mining counties without a gold rush discovery. Appendix Figure 1.12 shows the "Other Mineral" counties, i.e., counties with at least one discovery of either copper, iron, nickel or silver during the time period 1848–1899 and no gold discoveries.

1.5 Identification Strategy

1.5.1 Identification Strategy

Our identification strategy consists of comparing gold rush counties to counterfactual counties. In our main empirical specification, we compare gold rush counties to a set of gold mining counties in which gold was not discovered during one of the gold rushes. The main concern is the omitted-variable bias. Gold rush counties could be in locations that have unobservable characteristics that are associated with the presence of the LGBT population. We investigate in Table 1.6 whether observable characteristics at the county-level predict whether a gold mining county is a gold rush county or not. We thus restrict the sample to counties with gold discoveries and directly compare gold rush and non-gold rush mining counties.[12]

We show the results from linear probability models for two specifications. We include in both columns Census division fixed effects and the following geographic variables: latitude, longitude, the natural log of total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties. We also include the natural log of population in 2010, the percentage of foreign born in 2010 and a dummy for the presence of military bases. In column 2, we add contemporaneous county variables such as the natural log of median personal income and median house value of owner-occupied housing units, an ethno-linguistic fractionalization index, the unemployment rate, the urbanization rate and the percentage of residents who graduated high school.

[10] Including the 17 counties in the South and Northeast Census regions with Gold discoveries during the gold rushes period (1848–1899) in the analysis has no effect on our main conclusions (Brodeur and Haddad, 2018).

[11] Appendix Figures 1.9, 1.10 and 1.11 provide illustrations of gold rush discoveries for the time periods 1848–1859, 1860–1869 and 1870–1899.

[12] Throughout our analysis, we exclude from our sample 17 counties in the South and Northeast Census regions that had gold discoveries between the years 1848–1899 and no gold discoveries before or after the gold rushes.

In column 1, none of the ten variables is statistically significant at the 10 percent level. In column 2, only three variables are significantly (two at the 10 percent level and one variable at the five percent level) related to whether the county is a gold rush county: unemployment rate, natural log of income and urbanization. When considering the set of geographic variables jointly, the joint p-value is about 0.31. Our findings suggest that the sample of gold rush and non-gold rush mining counties is balanced across a wide range of covariates and that the identification assumption is credible. In the empirical analysis, we include state fixed effects and the geographic variables to relax the identification assumption.

1.5.2 Model Specification

The objective is to investigate the impact of the gold rushes. Our basic model compares counties that experienced a gold rush to non-gold rush mining counties as a direct comparison group controlling for covariates susceptible to affect the location of gold mines and the spatial distribution of the LGBT population.

We estimate the following specification:

$$Y_{cs} = \alpha + \gamma Gold_c + \phi GoldRushes_c + X_c'\zeta + \delta_s + \varepsilon_{cs} \tag{1.1}$$

where Y_{cs} is the natural log of one plus the number of unmarried partner households of the same sex in county c and state s in 2010. $GoldRushes_c$, our variable of interest, is a dummy indicating whether the county experienced a gold rush during the gold rushes period (1848–1899). ϕ denotes the coefficient of $GoldRushes_c$. γ captures the average relation between gold discoveries and the natural log of unmarried partner households of the same sex, whereas ϕ reports whether the effect is different whether or not discoveries occurred during one of the gold rushes.

δ_s are state fixed effects and X_c' is a set of control variables at the county level. We control for a number of county geographic controls including latitude, longitude, land area in miles, mean elevation, standard deviation of elevation, distance from the county to the state capital, and lastly whether or not it is a coastal county. We also control for the natural log of population in 2010. In some specifications, we further control for a set of (endogenous) contemporary controls to proxy for demography, economic activity, education and urbanization. We cluster standard errors at the state level. We also report wild cluster bootstrap p-values given the small number of Western states contributing to the identifying variation (MacKinnon and Webb, 2017; Roodman et al., 2019). As a robustness check, we use the estimation method developed by Conley, 1999; Conley, 2008 to take into account spatial correlation (Hsiang, 2010).

1.6 Historical Environment

In this section, we show that our treated and control groups differ in few key dimensions.

1.6.1 Selective Mass Migration

Table 1.1 reports summary statistics for counties that had gold discoveries during one of the gold rushes (1848–1899), counties that had gold discoveries before this

time period (and no gold rush discoveries), counties that had their first gold discovery after this time period and the rest of the U.S. The main variables of interest refer to population and population density per square mile. Population data is based on the first census before the year of earliest gold discovery. This table confirms that gold rush counties were mostly uninhabited at the moment of the first gold discovery, whereas counties that had their first gold discovery before or after the gold rushes (1848–1899) were already settled. The average number of inhabitants for gold rush counties was 1,976 (population density of 1.2). This is in stark contrast to counties that had gold discoveries before 1848–1899 (average population of 9,455 and population density of 19.6) and to counties that had their first gold discoveries after 1848–1899 (average population of 38,214 and population density of 27.9). The rest of U.S. counties had on average 5,700 inhabitants (population density of 18.3) in 1840.

Table 1.1 also provides descriptive statistics for our second comparison group, "Other Minerals" counties (i.e., counties with at least one discovery of either copper, iron, nickel or silver during the gold rushes and no gold discoveries). Based on the first census before the earliest non-gold mineral discovery, "Other Minerals" counties had an average number of inhabitants of 4,320 (population density of 5.95). This is about twice as many inhabitants as gold rush counties and slightly less than the rest of U.S.

Table 1.1 also confirms Clay and Jones, 2008's findings that age and literacy are important factors in determining migration during the California Gold Rush. Using data from the first census following the first gold discovery, we show that 82 percent of the white population were prime age adults (15 to 49 years old) in gold rush counties and that only 10 percent were illiterate. We also document the presence of unbalanced sex ratios in gold rush counties in Appendix Figures 1.13 and 1.14. These figures illustrate sex ratios in 1860, 1870 and 1960 across U.S. counties. Sex ratios in the West were imbalanced during the 1860–70s, whereas counties in the South and Northeast regions had more balanced sex ratios. In contrast, sex ratios were nearly balanced across all counties in more recent years (see Appendix Figures 1.15 and 1.16).

To sum up, we find suggestive evidence that gold rush counties were not yet settled at the moment of gold discovery and that the gold rushes led to mass migration. Migration patterns to gold rush counties were characterized by a selective migration of young literate men.

1.6.2 Places of Worship During the Gold Rushes

We now provide empirical evidence that gold rush counties lacked strong institutions at the moment of gold discovery in comparison to the control group. We focus on religious institutions and rely on a novel data set of notable cathedrals, churches and missions (see Appendix Section 1.12 for more details on this data set). We first plot the share of gold mining counties with at least one notable place of worship by decade. As shown in Fig. 1.4, the percentage of gold mining counties with at least one notable place of worship at the time of gold discovery is much lower during the gold rushes (1848–1899) than during any other decade for the time period 1830–2000. We also check whether gold rush counties were less likely to have a notable place of worship at the time of (the first) gold discovery. We find that only 9 percent of gold rush counties had a notable place of worship at the time of gold discovery. This is in stark contrast to gold mining counties *without* gold rush discoveries. For this group, about 44 percent had a notable place of worship at the time of gold discovery.

We formally test whether gold rush counties were less likely to have a notable place of worship at the time of gold discovery than other gold mining counties in Table 1.2 and Appendix Table 1.7. We present estimates of a model in which the dependent variable is a dummy that takes the value of one for counties with a notable place of worship at the time of (the first) gold discovery and zero for counties without a notable place of worship at the time of discovery. We rely on probit estimation and report marginal effects. We restrict our sample to counties with at least one gold discovery at any time. We thus directly compare gold rush counties to non-gold rush mining counties.

The estimated relationship between having a notable place of worship at the time of gold discovery and being a gold rush mining county is negative and statistically significant at the 1 percent level. The estimate in column 1 suggests that gold rush counties were about 36 percent less likely to have a notable place of worship.

Overall, our results confirm that gold rush counties lacked strong local (religious) institutions.

1.7 Present-Day Outcomes

1.7.1 Baseline Specification: LGBT Population

The results derived from estimating Eq. (1.1) are reported in Table 1.3. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. What clearly emerges is that counties with at least one gold discovery during the gold rushes time period, i.e., 1848–1899, currently have statistically significantly more unmarried partner households of the same sex than their counterfactual counties. We include only the natural log of population in 2010 in columns 1 and 2. We then progressively add control variables and fixed effects in the remaining columns. Column 1 first shows the relationship between being a gold mining county and the number of unmarried partner households of the same sex. We find that there are about 11 percent more unmarried partner households of the same sex in 2010 in gold mining counties than in the rest of the U.S.

In column 2, we include our variable of interest *GoldRushes*. Column 2 shows that the estimated relationship between the number of unmarried partner households of the same sex and being a gold rush mining county is positive and statistically significant at the 2 percent level. The estimate suggests that there are about 13 percent more unmarried partner households of the same sex in 2010 in gold rush mining counties than in their counterfactual counties.

In all regressions, the coefficient of interest, ϕ, is positive and the estimated coefficients range from 0.13 to 0.17. The estimates are statistically significant at the 2 percent level in all columns. This means that counties with at least one gold discovery which occurred during one of the gold rushes in the West currently have approximately 13–17 percent more unmarried partner households of the same sex.

In Appendix table 1.8, we explore a measure of intensity for the gold rushes: the number of gold discoveries during the gold rushes (1848–1899).[13] There were approximately 8 gold rush discoveries on average for counties with at least one gold rush discovery. Furthermore, there is a lot of variation in gold discoveries across treated counties (std. dev. 11). See Appendix Figure 1.17. As we can see, the estimated coefficients of gold rushes intensity are positive and statistically significant.

[13] We explore another measure of intensity in the Appendix Section 1.12: neighboring a county with at least one gold discovery during the gold rushes.

In contrast, the evidence presented suggests that the number of gold discoveries before or after the gold rushes is not related to the number of unmarried partner households of the same sex.

In order to gain insights on why more gold discoveries during the gold rushes would lead to larger effects on the size of the LGBT population, we split the treated set of counties in two subgroups: below and above the mean number of gold rush discoveries. We then predict (using a probit model) whether a gold rush county has more than the mean number of gold discoveries with our set of geographic controls (estimates available upon request). We find that gold rush counties with eight or more discoveries are significantly further away from the coast and have higher elevation. This result suggests that treated counties with many gold discoveries are relatively more isolated than treated counties with less than eight discoveries. We illustrate this pattern in Appendix Figure 1.18. We also find that gold rush counties with higher than the mean number of gold rush discoveries were less likely to host a notable place of worship. About 12 percent of gold rush counties with at least eight gold rush discoveries had a notable place of worship at the moment of earliest discovery in comparison to 19 percent for treated counties with less than eight gold rush discoveries.

The fact that the abundance of gold discoveries during the gold rush is related to the number of unmarried partner households of the same sex in 2010 provides suggestive (indirect) evidence that isolation and remoteness played an important role in shaping social norms.

1.7.2 Robustness Checks

In the Appendix Section 1.12, we test the sensitivity of our results to the choice of covariates, clustering, the definition of the dependent variable and to the use of a second control group, i.e, counties with at least one discovery of either copper, iron, nickel or silver during the gold rushes and no gold discoveries and zero otherwise. We also show that our results are robust to controlling for non-gold rush mining counties with discoveries before and after the gold rushes separately. Results from these robustness exercises are found in Appendix Tables 1.9, 1.10, 1.11, 1.12 and 1.13.

So far, our findings suggest that there are currently more unmarried partner households of the same sex in counties in which gold was discovered during one of the gold rushes.[14] We now test whether there are currently more same-sex *married* couples in these areas. We rely on ACS data for the years 2013–2017.

Appendix Table 1.14 presents estimates of Eq. (1.1) where the dependent variable is the natural log of one plus the number of same-sex married couples. The analysis is at the PUMA level and we have 2,323 observations. In our sample of respondents aged 18 to 65, the mean of the variable "Same-Sex Married Couples" is 4.3 per 1,000 inhabitant. In the specifications without Census region, Census division or state fixed effects, the estimates are large, positive, significant at the 1 and 5 percent levels (with geographic controls) and range from 20 to 27 percent. In contrast, the estimates are much smaller (coeff. 0.02 and std. err. 0.052) and statistically insignificant when we include state fixed effects. These results suggest that same-sex married couples are much more likely to be living in gold-rush counties and more generally in Western states. But we provide only weak evidence that the gold rushes

[14] In Appendix Section 1.12, we repeat the analysis but for male and female same-sex unmarried partners separately. Results are reported in Appendix Table 1.23.

are related to the number of same-sex married couples when we compare counties within a census region, census division or a state.[15]

The discrepancy in the size of the estimates for the different measures of the LGBT population, i.e., same-sex *married* couples and *unmarried* partner households of the same sex, may be due to many reasons. First, the analysis is at the PUMA level for the analysis using same-sex married couples data, whereas it is at the county level for our main outcome variable. We thus have less observations per state at the PUMA level. Second, the analysis covers a time period in which the number of same-sex married couples was growing and in which the Supreme Court ruled that states must license and recognize same-sex marriages.[16]

1.7.3 Socioeconomic Characteristics

We test in Table 1.4 whether current and historical local amenities and socioeconomic characteristics may explain the geographic distribution of the LGBT population in gold rush and non-gold rush mining counties. Table 1.4 presents estimates of Eq. (1.1). The estimates suggest that our main results are robust to controlling for the racial composition (Black, Sanders, and Taylor, 2007; Christafore, Leguizamon, and Leguizamon, 2013) and *historical* controls for isolation and geography. By contrast, our estimates slightly decrease when controlling for the local economic environment and current education and urbanization. When all the controls are included simultaneously, our estimates decrease in size by about one-sixth and are statistically significant at conventional levels.[17]

To sum up, our estimates suggest that the persistence in the geographical distribution of LGBT population is (mostly) not driven by current and historical socioeconomic characteristics and isolation/geography.

1.7.4 Attitudes Regarding Homosexuality

In this subsection, we present the results on the impact of the gold rushes on current attitudes regarding homosexuality. We rely on data from the General Social Survey (GSS) over the years 1993–2014 and focus on the following question: "What about sexual relations between two adults of the same sex–do you think it is always wrong, almost always wrong, wrong only sometimes, or not wrong at all?" In our sample, about 55 percent of respondents report that it is always wrong, whereas a third of the respondents report it is not wrong at all.

The model is similar to Eq. (1.1) with the exception that the unit of observation is now the individual. We also control for the individual's demographic characteristics. Specifically, we estimate:

$$Y_{icst} = \alpha + \gamma Gold_c + \phi GoldRushes_c + X_c'\zeta + Z_{it}'\theta + \delta_s + \eta_t + \varepsilon_{icst} \tag{1.2}$$

where Y_{icst} is the answer to the question of whether sexual relations between two adults of the same sex is always wrong, almost always wrong, wrong only sometimes, or not wrong at all for individual i in county c, state s and year t. We rely

[15] The estimates are virtually the same if we instead rely on the natural log of one plus the number of same-sex married couples per capita as a dependent variable. See Appendix Table 1.15.

[16] Over 30 states had legalized same-sex marriage before nationwide legalization on June 26, 2015.

[17] Appendix Table 1.16 report the estimates for some of these controls. We find a positive and statistically significant association between the number of unmarried partner households of the same sex and the median personal income, the percentage of units with a value of more than $500,000 and educational attainment.

on ordered probit response models where the dependent variable is discrete and defined on a finite ordinal scale. Z'_{it} is a vector of individual characteristics. These characteristics include the individual's gender, age, age squared, six education dummies, three race dummies and five marital status dummies. X'_c include our set of county geographic controls.

Table 1.5 shows the results from ordered probit regressions. See Appendix Table 1.17 for the estimates of the individual control variables.[18] We include year fixed effects throughout. We control for respondents' characteristics in columns 3–6 and include our set of geographic controls in columns 4–6. The estimates are all positive and statistically significant suggesting that the gold rushes had a persistent (positive) effect on attitudes toward homosexuality. We check the robustness of our results in the Appendix (**??**) to other questions concerning attitudes toward LGBT.

One way to gauge the size of our estimates is to standardize the dependent variable for all respondents (within each year) to have a mean of zero and a standard deviation of one. We then run OLS regressions using Eq. 1.2. Our estimates suggest that respondents in counties with gold discoveries during one of the gold rushes are about 12 to 33 percent of a standard deviation more likely to report that sexual relations between two adults of the same sex is not wrong. Put differently, the estimated effect of the gold rushes is about half of the coefficient of graduating college in comparison to having a high school diploma and approximately the size of being a woman (in comparison to men).

We test in Appendix Table 1.18 whether the results on pro-LGBT attitudes are driven by individuals who grew up in gold rush counties or by migrants who are now living in these counties. We restrict the sample respectively to respondents who are living in the same city since the age of 16 and to individuals who moved to a different state or city since the age of 16. The estimates for migrants are all positive and statistically significant, suggesting that respondents who migrated to gold rush counties are more likely report pro-LGBT attitudes. Similarly, we find that respondents who grew up in gold rush counties are more likely to report pro-LGBT attitudes. The estimates are statistically significant in two columns out of three and of similar magnitude than for migrants when state fixed effects are included. These findings provide suggestive evidence that pro-LGBT attitudes have persisted partly through selective migration.

To sum up, we find that gold rushes had persistent effects on pro-LGBT attitudes and that both gays and lesbians are more likely to live in former gold rush counties than in counterfactual locations. Our results are consistent with the historic accounts of the gold rushes describing pro-LGBT attitudes. These findings could explain the positive relationship between the number of same-sex couples and gold rush discoveries for two reasons. First, LGBT individuals could migrate to gold rush counties because of the local population's pro-LGBT attitudes or because of the thickness of the LGBT marriage market. We confirm in Appendix Figure 1.19 that respondents living in counties with more same-sex couples are more likely to report pro-LGBT attitudes. This figure plots the relationship between unmarried partner households of the same sex per capita and the percentage of respondents reporting "Not wrong at all" at the county level. The relationship is positive and has a concave shape. Second, pro-LGBT attitudes could be related to the likelihood to self-disclose one's sexual orientation or gender identity. We are unfortunately unable to disentangle these two mechanisms.

[18] Females and educated individuals tend to be more likely to report pro-LGBT attitudes. Black or African Americans and other races (in comparison to white respondents) are less likely to report pro-LGBT attitudes.

1.8 Possible Channels of Causality

In what follows, we provide two pieces of evidence that the gold rushes may have affected the size of the LGBT population through the quality and existence of local institutions.

1.8.1 Religiosity Since the Gold Rushes

We first compute the number of years from gold discovery to having at a notable place of worship. Recall that only 9 percent of gold rush counties had a notable place of worship at the time of discovery. We find that about 42 percent of gold rush counties still do not have a notable place of worship in 2000. For the remaining 49 percent, it took on average 28 years before a notable place of worship was built. This figure confirms that residents of gold rush counties were not subject to the presence of strong religious institutions in the decades following gold discovery. While the presence of a notable place of worship at the time of discovery may be plausibly exogenous, it is likely that the lack of a notable place of worship in the following decades is an outcome of the gold rush.

We also test whether residents of gold rush counties are currently less fundamentalist using answers to the following question in the GSS: "How fundamentalist is r currently?" where respondents have three choices (1=Fundamentalist, 2=Moderate, and 3=Liberal). We present estimates from ordered probit regressions in Table 1.20. The structure of the table is the same as Table 1.5. The estimates are all positive and statistically significant at the 1 percent level, confirming that individuals in gold rush counties are less likely to be fundamentalist. Our OLS estimates (not shown for space consideration) suggest that the gold rushes decreased by about 13 to 40 percent of a standard deviation self-reported fundamentalism.[19]

Overall, our findings paint a clear picture of religiosity during and following the gold rushes. Individuals living in gold rush counties remained less religious and more tolerant toward the LGBT population up to nowadays.

1.8.2 Other Attitudes

It remains unclear whether attitudes in gold rush counties are more liberal in general or only toward the LGBT population. Previous work has documented that temporary imbalanced sex ratios and historical mining activity are associated with greater opposition to redistribution and economically conservative positions (e.g., Bazzi, Fiszbein, and Gebresilasse, 2017; Couttenier and Sangnier, 2015).

We thus test the effect of the gold rushes on other attitudes and values. We scoured the GSS questionnaires and found four variables covering diverse aspects of traditional American values that were asked repeatedly. The questions are "Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if she is married and does not want any more children", "Do you think the use of marijuana should be made legal or not?", "Do you favor or oppose the death penalty for persons convicted of murder?" and "Would you favor or oppose a law which would require a person to obtain a police permit before he or she could buy a gun?" We code the variables as dummies. We also rely on answers

[19] In another set of regressions, we also provide evidence that respondents in gold rush counties are significantly more likely to approve on the United States Supreme Court ruling that no state or local government may require the reading of the Lord's Prayer or Bible verses in public schools. We present the estimates in Appendix Table 1.21.

to a question in which respondents are asked whether they think of themselves as liberal or conservative. Respondents are offered seven choices (from 7 =extremely liberal to 1 =extremely conservative).

We present estimates of Equation 1.2 in Appendix Table 1.22. We include individual and geographic controls, and year and state fixed effects in all columns. In column 1, we present the estimate from an ordered probit regression where the dependent variable is whether respondents think they are liberal or conservative. The point estimate in column 1 is close to zero and statistically insignificant. This question from the GSS provides weak evidence of the relationship between the gold rushes and the likelihood to report being liberal.

In columns 2–5, we present estimates from probit regressions for the other four variables. We find a positive and statistically significant relationship between gold rush mining counties and the likelihood to be in favor of legalizing marijuana. The estimate suggests that residents in gold rush counties are about fifteen percent more likely to support legalizing marijuana. On the other hand, the estimates are not statistically significant for death penalty, gun laws and abortion.

To sum up, we find (very) weak evidence that gold rush counties are more liberal in general. The gold rushes led to pro-LGBT attitudes and to slightly more liberal attitudes.

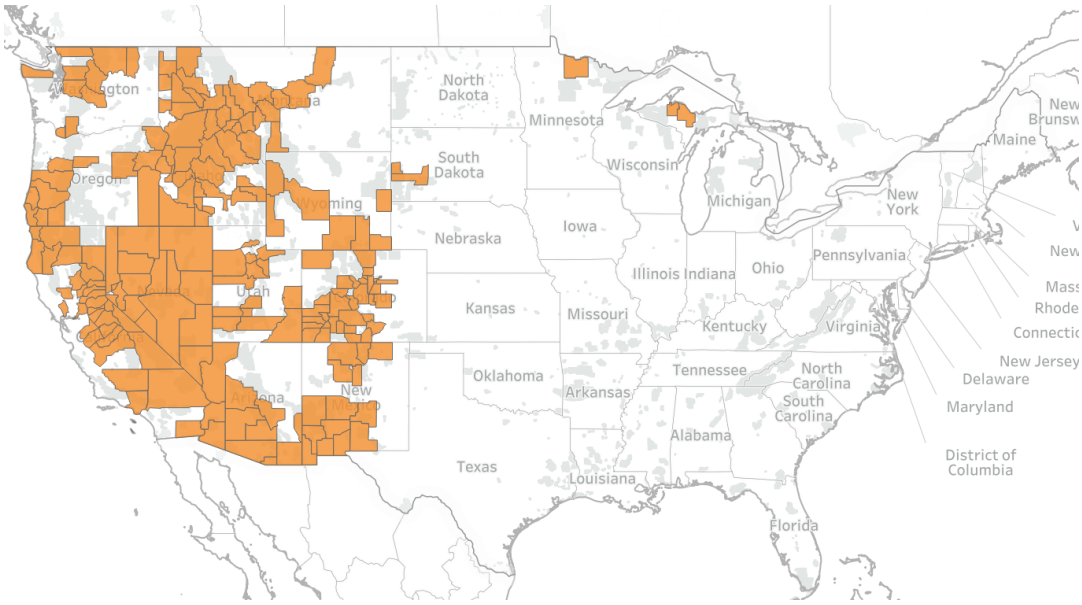
1.9 Conclusion

This chapter provides empirical evidence that the current geographic distribution of the LGBT population is related to the gold rushes during the mid-to-late nineteenth century. We rely on census estimates of the LGBT population to analyze the association between the number of unmarried partnered households of the same sex and gold discovery in the West during the gold rushes. We find that gold mining counties with at least one discovery during a gold rush have significantly more LGBT individuals nowadays than other mining counties. We then provide evidence that residents of gold rush counties are currently more likely to report pro-LGBT attitudes.

Our empirical findings have interesting policy implications. First, we provide evidence that a historical shock may have persistent effects on where people live, but also on attitudes. These findings suggest that economic policies or shocks such as migration of workers may shape social norms in the long-run. Second, our results point to a role for early social and religious institutions in shaping norms and attitudes (Cagé and Rueda, 2017; Nunn, 2010). That the initial institutional environment would affect norms and values is of particular importance for economic development (Bisin and Verdier, 2011).

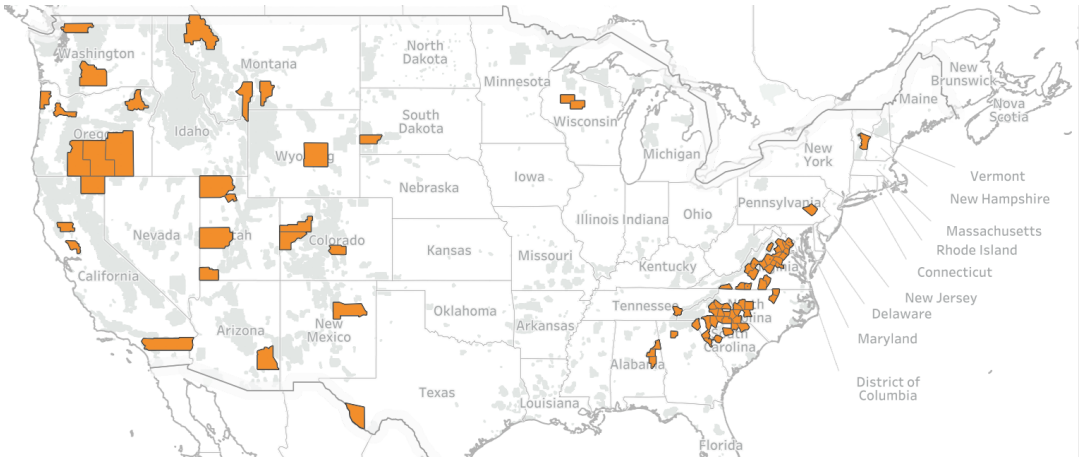
1.10 Figures

FIGURE 1.1: Gold Discoveries During the Gold Rushes



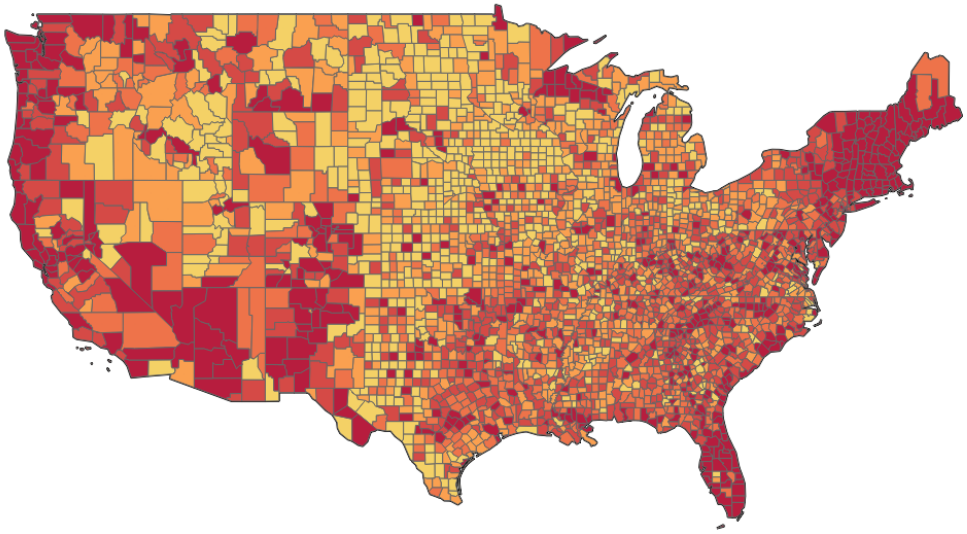
Notes: Based on data from the Mineral Resources Data System. Orange indicates that a county had at least one gold discovery during the gold rushes (1848–1899).

FIGURE 1.2: Gold Mining Counties *Without* Gold Rush Discoveries



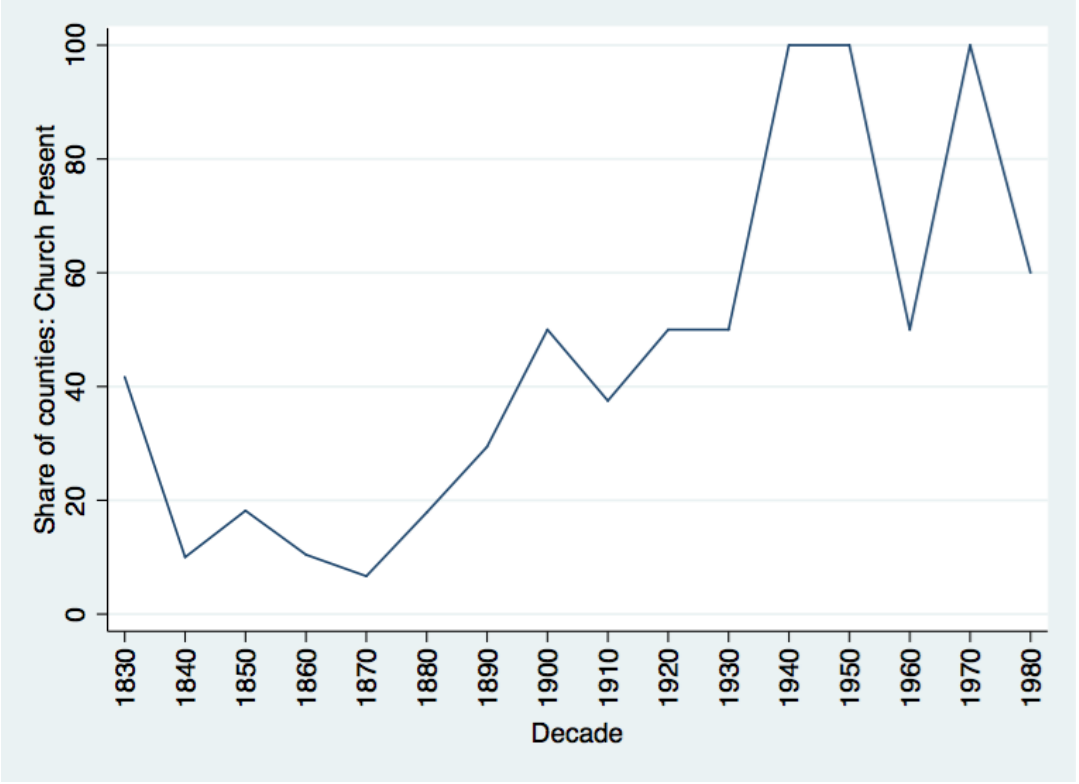
Notes: Based on data from the Mineral Resources Data System. Orange indicates that a county had at least one gold discovery before or after the gold rushes (1848–1899) and no gold discoveries during one of the gold rushes.

FIGURE 1.3: Unmarried Partner Households of the Same Sex



Notes: Unmarried Partner Households of the Same Sex per Capita in 2010. Light yellow indicates the lower quintile and dark red the top quintile.

FIGURE 1.4: Gold Mining Counties and Notable Place of Worship



Notes: Share of gold mining counties with at least one notable place of worship at the time of gold discovery by decade. Based on data from National Register of Historic Places.

1.11 Tables

TABLE 1.1: Descriptive Statistics at the Moment of Discovery

	Mean (1)	SD (2)	Counties (3)
Gold (Before Gold Rushes)			
Population	9,455	(5173.6)	33
Population Density	19.56	(9.835)	33
Sex Ratio	1.009	(0.027)	32
Prime Age Adult Share Among Whites	0.450	(0.030)	29
Immigrant Share	0.001	(0.001)	17
Illiteracy Among Whites	0.180	(0.104)	17
Gold Rushes			
Population	1,976	(5,625)	195
Population Density	1.21	(3.396)	195
Sex Ratio	6.937	(10.84)	188
Prime Age Adult Share Among Whites	0.810	(0.189)	52
Immigrant Share	0.269	(0.172)	188
Illiteracy Among Whites	0.096	(0.251)	22
Gold (After Gold Rushes)			
Population	38,214	(101,384)	43
Population Density	27.92	(1356)	43
Sex Ratio	1.147	(0.023)	34
Prime Age Adult Share Among Whites			
Immigrant Share	0.066	(0.07)	34
Illiteracy Among Whites			
Other Counties (1840–50)			
Population	5,719	(12,764)	2,811
Population Density	18.32	(260.20)	2,811
Sex Ratio	1.287	(1.653)	2,418
Prime Age Adult Share Among Whites	0.498	(0.075)	2,418
Immigrant Share	0.099	(0.145)	2,418
Illiteracy Among Whites	0.200	(0.173)	2,418
Other Minerals (During Gold Rushes)			
Population	4,320	(2,797)	11
Population Density	5.95	(5.98)	11
Sex Ratio	1.235	(0.333)	11
Prime Age Adult Share Among Whites	0.552	(0.126)	2
Immigrant Share	0.109	(0.138)	11
Illiteracy Among Whites			

Notes: Data from the Minnesota Population Center (2011). *Gold Rushes* are counties with at least one gold discovery during the gold rushes. *Gold (Before Gold Rushes)* are counties with at least one gold discovery before the gold rushes and no gold rush discoveries. *Gold (After Gold Rushes)* are counties with a first gold discovery after the gold rushes. *Other Counties* are counties that never had a gold discovery before, during and after the gold rushes. *Other Minerals (During Gold Rushes)* are counties with at least one discovery of either copper, iron, nickel or silver during the gold rushes and no gold discoveries. Population data is based on the first census before the year of earliest gold discovery for gold counties and for the year 1840 for *Other Counties*. The data on population density per square mile is based on the census year of the earliest gold discovery for gold counties and for the year 1840 for *Other Counties*. Sex ratio is measured for the first census after the year of earliest gold discovery for gold counties and for the year 1850 for *Other Counties*. Prime age adult (15 to 49 years old) share among whites data, illiteracy among whites over 20 years old and immigrant share is measured for the first census after the year of earliest discovery for gold counties and for the year 1950 for *Other Counties*.

TABLE 1.2: Gold Rushes and Notable Places of Worship

Probit	Place of Worship at the Moment of Gold Discovery				
	(1)	(2)	(3)	(4)	(5)
Gold Rushes	-0.364	-0.339	-0.381	-0.381	-0.333
	(0.076)	(0.079)	(0.083)	(0.097)	(0.119)
ln(population)	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	Yes	No	No
Census Division FE	No	No	No	Yes	No
State FE	No	No	No	No	Yes
Geographic Controls	No	Yes	Yes	Yes	Yes
Observations	294	294	292	283	279
Pseudo R-Squared	0.228	0.261	0.267	0.273	0.332

Notes: The unit of observation is a county. The sample is restricted to counties with at least one gold discovery at any time. The dependent variable is a dummy which equals one if a notable place of worship existed at the time of gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Marginal effects are reported. Standard errors clustered by state are reported between parentheses. Columns 1–5 include the (log of) 2010 county population. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.3: Baseline Specification: Main Estimates

	Ln(1 + Unmarried Partner Households Same Sex)					
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	0.106	0.017	0.008	-0.032	-0.068	-0.0434
	(0.047)	(0.027)	(0.027)	(0.021)	(0.019)	(0.014)
	[0.041]	[0.541]	[0.754]	[0.128]	[0.004]	[0.008]
Gold Rushes		0.127	0.170	0.133	0.148	0.140
		(0.049)	(0.047)	(0.047)	(0.049)	(0.050)
		[0.025]	[0.002]	[0.006]	[0.005]	[0.014]
ln(population)	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	Yes	No	No
Census Division FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Geographic Controls	No	No	Yes	Yes	Yes	Yes
Observations	3,111	3,111	3,111	3,111	3,111	3,111
Adjusted R-Squared	0.955	0.955	0.958	0.961	0.962	0.964

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap *p*-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). Columns 1–6 include the (log of) 2010 county population. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.4: Socioeconomic Characteristics

	Ln(1 + Unmarried Partner Households Same Sex)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Gold	-0.050	-0.043	-0.034	-0.058	-0.042	-0.052	-0.040
	(0.018)	(0.018)	(0.018)	(0.024)	(0.018)	(0.019)	(0.024)
	[0.035]	[0.058]	[0.103]	[0.068]	[0.065]	[0.034]	[0.163]
Gold Rushes	0.143	0.139	0.133	0.145	0.132	0.142	0.120
	(0.051)	(0.050)	(0.047)	(0.040)	(0.046)	(0.050)	(0.036)
	[0.015]	[0.014]	[0.015]	[0.003]	[0.013]	[0.012]	[0.006]
Additional Controls							
Demographic Mechanism:							
% Black or African American		Yes					Yes
Fractionalization Index		Yes					Yes
Economic Mechanism:							
Median Personal Income			Yes				Yes
Poverty Rate			Yes				Yes
Unemployment rate			Yes				Yes
Urban Mechanism:							
Median House Value				Yes			Yes
% House Less $100K				Yes			Yes
% House More $500K				Yes			Yes
% Urban				Yes			Yes
Population Density				Yes			Yes
Education Mechanism:							
% High School Graduate					Yes		Yes
Isolation Mechanism							
Distance to Lake						Yes	Yes
Distance to River						Yes	Yes
Portage Site						Yes	Yes
Access to Railroad						Yes	Yes
Average Agricultural Yield						Yes	Yes
ln(population)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,860	2,860	2,860	2,860	2,860	2,860	2,860
Adjusted R-Squared	0.966	0.966	0.967	0.968	0.967	0.966	0.969

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap *p*-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). Columns 1–7 include the (log of) 2010 county population. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties. We also control for other minerals (dummy equals one for counties with at least one discovery of either copper, iron, nickel or silver during the gold rushes and no gold discoveries), the share of foreign born and for whether the county is housing a U.S. military base.

TABLE 1.5: Attitudes Toward LGBT: GSS

	Sexual Relations Between Two Adults of the Same Sex: Not Wrong					
	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	0.219	-0.091	-0.179	-0.205	-0.107	-0.019
	(0.091)	(0.135)	(0.126)	(0.055)	(0.073)	(0.073)
Gold Rushes		0.411	0.409	0.353	0.198	0.124
		(0.141)	(0.129)	(0.075)	(0.078)	(0.062)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Individual Controls	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes
Observations	15,916	15,916	15,916	15,916	15,916	15,916

Notes: The unit of observation is a respondent. The period covered is 1993–2014. Attitudes toward LGBT is assessed through the following question: "What about sexual relations between two adults of the same sex–do you think it is" where respondents have four choices (4=not wrong at all, 3=wrong only sometimes, 2=almost always wrong and 1=always wrong). *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Columns 1–6 include year fixed effects. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties. The set of individual controls include the following variables: gender, age, age squared, six education dummies, three race dummies and five marital status dummies.

1.12 Appendix

Channels of Persistence and Transmission of Attitudes

A plausible channel of persistence is situational homosexuality as a result of elevated sex ratios and the presence of a large LGBT community. We think situational homosexuality is unlikely to be a channel of persistence because sex ratios have been balanced for more than 75 years in former gold rush counties and that both gay men and lesbians are currently living in these areas. The fact that we find estimates of similar magnitudes for gay men and lesbians tells us, arguably, that the lasting attitudinal effects were pro-LGBT rather than pro-gay.

LGBT Data

In this chapter, we examine a subset of the LGBT population: same-sex cohabiting (male and female) couples who describe their relationship as "unmarried partnership" and same-sex cohabiting (male and female) married couples. Census data at the county level is unfortunately not available for gay men and lesbians not in a cohabiting relationship (Black et al., 2000; Carpenter and Gates, 2008), nor bisexuals and transgenders not in a same-sex cohabiting relationship.[20] Nonetheless, we believe that census data is informative on the whole LGBT population due to the high correlation of this measure with other surveys. Using Census data and self-reported data from Gallup's Daily Tracking survey, Gates, 2015 shows that the geographic patterns of same-sex couples and adults self-identifying as LGBT are similar. For instance, three metropolitan statistical areas are in the top five of the 1990 Census same-sex couple rankings and the Gallup rankings. Moreover, Carpenter and Gates, 2008 compare individual level surveys (California's 2003 Lesbian, Gay, Bisexual, and Transgender Tobacco Survey and California Health Interview Survey (CHIS)) data and 2000 Census data. They find that the Census 2000 sample is very similar across a variety of demographic characteristics to both the CHIS and Tobacco Survey.[21]

The spatial distribution of LGBT in the Census and ACS is similar to the data compiled by Fisher, Gee, and Looney, 2018. In their paper, they provide estimates of the population of married same-sex tax filers for the years 2013–2014. They estimate that 0.35 percent of all joint filers were same-sex couples in 2014. Appendix Figure 1.7 illustrates the spatial distribution of married same-sex tax filers using their data. The data is available at the 3-digit ZIP code. Among the 100 largest commuting zones, San Francisco and Santa Rosa (California) had the highest rate of same-sex couples in married filing jointly returns in 2014.

Churches and Other Places of Worship

We use the National Register of Historic Places (NRHP) to create a database on notable cathedrals, churches and missions. NRHP provides information on the United States historic places deemed worthy of preservation. It categorizes five types of properties: district, site, structure, building, or object. This database includes data

[20] Carpenter and Gates, 2008 provide estimates on the number of gay men and lesbians not in a cohabiting relationship. They find that more than half of gay men and a third of lesbians were not cohabiting.

[21] Carpenter and Gates, 2008 document that gay and lesbian couples living in California are 39 years old on average. About 5 percent of gay and lesbian couples are Black, non-Hispanic, and 15 percent are Hispanic. Approximately 18 percent had a high school diploma or less.

on more than 1 million property out of which 80,000 are individually listed and the rest are contributing resources within historic districts.

Additional Robustness Checks

In Table 1.9 and Appendix Tables 1.10 and 1.11.[22] Table 1.9 presents estimates of Eq. (1.1). In columns 1–5, we include our full set of geographic controls and state fixed effects. In column 1, we add a variable for other mineral discoveries. This serves as a placebo treatment. The variable "Other Minerals" equals one for counties with at least one discovery of either copper, iron, nickel or silver during the gold rushes and no gold discoveries and zero otherwise. The point estimate for the variable "Other Minerals" is very small and statistically insignificant. Moreover, the estimated effect of the gold rushes is significantly larger than the effect for the variable "Other Minerals." This result is in line with the fact that other mineral discoveries did not lead to mass migration and occurred in many areas already inhabited (Table 1.1 and Appendix Figure 1.12).

Columns 2 and 3 include controls for the percentage of foreign born and a dummy for whether the county is housing a U.S. military base.[23] Arguably, military bases may lead to male biased sex ratios.[24] In column 4, we include these additional controls simultaneously. The point estimates are virtually the same as in Table 1.3, column 6, confirming the robustness of our results. Last, we test whether our estimates are similar if we keep only counties in the West, i.e., we drop counties in the other three Census regions. Restricting the sample to the West has no effect on our main conclusions.

In Appendix Table 1.10, we check whether defining the prevalence of LGBT individuals as a share of the population rather than as a level, conditional on population, affects our main conclusions. We present estimates of Eq. (1.1) in which the dependent variable is the natural log of unmarried partner households of the same sex per capita in 2010. The structure of the table is the same as Table 1.3. The estimates are all positive and of similar magnitudes than for the specifications using the size of the LGBT population as a level.

We show that our findings are also robust to alternative definitions for our main comparison group in Appendix Table 1.11. In column 1, we report our preferred estimates which are reported in Table 1.3, column 6. Our estimates are robust to the exclusion of gold discoveries before the gold rushes (column 2), during the years 1900–1950 (column 3) and since 1950 (column 4). For example, in column 2, *Gold* is now equal to zero for gold discoveries before the gold rushes. The estimates presented in these four columns are all positive and of similar magnitudes than in Table 1.3, column 6. Counties with at least one gold discovery which occurred during

[22] In Appendix Table 1.12, we report the standard errors using the estimation method developed by Conley, 1999; Conley, 2008 to take into account spatial correlation (Hsiang, 2010). In columns 1–2, standard errors are adjusted for spatial correlation within a 100-km radius. Columns 3–4 report standard errors adjusted for spatial correlation within a 300km radius. In columns 5–6, we rely on a 500-km radius. Standard errors are slightly smaller than for the corresponding columns (5–6) of Table 1.3.

[23] We collect information on U.S. military bases from this website `https://militarybases.com/`. The military base variable is equal to one if there is an air force, army, coast guard, marine corps, navy or joint operations military bases and zero otherwise. Including dummies for the different categories of military bases has no effect on our main conclusions. About seven percent of counties have at least one military base.

[24] We find no evidence that military bases increase the size of the LGBT population. The estimate is a well-estimated zero and statistically insignificant.

one of the gold rushes in the West currently have approximately 11–16 percent more unmarried partner households of the same sex.

We further document in Appendix Table 1.13 that our results are robust to controlling for non-gold rush mining counties with discoveries before and after the gold rushes separately. We generate an indicator "Gold (Before Gold Rushes)" that equals one for counties with at least one gold discovery before the gold rushes and zero otherwise. We generate a second indicator "Gold (After Gold Rushes)" equals one for counties with at least one gold discovery after the gold rushes and zero otherwise. Appendix Table 1.13 allows us to investigate whether counties that had discoveries before, during or after the gold rushes (1848–1899) have more unmarried partner households of the same sex in comparison to the rest of the U.S. (i.e., non mining counties). We find a positive and statistically significant effect and of similar magnitudes of the gold rushes as in Table 1.3. In contrast, the estimates for counties with gold discoveries before or after the gold rushes are smaller and statistically insignificant. The effects of the gold rushes are statistically significantly different from the estimates for counties with gold discoveries before and after the gold rushes. These results are in line with the fact that counties with gold discoveries before or after the gold rushes were already settled (Table 1.1).

We also check whether our attitudes results are robust to the omission of homosexual and bisexual respondents. The GSS includes a sexual orientation question since 2008. About 1.7 percent of respondents report being homosexual and 2.1 percent report being bisexual. Restricting the analysis to the time period 2008–2014 decreases our sample size from 15,916 respondents to only 5,200. Excluding LGBT respondents drops down our sample size to 5,011. Our estimates are of similar magnitudes than for the analysis on the whole sample (Appendix Table 1.5) and remain statistically significant in most specifications. (Estimates available upon request.) These findings confirm that non-LGBT residents are more likely to report pro-LGBT attitudes in gold rush counties.

We also check the robustness of our results to other questions concerning attitudes toward LGBT. We create a pro-LGBT index using answers to the question of whether sexual relations between two adults of the same sex is wrong or not and the following three questions: "A man admits to be homosexual: should he be allowed to make a speech in your community?", "A man admits to be homosexual: should he be allowed to teach in a college or university?" and "If some people in your community suggested that a book he wrote in favor of homosexuality, should it be taken out of your public library?" We code the variables as dummies. They are equal to one if respondents report "allowed/not removed" and zero if they report "not allowed/removed" to the last three questions. We also create a dummy for the question of whether sexual relations between two adults of the same sex is not wrong. The dummy variable is equal to one if the respondents report that it is "not wrong at all" and zero otherwise.

We use these four dummy variables to create the pro-LGBT index. The index is equal to the sum of the four dummies. For instance, the index takes the value zero if respondents reported that it is "always wrong," "almost always wrong" or "wrong only sometimes" to the first question and answered "not allowed/removed" to the last three questions. Approximately 11 percent of respondents have a score of zero. In contrast, the index takes the value four if respondents answered "not wrong at all" to the first question and "allowed/not removed" to the last three questions. About 31 percent of respondents have a value of four. The mean of this index is 2.7 (std. dev. 1.27). Our estimates are presented in Appendix Table 1.19. The estimates are all positive and of similar magnitudes than the estimates in Table 1.5.

Male and Female Same-Sex Unmarried Partners

In Appendix Table 1.23, we re-estimate Eq. (1.1), but for male and female same-sex unmarried partners, respectively. This allows us to check whether present-day attitudes in gold rush counties are more favorable towards gay men specifically or homosexuals in general and thus attracting both homosexual men and women. In column 1, the dependent variable is now the natural log of one plus the number of male unmarried partner households of the same sex in 2010. In contrast, the dependent variable is the natural log of one plus the number of female unmarried partner households of the same sex in column 4. We include state fixed effects and our set of geographic controls in all columns. The estimates for *Gold Rushes* are positive and of similar magnitudes for female and male unmarried partner households of the same sex (statistically significant at the 19% level for homosexual men and at the 1% level for homosexual women, respectively). In other words, the gold rushes did not have differential effects on gays and lesbians.[25] These findings provide suggestive evidence that the gold rushes left an ethos of homosexual tolerance, rather than of gay tolerance.

We further investigate whether the gold rushes had heterogeneous effects by demographic characteristics in columns 2–3 and 5–6. In columns 2 and 5, the dependent variables are the number of male and female unmarried partner households of the same sex with at least one child, respectively.[26] In columns 3 and 6, the dependent variables are the number of male and female unmarried partner households of the same sex without a child, respectively. As of 2010, about 25% of same-sex unmarried partners had at least one child in comparison to 41% for opposite sex unmarried partners. Female same-sex unmarried partners were more likely to have at least one child (31%) in comparison to male same-sex unmarried partners (20%). The estimates for *Gold Rushes* are positive and statistically significant for both male and female unmarried partner households of the same sex without a child. In contrast, the estimates are smaller, negative and statistically insignificant for same-sex unmarried partners with children.

Male-to-Female Sex Ratio

We empirically test the effect of the gold rushes on male-to-female ratios in Appendix Table 1.24 using the following specification:

$$MFRatio_{cd} = \alpha + \beta GoldRushes_{cd-10} + \delta_c + \gamma_d + \varepsilon_{cd} \tag{1.3}$$

where $MFRatio_{cd}$ is the male-to-female ratio in county c and decade d. $GoldRushes_{cd-10}$ is a dummy variable that assumes the value of one for county-decade observations with the earliest gold rush discovery during the 10 years preceding a decennial census and zero otherwise. Gold rush counties are counties with at least one gold discovery during one of the gold rushes (1848–1899). County and decade fixed effects are included in the model. β thus captures the average relation between the first

[25] The remoteness of gold rush counties may also explain the fact that the size of the estimate is not larger for gays than for lesbians. Black, Sanders, and Taylor, 2007 document that 90% of gay male partners live in urban areas in comparison to 85% of lesbian partners (75% of heterosexual couples). Including larger cities such as San Francisco, Oakland and Seattle increases the size of our estimates for gay male partners. These cities did not have gold rush discoveries, but were greatly affected by the gold rushes. Results are available upon requests.

[26] The U.S. Census identifies unmarried partner households of the same sex with related (own and not own) children under 18 years if at least one person is related to the householder by birth or adoption.

gold discovery and the male-to-female ratio in the next decennial census for gold rush counties. The time period for this analysis is 1820–2000. The results are reported in Appendix Table 1.24). The estimates in column 1 suggests that the male-to-female ratio increases by about 5.8 following a gold discovery during one of the gold rushes. Our estimate is statistically significant at the 1 percent level. The mean of the dependent variable is about 1.1 suggesting that there were 7 men for each women on average in gold rush counties. In columns 2 and 3, we also include in the model two variables capturing the medium- and long-term effect of the gold rushes on the male-to-female ratio. The estimates suggest that the male-to-female ratio was approximately 1.5 (1.0) 10 to 20 years (20 to 30 years) after the first gold rush discovery. These results suggest that unbalanced sex ratios in gold rush counties did not last.

As a comparison group, we rely on non-gold rush discoveries. We present the estimates in column 4. We find no evidence that non-gold rush discoveries had a positive impact in the short-, medium- and long-run on the male-to-female ratio in these localities. The estimates are all very small and negative and range from -0.01 to -0.07. These results are consistent with the idea that only gold rush discoveries led to a mass migration of men and that gold rush counties were virtually uninhabited at the moment of the first gold discovery. In other words, non-gold rush discoveries occurred in counties already populated by single men, women and families.

Our results are robust to alternative definitions of the variable *GoldRushes*. For example, we show that subsequent gold discoveries in gold rush counties, i.e., after the first gold discovery, also increased the male-to-female sex ratio, but to a lesser extent (see Appendix Table 1.25).

Neighboring a Gold Rush County

Appendix Table 1.26 explores another measure of intensity: neighboring a county with at least one gold discovery during the gold rushes.[27] In column 1, we show that counties neighboring a gold rush county do not have significantly more unmarried partner households of the same sex than other counties in the U.S. The coefficient is positive, but not statistically significant.

In columns 2–6, we interact our variable $GoldRushes_c$ with an indicator for whether the county is neighboring a gold rush county. The interaction is positive and statistically significant at conventional levels in most specifications. The estimates range from 0.10 to 0.30, suggesting that gold rush counties neighboring another gold rush county have approximately 10 to 30 percent more unmarried partner households of the same sex. These results suggest that the long-run effects of the gold rushes are somewhat local, but amplify when many neighboring counties had gold discoveries.

Appendix Figures and Tables

[27] We estimate the following specification:

$$Y_{cs} = \alpha + \beta GoldRushes_c + \gamma Neighbor_c + \mu GoldRushes \times Neighbor_c + X_c'\zeta + \delta_s + \varepsilon_{cs} \quad (1.4)$$

where Y_{cs} is the natural log of one plus the number of unmarried partner households of the same sex in county c in 2010. The variable $Neighbor_c$ is equal to one if the county neighbors a county for which $GoldRushes_c$ is equal to one.

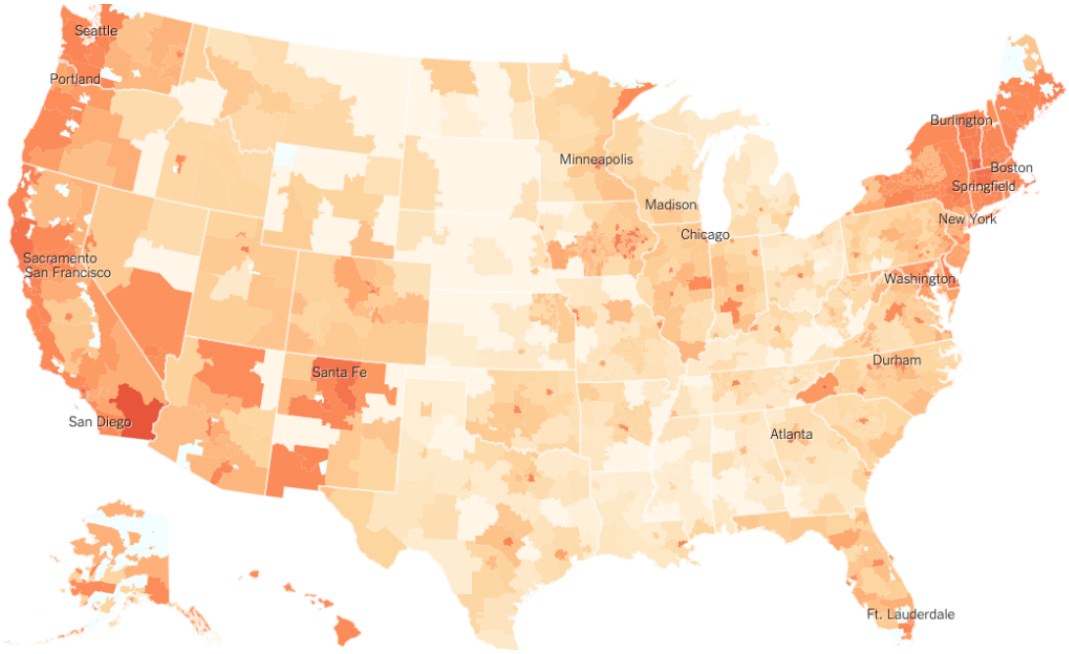
FIGURE 1.5: Imaginary Men Ball During the 1849 California Gold Rush

Source: 1891 etching by Andre Castaigne.

FIGURE 1.6: Men Dancing in the Old West

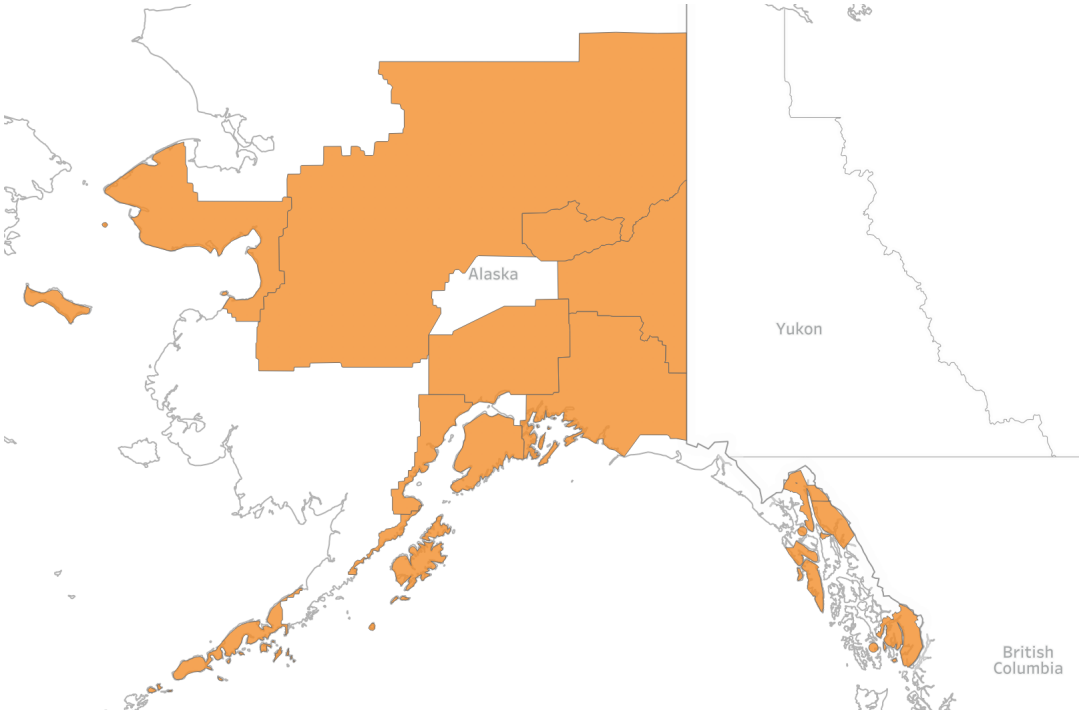
Source: True West Archive.

FIGURE 1.7: Married Couples who Have Filed Jointly



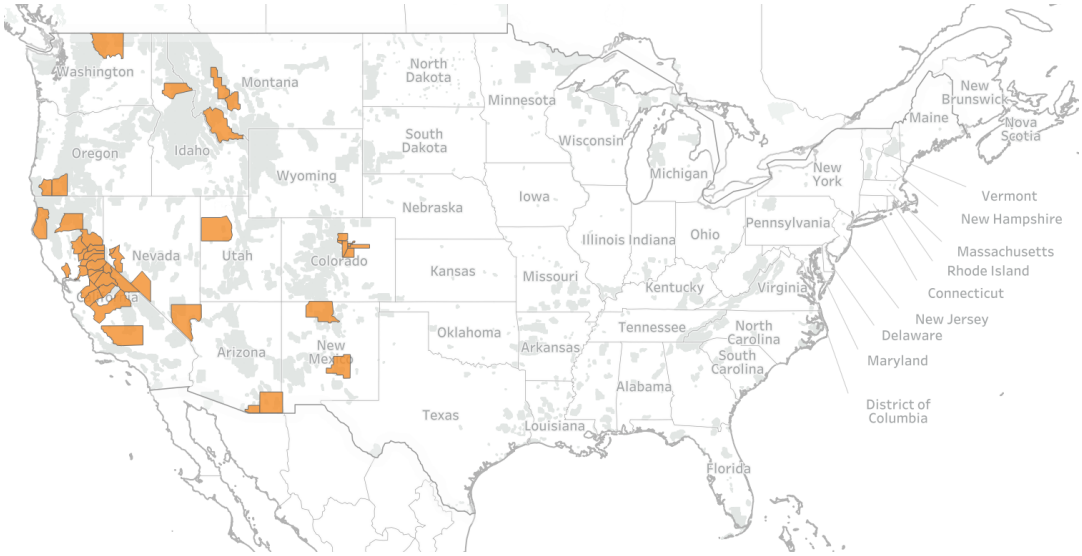
Notes: Same-sex marriage as a percentage of all marriages for married couples who have filed jointly on their tax return. Data presented at the three-digit zip code area. No data in areas where there are fewer than 500 married couples.

FIGURE 1.8: Gold Rush Discoveries in Alaska: 1848–1899



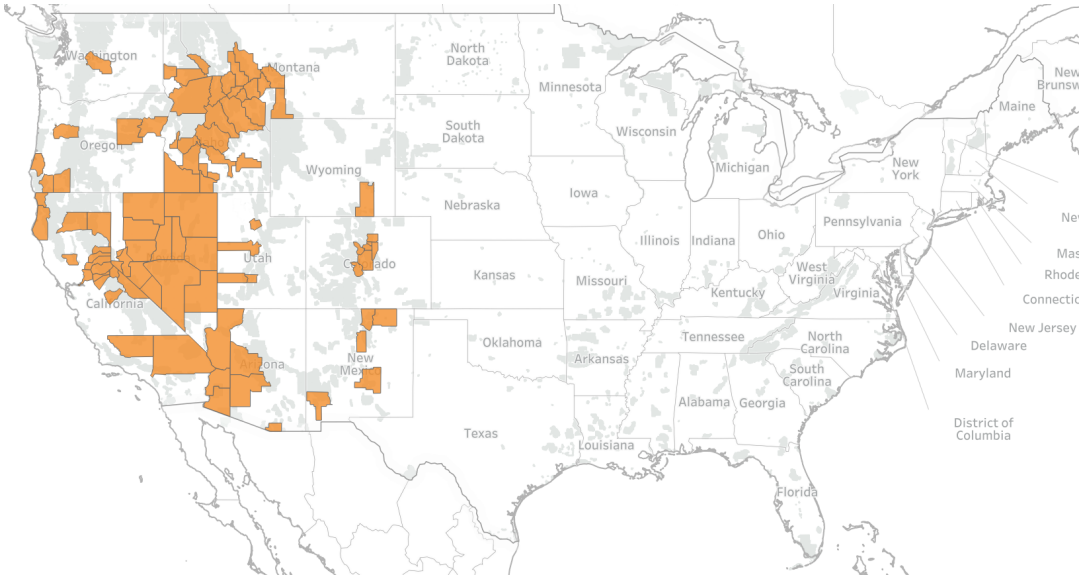
Notes: Based on data from the Mineral Resources Data System. Orange indicates gold discoveries during the gold rushes, only 1848–1899.

FIGURE 1.9: Gold Rush Discoveries: 1848–1859



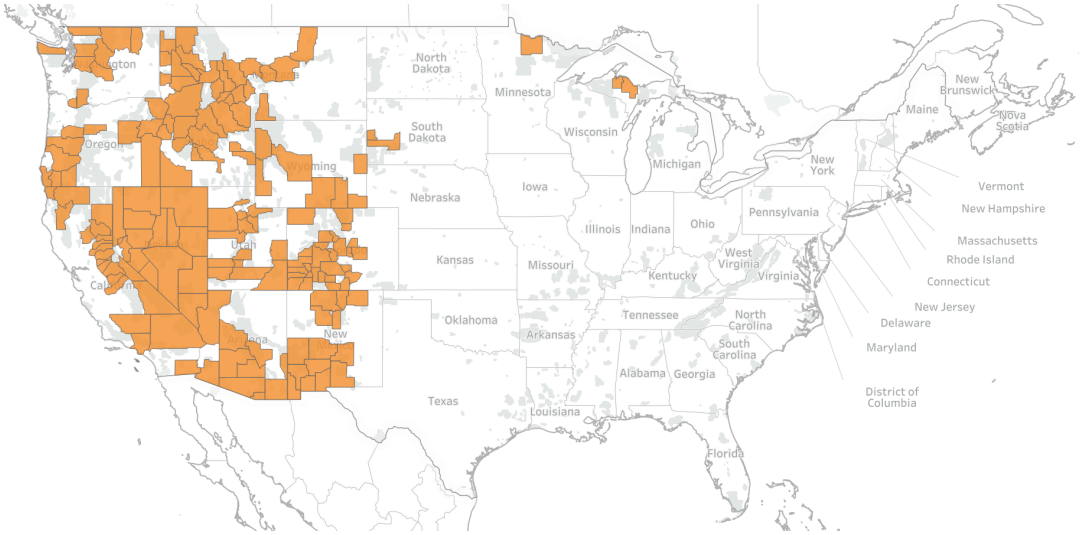
Notes: Based on data from the Mineral Resources Data System. Orange indicates gold discoveries during the gold rushes, only 1848–1859.

FIGURE 1.10: Gold Rush Discoveries: 1860–1869



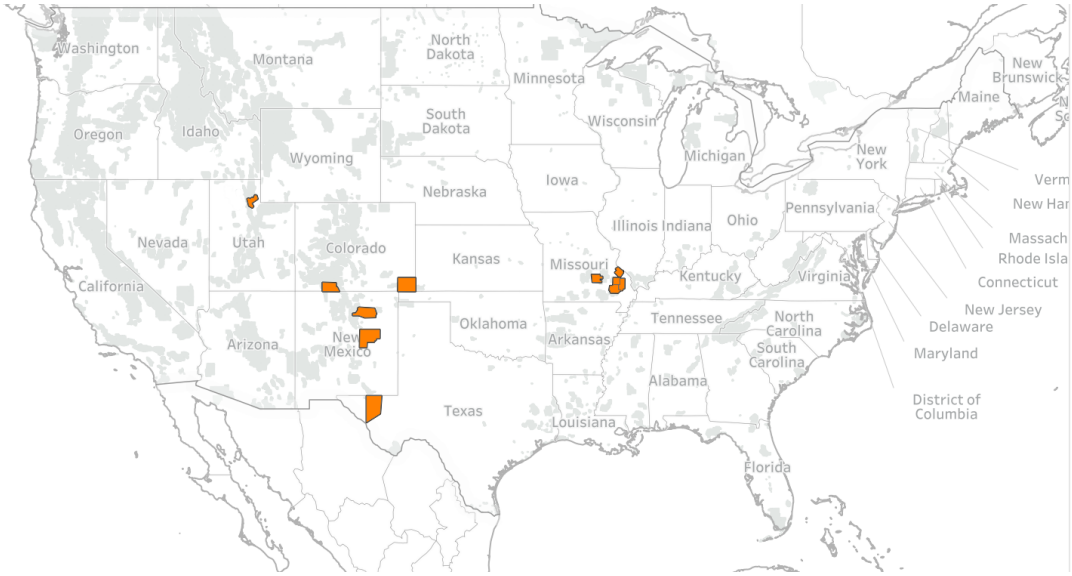
Notes: Based on data from the Mineral Resources Data System. Orange indicates gold discoveries during the gold rushes, only 1860–1869.

FIGURE 1.11: Gold Rush Discoveries: 1870–1899



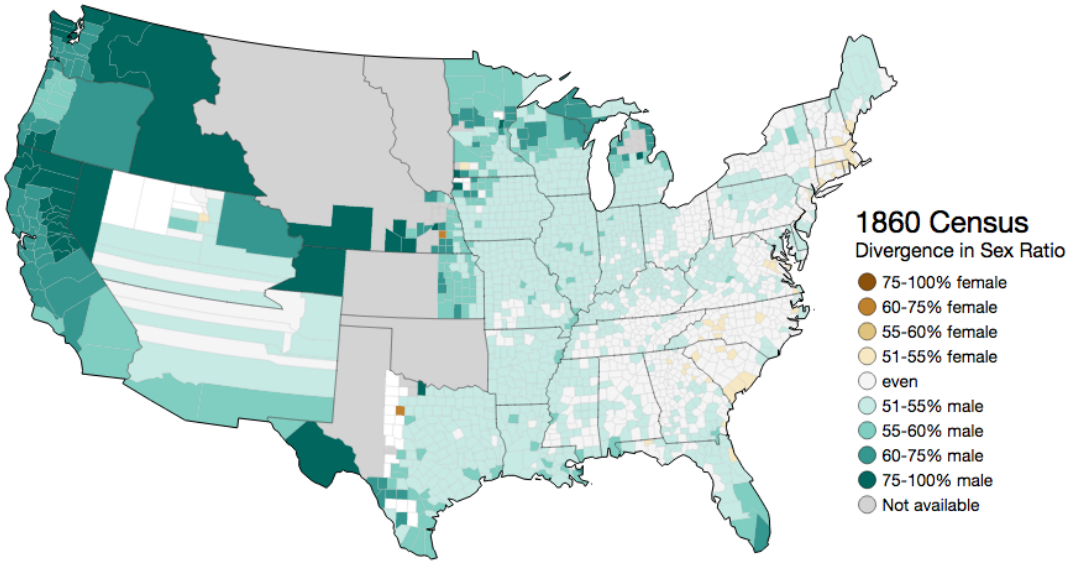
Notes: Based on data from the Mineral Resources Data System. Orange indicates gold discoveries during the gold rushes, only 1870–1899.

FIGURE 1.12: Other Non-Gold Mining Counties during the Gold Rushes



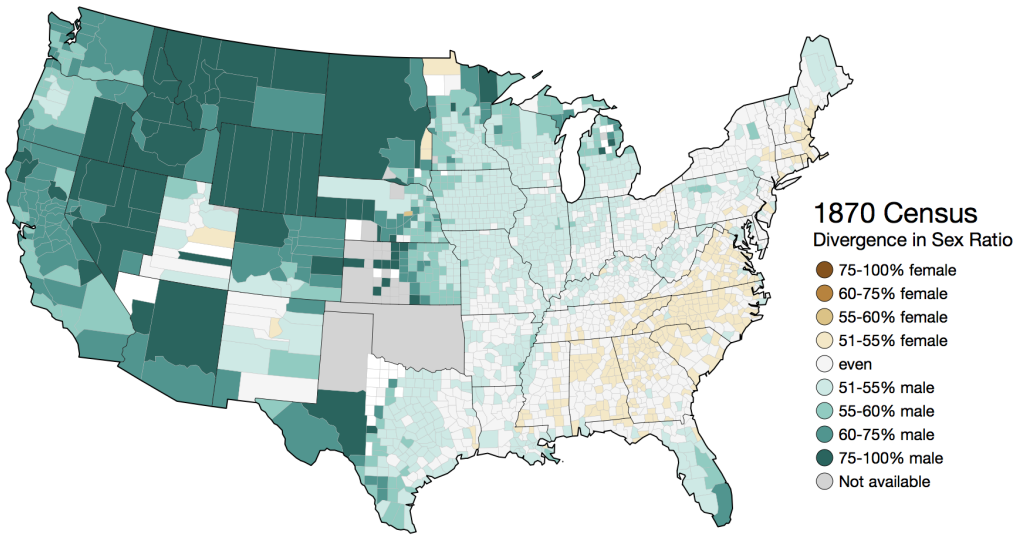
Notes: Based on data from the Mineral Resources Data System. Orange indicates counties with at least one discovery of either copper, iron, nickel or silver during the gold rushes (1848–1899) and no gold discoveries.

FIGURE 1.13: Male-to-Female Ratio in 1860



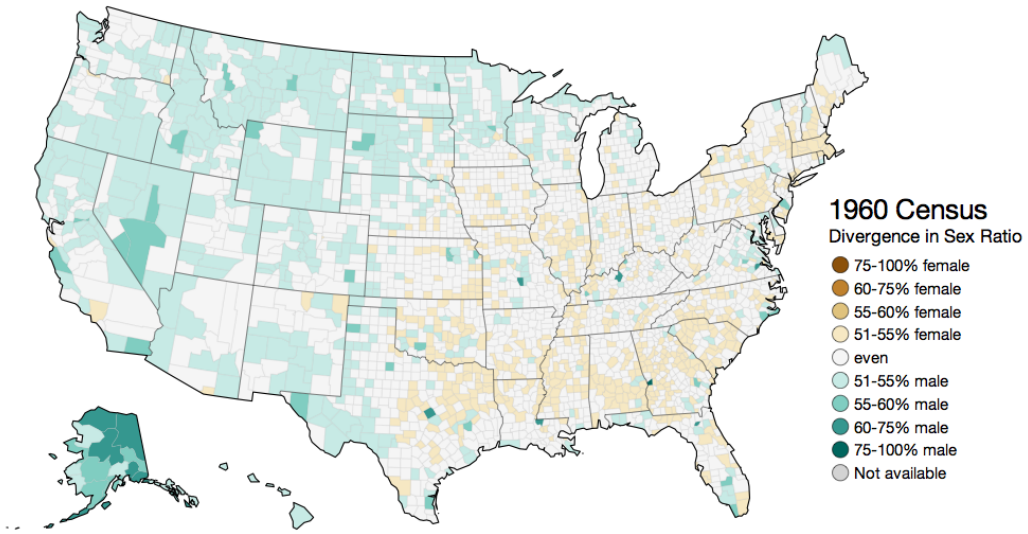
Notes: Figure from Mullen, 2018. Data from the NHGIS. Male-to-female ratio in 1860.

FIGURE 1.14: Male-to-Female Ratio in 1870



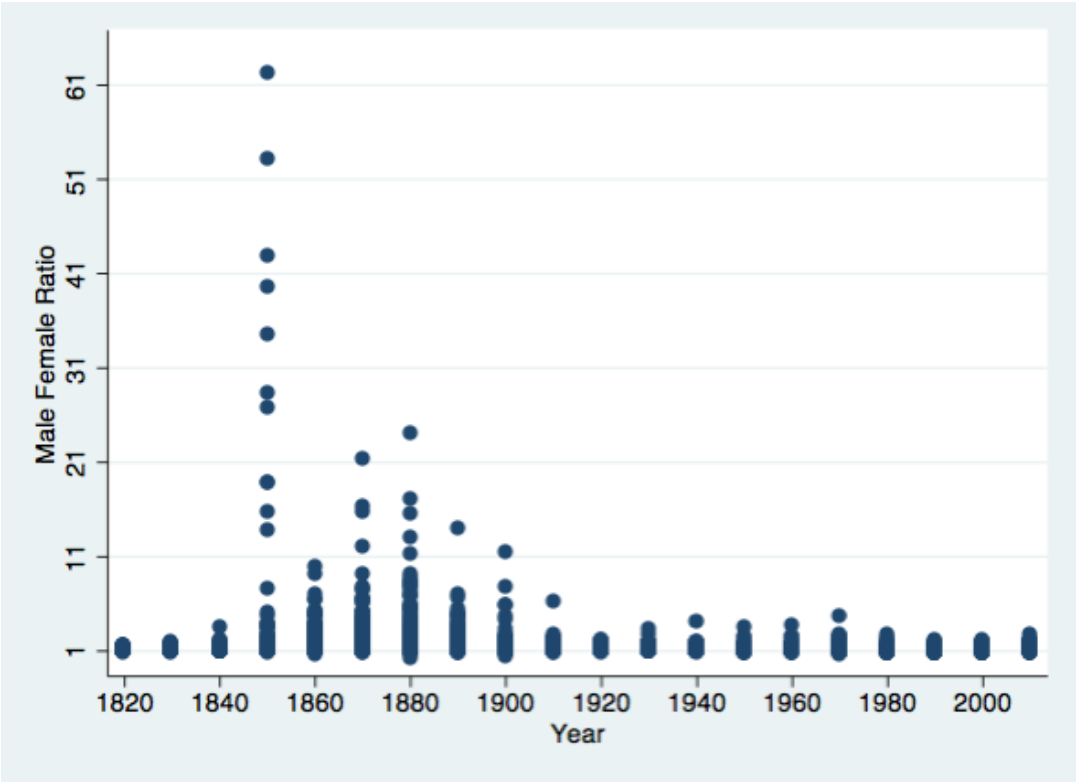
Notes: Figure from Mullen, 2018. Data from the NHGIS. Male-to-female ratio in 1870.

FIGURE 1.15: Male-to-Female Ratio in 1960



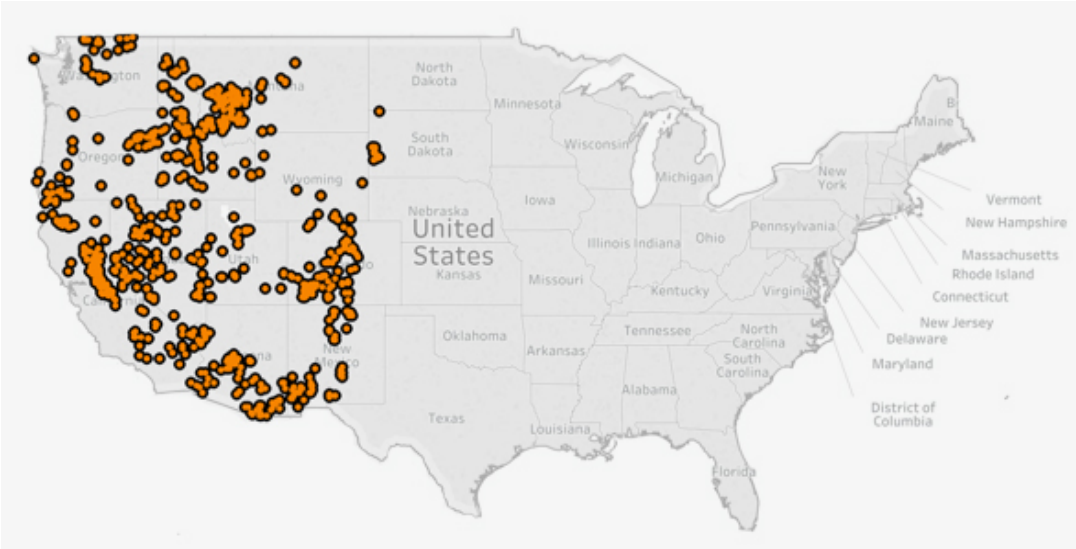
Notes: Figure from Mullen, 2018. Data from the NHGIS. Male-to-female ratio in 1960.

FIGURE 1.16: Male-to-Female Ratio



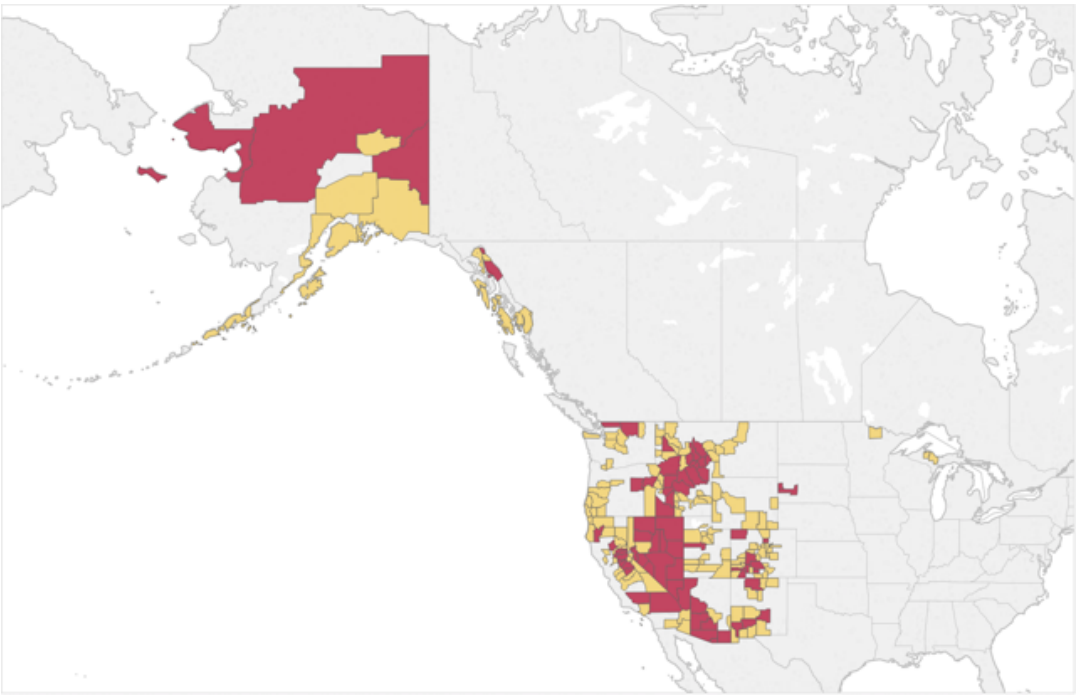
Notes: Male-to-Female Ratio across all counties.

FIGURE 1.17: Gold Discoveries During the Gold Rushes: Intensity



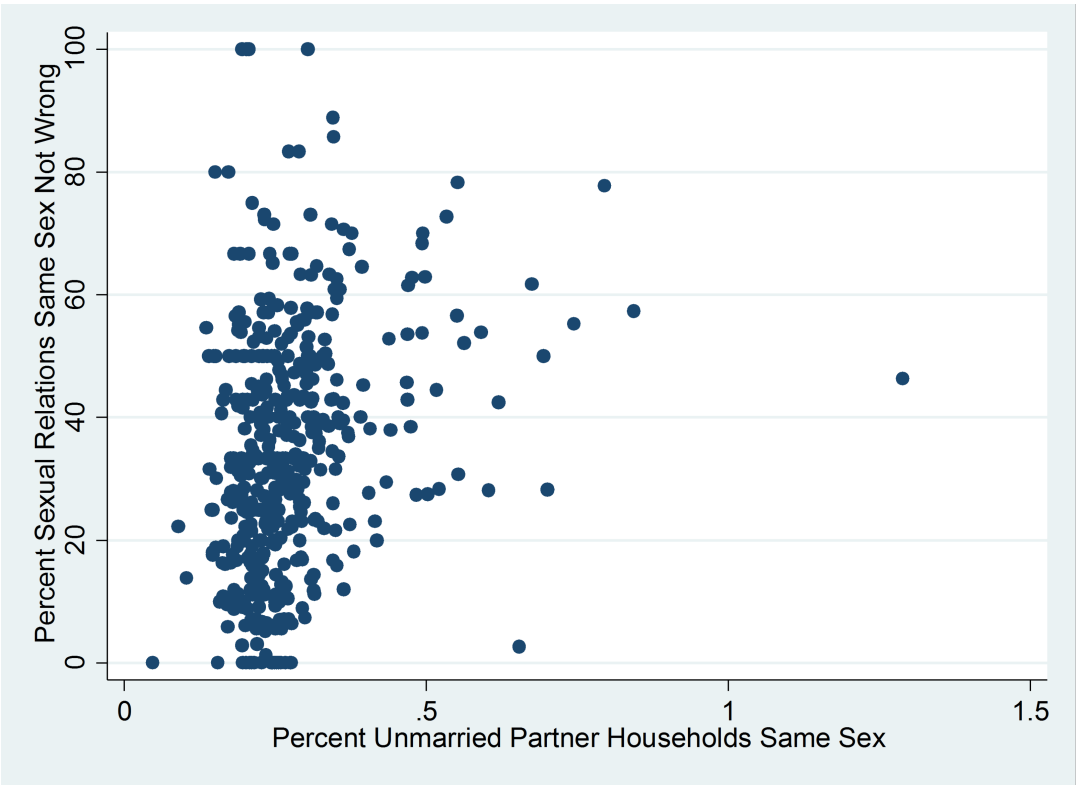
Notes: Based on data from the Mineral Resources Data System. Dots indicate gold discoveries during the gold rushes (1848–1899).

FIGURE 1.18: Gold Discoveries During the Gold Rushes: Intensity



Notes: Based on data from the Mineral Resources Data System. There were eight gold rush discoveries on average for counties with at least one gold rush discovery. Red indicates counties with at least eight gold rush discoveries (1848–1899). Yellow indicates counties with less than eight gold rush counties. Non-gold rush mining counties are in gray.

FIGURE 1.19: Sexual Relations Between Same Sex Adults Not Wrong at All



Notes: Based on data from the General Social Survey. Relationship between unmarried partner households same sex per capita and the percentage of respondents reporting "Not wrong at all" to the following question: "What about sexual relations between two adults of the same sex–do you think it is always wrong, almost always wrong, wrong only sometimes, or not wrong at all?" The unit of observation is a county.

TABLE 1.6: Predict Gold Rush Counties

	Gold Rush County Compare to Other Gold Counties	
	(1)	(2)
Latitude	-0.004	-0.003
	(0.005)	(0.005)
Longitude	-0.004	-0.004
	(0.003)	(0.004)
ln(Total Land Area)	-0.012	-0.027
	(0.024)	(0.023)
Mean of Elevation	0.0000	0.0000
	(0.000)	(0.000)
Standard Deviation of Elevation	-0.0003	-0.0006
	(0.000)	(0.000)
Distance State Capital	-0.0001	-0.0000
	(0.000)	(0.000)
Coastal County	0.079	0.061
	(0.068)	(0.071)
% Foreign Born	0.002	0.0006
	(0.003)	(0.006)
Military Base	0.061	0.048
	(0.050)	(0.062)
% Black or African American		-0.406
		(0.294)
Fractionalization Index		0.344
		(0.275)
ln(Median Personal Income)		0.170
		(0.098)
Poverty Rate		0.0063
		(0.007)
Unemployment rate		0.0107
		(0.006)
ln(Median House Value)		0.002
		(0.096)
% Urban		-0.0022
		(0.001)
% High School Graduate		0.0069
		(0.010)
ln(Population)	-0.019	-0.158
	(0.012)	(0.105)
Census Division FE	Yes	Yes
Observations	294	286
Adjusted R-Squared	0.560	0.539

Notes: The unit of observation is a county. The dependent variable is a dummy for whether the county had at least one gold discovery during the gold rushes. The sample is restricted to counties with at least one gold discovery. Standard errors clustered by state are reported between parentheses.

TABLE 1.7: Gold Rushes and Notable Places of Worship: Robustness Checks

Probit	Place of Worship at the Moment of Gold Discovery			
	(1)	(2)	(3)	(4)
Gold Rushes	-0.331	-0.334	-0.332	-0.331
	(0.118)	(0.130)	(0.129)	(0.116)
ln(population)	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
% of Foreign Born	Yes	No	Yes	No
Military Base	No	Yes	Yes	No
Observations	267	267	267	216
Pseudo R-Squared	0.333	0.332	0.333	0.299

Notes: The unit of observation is a county. The sample is restricted to counties with at least one gold discovery at any time. The dependent variable is a dummy which equals one if a notable place of worship existed at the time of gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Marginal effects are reported. Standard errors clustered by state are reported between parentheses. Columns 1–4 include the (log of) 2010 county population and the following geographic controls: latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties. In column 1, we add a control for the share of foreign born. Column 2 includes a dummy for whether the county is housing a U.S. military base. Column 3 adds indicators for military basis and for the share of foreign born simultaneously. Column 4 restricts the sample to the Western region.

TABLE 1.8: Intensive Margin: Number of Gold Rushes Discoveries

	Ln(1 + Unmarried Partner Households Same Sex)					
	(1)	(2)	(3)	(4)	(5)	(6)
Num. Gold Rushes	0.0082	0.0084	0.0076	0.0056	0.0048	0.0044
β	(0.0027)	(0.0028)	(0.0023)	(0.0022)	(0.0019)	(0.0021)
	[0.041]	[0.067]	[0.056]	[0.108]	[0.093]	[0.186]
Num. Gold (Not Gold		-0.0003	-0.0007	-0.0011	-0.0011	-0.0011
Rushes) Discoveries		(0.0012)	(0.0013)	(0.0010)	(0.0008)	(0.0015)
λ		[0.751]	[0.573]	[0.488]	[0.397]	[0.570]
ln(population)	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	Yes	No	No
Census Division FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Geographic Controls	No	No	Yes	Yes	Yes	Yes
$P(\beta \neq \lambda)$		0.011	0.003	0.012	0.007	0.024
Observations	3,111	3,111	3,111	3,111	3,111	3,111
Adjusted R-Squared	0.955	0.955	0.958	0.961	0.962	0.964

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. *Num. Gold Rushes* is equal to the number of gold rush discoveries. *Num. Gold (Not Gold Rushes)* is equal to the number of gold discoveries before or after the gold rushes. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap p-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). Columns 1–6 include the (log of) 2010 county population. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.9: Robustness Checks

	Ln(1 + Unmarried Partner Households Same Sex)				
	(1)	(2)	(3)	(4)	(5)
Gold	-0.043	-0.043	-0.043	-0.042	-0.022
	(0.014)	(0.014)	(0.014)	(0.014)	(0.043)
	[0.010]	[0.008]	[0.008]	[0.011]	[0.576]
Gold Rushes	0.140	0.140	0.140	0.139	0.123
	(0.050)	(0.049)	(0.049)	(0.049)	(0.044)
	[0.016]	[0.016]	[0.014]	[0.014]	[0.006]
Other Minerals	0.002			0.002	
	(0.054)			(0.054)	
ln(population)	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes
% of Foreign Born	No	Yes	No	Yes	No
Military Base	No	No	Yes	Yes	No
Observations	3,111	3,111	3,111	3,111	434
Adjusted R-Squared	0.964	0.964	0.964	0.964	0.963

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap *p*-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). Columns 1–5 include the (log of) 2010 county population and our set of geographic controls (see Table 1.3). In columns 1 and 4, we add a control for other minerals (dummy equals one for counties with at least one discovery of either copper, iron, nickel or silver during the gold rushes and no gold discoveries). Columns 2 and 4 include a control for the share of foreign born. Columns 3 and 4 include a dummy for whether the county is housing a U.S. military base. Column 5 restricts the sample to the Western region.

TABLE 1.10: Baseline Specification: Unmarried Partner Households Same Sex per Capita

	Ln(1 + Unmarried Partner Households Same Sex per Capita)					
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	0.122	0.056	0.036	-0.014	-0.049	-0.023
	(0.048)	(0.030)	(0.030)	(0.021)	(0.019)	(0.015)
	[0.030]	[0.109]	[0.234]	[0.504]	[0.016]	[0.096]
Gold Rushes		0.093	0.183	0.134	0.147	0.137
		(0.052)	(0.051)	(0.047)	(0.049)	(0.051)
		[0.118]	[0.001]	[0.005]	[0.009]	[0.020]
Census Region FE	No	No	No	Yes	No	No
Census Division FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Geographic Controls	No	No	Yes	Yes	Yes	Yes
Observations	3,111	3,111	3,111	3,111	3,111	3,111
Adjusted R-Squared	0.010	0.011	0.115	0.187	0.215	0.252

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus unmarried partner households of the same sex per capita in 2010. *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap *p*-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.11: Robustness Checks: Definition of the Control Group

	Ln(1 + Unmarried Partner Households Same Sex)			
	Baseline Specification	Exclude Discoveries Pre-1848	Exclude Discoveries 1900-1950	Exclude Discoveries Since 1950
	(1)	(2)	(3)	(4)
Gold	-0.043	-0.041	-0.007	-0.061
	(0.014)	(0.025)	(0.023)	(0.017)
	[0.008]	[0.073]	[0.746]	[0.005]
Gold Rushes	0.140	0.139	0.111	0.157
	(0.050)	(0.050)	(0.057)	(0.046)
	[0.014]	[0.019]	[0.116]	[0.003]
ln(population)	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Observations	3,111	3,111	3,111	3,111
Adjusted R-Squared	0.964	0.964	0.964	0.964

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. *Gold* equals one for counties with at least one gold discovery and zero otherwise. In column 2, *Gold* equals zero for gold discoveries before the gold rushes. In column 3, *Gold* equals zero for gold discoveries during the years 1900–1950. In column 4, *Gold* equals zero for gold discoveries since 1950. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap *p*-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). Columns 1–4 include the (log of) 2010 county population and state fixed effects. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.12: Conley's 100-, 300-, and 500-KM Spatial Adjustments

	Ln(1 + Unmarried Partner Households Same Sex)					
	Spatial Adjust. 100-km	Spatial Adjust. 100-km	Spatial Adjust. 300-km	Spatial Adjust. 300-km	Spatial Adjust. 500-km	Spatial Adjust. 500-km
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	-0.068	-0.0434	-0.068	-0.0434	-0.068	-0.0434
	(0.032)	(0.035)	(0.025)	(0.025)	(0.013)	(0.019)
Gold Rushes	0.148	0.140	0.148	0.140	0.148	0.140
	(0.043)	(0.044)	(0.047)	(0.046)	(0.035)	(0.045)
ln(population)	Yes	Yes	Yes	Yes	Yes	Yes
Census Division FE	Yes	No	Yes	No	Yes	No
State FE	No	Yes	No	Yes	No	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,111	3,111	3,111	3,111	3,111	3,111

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors are computed following Conley, 1999; Conley, 2008 and Hsiang, 2010 and using 100-, 300-, and 500-km spatial adjustments. Columns 1–6 include the (log of) 2010 county population. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.13: Before, During and After Gold Rushes

	Ln(1 + Unmarried Partner Households Same Sex)					
	(1)	(2)	(3)	(4)	(5)	(6)
Gold Rushes	0.143	0.149	0.176	0.123	0.120	0.128
β	(0.056)	(0.062)	(0.044)	(0.047)	(0.050)	(0.050)
	[0.030]	[0.048]	[0.000]	[0.008]	[0.039]	[0.032]
Gold (Before Gold Rushes)		0.071	0.012	0.028	0.019	0.024
λ_1		(0.052)	(0.055)	(0.055)	(0.054)	(0.047)
		[0.215]	[0.837]	[0.634]	[0.754]	[0.618]
Gold (After Gold Rushes)		-0.011	0.001	-0.034	-0.061	-0.055
λ_2		(0.034)	(0.029)	(0.029)	(0.027)	(0.025)
		[0.747]	[0.945]	[0.258]	[0.042]	[0.038]
ln(population)	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	Yes	No	No
Census Division FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Geographic Controls	No	No	Yes	Yes	Yes	Yes
$P(\beta \neq \lambda_1)$		0.133	0.015	0.083	0.057	0.029
$P(\beta \neq \lambda_2)$		0.063	0.001	0.012	0.005	0.001
Observations	3,111	3,111	3,111	3,111	3,111	3,111
Adjusted R-Squared	0.955	0.955	0.958	0.961	0.962	0.964

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. *Gold (Before Gold Rushes)* equals one for counties with at least one gold discovery before the gold rushes and zero otherwise. *Gold (After Gold Rushes)* equals one for counties with at least one gold discovery after the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap *p*-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). Columns 1–6 include the (log of) 2010 county population. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.14: Same-Sex Married Couples

	Ln(1 + Same-Sex Married Couples)					
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	0.190	-0.015	0.022	-0.048	-0.131	-0.021
	(0.080)	(0.095)	(0.066)	(0.081)	(0.078)	(0.058)
	[0.053]	[0.885]	[0.737]	[0.584]	[0.116]	[0.721]
Gold Rushes		0.268	0.198	0.013	0.107	0.028
		(0.089)	(0.094)	(0.081)	(0.077)	(0.053)
		[0.010]	[0.020]	[0.879]	[0.197]	[0.574]
ln(population)	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	Yes	No	No
Census Division FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Geographic Controls	No	No	Yes	Yes	Yes	Yes
Observations	2,323	2,323	2,323	2,323	2,323	2,323
Adjusted R-Squared	0.107	0.110	0.155	0.190	0.211	0.184

Notes: The unit of observation is a PUMA. The dependent variable is the natural log of one plus the number of same-sex *married* couples. The time period is 2013–2017. *Gold Rushes* equals one for PUMAs with at least one gold discovery during the gold rushes and zero otherwise. *Gold* equals one for PUMAs with at least one gold discovery and zero otherwise. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap *p*-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). Columns 1–6 include the (log of) 2010 county population. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.15: Same-Sex Married Couples per Capita

	Ln(1 + Same-Sex Married Couples Per Capita)					
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	0.194	0.004	0.036	-0.038	-0.123	-0.022
	(0.080)	(0.101)	(0.068)	(0.083)	(0.080)	(0.062)
	[0.055]	[0.969]	[0.618]	[0.679]	[0.155]	[0.720]
Gold Rushes		0.248	0.180	-0.004	0.095	0.022
		(0.094)	(0.094)	(0.082)	(0.078)	(0.054)
		[0.033]	[0.034]	[0.964]	[0.253]	[0.670]
ln(population)	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	Yes	No	No
Census Division FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Geographic Controls	No	No	Yes	Yes	Yes	Yes
Observations	2,323	2,323	2,323	2,323	2,323	2,323
Adjusted R-Squared	0.007	0.010	0.063	0.102	0.125	0.184

Notes: The unit of observation is a PUMA. The dependent variable is the natural log of one plus the number of same-sex *married* couples per capita. The time period is 2013–2017. *Gold Rushes* equals one for PUMAs with at least one gold discovery during the gold rushes and zero otherwise. *Gold* equals one for PUMAs with at least one gold discovery and zero otherwise. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap *p*-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). Columns 1–6 include the (log of) 2010 county population. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.16: Socioeconomic Covariates

	Ln(1 + Unmarried Partner HH Same Sex)				
	(1)	(2)	(3)	(4)	(5)
% Black or African American	0.020				
	(0.086)				
Fractionalization Index	0.195				
	(0.122)				
ln(Median Personal Income)		0.349			
		(0.059)			
Poverty Rate		0.006			
		(0.012)			
Unemployment Rate		0.002			
		(0.005)			
ln(Median House Value)			-0.240		
			(0.122)		
% House Value Less $100K			-0.006		
			(0.002)		
% House Value More $500K			0.010		
			(0.002)		
% Urban			-0.001		
			(0.000)		
Population Density			Omitted		
% High School Graduate				0.010	
				(0.002)	
Distance to River					-3.75e-08
					(2.62e-07)
Distance to Lake					Omitted
Portage Site					0.022
					(0.035)
Access to Railroad					0.0003
					(0.0005)
Average Agricultural Yield					0.317
					(0.186)
ln(population)	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Other Minerals	Yes	Yes	Yes	Yes	Yes
% of Foreign Born	Yes	Yes	Yes	Yes	Yes
Military Base	Yes	Yes	Yes	Yes	Yes
Observations	2,860	2,860	2,860	2,860	2,860

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. Columns 1–5 include the (log of) 2010 county population, state fixed effects and our set of geographic controls. We also control for other minerals (dummy equals one for counties with at least one discovery of either copper, iron, nickel or silver during the gold rushes and no gold discoveries), the share of foreign born and for whether the county is housing a U.S. military base.

TABLE 1.17: Attitudes Toward LGBT: Socioeconomic Covariates

	Sex Between Two Adults of the Same Sex		
	(1)	(2)	(3)
Male	-0.209	-0.211	-0.212
	(0.022)	(0.022)	(0.022)
Age	0.005	0.003	0.004
	(0.004)	(0.004)	(0.004)
Age Squared	-0.015	-0.014	-0.016
	(0.004)	(0.004)	(0.004)
Elementary School	-0.992	-0.963	-0.945
	(0.062)	(0.060)	(0.067)
Attended High School	-0.846	-0.822	-0.799
	(0.041)	(0.041)	(0.041)
Grad. High School	-0.663	-0.653	-0.637
	(0.042)	(0.042)	(0.039)
Attended College	-0.446	-0.440	-0.430
	(0.041)	(0.040)	(0.039)
College	-0.181	-0.180	-0.179
	(0.043)	(0.043)	(0.044)
Post Graduate	Omitted	Omitted	Omitted
White	0.170	0.216	0.245
	(0.057)	(0.049)	(0.050)
Black or African American	-0.383	-0.289	-0.262
	(0.057)	(0.056)	(0.058)
Other Race	Omitted	Omitted	Omitted
Married	Omitted	Omitted	Omitted
Widow	0.097	0.087	0.095
	(0.039)	(0.039)	(0.040)
Separated	0.297	0.284	0.271
	(0.061)	(0.063)	(0.060)
Divorced	0.293	0.304	0.304
	(0.035)	(0.036)	(0.036)
Never Married	0.489	0.466	0.457
	(0.029)	(0.029)	(0.029)
Census Division FE		Yes	
State FE			Yes
Geographic Controls	Yes	Yes	Yes
Observations	16,084	16,084	16,084
Pseudo R-Squared	0.082	0.090	0.098

Notes: The unit of observation is a respondent. The period covered is 1993–2014. Attitudes toward LGBT is assessed through the following question: "What about sexual relations between two adults of the same sex–do you think it is" where respondents have four choices (4=not wrong at all, 3=wrong only sometimes, 2=almost always wrong and 1=always wrong). Standard errors clustered by state are reported between parentheses. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

TABLE 1.18: Attitudes Toward LGBT: Geographic Mobility

	Sexual Relations Between Two Adults of the Same Sex: Not Wrong					
	Same City Since 16			Moved Since 16		
	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	-0.237	-0.300	0.053	-0.054	-0.174	-0.037
	(0.291)	(0.171)	(0.112)	(0.075)	(0.062)	(0.082)
Gold Rushes	0.666	0.477	0.101	0.308	0.311	0.135
	(0.307)	(0.166)	(0.110)	(0.085)	(0.081)	(0.069)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	No	Yes	No	No	Yes
Individual Controls	No	Yes	Yes	No	Yes	Yes
Geographic Controls	No	Yes	Yes	No	Yes	Yes
Observations	6,046	6,046	6,046	9,870	9,870	9,870

Notes: The unit of observation is a respondent. The period covered is 1993–2014. Attitudes toward LGBT is assessed through the following question: "What about sexual relations between two adults of the same sex–do you think it is" where respondents have four choices (4=not wrong at all, 3=wrong only sometimes, 2=almost always wrong and 1=always wrong). Columns 1–3 restrict the sample to respondents who are living in the same city since age 16. Columns 4–6 restrict the sample to individuals who moved to a different city or state since the age of 16. *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Columns 1–6 include year fixed effects. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties. The set of individual controls include the following variables: gender, age, age squared, six education dummies, three race dummies and five marital status dummies.

TABLE 1.19: Attitudes Toward LGBT: Pro-LGBT Index

	Pro-LGBT Index					
	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	0.191	-0.131	-0.221	-0.263	-0.158	-0.070
	(0.089)	(0.135)	(0.125)	(0.043)	(0.057)	(0.076)
Gold Rushes		0.435	0.447	0.366	0.235	0.136
		(0.144)	(0.136)	(0.066)	(0.069)	(0.080)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Individual Controls	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes
Observations	15,200	15,200	15,200	15,200	15,200	15,200

Notes: The unit of observation is a respondent. The period covered is 1993–2014. We use answers to the following four questions to create a pro-LGBT index: "What about sexual relations between two adults of the same sex–do you think it is, not wrong at all, wrong only sometimes, almost always wrong and always wrong?", "A man admits to be homosexual: should he be allowed to make a speech in your community?", "A man admits to be homosexual: should he be allowed to teach in a college or university?" and "If some people in your community suggested that a book he wrote in favor of homosexuality, should it be taken out of your public library?" The index goes from 0 to 4. *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Columns 1–6 include year fixed effects. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties. The set of individual controls include the following variables: gender, age, age squared, six education dummies, three race dummies and five marital status dummies.

TABLE 1.20: Fundamentalist or Liberal?: GSS

	Currently Fundamentalist, Moderate or Liberal?					
	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit	Ordered Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	0.168	-0.177	-0.216	-0.239	-0.183	-0.050
	(0.082)	(0.162)	(0.149)	(0.089)	(0.113)	(0.098)
Gold Rushes		0.458	0.406	0.405	0.289	0.153
		(0.161)	(0.149)	(0.091)	(0.105)	(0.087)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Individual Controls	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes
Observations	28,153	28,153	28,153	28,153	28,153	28,153

Notes: The unit of observation is a respondent. The period covered is 1993–2014. Fundamentalism is assessed through the following question: "How fundamentalist is r currently" where respondents have three choices (1=Fundamentalist, 2=Moderate, and 3=Liberal). *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Columns 1–6 include year fixed effects. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties. The set of individual controls include the following variables: gender, age, age squared, six education dummies, three race dummies and five marital status dummies.

TABLE 1.21: Reading the Bible in Public Schools: GSS

	Respondent Approves Court Ruling					
	Probit	Probit	Probit	Probit	Probit	Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	0.115	-0.008	-0.026	-0.048	-0.009	0.002
	(0.032)	(0.045)	(0.034)	(0.016)	(0.020)	(0.022)
Gold Rushes		0.164	0.147	0.109	0.066	0.052
		(0.051)	(0.042)	(0.029)	(0.032)	(0.025)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Individual Controls	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes
Observations	16,209	16,209	16,209	16,209	16,209	16,209

Notes: The unit of observation is a respondent. The period covered is 1993–2014. The dependent variable is a dummy for whether the respondent approves the United States Supreme Court ruling that no state or local government may require the reading of the Lord's Prayer or Bible verses in public schools. *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Marginal effects are reported. Standard errors clustered by state are reported between parentheses. Columns 1–6 include year fixed effects. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties. The set of individual controls include the following variables: gender, age, age squared, six education dummies, three race dummies and five marital status dummies.

TABLE 1.22: Other Attitudes and Values: GSS

	Rs is Liberal Ordered Probit	Abortion Legal Probit	Marijuana Legal Probit	Death Penalty Probit	Permit Gun Probit
	(1)	(2)	(3)	(4)	(5)
Gold	0.039	0.019	-0.057	0.038	0.071
	(0.022)	(0.062)	(0.093)	(0.064)	(0.052)
Gold Rushes	0.0001	0.073	0.146	-0.030	0.045
	(0.022)	(0.059)	(0.087)	(0.072)	(0.076)
Year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Observations	25,746	16,130	15,838	23,676	16,586

Notes: The unit of observation is a respondent. The period covered is 1993–2014. In column 1, the dependent variable is whether respondents think of themselves as liberal or conservative. Respondents are offered seven choices (7 =extremely liberal, 6 =liberal, 5 =slightly liberal, 4 =moderate, 3 =slightly conservative, 2 = conservative and 1 =extremely conservative). In columns 2–5, the dependent variables are answers to the following questions "Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if she is married and does not want any more children", "Do you think the use of marijuana should be made legal or not?", "Do you favor or oppose the death penalty for persons convicted of murder?" and "Would you favor or oppose a law which would require a person to obtain a police permit before he or she could buy a gun?" *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Columns 1–5 include year fixed effects. Marginal effects are reported in columns 2 to 5. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties. The set of individual controls include the following variables: gender, age, age squared, six education dummies, three race dummies and five marital status dummies.

TABLE 1.23: Male and Female Same Sex Unmarried Partners: With and Without Children

	Ln(1 + Male Same Sex)			Ln(1 + Female Same Sex)		
	(1)	(2)	(3)	(4)	(5)	(6)
Gold	-0.058	-0.073	-0.043	-0.035	-0.078	-0.024
	(0.023)	(0.050)	(0.032)	(0.029)	(0.048)	(0.030)
	[0.039]	[0.168]	[0.243]	[0.268]	[0.101]	[0.429]
Gold Rushes	0.111	-0.081	0.153	0.128	-0.031	0.165
	(0.073)	(0.073)	(0.082)	(0.039)	(0.587)	(0.048)
	[0.194]	[0.327]	[0.109]	[0.0009]	[0.587]	[0.002]
ln(population)	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,111	3,111	3,111	3,111	3,111	3,111
Adjusted R-Squared	0.934	0.880	0.917	0.957	0.917	0.938

Notes: The unit of observation is a county. In column 1, the dependent variable is the natural log of one plus the number of male unmarried partner households of the same sex in 2010. In columns 2 and 3, the dependent variable is the natural log of one plus the number of male unmarried partner households of the same sex with and without children respectively. Columns 4–6 have the same structure as columns 1–3, but for female unmarried partner households of the same sex. In columns 2 and 5, the dependent variables are the number of male and female unmarried partner households of the same sex with at least one child, respectively. In columns 3 and 6, the dependent variables are the number of male and female unmarried partner households of the same sex without a child, respectively. Columns 1–6 include state fixed effects, our set of geographic controls and the (log of) 2010 county population. *Gold* equals one for counties with at least one gold discovery and zero otherwise. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. Standard errors clustered by state are reported between parentheses. Wild cluster bootstrap *p*-values are reported between brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)).

TABLE 1.24: Gold Rushes and Male-to-Female Ratio: Unbalanced Panel

	Male-to-Female Ratio			
	(1)	(2)	(3)	(4)
Gold Rushes	5.806	5.892	5.892	5.876
(Within 10 Years)	(3.591)	(3.597)	(3.592)	(3.597)
Gold Rushes		0.519	0.519	0.504
(10-20 Years Later)		(0.056)	(0.063)	(0.066)
Gold Rushes			0.001	0.004
(20-30 Years Later)			(0.067)	(0.066)
Gold (Not Gold Rushes)				-0.070
(Within 10 Years)				(0.050)
Gold (Not Gold Rushes)				-0.011
(10-20 Years Later)				(0.023)
Gold (Not Gold Rushes)				-0.037
(20-30 Years Later)				(0.014)
Constant	1.128	1.125	1.125	1.126
	(0.010)	(0.010)	(0.011)	(0.011)
Decade FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	50,046	50,046	50,046	50,046
Adjusted R-Squared	0.135	0.139	0.139	0.139

Notes: The unit of observation is a county-decade. The dependent variable is the male-to-female ratio. The time period is 1820–2000. *Gold Rushes (Within 10 Years)* equals one for county-decade observations with the earliest gold rush discovery and zero otherwise. *Gold Rushes (10–20 Years Later)* and *Gold Rushes (20–30 Years Later)* equal one for county-decade observations with respectively the earliest gold rush discovery 10–20 years earlier and 20–30 years earlier. *Gold–Not Gold Rushes– (Within 10 Years)* equals one for county-decade observations with the earliest gold discovery before or after the gold rushes and zero otherwise. *Gold Rushes–Not Gold Rushes– (10–20 Years Later)* and *Gold Rushes–Not Gold Rushes– (20–30 Years Later)* equal one for county-decade observations with respectively the earliest non-gold rush discovery 10–20 years earlier and 20–30 years earlier. Standard errors clustered by state are reported between parentheses. Columns 1–4 include county and decade fixed effects.

TABLE 1.25: Gold Rushes and Male-to-Female Ratio: Unbalanced Panel

	Male-to-Female Ratio			
	(1)	(2)	(3)	(4)
Gold Rushes	1.956	2.234	2.254	2.249
(Within 10 Years)	(0.871)	(1.149)	(1.171)	(1.175)
Gold Rushes		-0.569	-0.475	-0.483
(10-20 Years Later)		(0.585)	(0.499)	(0.499)
Gold Rushes			-0.234	-0.239
(20-30 Years Later)			(0.234)	(0.236)
Gold (Not Gold Rushes)				-0.011
(Within 10 Years)				(0.034)
Gold (Not Gold Rushes)				0.033
(10-20 Years Later)				(0.035)
Gold (Not Gold Rushes)				-0.064
(20-30 Years Later)				(0.022)
Constant	1.126	1.128	1.129	1.130
	(0.010)	(0.011)	(0.011)	(0.011)
Decade FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	50,046	50,046	50,046	50,046

Notes: The unit of observation is a county-decade. The dependent variable is the male-to-female ratio. The time period is 1820–2000. *Gold Rushes (Within 10 Years)* equals one for county-decade observations with at least one gold rush discovery and zero otherwise. *Gold Rushes (10–20 Years Later)* and *Gold Rushes (20–30 Years Later)* equal one for county-decade observations with respectively at least one gold rush discovery 10–20 years earlier and 20–30 years earlier. *Gold–Not Gold Rushes– (Within 10 Years)* equals one for county-decade observations with at least one gold discovery before or after the gold rushes and zero otherwise. *Gold Rushes–Not Gold Rushes– (10–20 Years Later)* and *Gold Rushes–Not Gold Rushes– (20–30 Years Later)* equal one for county-decade observations with respectively at least one non-gold rush discovery 10–20 years earlier and 20–30 years earlier. Standard errors clustered by state are reported between parentheses. Columns 1–4 include county and year fixed effects.

TABLE 1.26: Intensive Margin: Neighboring Effect

	Ln(1 + Unmarried Partner Households Same Sex)					
	(1)	(2)	(3)	(4)	(5)	(6)
Gold Rushes	0.101	0.069	0.132	-0.037	-0.049	0.020
	(0.047)	(0.091)	(0.082)	(0.071)	(0.054)	(0.045)
Neighbor	0.046	0.045	0.182	0.023	0.002	0.053
	(0.044)	(0.045)	(0.063)	(0.064)	(0.062)	(0.062)
Gold Rushes ×		0.148	0.296	0.128	0.096	0.151
Neighbor		(0.059)	(0.073)	(0.080)	(0.082)	(0.088)
ln(population)	Yes	Yes	Yes	Yes	Yes	Yes
Census Region FE	No	No	No	Yes	No	No
Census Division FE	No	No	No	No	Yes	No
State FE	No	No	No	No	No	Yes
Geographic Controls	No	No	Yes	Yes	Yes	Yes
Observations	3,111	3,111	3,111	3,111	3,111	3,111
Adjusted R-Squared	0.955	0.955	0.958	0.961	0.962	0.964

Notes: The unit of observation is a county. The dependent variable is the natural log of one plus the number of unmarried partner households of the same sex in 2010. *Gold Rushes* equals one for counties with at least one gold discovery during the gold rushes and zero otherwise. *Neighbor* equals one for counties neighboring a *Gold Rushes* county and zero otherwise. Standard errors clustered by state are reported between parentheses. Columns 1–6 include the (log of) 2010 county population. The set of geographic controls include latitude, longitude, total land area, mean and standard deviation of elevation, distance to the state's capital and a dummy for coastal counties.

Chapter 2

Labor Legislation and Women's Economic Empowerment: Evidence from Night Work, Regulatory and Seating Laws

2.1 Abstract

This chapter examines the impact of gender focused labor legislation on women's labor force participation and economic empowerment. We rely on historical legislative acts passed by state legislatures and exploit whether or not states passed regulatory laws regulating overall and industry specific employment and work conditions for women, night work laws and labor laws requiring provision of seats for working women. We exploit the fact that not all states enacted these laws as well as the variation in the timing of enactment of such laws. Our results show that women in comparison to men in treated states are more likely to be in the labor force post introduction of night work laws in comparison to control states. We also document the effect of industry-specific labor policies on women's likelihood to be employed in the affected industry and in higher-wage occupations within the industry of interest. Policy implications of our findings endorse the adoption of labor laws in favor of women to advocate their empowerment through a higher involvement in the labor market and financial independence.

2.2 Introduction

Female labor force participation patterns differ significantly across countries. Economic development (Boserup, 1970; Goldin, 1995) and cultural differences in gender norms arising from historical agricultural practices (Alesina, Giuliano, and Nunn, 2013; Hansen, Jensen, and Skovsgaard, 2015) or family cultures (Alesina and Giuliano, 2010) or ties to religious roots (Nunn et al., 2014) have been shown to be crucial factors contributing to the understanding of these patterns. In addition, gender focused policies including maternity leave, child benefits and tax credits favoring women as well as policies granting them legal rights or financial incentives to invest in female education have been presented as key determinants of women's participation on the labor market.[1]

[1]See Jayachandran, 2015 and Heath and Jayachandran, 2017 for an overview of this literature for developing countries.

The purpose of this chapter is to assess whether public policies, specifically labor legislations for women, have a direct effect on female economic empowerment. We focus on the context of the United States and restrict our analysis to the period 1860 to 1940. This setting is attractive for at least two reasons.

First, it allows us to study a large number of legislative acts and gender based labor laws enacted during that period. More precisely, we examine three types of laws: regulatory, night work and seating laws. Regulatory laws include laws that aim at enhancing the well-being of employed women by improving their labor conditions and banning their employment in certain types of occupations that may include unfavourable work conditions. For instance, working in mines, especially coal mines, involves detrimental settings potentially affecting maternal capacity or women's health. It was thus deemed illegal for women to occupy these types of occupations. Other laws organized work conditions in specific occupations by banning women from carrying heavy burdens (see Section 2.3 for details). Night work laws are labor regulations enacted by state legislatures that prohibit work at night for women in certain industries or occupations or restrict the number of hours that women are allowed to work at night. Lastly, seating laws made it compulsory that "suitable" seating accommodations were provided to female employees.

Second, focusing on the 1860–1940 U.S. context is interesting because the long time period studied allows us to examine the short and medium run effects of these labor laws targeting women. This is especially important given that gender norms are hard to change and it might take years or decades for the enactment and implementation of labor legislation to impact women's economic empowerment.

Our aim is to examine whether the introduction of labor laws targeting women had an effect on their likelihood to be in the labor force, to occupy jobs in targeted industries and to be employed in higher wage occupations within affected industries. We also examine whether gender focused labor laws have indirect policy effects on women's marital, divorce and fertility status.

To examine causal impacts of these labor laws, we rely on a differences in differences research design that allows us to compare changes in female labor force participation before and after labor laws were introduced in treated states (i.e. those that passed labor legislation for women) in comparison to control states (i.e. those that did not pass any of the labor laws targeting women). This strategy exploits geographical and temporal variation in the adoption of labor laws favoring women. This is because not all states adopted regulatory, night work or seating laws. Moreover, the states that enacted any of these labor laws, did not adopt them at the same time. The enactment of female focused labor laws happened rather gradually across states between 1872 and 1931.

A differences in differences estimation strategy requires assumptions that may not hold for a number of reasons. First, states that passed labor legislation for women might be systematically different than states that did not. Second, female labor force participation might have upward general trends in both types of states. Finally, other state specific shocks might have occurred and that could explain changes in female labor force participation.

We estimate leads and lags of the treatment over five decades which leverages variation in the timing of gender based labor laws' introduction across different states. We investigate regulatory, seating and night work laws simultaneously, and classify states in each decade into treated or control states depending on whether or not the state passed one of the labor laws targeting women in that specific decade. Parallel pre-trends provide suggestive evidence in support of parallel trends in the absence of the treatment, i.e. identification assumption.

When we examine women and men separately, we provide weak evidence in support of higher female labor force participation in the states that implemented these policies in comparison to the states that did not. The estimates for men are close to zero indicating that labor legislation targeting women does not impact men's labor force participation. This is suggestive that men can be used as an alternative counterfactual group in examining the impact of gender focused labor laws. Our findings from carrying out this analysis indicate that night work laws increased women's labor force participation by around 2 percentage points in comparison to men in treated in comparison to control states. Our results are robust to the inclusion of potential state specific confounding factors including laws granting women voting, property and earning rights.

We also provide a heterogeneity analysis by exploiting variation in electricity access and urbanization level. We document a positive and statistically significant impact of seating laws on female labor force participation, in comparison to their male counterparts, with higher electricity access and urbanization level.

Next, we consider industry specific labor laws' effects. Our results show that women are more likely to be employed in the manufacturing industry and less likely to occupy personal services related jobs when laws prohibiting night work shifts for women are introduced. We also document a higher likelihood of employment for women in manufacturing related occupations as regulatory laws enhancing work conditions for women are enacted. The decrease in women's employment in the personal services industry is likely due to the fact that this industry requires night work shifts which are prohibited to women.

In addition to examining industry specific employment effects, we investigate whether women are more likely to be employed in higher wage occupations within affected industries. We find evidence in support of this for the personal services industry. When we examine the impact of labor laws favoring women on women's marriage and fertility status, our results imply that there were no such effects suggesting that direct policies might be more efficient.

Our research offers an important contribution to the debate on public policies targeting female economic empowerment by documenting positive effects induced by the introduction of labor laws on female participation in the labor market through an amelioration of their working conditions. Our findings appear to provide evidence in support of emerging policies in developing countries where, for instance, India just implemented a seating law in 2018. This policy is believed to grant women greater freedom in taking factory jobs. Moreover, our research is also relevant to the controversial issues in the garment industry in Bangladesh and other developing countries. According to the World Bank, approximately 80 percent of garment workers in 2017 were women, most of them working in detrimental conditions and having serious difficulties finding alternative employment. Furthermore, numerous developing countries have gender-unequal laws and policies restricting gender equal legal rights, including property or inheritance rights, and that hinder female employment. Additional restrictions to female employment include husbands' right to prevent their wives from working or women not being allowed to head households. Gonzales et al., 2015 documents strong evidence in support of a cross-country correlational relationship between removing such discriminatory laws and increases in female employment. Causal evidence on the effect of gender based laws is still, however, weak.

This chapter contributes to the literature in three ways. First, to our knowledge, we are the first study to formally investigate the causal impacts of gender based labor legislation on female labor force participation. Second, we provide suggestive

evidence that such labor laws regulating formal labor market conditions in favor of women fail to have significant indirect effects on marriage or fertility choices for women. Finally, our results have policy implications endorsing the adoption of labor laws in favor of women to increase female empowerment through a higher involvement in the labor market and financial independence.

The rest of the analysis is structured as follows. In Section 2.3, we provide a detailed overview of labor legislations for women in the United States. In Section 2.4, we describe our data sources. Section 2.5 outlines our empirical strategy. In Section 2.6, we discuss our results and present a battery of robustness checks. We briefly conclude in Section 2.7.

2.3 Development of Labor Legislation for Women

"When the spinning, the weaving, the making of clothes and shoes, and all the other industrial pursuits once followed in the home were taken away and developed into factory industries, women were called upon still to conduct many of the operations, and thereby became part of the army of those gainfully employed" (Women's Bureau U.S. Department of Labor, Bulletin No. 3).

Prior to World War I, the number of women employed in industries gradually increased, but the need to recruit women then peaked as men vacated jobs to fight in the war. The urgent need for workers allowed women to fulfill jobs traditionally held by men. There was also a necessity for laws regulating the efficiency and the health of working women. *"The physical well-being of woman becomes an object of public interest and care in order to preserve the strength and vigor of the race"* (Cushman, 1998). Additionally, because women had been in a weaker economic position with respect to men, there was a great need for control of the standards of women's employment. During war, both federal and state agencies were responsible for building up standards for working women. After WWI, U.S. state's responsibilities increased and the need for protecting the health of female workers was recognized locally as a vital economic and social measure.

Standards for women's employment in industries, drawn up with the advice of both employers and workers, were published by the "Women in Industry Group" created by the Department of Labor which then became the United States Department of Labor Women's Bureau. These standards were eventually incorporated into labor laws at the state and federal levels. *"It was the first time the federal government had taken a practical stand on conditions of employment for women, and although the standards were only recommendations and had no legal force, they were a very important statement of policy and were widely used in all parts of the country"* (Mary Anderson, the first Director of the Women's Bureau).

Recommendations included posture at work: *"Continuous standing and continuous sitting are both injurious. A chair should be provided for every woman and its use encouraged."* Other recommendations covered comfort and sanitation: *"State labor laws and industrial codes should be consulted with reference to provisions for comfort and sanitation"* (Women's Bureau U.S. Department of Labor, Bulletin No. 173).

Governments of different states passed legislative acts regulating women's labor on the basis of these standards.[2] Our data on the chronological development

[2]The first law regulating the hours of work for both men and women was enacted in 1847 in New Hampshire. While hours of work and minimum wage laws were controversial issues in the U.S., regulations related to work conditions for women (except minimum wage laws) were enacted and implemented rather easily. Woloch, 2015 offers a good historical narrative for the roots of protective laws

of state level labor legislation for women is based on a digitalized archived bulletin published by the United States Department of Labor Women's Bureau. The establishment of this bureau in the Department of Labor was enacted during a congress assembly of the Senate and House of Representatives of the United States. The role of the Women's Bureau as dictated in the bulletin is to *"formulate standards and policies which shall promote the welfare of the wage-earning women, improve their working conditions, increase their efficiency, and advance their opportunities for profitable employment"* (Public Law 66-259, 66th Congress, H.R. 13229). Moreover, this bureau had investigative authority over any issue concerning women's welfare. The bureau subsequently sent its report to the Department of Labor, which had the authority to make the results of investigations available to the public.

For our empirical analysis, we rely on Bulletin number 66 Part II revised in December 1931 and published in 1932 that was prepared and published by the United States Department of Labor Women's Bureau. It presents the history of the evolution of all labor laws for women enacted from 1872 to 1931. One part of this bulletin (Part I) provides a detailed overview of women's labor legislation in only three states (California, Massachusetts and New York). The second part (Part II) provides a compressed overview of labor laws for women across all states.

We are the first to rely on this data to formally examine the effect of gender focused labor legislation favoring women. We are interested in three types of laws that affect women's labor (and not men's). First, laws prohibiting certain kinds of employment for women or regulating the conditions under which they may be performed. Legislative acts passed by legislatures of some states exclude women from employment in specific occupations. These acts mainly prohibit women from working in mines, specifically coal mines. Other bans forbid women from working as baggage handlers for instance, or working in the manufacturing or handling of nitro and amido compounds or any other dry substance (i.e. highly solid based substances). More broadly, labor laws banned women from working in occupations which involved harmful and detrimental work conditions or affected women's potential capacity for motherhood or health. Other laws banned employment for women immediately before and after childbirth. Additional regulatory laws organized and controlled the conditions under which specific occupations could be performed rather than banning them for women. For instance, women were not permitted to carry heavy burdens, clean out moving machinery, etc.

In summary, regulatory legislation included laws with the stated aim to improve the well-being of women at work and ban their employment in certain types of occupations that could include detrimental work conditions for them. The majority of these laws were enacted in a narrow range of only a few states. The distribution of regulatory laws by state is as follows: about 21 states did not enact any regulatory legislation as of 1931 whereas 13 states passed a single regulation.[3] A total of six states passed two regulatory regulations, two states passed three and two states enacted four regulations. Finally, the highest number of regulations enacted were passed by three states where one state passed six regulations, one passed 13 and

for working women such as maximum hours laws, minimum wage laws, and night work laws, including popular reformers' rationales and the legislative and judicial progression of protective policies in the United States.

[3] A total of seven out of these 13 states banned employment of women in coal mines, one state prohibited their employment in messenger services, two states regulated unemployment periods for women before and after childbirth, one state forbid heavy lifting for women, one state banned the cleaning of moving machinery and one state regulated their work on moving abrasives.

one passed 23 regulations respectively. Only two states had an exhaustive list of restrictions prohibiting women from working in many occupations that were deemed legal and women were free to occupy in all other states.

Second, we are interested in seating laws. These laws made it mandatory that "suitable" seats should be provided to female employees. The extent to which these seats could be used varied across states. Some states specified the purpose as to allow working women to *"sit when not actively engaged in their duties or when sitting does not interfere with the proper discharge of duties"* (Bulletin number 66 Part II, p. 173). Other states dictated that such seats could be used as it is necessary to preserve the health of workers. Almost all states except Mississippi introduced seating laws, however, the occupations or industries in which these laws were applicable varied significantly. For instance, seating laws were applicable to mercantile occupations exclusively in Alabama, Maryland, North Dakota and South Carolina. Moreover, the number of seats required varied across states and/or occupations. In this chapter, we only focus on whether or not states passed seating laws, and the type of industries in which these laws were applicable.

Lastly, we focus on night work laws. These laws prohibit night work for women in certain industries or occupations or limit the number of hours that women are allowed to work at night. In total, 16 states banned female work at night in specific industries or occupations. Three of those states prohibited night work shifts for women in manufacturing industries and one state in mercantile; two states ban their night work in one occupation each. The remaining ten states specified at least two targeted industries or occupations for night work laws. Few states allowed women to work at night but instead limited shifts to eight or ten hours. Some states combined these two types of laws by imposing a restrictive number of hours for female night work in certain occupations while banning work at night for women in other occupations. Most states specified 10 p.m. to 6 a.m. as the period during which night work is banned.

Appendix Figures 2.3, 2.4 and 2.5 describe the chronological timing of passage of regulatory, night work and seating laws across states. States enacted regulatory laws between 1880 and 1940, while night work laws and seating laws were enacted between 1900 and 1940. Appendix Figure 2.3 documents that only one state had enacted regulatory laws in 1880, and by 1940, about 21 states had not passed any, while the remaining states had passed at least one regulatory law. Appendix Figure 2.4 describes the timeline of passage of night work laws showing that by 1940 few states had enacted such laws. Appendix Figure 2.5 shows that by 1940, all but two states had passed seating laws.

We disregard other type of labor laws enacted, such as hours laws that provided daily/weekly limits for female hours worked, because these laws varied significantly across different states, making it hard to measure and identify relevant variation. This is especially true given that almost all states passed such laws but regulated the hours of work for women differently.[4] We also do not examine the equal-pay law that made it unlawful to discriminate in any way in the payment of wages between sexes (Megdal and Ransom, 1985). This is because during our period of study, only two states (Michigan and Montana) passed such a law (in 1919).

[4]Some states regulated hours worked for women to eight hours, other states enacted 8.5 hour laws, etc. Few states passed weekly hours laws. In many states, the industries or occupations targeted by these hours laws were very limited.

2.4 Data

In this section, we describe our data sources for state level labor legislation, female labor force participation, and other outcomes of interest. We also provide some detailed descriptive statistics.

2.4.1 Data on State Level Labor Legislation for Women

In order to obtain data on state level labor legislation for women in the United States, we rely on a digitalized archived bulletin published by the United States Department of Labor Women's Bureau. Bulletin number 66 part II published in 1932 presents an overview of the evolution of all labor laws for women enacted from 1872 to 1931.

When we focus on regulatory laws, we consider a state that passed a regulatory law of any kind targeting any industry as a treated state in the corresponding decade. By 1940, 28 states had passed regulatory laws while 21 had not (Appendix Figure 2.3). Regarding night work laws, 20 states enacted such laws by 1940 and 29 states did not (Appendix Figure 2.4). As discussed above, while night work laws in some states banned women from working at night in specific industries, other states dictated the number of hours during which women could work at night. In our analysis, we do not differentiate between different types of night work laws. We label any state passing night work laws of any kind as a treated state in our sample. Regulatory laws were enacted across U.S. states between 1872 and 1919 and night work laws were implemented between 1889 and 1921.

Finally, regarding seating laws, i.e. laws requiring provision of seats for working women, our treated group in 1940 (the last year in our sample) comprises 47 states and our control group includes two states. For the purpose of this chapter, we do not focus on the exact wording of these seating laws, which diverges across different states, nor, for example, the specific number of seats required. In our sample, a state is considered as treated if a seating law of any kind is enacted. The first seating law was implemented in 1881 and the last one in 1931 (see Appendix Figure 2.3).

2.4.2 IPUMS Data - Individual Level

We rely on 1% weighted samples from the United States Census data (1860–1940) available from the Integrated Public Use Microdata Series (IPUMS).[5] IPUMS provides access to U.S. census microdata and includes a wide range of variables related to education, work, income and demographic characteristics. We carry out our analysis at the individual level and we collect information about an individual's age, race and ethnicity. We gather data on labor force status using a dichotomous variable that indicates whether the person is in the labor force or not.

We also rely on information about the occupational income score for women from IPUMS. This is a constructed two digit numeric variable that allocates a value representing the median total income (in hundreds of 1950 dollars) to occupations in all years. This variable allows for the classification of higher wage occupations. For our analysis, we restrict the sample to women aged 16–65 years old. IPUMS also contains information about the marital and fertility status of women.

Table 2.1 shows summary statistics by decade for our sample of women. We report statistics for the following decades: 1860, 1880, 1900, 1920 and 1940, in columns

[5]Note that Census data is missing for 1890 from all sources. We are thus excluding this decennial year from our analysis. Also, we rely on a person weight variable available from IPUMS which gives the population represented by each individual in the sample.

(1) to (5) of Table 2.1 respectively. The share of women in the labor force increased from about 14% in 1860 to almost 29% in 1940. The average age in our sample of females increases from 33 to 37 between 1860 and 1940. The majority of women in our sample are married and the average number of own children decreased from about two to one child between 1860 and 1940.[6] In column (5), we compute descriptive statistics across all decades. On average, about 24% of women are in the labor force.

2.5 Identification Strategy

In this section, we describe in detail our identification strategy, identification assumption, main specifications and controls. The chronological variation in implementation of state-level policies, in addition to the large span of data we have on female labor force participation covering both pre and post policy introduction periods, provide us with an ideal framework to carry a differences in differences analysis with many treatment and control groups.

2.5.1 Outline of Identification Strategy

Identifying the causal effect of labor laws targeting women on their labor force participation is challenging for a number of reasons. First, the states that passed labor legislation for women might be systematically different than states that did not. Moreover, female labor force participation might have upward general trends in both type of states irrespective of such policies. Finally, other state specific shocks (policies) might have occurred (been introduced) that can explain the increase in women's labor force participation.

We thus rely on an identification strategy that exploits across states variation in the type of labor laws targeting women that were enacted. This is because not all states prohibited night work for women for instance. Moreover, while some states had no laws regulating occupation specific employment for women, other states had regulations prohibiting women from working in specific occupations. Finally, not all states enacted laws requiring seats provision to working women. We furthermore exploit the chronological nature of implementation. This is because not all states enacted labor laws at the same time, rather, implementation happened gradually between 1872 and 1931. This gradual state-level policy implementation, along with the fact that not all states enacted women labor laws, allows us to carry a differences in differences analysis with many treatment and control groups varying by decade.

We examine the impact of the introduction of seating, regulatory and night work laws by categorizing states in each decade into treated or control states depending on whether or not the state had any of the labor laws targeting women. We thus generate binary indicator variables that classify states into treated and control states. The indicators take the value 1 for treated states, i.e. states that had a given type of labor law passed in a given decade, as of that specific decade. The indicators take the value 0 for control states i.e. states that never had a labor law targeting women or treated states in decades before the passage of these laws. This research design allows us to identify the causal effect of introducing labor laws targeting women by comparing changes in female labor force participation before and after the introduction of labor legislations in treated states in comparison to control states.

[6] Marital status variable is not included in the U.S. Census before 1880.

2.5.2 Model Specification

Differences in Differences Estimation

The objective is to investigate the impact of the passage of gender based labor laws. Our identification strategy relies on a direct comparison between states that passed one of the labor laws targeting women and those that did not. We carry out our analysis at the individual level which allows us to control for individuals' characteristics. We also account for decade specific and time-invariant state specific characteristics. We estimate the following differences in differences specification:

$$y_{isd} = \alpha_s + \beta_d + \gamma SeatingLaw_{sd} + \delta RegulatoryLaw_{sd} + \omega NightWorkLaw_{sd} + \lambda X'_{isd} + \theta_{isd} \tag{2.1}$$

where y_{isd} is the labor force status of woman i residing in state s in decennial year d. It is a binary variable which takes the value one if woman i residing in state s in decennial year d is in the labor force and zero otherwise. $Seating\ Law_{sd}$, $Regulatory\ Law_{sd}$ and $Night\ Work\ Law_{sd}$ are indicators that equal one for states that had any seating, regulatory or night work laws respectively in decennial year d and zero otherwise. These are binary indicator variables that classify states into treated and control states. Our coefficients of interest are γ, δ and ω which capture our treatment effects. α_s and β_d are state and decade fixed effects, X'_{isd} is a list of individual level controls including age, ethnicity and race and θ_{isd} is our error term. Our standard errors are clustered at the state level. We also report wild cluster bootstrap p-values given the relatively small number of unbalanced clusters contributing to the identifying variation (MacKinnon and Webb, 2017, Roodman et al., 2019).

Our identification assumption is the parallel trends in outcome variables between treated and control states in the absence of the treatment. In other words, trends in female labor force participation would have continued to be the same in treated and control states had these labor laws targeting women not been introduced. We thus compare states that implemented women's labor laws to control states, i.e. those that had not implemented such laws in a given decade and we argue that these form a valid counterfactual group.

In an alternative specification, we rely on men as our counterfactual group and conduct a triple differences analysis that allows us to compare women's labor force participation in comparison to men in treated relative to control states. More precisely, our sample in this case is pooled to include both men and women but our dependent variable of interest remains labor force status. We interact our independent variables of interest $Seating\ Law_{sd}$, $Regulatory\ Law_{sd}$ and $Night\ Work\ Law_{sd}$ with a dummy "woman" that takes the value 1 if the individual is a female and zero otherwise.

When we investigate industry specific labor laws' effects, a state is considered treated in a given decennial year only if it had one of the three types of labor laws for the industry of interest. For instance, when we focus on targeted regulatory laws effects on the manufacturing industry, only a state that passed a law regulating female employment in manufacturing occupations is considered a treated state. Thus, our analysis examines the impacts of gender based labor laws on female employment in targeted industries. The same analysis applies to night work laws or seating laws targeting only specific occupations, which we classify into the corresponding industry. The dependent variable of interest, female employment, is a dummy variable that takes the value of 1 if a woman is employed in the indicated industry, and 0 otherwise. This type of analysis allows us to investigate whether there exists a direct

effect on female employment in specific industries from targeted policies rather than examining the impact on the overall labor force participation of women.

Leads and Lags of the Treatment

In an alternative specification, we estimate treatment impacts at different dates. We thus generate dummies for leads and lags of the binary indicator variables for the passage of different labor laws for two decades before and three decades after the passage of laws. For the analysis, we omit the first decade preceding the year of the policy implementation *(d-1)*.

We thus estimate the following specification:

$$y_{isd} = \lambda_s + \gamma_d + \Sigma_{\kappa=-2}^{3} \alpha_\kappa StateLaw_{s(d+\kappa)} + \beta OtherStateLaw_{sd} + \delta X'_{isd} + \varepsilon_{isd} \quad (2.2)$$

where y_{isd} is the labor force status of woman i residing in state s in decennial year d. It is a binary variable which takes the value one if woman i residing in state s in decennial year d is in the labor force and zero otherwise. *State Law*$_{sd}$ is set to equal one for states that had a given labor law in decennial year d and zero otherwise. The leads provide a falsification test of the main identification assumption, which is that trends in female labor force participation would have continued to be the same in treated and control states in the absence of the laws' passage by examining pre-treatment outcomes (i.e. for κ equals -2). The lags in our model allow us to examine whether the effect of the given labor law fades away or grows over time (for κ being equal to 1, 2, and 3, respectively). Because we do this for one type of law at a time, *Other State Law*$_{sd}$ indicators are two sets of dummies that control for the two other types of labor laws. These dummies take the value one if other laws are in place in state s in decennial year d and zero otherwise. Our coefficients of interest are α_κ which capture treatment impacts at different dates. λ_s and γ_d are state and decade fixed effects, X'_{isd} is a list of individual level controls including age, ethnicity and race and ε_{isd} is our error term. Our standard errors are clustered at the state level.

2.6 Main Results

2.6.1 Main Findings

Table 2.2 reports the effects of the three types of state level labor legislation for women estimated using Equation 2.1. We first examine seating laws in columns (1) and (2); then regulatory laws in columns (3) and (4) and lastly we report the analysis for night work laws in columns (5) and (6). All state laws are included in the same specification to account for the overlap that might occur between the three different policies. Thus, in column (1) we examine the impact of seating laws while controlling for regulatory and night work laws passage as well as state and decade fixed effects. Column (2) further controls for individual level exogenous characteristics including age, race and ethnicity. In column (3), we examine the impact of regulatory laws while controlling for seating and night work laws passage then add individual level controls in column (4). Finally, in column (5), we report our estimates for night work laws and we add individual level controls in column (6).

In Panel A, we analyze the impact of gender focused labor laws on labor force participation for both women and men separately. We find weak evidence that female labor force participation in states that implemented these policies is higher

than in those that did not. Our estimates on women's LFP are positive but not statistically significant. While investigating the effect on men's labor force participation, we find that our estimates are close to zero with a slightly negative impact of night work laws. These results indicate that labor legislation for women did not impact male labor force participation which suggests that men can be used as a potential counterfactual in studying the impact of these policies on women's LFP.

Therefore, we mimic Equation 2.1 but we pool men and women and add a "women" dummy - that takes the value of one if individual i is a woman - as an additional interaction term to each of our state law dummy indicator variables. The triple differences specification in Panel B allows us to assess the impact of state law policies on the change in labor force participation of women compared to men in treated compared to control states. Our results suggest that night work laws increased women's labor force participation by around 2 percentage points compared to men, respectively.

Given that the states of Michigan and Montana passed equal pay laws during our period of study (in 1919), we exclude them from our analysis as a robustness exercise. Our results, reported in Appendix Table 2.7, remain robust to the exclusion of these states and are of a similar magnitude as in Table 2.2.

Moreover, we carry a robustness analysis where we additionally control for state specific time trends. Appendix Table 2.8 displays the results we obtain from doing this specification. Our findings for the triple differences estimates are of similar magnitude to those in Table 2.2, confirming our previous conclusions.

Figures 2.1 and 2.2 show our analysis for the leads and lags of the treatment effect on women's and men's labor force participation respectively. These results are computed using Equation 2.2 where we estimate treatment impacts at different dates. Our leads estimates provide a falsification test of our main identification assumption. The fact that our estimates are close to zero and not statistically significant before policy enactment bolsters the credibility of our identification strategy.

The lags in our model allow us to examine how the effects of the different labor laws changed over time. This is especially important in our context since it may have taken several years for the policies to have an effect on gender roles and labor markets. Figure 2.1 shows that while the impact of seating laws remain statistically insignificant over the full period analyzed, the effect of regulatory laws became significantly positive on women's LFP one decade after the policy implementation *(d+1)*. The effects of night work laws were also not immediate; rather, we document a positive statistically significant estimate three decades after its implementation *(d+3)*. The increasing effect over time is probably due to the fact that these state laws were being reinforced over time. One way to reinforce these laws was by increasing the penalties for violations.

Figure 2.2 shows our leads and lags treatment effects on male labor force participation. Results in this figure confirm that female specific labor laws did not have an impact on men's labor outcomes.

2.6.2 Industry Specific Analysis

We now consider industry specific labor laws' effects. We use definitions for industries as they appear in the digitalized archived bulletin published by the United States Department of Labor Women's Bureau. These industry classifications appear in the same way in our data source, IPUMS. Our analysis includes four industries mentioned in the bulletin: mercantile, manufacturing, , personal services and mining. To gain more insight on the effect in the manufacturing industry, we divide

it into two subcategories: manufacturing of durable goods and manufacturing of non-durable goods.

We rely on a specification similar to Equation 2.1 but instead we change our definition of treated and control groups to become industry based. A state is considered treated in a given decennial year only if it had one of the three labor laws and the law targets the industry of interest. Our analysis thus allows us to investigate the impact of labor laws on women's employment by targeted industry, and therefore we are able to examine whether there exists a direct effect on women's employment in specific industries using targeted policies. Recall that, regulatory laws include laws that prohibit women from working in specific occupations (specifically related to the mining industry) and regulations that enhance work conditions by banning women from doing specific tasks such as heavy lifting which might have detrimental consequences for their health. Since all occupation bans are related to the mining industry, carrying out an industry specific analysis and examining labor laws' effects by targeted industries is one way to distinguish between regulatory laws and examine the impacts of occupation and within-occupation task bans separately.

We investigate the impact of each labor law and our results are reported in Table 2.3. While women's employment is not significantly higher in states that implemented seating laws in specific industries, we document that women are more likely to be employed in the manufacturing industry (more precisely in durable goods manufacturing) when regulatory laws are implemented. The increase in employment likelihood for women in manufacturing is likely due to the fact that regulatory laws enhanced work conditions for women by forbidding heavy weight lifting for instance. The effect of night work laws vary based on the industries they were implemented in. We find that women are more likely to work in manufacturing (specifically in non-durable goods manufacturing) but less likely to occupy personal services related jobs when laws prohibiting night work shifts for women are introduced. We argue that this decrease in women's likelihood in occupying personal services jobs is due to the nature of this industry that might require night work shifts which are prohibited for women.

In Appendix Table 2.9, we repeat the same analysis to examine the effect of these policies on men's employment in different industries. All our estimates are not statistically significant. These findings provide suggestive evidence that gender focused policies targeting women's employment in specific industries are not crowding out men from specific jobs in targeted industries.

As an additional exercise, we analyse the effect of each type of labor legislation on women's occupational score within industries (i.e. the likelihood to be employed in higher wage occupations within affected industries). We thus focus on each industry separately, and restrict our sample to women occupying jobs in the industry of interest. Our aim is to check whether such policies improve women's position within industries, i.e. whether women are more likely to be occupying high paying jobs. While seating laws that made it mandatory to provide seating accommodations to women improve their work conditions, in Appendix Table 2.10, we show that this policy actually led to a 0.08 standard deviation (SD) decrease in women's occupational score in manufacturing and manufacturing of non-durable goods industries.

Night work laws however increased this score by about 0.02 SD in the personal services industry. We previously showed, in our industry specific analysis, that women's overall employment in personal services jobs decreased after night work laws' enactment. However, our results on women's occupational income score within the personal services industry show that women are now more likely to be employed in higher wage occupations.

2.6.3 Marriage and Fertility Outcomes

In Table 2.4, we present results from estimating Equation 2.1, to examine the impact of labor laws for women on marriage, divorce and fertility outcomes. We refer to these effects as indirect policy impacts, with the direct effect of these policies being on women's labor force participation. Similar to Table 2.2, we first examine seating laws in columns (1) and (2); then regulatory laws in columns (3) and (4) and lastly we report the analysis for night work laws in columns (5) and (6). All three types of labor laws are included in the same regression. Each row displays a different outcome. First, we analyze the impact of gender focused labor laws on the likelihood of marriage for women, then the likelihood of being divorced. To study fertility choices, we focus on the number of children and the age of having the first child for women (i.e. the age of being a first time mother). In columns (1)–(6), we include state and decade fixed effects. We add individual level controls including age, race and ethnicity in columns (2), (4) and (6).

When we examine indirect policy effects, we find that all state law policies have no effect on marriage, divorce or fertility choices. This is suggestive that more direct policies might be more successful in influencing marriage or fertility choices than policies affecting women's labor.

2.6.4 Robustness Checks and Heterogeneity Analysis

Women's Suffrage and Financial Liberation

In this section, we examine the robustness of our analysis to the inclusion of other potential state specific confounding factors. We first consider the passage of suffrage laws enfranchising women at different points in time across difference states. Some states passed suffrage rights for women prior to the 1920 passage of the Nineteenth Amendment, a federal mandate for women's voting rights which made it mandatory for all states. We thus exploit variation in the timing of women's suffrage laws across states and examine whether female labor laws' effects remain robust to controlling for this shock. We rely on data on suffrage laws' passage from Lott and Kenny, 1999 and Miller, 2008.

We also examine the timing of granting property rights and earnings rights to women in the United States. Data on the timing of women's financial liberation by state is obtained from Geddes and Dean, 2002. This can be perceived as a positive shock increasing women's incentive to work.

Table 2.5 shows the effect of state labor laws on our main outcome variables after controlling for women's suffrage and financial liberation. Panel A shows the effect of seating laws on each of our outcomes of interest: women's labor force participation, men's labor force participation, and labor force participation in a triple differences framework. In Panel B, we investigate regulatory laws and Panel C reports the analysis for night work laws. Column (1) of Table 2.5 repeats our findings from column (2) of Table 2.2 with individual level controls and state and decade fixed effects. In column (2) of Table 2.5, we control for women's financial liberation. We then introduce women's suffrage in column (3) and lastly in column (4) of Table 2.5, we simultaneously control for women's suffrage and financial liberation. Comparing our estimates in column (4) to column (1) suggests that our results are robust to the inclusion of these state level shocks.

Electrification and Urbanization

In this section, we account for electricity access in the early 1900s as an important potential confounding factor that could have affected female labor force participation during that era. Electrification can have an impact on female employment and employment characteristics through many channels. It has been argued that electricity access reduces the time burden of housework stemming from the use of electrical appliances and thus results in an increase in female labor force participation (see Gordon, 2016, Greenwood, Seshadri, and Yorukoglu., 2005 and Ramey, 2009). This channel is particularly important in our context.

We rely on electrification data from Vidart, 2020.[7] To examine whether the effect of labor policies implemented by state legislatures differs across states based on their access to electricity during that era, we modify Equation 2.1 to carry out a heterogeneity analysis. Precisely, we add an additional term to our $Seating\ Law_{sd}$, $Regulatory\ Law_{sd}$ and $Night\ Work\ Law_{sd}$ variables, that is the natural log of total capacity in 1919 (total generation capacity in Kilowatt within state boundaries).[8]

Estimates in Panel A of Table 2.6 show the effect of the three types of policies interacted with the natural log of total electrical capacity on our main outcomes. Our differences in differences results suggest that electricity access had no impact on the change in female labor force participation in states that implemented gender based labor laws compared to those that didn't. Additional effects on male labor force participation remain absent. When we carry out our heterogeneous triple differences analysis, depending on electricity access, we document an impact of seating laws on women's labor force participation. Our estimates in column (2) Panel A of Table 2.6 show that a one percentage point increase in electricity access leads to a positive and statistically significant additional effect of seating laws on female labor force participation compared to their males' counterpart, with a magnitude around 1 percentage point. Of note, the average of the natural logarithm of total electrical capacity is close to 12.

Additionally, we argue that urbanization can have an effect on women's labor force status for structural and social reasons. Besides the expansion of the industrial and service sectors in metropolitan areas, urbanization has been seen as one part of the social transformation that has affected the occupational prospects of women.[9] To further investigate the additional effect of urbanization, we interact the $Seating\ Law_{sd}$, $Regulatory\ Law_{sd}$ and $Night\ Work\ Law_{sd}$ variables with the share of urban population at the state level in 1920.

Panel B of Table 2.6 shows the estimates for our interaction terms. Results suggest that a rise in the share of urban population increases the effect of seating laws on women's labor force participation. Moreover, the effect of seating laws policies on women's LFP compared to men's LFP is larger in urban areas. More precisely, we document that a one percentage point increase in the share of urban population leads to an additional effect of around 7 percentage points on women's LFP compared to men's LFP, in treated states compared to control states. The average share of urban population is around 45%.

[7] We thank Daniela Vidart for sharing this data with us. Vidart, 2020 relies on the "Central station directory: a complete list of electric light and power companies with data" (McGraw Publishing Company (1911, 1919)) to build a new dataset in 1911 and 1919 by digitalizing historical documents containing information on 5,409 and 5,631 generators, respectively.

[8] As Table 2.1 shows, almost all state policies were implemented by 1920.

[9] Women's employment was even used as an index of urbanism in Shevsky and Bell, 1955 in their social area analysis.

2.7 Conclusion

This chapter assesses whether public policies, specifically, labor legislations targeting women had direct effects on their likelihood to be in the labor force, to occupy jobs in targeted industries and to be employed in higher wage occupations within affected industries. We examine three types of laws targeting women's labor conditions: regulatory, night work and seating laws. These laws were manifested in regulations that banned women from performing certain tasks that might involve detrimental work conditions, laws that banned them from working at night and regulations that made it mandatory to provide seating accommodations for women.

Our differences in differences estimates provide weak evidence in support of higher women's involvement in the formal labor market after introduction of labor laws by state legislatures in treated states in comparison to control states. However, when we pool our sample to include both men and women and carry out a triple differences analysis, we find that night work laws increased women's labor force participation in comparison to men in treated relative to control states. Our results remain robust to the inclusion of other state level laws granting women voting, property and earning rights.

We document a positive impact on women's employment in the manufacturing industry, and a negative impact on employment in personal services related jobs, when laws prohibiting night work shifts for women in these specific industries are introduced. Our estimates also show a higher likelihood of employment for women in manufacturing related occupations as regulatory laws enhancing industry-specific work conditions for women are enacted. Moreover, we examine whether women are more likely to be employed in higher wage occupations within targeted industries and we find evidence in support of this for the personal services industry.

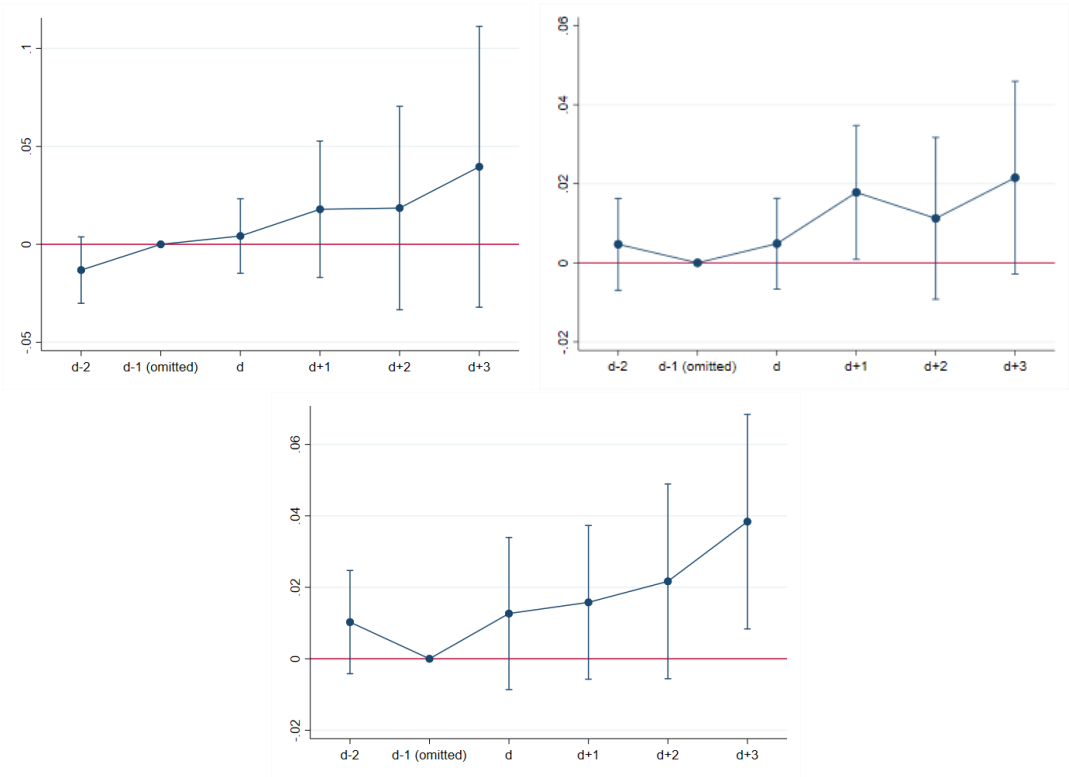
When we carry a heterogeneity analysis, we document a positive and statistically significant impact of seating laws on female labor force participation, in comparison to their male counterparts, with higher electricity access and urbanization level.

Our analysis of policy repercussions on marriage, divorce or fertility choices shows no significant impacts suggesting that more direct policies might be more successful for marriage or fertility choices than policies affecting women's labor.

Our empirical findings make important contributions. First, to our knowledge, we are the first study to formally document the causal impacts of gender focused labor laws favoring women on female labor force participation. Second, we provide suggestive evidence that such labor laws regulating formal labor market conditions in favor of women fail to have significant indirect effects on their marriage or fertility choices. Finally, our results have policy implications endorsing the adoption of labor laws in favor of women to increase women's economic empowerment through higher involvement in the labor market and financial independence, especially in countries where work conditions for women are still detrimental or their involvement is still low.

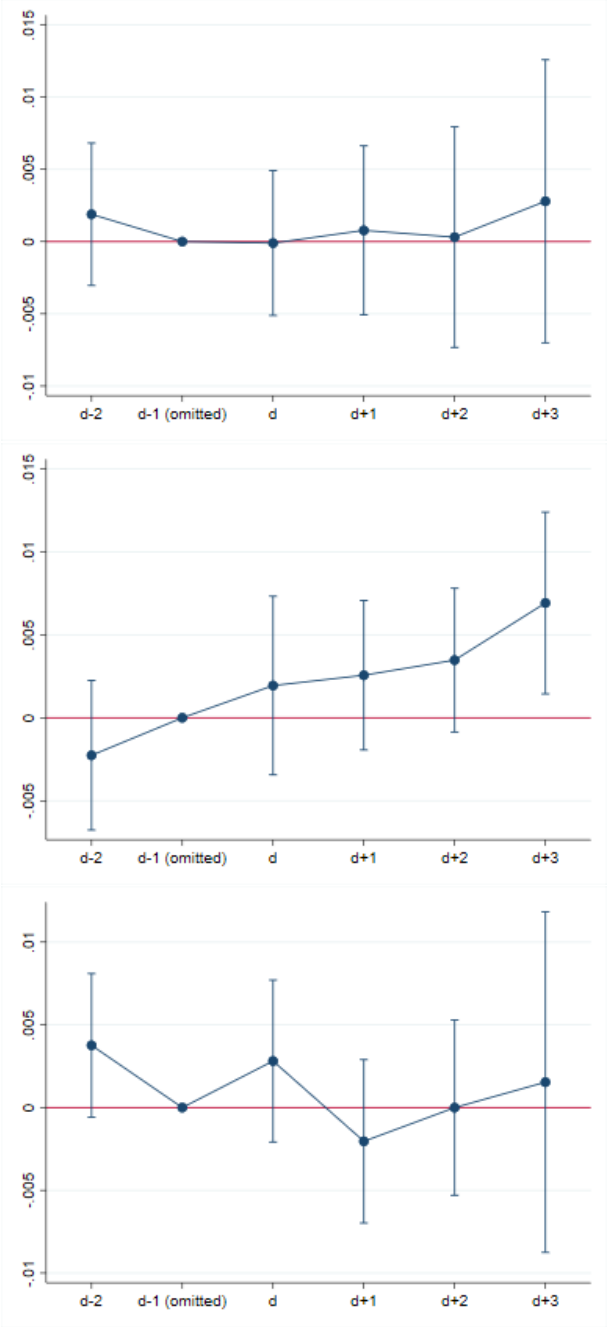
2.8 Figures

FIGURE 2.1: Leads and lags regression on Female LFP



Notes: This figure plots the estimated coefficients of the leads and lags regression on Female LFP for seating, regulatory, and night work laws, respectively. All three policies are included in the same specification. The first decade before policy implementation is omitted from the regression, i.e. the omitted category. *Source*: Individual level data are from Census IPUMS 1% samples for decades between 1860 and 1940.

FIGURE 2.2: Leads and lags regression on Male LFP



Notes: This figure plots the estimated coefficients of the leads and lags regression on Male LFP for seating, regulatory, and night work laws, respectively. All three policies are included in the same specification. The first decade before policy implementation is omitted from the regression, i.e. the omitted category. *Source*: Individual level data are from Census IPUMS 1% samples for decades between 1860 and 1940.

2.9 Tables

TABLE 2.1: Descriptive Statistics: Females' Sample

	1860 (1)	1880 (2)	1900 (3)	1920 (4)	1940 (5)	All (6)
In the labor force	0.144	0.156	0.202	0.237	0.289	0.236
	(0.351)	(0.363)	(0.401)	(0.425)	(0.453)	(0.425)
Age	33	34	34	35	37	35
	(12)	(13)	(13)	(13)	(13)	(13)
White	0.980	0.875	0.886	0.898	0.899	0.896
	(0.141)	(0.331)	(0.318)	(0.303)	(0.301)	(0.305)
Hispanic	0.007	0.007	0.006	0.010	0.015	0.011
	(0.082)	(0.083)	(0.079)	(0.100)	(0.120)	(0.101)
Married	.	0.735	0.728	0.767	0.773	0.759
	(.)	(0.441)	(0.444)	(0.423)	(0.419)	(0.428)
Number of Children	1.980	1.866	1.690	1.536	1.281	1.537
	(2.194)	(2.126)	(2.056)	(1.911)	(1.734)	(1.929)
N	67,512	129,316	203,045	292,626	416,559	1,811,437
Age First Time Mother	24.052	24.213	24.370	24.390	24.503	24.397
	(5.750)	(5.885)	(5.646)	(5.580)	(5.662)	(5.677)
N	40,881	76,769	114,326	163,991	217,901	1,006,066
Nb of States	49	49	49	49	49	49
States: Seating Law	0	0	27	45	47	47
States: Regulatory Law	0	1	12	28	28	28
States: Night work Law	0	0	5	17	20	20

Notes: Individual level data are from Census IPUMS 1% samples for decades between 1860 and 1940 (except for 1890 due to missing census data). Data covers all U.S. states except for Hawaii and Alaska. We report means and standard deviations in parentheses for each variable of interest for the decades 1860, 1980, 1900, 1920, 1940, and all decades combined. The labor force status and individual characteristics are for women between the age of 16 and 65. "Married" is a dummy variable that takes the value of one if a woman over the age of 16 ever got married and zero if she never got married. The "Number of Children" variable counts the number of own children (of any marital status) residing with each individual. "Age of First Time Mother" is computed by subtracting the age of the eldest child from the mother's age. We restrict the sample to the "age of first time mother" values between 15 and 45 to account for measurement errors. Census person weights are used in all calculations. The table also reports the number of states that implemented each policy in a specific decade.

TABLE 2.2: Effects of State Law Policies

	Seating Laws		Regulatory Laws		Night work Laws	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: DID estimates on $StateLaw_{sd}$						
Women's LFP	-0.003	0.001	0.007	0.008	0.015	0.013
	(0.008)	(0.008)	(0.007)	(0.006)	(0.010)	(0.009)
	[0.768]	[0.888]	[0.352]	[0.246]	[0.156]	[0.177]
N			1,811,237			
Men's LFP	-0.001	-0.001	0.003	0.003	-0.004*	-0.004*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	[0.638]	[0.731]	[0.188]	[0.171]	[0.100]	[0.086]
N			1,818,771			
Panel B: DDD estimates on $StateLaw_{sd} \times Women$						
LFP	0.068***	0.067***	-0.007	-0.007	0.040***	0.041***
	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)
	[0.001]	[0.001]	[0.642]	[0.657]	[0.028]	[0.023]
N			3,630,008			
States	49	49	49	49	49	49
Individual Controls	No	Yes	No	Yes	No	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Panel A displays estimates for the State Law Policies for men and women separately. In Panel B, we pool the sample of men and women and show estimates for the interaction between each State Law Policy and a dummy variable "women" that takes the value of one if the individual "i" is a woman. The dependent variables thus are "Women's labor force participation", "Men's labor force participation" and "Labor force participation" respectively. Each row displays a separate weighted estimation and each value displays an estimate for a different labor law. Standard errors in parentheses are clustered at the state-level. Wild cluster bootstrap p-values are reported in square brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). We report estimates for our variables of interest $Seating\ Law_{sd}$, $Regulatory\ Law_{sd}$ and $Night\ Work\ Law_{sd}$ that equal one for states that had any seating, regulatory or night work laws respectively as of decennial year d and zero otherwise. Individual controls include age, race and ethnicity. State and decade fixed effects are included in all regressions. As for the significance of the results: *** p<0.01, ** p<0.05, * p<0.1.

TABLE 2.3: Effects of State Law Policies on Women's LFP in Affected Industries

	Mercantile	Manufacturing All	Manufacturing Durable Goods	Manufacturing Non-durable	Personal Services	Mining
	(1)	(2)	(3)	(4)	(5)	(6)
Seating Laws	-0.003	0.000	-0.000	0.000	-0.005	
	(0.004)	(0.004)	(0.002)	(0.003)	(0.006)	
	[0.696]	[0.968]	[0.923]	[0.906]	[0.424]	
Regulatory		0.002	0.006**	-0.004		0.000
Laws		(0.004)	(0.003)	(0.003)		(0.000)
		[0.596]	[0.155]	[0.247]		[0.414]
Night work	-0.003	0.010***	0.000	0.010***	-0.020***	
Laws	(0.002)	(0.004)	(0.003)	(0.003)	(0.007)	
	[0.343]	[0.009]	[0.923]	[0.007]	[0.036]	
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
N	1,811,238	1,811,238	1,811,238	1,811,238	1,811,238	1,811,238
Nb of States	49	49	49	49	49	49

Notes: The sample is restricted to women aged 16–65. Each column shows estimates from separate weighted regressions for a different indicated industry. The dependent variable is thus women's employment, a dummy variable that takes the value of 1 if a woman is employed in the indicated industry, and 0 otherwise. We report estimates for our variables of interest $Seating\ Law_{sd}$, $Regulatory\ Law_{sd}$ and $Night\ Work\ Law_{sd}$. These are now dummy variables that take the value of one if the state had a labor law targeting the specific industry. Missing estimates indicate that the policy in question did not target the industry of interest. Individual controls include age, race and ethnicity. Standard errors are clustered at the state level. Wild cluster bootstrap p-values are reported in square brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). State fixed effects and decade fixed effects are included in all regressions. As for the significance of the results: *** p<0.01, ** p<0.05, * p<0.1

TABLE 2.4: Effects of State Law Policies on Women's Marriage and Fertility

	Seating Laws		Regulatory Laws		Night work Laws	
	(1)	(2)	(3)	(4)	(5)	(6)
Marriage	-0.003	-0.004	0.004	0.001	0.006	0.003
	(0.003)	(0.003)	(0.005)	(0.005)	(0.004)	(0.005)
	[0.360]	[0.236]	[0.474]	[0.778]	[0.230]	[0.543]
N			1,646,586			
Divorce	-0.000	-0.000	0.002	0.002	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	[0.961]	[0.983]	[0.215]	[0.217]	[0.991]	[0.887]
N			1,121,360			
Nb of Children	-0.073	-0.075	-0.005	-0.005	0.038	0.039
	(0.046)	(0.047)	(0.040)	(0.041)	(0.048)	(0.048)
	[0.176]	[0.167]	[0.909]	[0.906]	[0.479]	[0.471]
N			1,811,238			
Age at First Birth	0.040	0.012	-0.035	-0.015	-0.130	-0.040
	(0.061)	(0.040)	(0.074)	(0.042)	(0.082)	(0.055)
	[0.569]	[0.796]	[0.639]	[0.749]	[0.190]	[0.566]
N			1,005,972			
States	49	49	49	49	49	49
Individual Controls	No	Yes	No	Yes	No	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variables are "Marriage", "Divorce", "Number of Children" and "Age at First Birth". See notes of Table **??** for details of sample and variables. Each row shows estimates from separate weighted regressions for different indicated outcomes. Standard errors in parentheses are clustered at the state level. We report estimates for our variables of interest $Seating\ Law_{sd}$, $Regulatory\ Law_{sd}$ and $Night\ Work\ Law_{sd}$. Each law is a dummy variable that takes the value of one if the state passed the indicated law. Individual controls include age, race and ethnicity. Wild cluster bootstrap p-values are reported in square brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). State fixed effects and decade fixed effects are included in all regressions. As for the significance of the results: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 2.5: Effects of State Law Policies - Women Liberation and Suffrage Controlled

	(1)	(2)	(3)	(4)
Panel A: Seating Law				
Women's LFP	0.001	0.002	0.001	0.002
	(0.008)	(0.008)	(0.008)	(0.008)
	[0.888]	[0.816]	[0.874]	[0.792]
N		1,811,237		
Men's LFP	-0.001	-0.001	-0.001	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)
	[0.731]	[0.580]	[0.726]	[0.589]
N		1,818,771		
LFP	0.067***	0.067***	0.067***	0.067***
	(0.015)	(0.015)	(0.015)	(0.015)
	[0.001]	[0.001]	[0.001]	[0.001]
N		3,630,008		
Panel B: Regulatory Law				
Women's LFP	0.008	0.008	0.007	0.007
	(0.006)	(0.006)	(0.006)	(0.006)
	[0.246]	[0.261]	[0.277]	[0.272]
N		1,811,237		
Men's LFP	0.003	0.003	0.003	0.003
	(0.002)	(0.002)	(0.002)	(0.002)
	[0.171]	[0.247]	[0.173]	[0.147]
N		1,818,771		
LFP	-0.007	-0.007	-0.007	-0.007
	(0.014)	(0.014)	(0.014)	(0.014)
	[0.657]	[0.651]	[0.657]	[0.658]
N		3,630,008		
Panel C: Night work Law				
Women's LFP	0.013	0.013	0.013	0.013
	(0.009)	(0.010)	(0.009)	(0.010)
	[0.177]	[0.247]	[0.170]	[0.244]
N		1,811,237		
Men's LFP	-0.004*	-0.003	-0.004*	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)
	[0.086]	[0.168]	[0.085]	[0.167]
N		1,818,771		
LFP	0.041***	0.041***	0.041***	0.041***
	(0.014)	(0.014)	(0.014)	(0.014)
	[0.023]	[0.023]	[0.026]	[0.021]
N		3,630,008		
Financial Liberation	No	Yes	No	Yes
Suffrage Rights	No	No	Yes	Yes
States	49	49	49	49
Individual Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes

Notes: See notes of Table 2.2. For the Labor Force Participation (LFP) outcome, we show estimates for the interaction between State Law policy and a dummy variable "women" that takes the value of one if the individual "i" is a woman. In addition to previous controls, we control for Women's Suffrage and Women's Financial Liberation. We rely on data on suffrage laws passage from Lott and Kenny, 1999 and Miller, 2008. Data on the timing of women's financial liberation by state is obtained from Geddes and Dean, 2002. We control for women's suffrage by adding a dummy variable that takes the value of one if a given state s had a law allowing women to vote as of a given decade d, and zero otherwise. Similarly for women's financial liberation.

TABLE 2.6: Effects of State Law Policies - Heterogeneity Analysis

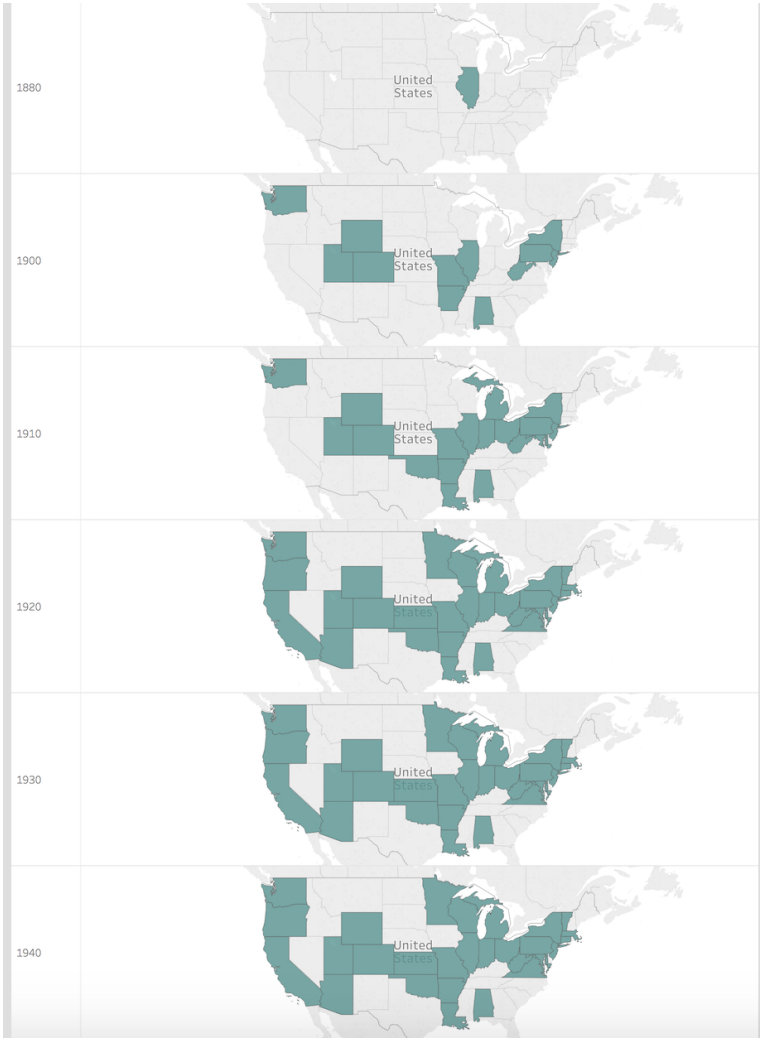
	Seating Law		Regulatory Law		Night work Law	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Log of Total Electrical Capacity						
Women's LFP	0.009	0.010	0.001	-0.001	-0.001	-0.003
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
N			1,811,237			
Men's LFP	-0.000	-0.000	-0.001	-0.001	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
N			1,818,771			
LFP	0.009	0.009	0.015	0.015	0.024**	0.024**
	(0.014)	(0.014)	(0.013)	(0.013)	(0.011)	(0.011)
N			3,630,008			
Panel B: Share of Urban Population						
Women's LFP	0.052	0.062**	0.002	-0.005	0.030	0.010
	(0.036)	(0.029)	(0.027)	(0.028)	(0.071)	(0.068)
N			1,811,237			
Men's LFP	-0.005	-0.005	0.003	0.003	-0.022*	-0.018
	(0.011)	(0.010)	(0.008)	(0.007)	(0.013)	(0.011)
N			1,818,771			
LFP	0.081	0.084	0.055	0.053	0.142*	0.139*
	(0.089)	(0.088)	(0.076)	(0.075)	(0.075)	(0.074)
N			3,630,008			
States	49	49	49	49	49	49
Individual Controls	No	Yes	No	Yes	No	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See notes in Table 2.2. For the Labor Force Participation (LFP) outcome, we show estimates for the interaction between State Law policy and a dummy variable "women" that takes the value of one if the individual "i" is a woman. Each row displays a different estimate for the outcome of interest. In Panel A, we report coefficients on the interaction term between each of our variables of interest *Seating Law*$_{sd}$, *Regulatory Law*$_{sd}$ and *Night Work Law*$_{sd}$ and the natural log of Total Electrical Capacity in 1919. Each law is a dummy variable that takes the value of one if the state passed the indicated law. Total capacity is calculated as the average capacity across counties in a given state. In Panel B, we report coefficients for the interaction term between each of our variables of interest *Seating Law*$_{sd}$, *Regulatory Law*$_{sd}$ and *Night Work Law*$_{sd}$ and the share of urban population in 1920 for each state. Individual controls include age, race and ethnicity.

2.10 Appendix

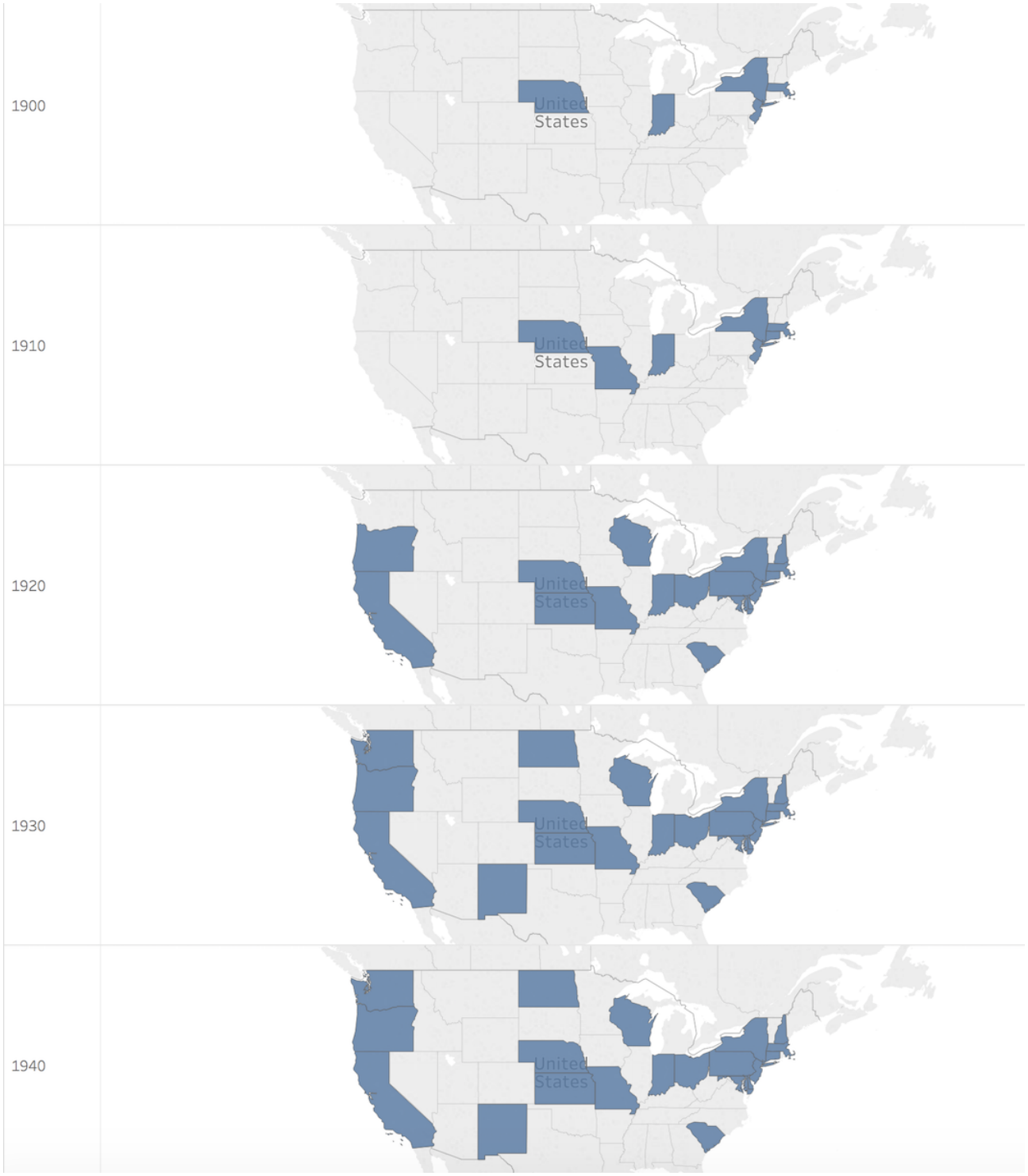
Appendix Figures and Tables

FIGURE 2.3: Timing of passage of regulatory laws



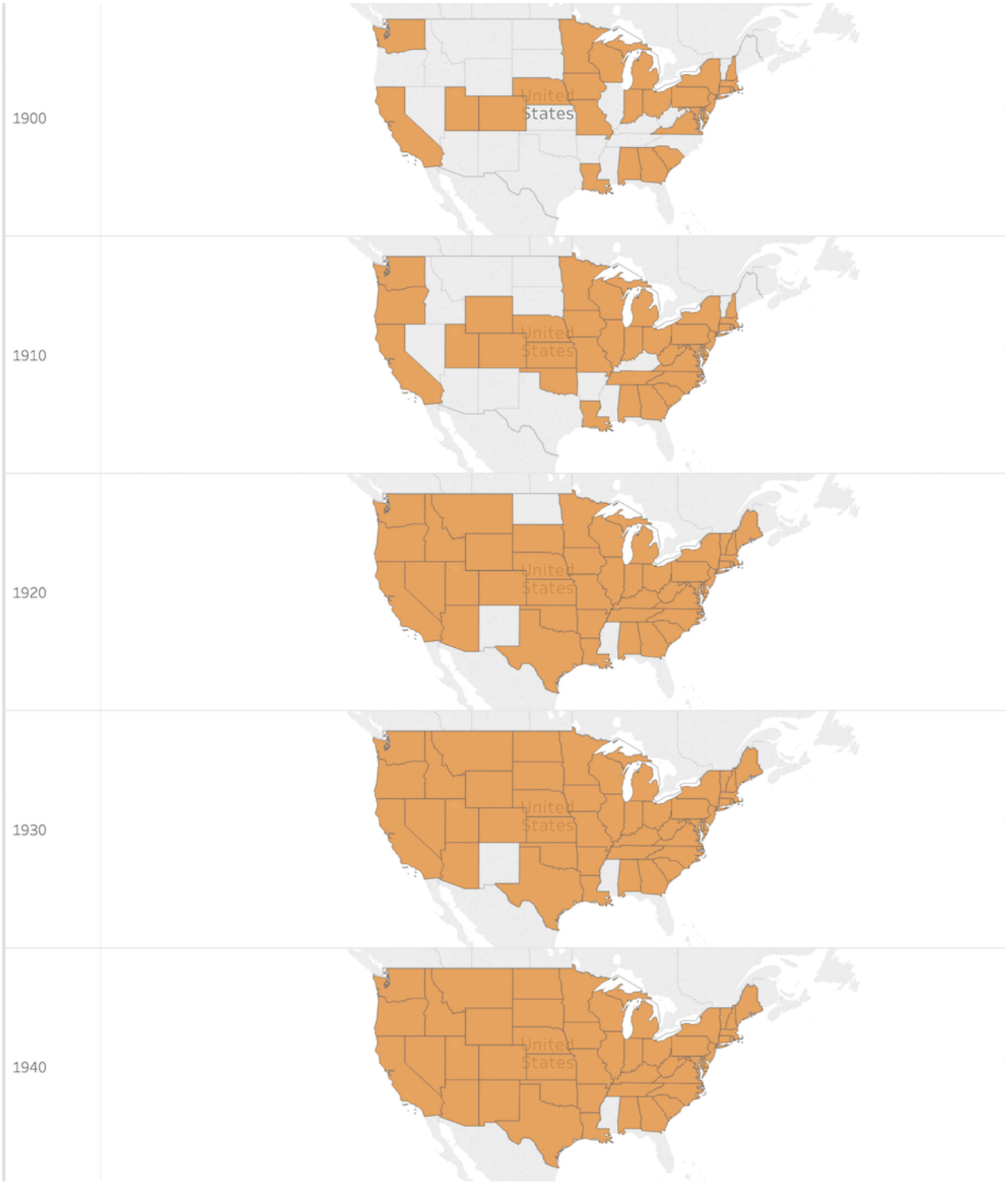
Notes: Chronological timing of regulatory laws' passage across states. Green indicates that a state had a regulatory law. *Source*: Authors' compilation.

FIGURE 2.4: Timing of passage of night work laws



Notes: Chronological timing of night work laws' passage across states. Blue indicates that a state had a night work law. *Source*: Authors' compilation.

FIGURE 2.5: Timing of passage of seating laws



Notes: Chronological timing of seating laws' passage across states. Orange indicates that a state had a seating law. *Source*: Authors' compilation.

TABLE 2.7: Effects of State Law Policies – Excluding Michigan and Montana

	Seating Laws		Regulatory Laws		Night work Laws	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: DID estimates on $StateLaw_{sd}$						
Women's LFP	-0.002	0.001	0.006	0.008	0.017	0.014
	(0.008)	(0.008)	(0.007)	(0.007)	(0.010)	(0.010)
	[0.806]	[0.904]	[0.409]	[0.269]	[0.157]	[0.202]
N			1,742,091			
Men's LFP	-0.001	-0.000	0.003	0.003	-0.004	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
	[0.680]	[0.835]	[0.222]	[0.214]	[0.147]	[0.140]
N			1,744,258			
Panel B: DDD estimates on $StateLaw_{sd} \times Women$						
LFP	-0.001	0.003	0.004	0.005	0.020*	0.017*
	(0.008)	(0.008)	(0.007)	(0.007)	(0.010)	(0.010)
	[0.877]	[0.759]	[0.631]	[0.516]	[0.077]	[0.099]
N			3,486,349			
Seating Law	No	No	Yes	Yes	Yes	Yes
Regulatory Law	Yes	Yes	No	No	Yes	Yes
Night work Law	Yes	Yes	Yes	Yes	No	No
States	47	47	47	47	47	47
Individual Controls	No	Yes	No	Yes	No	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See notes of Table 2.2. We exclude the states of Michigan and Montana from the analysis.

TABLE 2.8: Effects of State Law Policies – With State Trends

	Seating Laws		Regulatory Laws		Night work Laws	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: DID estimates on $StateLaw_{sd}$						
Women's LFP	-0.012*	-0.009	-0.003	-0.001	0.005	0.003
	(0.007)	(0.006)	(0.006)	(0.005)	(0.008)	(0.007)
	[0.099]	[0.160]	[0.590]	[0.821]	[0.564]	[0.686]
N			1,811,237			
Men's LFP	-0.003	-0.002	0.007***	0.008***	0.002	0.002
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
	[0.334]	[0.416]	[0.020]	[0.019]	[0.454]	[0.248]
N			1,818,771			
Panel B: DDD estimates on $StateLaw_{sd} \times Women$						
LFP	-0.002	0.003	0.004	0.005	0.019*	0.016*
	(0.008)	(0.008)	(0.007)	(0.007)	(0.009)	(0.009)
	[0.838]	[0.767]	[0.582]	[0.512]	[0.079]	[0.093]
N			3,630,008			
Seating Law	No	No	Yes	Yes	Yes	Yes
Regulatory Law	Yes	Yes	No	No	Yes	Yes
Night work Law	Yes	Yes	Yes	Yes	No	No
States	49	49	49	49	49	49
Individual Controls	No	Yes	No	Yes	No	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
State Trends	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See notes of Table 2.2. Additionally, we control for state trends.

TABLE 2.9: Effects of State Law Policies on Men's LFP in Affected Industries

	Mercantile	Manufacturing All	Manufacturing Durable Goods	Manufacturing Non-durable	Personal Services	Mining
	(1)	(2)	(3)	(4)	(5)	(6)
Seating Laws	0.001	-0.001	-0.001	0.000	0.000	
	(0.006)	(0.007)	(0.005)	(0.005)	(0.001)	
	[0.810]	[0.874]	[0.774]	[0.966]	[0.817]	
Regulatory		0.0234	0.031	-0.007		0.009
Laws		(0.023)	(0.022)	(0.006)		(0.008)
		[0.446]	[0.377]	[0.272]		[0.244]
Night work	-0.004	-0.025	-0.024	-0.001	0.002	
Laws	(0.004)	(0.021)	(0.020)	(0.005)	(0.001)	
	[0.488]	[0.396]	[0.387]	[0.904]	[0.396]	
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
N	1,818,771	1,818,771	1,818,771	1,818,771	1,818,771	1,818,771
Nb of States	49	49	49	49	49	49

Notes: See notes of Table 2.3. The sample here is restricted to men aged between 16 and 65.

TABLE 2.10: Effects of State Law Policies on Women's Occupational Income Score within Affected Industries

	Mercantile	Manufacturing All	Manufacturing Durable Goods	Manufacturing Non-durable	Personal Services	Mining
	(1)	(2)	(3)	(4)	(5)	(6)
Seating Laws	0.041	-0.076**	-0.044	-0.081**	0.009	0.017
	(0.034)	(0.035)	(0.042)	(0.036)	(0.008)	(0.015)
	[0.256]	[0.083]	[0.371]	[0.064]	[0.316]	
Regulatory		0.0021	0.021	-0.000		-0.002
Laws		(0.017)	(0.023)	(0.015)		(0.123)
		[0.920]	[0.451]	[0.994]		[0.987]
Night work	0.022	0.004	-0.007	0.006	0.024*	
Laws	(0.024)	(0.016)	(0.023)	(0.014)	(0.014)	
	[0.421]	[0.830]	[0.762]	[0.705]	[0.102]	
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes	Yes	Yes
N	60,277	82,316	17,815	64,501	132,145	466
Nb of States	49	49	49	49	49	49

Notes: The sample is restricted to women aged 16–65. Each column shows estimates from a separate weighted regression on women's occupational score (standardized) for different indicated industry. The occupational score variable is a constructed two digit numeric variable that allocates a value representing the median total income (in hundreds of 1950 dollars) to occupations in all years. In each column, we restrict the sample to the indicated industry. We report estimates for our variables of interest $Seating\ Law_{sd}$, $Regulatory\ Law_{sd}$ and $Night\ Work\ Law_{sd}$. Each law is a dummy variable that take the value of one if the state passed a labor law targeting the specific industry. Missing estimates indicate that the policy in question did not target the industry of interest. Individual controls include age, race and ethnicity. Standard errors are clustered at the state level. Wild cluster bootstrap p-values are reported in square brackets. Bootstrap errors are clustered by state (the level at which they are clustered in the cluster-robust variance estimator (CRVE)). State fixed effects and decade fixed effects are included in all regressions. As for the significance of the results: *** p<0.01, ** p<0.05, * p<0.1

Chapter 3

Settlers and Norms

3.1 Abstract

The distinctive traits of early settlers at initial stages of institutional development may be crucial for cultural formation. In 1973, the cultural geographer Wilbur Zelinsky postulated this in his doctrine of "first effective settlement". There is however little empirical evidence supporting the role of early settlers in shaping culture over the long run. This chapter tests this hypothesis by relating early settlers' culture to within state variation in gender norms in the United States. I capture settlers' culture using past female labor force participation, women's suffrage and financial rights at their place of origin. I document the distinctive characteristics of settlers' populations and provide suggestive evidence in support of the spatial (across locations) and vertical (over time) transmission of gender norms. My results show that women's labor supply is higher, in both the short and long run, in U.S. counties that historically hosted a larger settler population originating from places with favorable gender attitudes. My findings shed new light on the importance of immigrants' characteristics and their countries/states of origin for cultural formation in hosting societies.

3.2 Introduction

Cultural variation is pervasive both within and across countries and is known to correlate strongly with economic and political development.[1] According to the cultural geographer Wilbur Zelinsky, "the dominant culture of a given nation is determined by the characteristics of the first group of settlers who came to an empty territory regardless of how small the initial band of settlers might have been".

Zelinsky, 1973's doctrine of "first effective settlement" argues that "the activities of this first group of people matter much more for the cultural geography of a place than the contribution of tens of thousands of new immigrants a few generations later". This is consistent with theories of persistence (via horizontal/spatial and vertical/across generations transmission of norms), path dependence and how initial conditions at critical junctures of institutional development play an important role in shaping social norms and attitudes in the short and long run (Alesina and Giuliano, 2015; Bisin and Verdier, 2017; Tabellini, 2008).

To understand the role of immigrants in shaping the cultural and institutional development of settler societies, I partially revisit Zelinsky's doctrine and focus on settlers' culture as one key set of their characteristics. The aim is to evaluate the role

[1] Nunn, 2020 for instance presents a detailed overview on the short and long run determinants of cultural traits.

of early settlers' culture in explaining within state variation in gender norms in the United States.[2]

To do so, I focus on county creation events and examine the culture of the population that inhabited counties at the time of their territorial government creation around the "Age of Mass Migration". This refers to the era of massive influx of diverse migrants to the United States between 1850 and 1940. Focusing on county creation events and restricting to U.S. counties created between 1840–1940 is informative for a number of reasons. First, this allows me to capture counties at their early stages of cultural and institutional development. Second, the era of mass migration provides an adequate setting for both across and within state variation in settlers' composition as a result of the diverse and heavy migrant inflows to the United States during that period.

This chapter focuses on gender norms given the large disparities in beliefs and values regarding women's role in society both across and within states. Survey based measures like the General Social Survey (GSS)[3] capturing respondents' views on gender issues in the United States are revealing of the great differences in gender roles and attitudes. Moreover, by focusing on gender norms, I am able to provide suggestive evidence in support of potential mechanisms related to gender values and belief formation and evolution in U.S. counties. These mechanisms relate to gender attitudes at the place of origin of settlers (See Sections 3.4 and 1.4 for details).

This research provides a framework that allows for the empirical examination of the doctrine of "first effective settlement". To investigate cultural formation in settler societies, I consider the context of the United States and capture the settlers population using information about the people that lived in U.S. counties around the time of their creation. Settlers include foreign-born, out-of-state and in-state born individuals. The data on the settler population is derived from the first U.S. Census available after the county creation date. I then explore the composition of this population using information on settlers' demographic characteristics, their birthplace, and most importantly gender related characteristics at their place of origin. This is because settlers' culture is proxied with values and beliefs from the country/state of origin.[4] The underlying assumption is that settlers internalize their culture before migrating to new places (i.e. there is a correspondence between settlers' culture and the dominant culture in their sending country/state). I provide suggestive evidence in support of this assumption in Section 3.8.

My main sample of newly established counties includes those that were not partitioned or subdivisioned from previously formed county(ies), but that were created from non-county areas. I refer to those as "new" counties. In an alternative analysis, I use counties that were partitioned from already settled places and other type of counties that were not created from non-county areas as a placebo treatment. I refer to those as "partitioned" and "other" counties. The rationale for focusing on

[2]See Giuliano, 2020 for an excellent review of the literature on determinants and persistence of gender norms.

[3]GSS asks respondents' views for example about the following: "It is better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and the family".

[4]People moving to U.S. counties from different places might be exposed to a different set of norms at their place of origin/birth, including gender related ones. Distance travelled by these settlers is not the intended underlying characteristic to be examined in this chapter. Von Berlepsch and Rodríguez-Pose, 2019 for instance exploit distance travelled and distinguish between internal migrants (what I refer to as out-of-state born individuals) and external migrants (foreign born individuals) in their examination of the impact of migrants on counties' long run economic development. Unlike Von Berlepsch and Rodríguez-Pose, 2019, I exploit migrants' culture using gender norms at their place of origin/birth as the underlying variation and not the distance travelled.

"new" counties is that these might be different from those that are subdivisions of already formed counties with regards to how established and developed the county, community, society, culture and institutions are. Thus, "new" counties better reflect the "empty" territories that Zelinsky, 1973's doctrine refers to. A major difference between "new" counties and "partitioned" and "other" counties is the fact that the latters were densely populated.

I exploit the data on settler population and provide a descriptive analysis which offers novel insights on the characteristics of early settlers living in U.S. counties around their creation time. I document that settlers were mostly literate men in their prime age. The majority of settlers were out-of-state born migrants followed by in-state and foreign borns. Examining settlers' characteristics by gender, as well as by origin and by gender, I find that male settlers were younger and more likely to be literate and single, in comparison to women settlers, especially among foreign born and out-of-state born individuals. With regards to settlers' culture, foreign born settlers came mostly from countries with high FLFP and out-of-state born individuals came from states where women had property and earning rights.

When I examine the role of settlers' culture in explaining within state variation in female labor force participation, my findings provide suggestive evidence in support of "cultural continuity via portability" of norms (spatial/horizontal) and transmission of norms over time (across generations/vertical) for my main sample of "new" counties. I document a positive and statistically significant relationship between female labor force participation in newly established U.S. counties in the short and long run and settlers' culture proxied by past female labor force participation, women's suffrage rights passage and financial liberation in their place of origin, as well as intensity measures capturing the length of time since the passage of these rights. I show that by restricting to "partitioned" and "other" counties, this relationship does not hold.

I test whether individuals currently living in U.S. counties that historically hosted a larger share of settlers with liberal gender attitudes, have pro-women working and achieving outside the home attitudes. I rely on data from the General Social Survey (GSS) and show that residents in counties with higher shares of early settlers from places with high FLFP, and from places where women could vote and had financial rights, are more likely to approve on women working and taking care of the country not just the home. This result suggests that liberal gender attitudes persisted in these counties.

This chapter contributes to several strands of the literature. First, I contribute to the literature on the roots and persistence of cultural traits, specifically gender norms (Alesina, Giuliano, and Nunn, 2013; Algan and Cahuc, 2006; Ashraf and Galor, 2011; Becker and Woessmann, 2008; Campa and Serafinelli, 2019; Grosjean and Khattar, 2019; Hansen, Jensen, and Skovsgaard, 2015; Nunn et al., 2014; Teso, 2019). Natural experiments in history affecting sex ratios, historical agricultural practices, historical institutions including religion and family structures, are documented as crucial determinants affecting gender norms formation. I contribute to this literature by showing that the cultural traits of settlers at early stages of institutional and cultural formation have lasting impacts on the prevailing culture. One relevant study is perhaps Bazzi, Fizbein, and Gebresilasse, 2018 that revisits Turner's thesis which argues that the American frontier fostered individualism in the United States. The chapter documents a more pervasive individualism and a greater opposition to redistribution in U.S. counties with greater frontier experience. Frontier locations had distinctive demographics and greater individualism.

Second, this chapter contributes to the literature on immigrants, immigrants' assimilation and gender norms (Alesina and Giuliano, 2010; Antecol, 2000; Blau, 1992; Blau, Kahn, and Papps, 2011; Blau and Kahn, 2015; Blau et al., 2020; Fernandez, Fogli, and Olivetti, 2004; Fernandez and Fogli, 2009; Fortin, 2005). A strand of literature examining culture and gender norms relies on an epidemiological approach which aims at seperating the impacts of culture from those of institutions and economic environments. This approach relies on the descendants of immigrants arguing that the latters transmit the values and beliefs of their country of origin in an institutional environment that is the same across all different immigrant groups (Fernandez, 2011). This chapter considers the first iteration of immigrants, which themselves played an important role in shaping the institutional environment that studies relying on epidemiological approaches treat as constant when examing immigrants types.
Third, my results support the theoretical models that highlight the importance of initial conditions in determining the long-run equilibrium as well as the modes of transmission (Akerlof and Kranton, 2000; Bisin and Verdier, 2011; Shayo, 2009) and to the emerging quantitative research that shows that culture matters for economic outcomes (Algan and Cahuc, 2006; Barro and McCleary, 2003; Fernandez and Fogli, 2009; Guiso, Sapienza, and Zingales, 2009; Giuliano, 2007; Tabellini, 2010). My results provide evidence in support of horizontal/spatial and vertical/over time (across generations) transmission of norms. Finally, this research sets the stage for future research to look at a host of other cultural traits.

The rest of the chapter is structured as follows. In Section 3.3, I provide a brief historical background on the process of county creation in the United States. Section 3.4 presents the conceptual framework. In Section 3.5, I describe my sample of U.S. counties and the various data sources I use to construct the settler population, settlers' characteristics and my outcomes of interest. Section 3.6 outlines my empirical strategy. In Section 3.7, I discuss my results and, then present a battery of robustness checks in Section 3.8. I briefly conclude in Section 3.9.

3.3 Historical Background

County creation events provide an adequate setting to focus on counties at early stages of their community, societal, cultural and institutional development. In this section, I provide a brief overview on the process of territorial expansion in the United States, as well as state incorporation and county creation.

On July 4th 1776, the United States of America was created out of the Thirteen British colonies[5] which declared their independence from the Kingdom of Great Britain and proclaimed themselves as free and independent states. It was not until 1873 with the Treaty of Paris, which put an end to the American Revolutionary War, that their independence was recognized by Great Britain.

The United States of America evolved from the Thirteen Colonies to its current form as a result of the following five largest territorial expansion events[6]. The first was the Louisiana Purchase (1803) which was a massive land purchase constituting almost 25% of the present day U.S. covering land from New Orleans up to Montana and North Dakota. The Adams-Onis Treaty or the Florida Purchase Treaty (1819) put

[5]The Thirteen British Colonies became New Hampshire, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Pennsylvania, Maryland, Delaware, Virginia, North Carolina, South Carolina, and Georgia states.

[6]Appendix Figure 3.9 displays these expansion events.

an end to lengthy negotiations between the U.S. and Spain, officially transferring Florida to the United States. The third largest territorial expansion was the Texas Annexation (1845) resulting in the annexation of the Republic of Texas, declaring its independence from Mexico and transferring it to a U.S. state that was admitted to the Union. In 1848, the Mexican Cession encompassed the region that Mexico ceded to the U.S. as a result of the Treaty of Guadalupe Hidalgo after the Mexican-American war. Finally, the Alaska Purchase in 1867 resulted in the acquisition of Alaska from the Russian Empire by the United States.

The Congress of the Confederation, known as the United States in Congress Assembled, had governing authority over the United States. Its authority was granted by The Articles of Confederation and Perpetual Union which was the first Constitution of the United States (an agreement among the 13 original states). The Congress of the Confederation enacted two key ordinances: the Land Ordinance of 1784 and Northwest Ordinance of 1787. These two ordinances organized the creation of territorial governance and dictated the protocols for state admission to the union, the division of land into administrative units and public use of land. The Land Ordinance (1784) was a standardized system for settling and selling land allowing frontier migrants moving westward to acquire land through direct sales from the federal government via the Public Lands Survey System (PLSS) of grids of square townships for the distribution and sale of land in definable parcels as a commodity. The Northwest Ordinance (1787) created the Northwest territory, the first organized incorporated territory of the U.S. beyond the thirteen original colonies.

The U.S. territorial expansion westward happened gradually and was largely driven by population pressures and external geopolitical forces (Gallman et al., 1972). The westward expansion however did not occur peacefully. With the arrival of more explorers, and as new settlers moved in, Native American tribes, previously occupying the west, were displaced and lands were violently taken from them. Treaties forced millions of Native Americans onto reservations which were then frequently broken leading to even larger shares of lands being acquired by settlers.

United States territories were administrative divisions overseen by the U.S. government, but they were not sovereign entities like U.S. states. They included both organized incorporated territories where governance was dictated through an organic act and that constituted integral parts of the United States (i.e. full constitutional rights were applicable) and unincorporated territories which were not integral parts of the United States (i.e. only partial application of the Constitution). The process of incorporation was under the authority of the U.S. Congress. The Admission to the Union Clause of the United States Constitution (preceded by the two ordinances), dictated how these territories would be admitted to the Union as U.S. states. A total of 31 out of 37 states admitted to the Union by Congress were established within U.S. organized incorporated territories. Sometimes an entire U.S. territory became a state and sometimes just part of it.

The Northwest ordinance (1787) authorized county creation by proclamation of the governor until the organization of the territorial general assembly, and thereafter by the latter. U.S. counties constituted administrative or political subdivisions of a state. In an organized incorporated territory (not yet granted statehood), the territorial legislative assembly had the authority to create counties. For example, Arizona territory established by the Arizona Organic act enacted the creation of Arizona's first four counties (Mohave, Pima, Yavapai and Yuma counties). Thus, U.S.

counties were in some cases formed prior to statehood.[7] In U.S. states (organized incorporated territory admitted to the Union or U.S. states not established within U.S. organized incorporated territories), county creation was under the authority of the state specific General Assembly of Senate and House representatives and conditions for county creation were dictated by state constitutions.[8] Appendix Figure 3.11 displays an act enabled by Alabama state to establish a new county as a subdivision of previously formed counties.

As counties are newly formed, and given that settlement is at its early stages, settlers who first inhabited these territories may influence the formation of local economic, social, political and cultural, both formal and informal, institutions in a way that shapes the social fabric of that given county.

3.4 Conceptual Background

Wilbur Zelinsky is an American cultural geographer with many geographical studies on the American popular culture. He famously argued in one of his books (Zelinsky, 1973) that the first settlers significantly impact the dominant culture of a given nation. His doctrine of "first effective settlement" (also known as the Zelinsky, 1973 doctrine) is a theory in cultural geography that served as basis for future theories linking American history to present day events. Woodard, 2011 for instance expanded the doctrine of "first effective settlement" and argued that the movement of people to new territories, bringing with them the culture of the society where they came from, resulted in the creation of multiple nations which together constitute the country. These multiple American nations are thus culturally segmented parts, each composed of a group of people that share a common culture and origin defined beyond legal states and international boundaries.

Woodard's argument, inspired by the doctrine of "first effective settlement", relates directly to cultural formation in settler societies. The question that arises from it is the following: did the migrants that moved to U.S. counties, in their early creation, carry their values and beliefs from the societies where they came from to the areas they moved to, affecting thus the way in which institutions and culture were formed in these newly settled places? In other words, the question is whether these early settlers shaped culture in a way that mirrored the culture of their country of origin or whether a new culture was formed in newly established U.S. counties. The implications of Woodard, 2011 and the Zelinsky, 1973 doctrine suggest that the culture of settlers of newly created U.S. counties has a lasting impact on the culture formed in these areas.

This hypothesis relates to both the spatial (horizontal/cultural continuity via portability) and the transmission of cultural beliefs over time (vertical/across generations). One possible outcome is that these settlers carried their cultural beliefs from their home country/state and moved to U.S. counties and shaped a culture that mirrors their home country/state culture and this persists to the present day. This would validate both the horizontal and vertical aspects of transmission of norms and values in newly formed U.S. counties. Another possible outcome is that these

[7] Appendix Figure 3.10 for instance displays the territorial act in 1818 enacted by the territorial legislature of Alabama which established Marengo county.

[8] Texas' state constitution for instance dictated that "no existing county could be reduced to less than 900 square miles without the consent of a two-thirds majority of the Legislature. In addition, the Legislature could continue to create counties without consent of the residents living on the land area being considered." Other conditions imposed that for new counties to be formed in Texas state from unorganized lands, these must be at least 900 square miles.

settlers moved to U.S. counties and shaped a culture that mirrors their home country/state culture but this did not persist over time/across generations. This would thus validate the horizontal transmission of norms and values only. Lastly, settlers may have arrived to U.S. counties and formed a "new" culture. This means that both horizontal and vertical transmissions of cultural beliefs are absent in newly formed counties.

3.5 Data

In this section, I describe the data sources for my sample of U.S. counties, settler population and characteristics, and other variables of interest. I also provide some detailed descriptive statistics.

3.5.1 Data on U.S. Counties

I focus on county creation events to capture counties at their early stages of cultural and institutional development. I disregard counties created pre-1840 and post-1940 and limit my analysis to U.S. counties formed between 1840–1940 for two reasons. First, the time period falls within the era of mass migration which provides an adequate setting for both across and within state variation in settler composition as a result of the diverse and heavy migrant inflows to the United States during that period. Second, given that full count U.S. Censuses are available only between 1850–1940, and counties are not identified in public-use microdata from 1950 onwards, I will not be able to examine the composition of settlers residing in newly formed U.S. counties any time before or after that period.

In order to construct my sample of U.S. counties created between 1840–1940, I rely on the ATLAS of Historical County Boundaries dataset[9] which offers information about the creation of every U.S. county, as well as the changes in administrative status, size (land area in square miles), shape and location of these counties.

Focusing on county creation events results in a total of 1,494 U.S. counties created between 1840 and 1940 (See Figure 3.1). About 75% of these counties were created before 1900 and about 57% are in the West and Midwest Census region, 41% in the South and the remaining counties are in the Northeast Census region. Due to the lack of data on inhabitants of 153 counties created between 1880 and 1889, these counties are excluded from my sample. This is because Census data is missing for 1890 from all sources. Figure 3.2 displays the chronological timing of U.S. counties' creation.

The 1,494 U.S. counties created between 1840 and 1940 include counties that were not subdivisioned or partitioned but that were created from non-county areas. I refer to those as "new" counties. Figure 3.3 displays these "new" U.S. counties. About 56% of these counties are in the Midwest Census region, about 25% are in the South Census region and the remaining 19% are in the West Census region. Close to 83% of these counties were formed before 1900 and the remaining 17% were formed sometime after 1900.

Additionally, new counties were created from a subdivision of a previously formed county or as a result of a combination of many established formed counties. I refer to those as "partitioned" counties. Finally, there are counties that were created from a combination of districts and non-county areas, those already created under territorial juridistriction which then changed from an organized incorporated territory to a

[9] Data is available from the following website: https://publications.newberry.org/ahcbp.

U.S. state, those created under a given territorial juridistriction which then came under another territorial juridistriction and counties created as a result of the passage of a new constitution converting all judicial districts to counties. I refer to those as "other" counties. Figure 3.4 displays "partitioned" and "other" counties.

In this chapter, my main analysis is limited to the sample of "new" counties, i.e. those that were not subdivisioned or partitioned but that were created from non-county areas. These sum up to 436 counties. The main sample excludes "partitioned" and "other" counties. This is crucial given that counties that are subdivisions of previously created counties could be different from "new" counties with regards to how established the county, community, society and institutions are. While the average number of inhabitants for "new" counties was 6,264 and 7,110 for "partitioned" and "other" counties, population density per square mile was more than 3 times larger in the latter group of counties (population density of 48.8) in comparison to "new" counties (population density of 12.4). Population and population density data are extracted from the first U.S. Census available post county creation date.

To fix ideas, consider Bullock county (Appendix Figure 3.11) for example, founded in 1866, in Alabama state. It was created as a combination of four previously established counties (Macon, Montgomery, Pike and Barbour). Bullock county is thus a "partitioned" county and it is excluded from my main sample of "new" counties.

3.5.2 Data on the Settler Population

In the following subsection, I describe how I construct my settlers' population as well as the data sources and variables I use to examine the composition of this population. I also provide a descriptive analysis which offers new insights on the characteristics of settlers living in newly created U.S. counties. I present summary statistics for my entire sample of settlers. I also report statistics by gender and by gender and category (foreign born, out-of-state and in-state born settlers) simultaneously.

Settler Population and Demographic Characteristics

I define the population of settlers as the people that inhabited U.S. counties at the time of the creation of their territorial government. Having information about the year of creation of each U.S. county, I construct the settler population using information about county identifiers from full count Census data. Specifically, I build a dataset of people living in these U.S. counties by relying on the first U.S. full count Census available after the county creation date. In Section 3.8, I examine the validity of settlers' definition given that Census data is not available for areas before they become politically organized as an administrative entity of the United States.

I rely on the complete count United States Census data (1850–1940) from the Integrated Public Use Microdata Series (IPUMS).[10] IPUMS provides access to U.S. Census microdata and includes a wide range of information about individual's education/literacy, labor force and fertility status, income and occupational score among other information. I carry out my analysis at the county level, so I generate county averages based on individual characteristics. County identifiers allow me to identify the county where the household was enumerated, and more importantly where

[10]Data is available from the following website: https://usa.ipums.org/usa/. Note that Census data is missing for 1890 from all sources. Thus, data on settlers of counties created between 1880 and 1889 is missing and therefore these counties are excluded from my sample.

individuals are residing. I generate the county level average age and gender composition of settlers, as well as their marital, fertility and literacy status.

Additionally, given that people coming from different places are exposed to a different set of values and beliefs, it is crucial to identify the country/state of origin of these settlers. To do so, I rely on a variable available from IPUMS which indicates the U.S. state or foreign country where the person was born. Using information about the birthplace of individuals allows me to divide the settler population into three different categories. The first category comprises foreign born individuals, i.e. those who were born in a country different from the United States. The second category includes out-of-state born individuals, i.e. those who were born in a U.S. state that is different from the state in which the household was located when the Census enumerator conducted the interview. Finally, in-state born individuals are those born in the same state as the one where the household is located.

In Table 3.1, I provide summary statistics of the characteristics of settlers living in my sample of "new" U.S. counties. I present statistics for my entire sample of settlers in column (1) of Table 3.1. I also report statistics by gender in columns (2) and (3). Figures from column (1) show that settlers that occupied newly formed counties were mostly literate men in their prime age. Settlers were mainly out-of-state born migrants (62%) followed by in-state borns (22%). Foreign born individuals constitute 15% of settler populations. Figure 3.5 illustrates the distribution of foreign born, out-of-state and in-state migrants out of the entire population respectively across my main sample of "new" counties. Appendix Figure 3.12 shows this distribution for the alternative sample of "partitioned" and "other" counties.
Columns (2) and (3) of Table 3.1 show that male settlers were more likely to be in their prime age, literate and single in comparison to female settlers. In Section **??**, I report these statistics for the entire sample of U.S. counties created between 1840–1940. The majority of male and female settlers were out-of-state born individuals, followed by in-state borns.

Table 3.2 repeats these descriptive statistics by gender for foreign born, out-of-state and in-state born settlers seperately in columns (1) and (2), (3) and (4), and (5) and (6) respectively. This table shows that among the population of foreign born settlers and out-of-state born settlers, men are more likely to be in their prime age, literate and single in comparison to women settlers. The characteristics of settlers by gender are similar between men and women with respect to the literacy level and distribution by age for in-state born individuals.

Settlers' Culture

To capture settlers' culture, I use various proxies that reflect values and beliefs from their place of origin. The underlying assumption is the correspondence between settlers' culture and the dominant culture in their sending country/state. I use a series of quantitative variables including female labor force participation, women's suffrage rights and women's financial liberation, rather than simply using the country or state of birth as a proxy variable for gender norms at the place of birth.

Specifically, I build three quantitative variables. The first captures female labor force participation by country of origin by decade. Second, I explore the chronological implementation and passage of women's suffrage rights across U.S. sending states and countries. Lastly, I exploit variation in the passage of women's financial rights as well as their intensity (number of years since passage of these rights) across U.S. sending states.

I construct a dataset of historical female labor force participation for sending countries (countries of birth) for foreign born settlers using a combination of at least three different sources.[11] I rely on data from the Integrated Public Use Microdata Series (IPUMS) International Historical Censuses.[12] I combine this with information on female labor force participation by country by decade extracted from (Olivetti, 2014) and (Olivetti and Petrongolo, 2016). The optimal choice of decade from which to construct historical women's labor force participation by sending country is not obvious. In this chapter, I use source countries' labor force participation from the same decade, or a decade or two earlier, depending on data availability, relative to when I capture the population of foreign born settlers. Given that data on the migration date of settlers is not available, I measure source countries' characteristics based on when I observe settlers, i.e. the same decade, or a decade or two earlier, depending on data availability, before county creation date. The assumption is that the cultural beliefs of these foreign born settlers are best reflected in what their counterparts were doing at the time in the country of origin (Fernandez and Fogli, 2009).

I aim to extract data on women's labor force participation from the same decade, or a decade or two earlier, relative to when I study the population of out-of-state born settlers. Such data is unavailable at the time for counties that had not yet been created or that were recently created. This means that using female labor force participation from sending states to capture out-of-state born settlers' cultural beliefs is not always possible.

Summary statistics related to settlers' culture show that foreign born migrants came mostly from places with high FLFP. In Table 3.3, I document that 54% of foreign born individuals from countries with known female labor force participation are from countries with *above* decade specific median female labor force participation.[13] Figure 3.6 shows the distribution, across my main sample of "new" counties, of the share of foreign born settlers from countries known to have *above* median female labor force participation. Appendix Figure 3.13 displays this distribution for the alternative sample of "partitioned" and "other" counties.

My second quantitative measure to proxy for settlers' gender norms explores the chronological implementation and passage of women's suffrage rights across U.S. sending states and countries, relative to county creation date. I rely on the variation in the timing of passage of suffrage laws enfranchising women at different points in time across different U.S. states, given that some states passed suffrage rights for women prior to the passage of the federal mandate (the Nineteenth Amendment of 1920). The data on passage of suffrage laws for U.S. states is obtained from Lott and Kenny, 1999 and Miller, 2008. I also collect data on the timeline of women's suffrage across countries. Table 3.3 shows that 16% and 0.3% of the foreign born settler and out-of state settler population respectively came from places where women could vote. Figure 3.7 displays the distribution of these shares across my main sample of "new" counties and Appendix Figure 3.14 shows this distribution for the alternative sample of "partitioned" and "other" counties.

[11] I also do a thorough web search using countries' official sources to append otherwise missing data on women's labor force participation. Foreign born individuals from countries with missing information on their historical labor workforce are excluded from the population of foreign born settlers with known FLFP. These are instead classified into a population of foreign born settlers with unknown FLFP. Foreign born settlers from countries with known and unknown FLFP add up to the entire foreign born settler population.

[12] Data is available from the following website: `https://international.ipums.org/international-action/samples`.

[13] Of note, using the mean instead of the median does not affect the analysis.

I also compute a measure of suffrage intensity, which is a weighted share of settlers coming from places where partial/full voting rights were granted to women, weighted by the number of years between the passage of the relevant suffrage law and the year of county creation. The mean of this variable for foreign born migrant is close to 6, and it is 0.07 for out-of-state born individuals.[14]

The third quantitative measure proxies for out-of-state born settlers' gender norms using the variation in the intensity and timing of passage and implementation of women's financial liberation, relative to county creation date. This measure explores the timing of granting of property and earnings rights to women across U.S. sending states. The data on the timing of women's financial liberation by state is obtained from (Geddes and Dean, 2002). Table 3.3 shows that almost 33% of out-of-state born settlers came from U.S. states where women had property and earnings rights (See Figure 3.8 for the distribution of this share across my main sample of "new" counties and Appendix Figure 3.15 for the sample of "partitioned" and "other" counties).

The measure of financial liberation intensity is generated using the share of out-of-state born settlers coming from U.S. states where property rights and earnings rights were granted to women, weighted by the number of years between the passage of financial liberation laws and the date of county creation. The mean of this variable is 6.7 as displayed in Table 3.3.

3.6 Empirical Strategy

In this section, I describe in detail my empirical strategy, the main specifications and the list of control variables I include in my estimations.

The objective is to examine the role of settlers' culture in explaining within state variation in gender norms in the United States. The analysis is carried at the county level for my sample of U.S. counties created between 1840–1940. While U.S. counties were created at different points in time, variation in the timing of county creation is not relevant for this research. Instead, because I control for decade of county creation fixed effects, the only relevant variation is the composition of settlers of these counties, i.e. the composition of the first inhabitants of these counties after their creation. Moreover, throughout the analysis, I also include state fixed effects given that the purpose is to compare counties within the same state. I thus estimate the following specification using Ordinary Least Squares (OLS) estimation:

$$y_{csd} = \alpha + \tau Settlers' Population_{csd} + \beta Settlers' Culture_{csd} + X'_{cs}\gamma + \theta_{FE(s)} + \phi_{FE(d)} + \varepsilon_{csd} \tag{3.1}$$

where y_{csd} is the county level female labor force participation. For the short (long) run analysis, data on female labor force is extracted from the first (tenth) U.S. full count Census available after county creation date. Settlers' Population$_{\mathbf{csd}}$ is the distribution of foreign, out-of-state and in-state born individuals out of total population. **Settlers' Culture$_{\mathbf{csd}}$** is my independent variable of interest. Data on the settler population is based on the first U.S. full count Census available after the county creation date. X'_{cs} includes a list of county level geographic controls such as latitude, longitude, mean county temperature and rainfall, elevation, distance to lakes and rivers from the county centroid and average potential agricultural yield. To these,

[14] The sufffrage intensity measure is much larger for foreign born settlers and this is likely driven by the fact that those came from places where at least partial rights were granted to women long before U.S. county creation.

I add a set of demographic controls including the share of prime age population, share of literate population, the sex ratio computed as the ratio of the male over the female population, the share of the single population and the child to women ratio computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000. Finally, I also include additional controls that capture counties' geography and isolation such as terrain ruggedness, rainfall risk, the distance to the nearest portage site, the distance to the nearest Indian battle site, the distance to the coast, the number of years that the county has been intersected by railroads since its creation date and the distance to the nearest mineral discovery site. θ_s and ϕ_d are state and decade of county creation fixed effects respectively. ε_{csd} is the error term. My standard errors are clustered on 60-square-mile grid cells (Bester, Conley, and Hansen, 2011).

My main analysis relies on the sample of "new" counties, i.e. those that were not subdivisioned or partitioned but that were created from non-county areas. In an alternative analysis, my sample is restricted to "partitioned" and "other" counties. This serves as a placebo treatment given that counties that are subdivisions of previously created counties could be different from "new" counties with regards to the level of societal, institutional and cultural development.

In Section 3.8, I examine potential threats to my identification. First, it might be argued that the timing of county creation might be a function of settlers' composition, where having a more homogenous population might speed up the process of county creation. Second, defining settlers as the first inhabitants of newly created counties using the first Census data available might be problematic if people resided in these counties long before their creation and therefore long before their first U.S. Census became available. Lastly, the correspondence assumption between settlers' culture and the dominant culture in their sending country/state might not hold if settlers have beliefs, preferences and values that are not representative of the norms of country/state of birth.

3.7 Main Results

3.7.1 Settler Population: Short Run Analysis

Table 3.4 reports the results from estimating a restricted version of Equation 3.1. I first examine settler population based on the distribution of foreign, out-of-state and in-state born individuals. The dependent variable in columns (1)–(6) is female labor force participation in U.S. counties in the short run. Data on labor force participation is based on the first U.S. Census available after county creation. Throughout the analysis, I include state and decade of county creation fixed effects. In columns (1)–(6) I control for the list of county level geographic controls. In columns (2)–(3) and (4)–(5), I also control for settlers' demographic characteristics including the share of prime age population, share of literate population, the sex ratio computed as the ratio of the male over the female population, the share of the single population and the child women to ratio computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000. I further include additional county level controls capturing geography and isolation in columns (3) and (6).

The analysis in columns (1)–(3) is based on my main sample of "new" counties, i.e. those that were not subdivisioned or partitioned but that were created from non-county areas. The estimates suggest that higher shares of foreign born and out-of-state born settlers in comparison to higher shares of in-state born settlers (the

omitted category) do not seem to be correlated with female labor force participation in U.S. counties in the short run. In columns (4)–(6), I repeat the analysis but restricting my sample to "partitioned" and "other" counties. The direction of the results remains very similar to the main analysis.

3.7.2 Settlers' Culture: Short Run Analysis

Next, rather than considering the settler population only based on the distribution of foreign, out-of-state and in-state born individuals, I exploit settlers' culture and rely on female labor force participation levels in sending countries to capture foreign born settlers' values and beliefs. In Table 3.5, I report the results from estimating Equation 3.1 where my outcome variable of interest is female labor force participation in U.S. counties in the short run. Data on labor force participation is extracted from the first U.S. Census data available after county creation. I control for state and decade of county creation fixed effects as well as my list of geographic county level variables in columns (1)–(4). I introduce settlers' demographic controls in columns (2)–(4) and additional county level controls related to geography and isolation in columns (3) and (4). My main sample of U.S. counties includes "new" counties created between 1840–1940 in columns (1)–(3). In column (4), I restrict my sample to "partitioned" and "other" counties.

Conditional on the shares of foreign and out-of-state born settlers in the total county level population, I examine whether having a higher share of foreign born settlers coming from countries known to have *above* median FLFP is correlated with women's labor force participation in U.S. counties in the short run. The findings for my main sample analysis in columns (1)–(3), restricted to "new" counties created between 1840–1940, show a positive and statistically significant correlation between the share of foreign born settlers from countries known to have *above* median FLFP and women's labor force participation in U.S. counties. The estimate in column (3) shows that a 1 percentage point increase in the share of foreign born settlers from countries known to have *above* median FLFP is associated with a 0.06 percentage point increase in female labor force participation in U.S. counties.[15]

My results for the sample of "partitioned" and "other" counties show weak evidence in support of this relationship. Interestingly, in column (4), I find that the estimate for the share of foreign born settlers from countries known to have *above* median FLFP is close to zero and not statistically significant.

In Table 3.6, I report the results from estimating Equation 3.1 using the variation in the timeline of passage of women's suffrage rights across countries and U.S. states to proxy for gender norms at the place of origin of foreign born and out-of-state born settlers. The structure of the table is the same as Table 3.5. Now instead of examining the share of foreign born settlers coming from countries known to have *above* median FLFP, I investigate the relationship between U.S. county level female labor force participation in the short run and the share of foreign born and out-of-state born settlers coming from countries/U.S. states were partial or full suffrage rights were granted to women anytime before the time I observe them. I do not find evidence in support of a relationship between having more settlers coming from places where women could vote and female labor force participation in U.S. counties in the short run. Estimates across columns (1)–(4) for my main sample of "new" counties and alternative sample restricting to "partitioned" and "other" counties

[15]The mean of the dependent variable of interest, female labor force participation, in my main sample of "new" counties is equal to 0.11 with a standard deviation of 0.12. The mean is equal to 0.13 with a SD of 0.12 for my alternative sample of "partitioned" and "other" counties.

are not statistically significant. In Appendix Table 3.19, I instead use my measure of suffrage intensity based on the number of years that suffrage laws had been in effect. I find that a longer exposure to suffrage rights is not related to women's labor force participation in U.S. hosting counties in the short run.

I then rely on an alternative quantitative measure that proxies for out-of-state born settlers' gender norms using the variation in the intensity and timing of passage and implementation of women's financial liberation. In columns (1)–(3) of Table 3.7, the positive and statistically significant estimates for the share of out-of-state born settlers coming from U.S. states were women had property and earnings rights document a positive association with women's labor force participation in U.S. hosting counties in the short run for my main sample of "new" counties, with a magnitude of about 0.07 percentage point. The relationship is not robust however to restricting to partitioned and subdivisioned counties, and other counties that are not created from non-county areas. The estimate in column (4) is about half the size of my estimates for the sample of "new" counties and not statistically significant.

In Table 3.8, I use my measure of financial liberation intensity based on the number of years that relevant laws had been in place. Positive and statistically significant estimates in columns (1)–(3) document a robust relationship between a longer exposure to women's financial liberation and female labor force participation for my main sample of "new" U.S. counties. Excluding those counties and restricting to partitioned and other type of counties decreases the size of my estimate by about one third the effect I find for my main sample analysis (column (4) of Table 3.8).

3.7.3 Long Run Analysis

In Table 3.9, I repeat the same analysis of Table 3.4 but using labor force participation data from 100 years after county creation. Now, instead of examining the relationship in the short run, I carry my analysis using data on labor force participation in the long run. This captures the relationship between having more foreign and out-of-state born settlers in comparison to in-state born settlers in U.S. counties in their early stages of cultural and institutional formation, and gender norms in the long run. The estimates suggest weak evidence in support of a relationship between settlers' population as categorized into shares of foreign, out-of-state and in-state born settlers and labor force participation outcomes in U.S. counties in the long run.

Results from both the short and long run analysis (Table 3.4 and Table 3.9 respectively) confirm that settlers' population per se as distributed between foreign, out-of-state and in-state born migrants does not relate to female labor force participation in hosting areas.

Exploiting settlers' culture instead, I document an increase in female labor force participation 100 years later with a higher share of foreign born settlers from countries known to have *above* median FLFP. The estimate reported in column (3) of Table 3.10 suggest that a 1 percentage point increase in this share is associated with about 0.02 percentage point increase in FLFP for the main sample of "new" U.S. counties in the long run. Findings in the long run go in the same direction of my results for women's involvement on the formal labor market few years after county creation (Table 3.5). This provides evidence in support of the persistent impact of settlers' culture. In column (4) of Table 3.10, I restrict to subdivisioned or partitioned counties from previously formed ones and I find that my estimate for the share of foreign born settlers from countries known to have high FLFP is smaller in magnitude in comparison to columns (1)–(3). While the estimate for settlers' culture is statistically significant for the alternative sample of partitioned and other counties, results from

the next two specifications using alternative quantitative proxies for settlers' culture confirm the weak evidence in support of this relationship for non "new" counties.

In Table 3.11, I estimate Equation 3.1 using the variation in the timeline of passage of women's suffrage rights across countries and U.S. states to capture settlers' culture. The structure of the table is the same as Table 3.6. My results for the main sample of "new" counties show that having more foreign born settlers from countries where women could vote is associated with an increase in women's labor force participation in the long run. This relationship is not robust to restricting to "partitioned" and "other" type of counties that resulted from already settled places.

Finally, in Table 3.12, I document a positive and statistically significant relationship between having more out-of-state born settlers from U.S. states where women were granted property and earning rights and women's labor force participation in "new" U.S. counties 100 years later. The estimates are close to 0.09 percentage point (column (3) of Table 3.12). The estimated effect is smaller in magnitude and not statistically significant for the sample of counties that were subdivisioned from already settled places.

3.7.4 Attitudes Regarding Women's Roles

In this subsection, I present the results on the impact of settlers' culture on current attitudes regarding women's roles in societies. I rely on data from the General Social Survey (GSS) over the years 1993–1998 and focus on the following two questions: "Do you approve or disapprove of a married woman earning money in business or industry if she has a husband capable of supporting her?" and "Do you agree or disagree with this statement? Women should take care of running their homes and leave running the country up to men?".

The model is similar to Equation 3.1 with the exception that the unit of observation is now the respondent. I also include controls for the respondent's demographic characteristics and a dummy for survey year. Specifically, I estimate:

$$y_{icsdt} = \alpha + \tau Settlers' Population_{csd} + \beta Settlers' Culture_{csd} + X'_{cs}\gamma + Z'_{it}\omega + \theta_{FE(s)} + \phi_{FE(d)} + \pi_{FE(t)} + \varepsilon_{icsdt} \tag{3.2}$$

where y_{icsdt} is the answer to the first question of whether women should work and the second question of whether women should take care of the country for individual i in county c, state s, decade of county creation d and GSS survey year t. The dependent variable is a binary dummy that takes the value 1 if respondents have liberal attitudes regarding women's roles in societies, i.e. if they approve on women working and if they disagree with the statement that women should take care of running homes and not the country. Z'_{it} is a vector of individual characteristics. These characteristics include the individual's gender, age, age squared, six education dummies, three race dummies and five marital status dummies. Settlers' Population$_{\mathbf{csd}}$ is the distribution of foreign, out-of-state and in-state born individuals out of total population. **Settlers' Culture$_{\mathbf{csd}}$** is my independent variable of interest. Data on the settler population is based on the first U.S. full count Census available after the county creation date. X'_{cs} includes my set of county level geographic, isolation and demographic controls. π_t, θ_s and ϕ_d are GSS survey year, state and decade of county creation fixed effects respectively. ε_{icsdt} is the error term. My standard errors are clustered on 60-square-mile grid cells (Bester, Conley, and Hansen, 2011).

Table 3.13 shows my OLS estimates. I include GSS survey year, state and decade of county creation fixed effects throughout. I control for county level geographic,

demographic, isolation and geography related characteristics in columns (1)–(6). I also include individual's characteristics. In columns (1) and (2), I capture settlers' culture using the share of foreign born settlers coming from places with high FLFP. In columns (3) and (4), settlers' culture is proxied for using the share of foreign born and out-of-state born settlers coming from places where women could vote. Lastly, in columns (5) and (6), I measure settlers' culture using the share of settlers coming from states that passed women's financial rights. The estimates are all positive and statistically significant (except one) suggesting that settlers with liberal gender norms had a persistent (positive) effect on attitudes toward women's roles in societies.

To examine the size of my estimates, I repeat this analysis using probit response models. Marginal effects suggest that respondents residing in counties that historically hosted a larger share of foreign born settlers from countries where women could vote, are about 30 percent more likely to approve on women working. I document an increase of about 34 percent in the likelihood of respondents approving on whether women should be allowed to work, for counties that historically hosted more settlers from states where women were granted their financial rights.

3.7.5 Discussion

This chapter provides evidence on the short and long run effects of settlers' culture on norms and values in newly established places. My findings are indicative of the long-lasting link between the composition of settlers, specifically settlers' culture, and cultural formation in hosting societies.

The mechanisms of transmission and persistence are related to having more people exposed and carrying gender liberal norms from their place of origin and moving to new places at their early stages of cultural, community, societal and institutional development. The weak cultural and institutional setting in the early stages of settlement in these newly established counties allowed settlers to impact the formation of culture, local institutions and social identity. The cultural fabric in the newly established and settled area is self-perpetuating, transmitted partly through generations (vertical/over time transmission of norms) and acquired by newcomers (later immigrants) that self select or that assimilate into the area's particular culture (spatial/horizontal transmission of norms). This is related to the idea of how initial conditions can determine long-run equilibria (Bisin and Verdier, 2011).

My empirical findings lie at the heart of Zelinsky, 1973's doctrine of "first effect settlement" and Woodard, 2011's argument of how culture in a given area is defined and determined by the people who first occupied it and the type of institutions that they establish.

In the next section, I examine potential threats to identification that might undermine the validity of the interpretation of my results.

3.8 Threats to Identification

In this section, I account for and address potential threats to identification. These include the potential caveats in the measurement of settlers' populations, the link between the timeline of county creation and settler population and lastly, the validity of the cultural correspondence assumption. Results that I present in the next subsections help address these concerns and support the validity of my findings and interpretation.

3.8.1 Settlers and County Creation

The historical background on U.S. territorial expansion, state incorporation and county formation helps support the claim that the composition of settlers is not a crucial determinant of county creation. Table 3.14 helps illustrate this claim, where I analyze the explanatory power of different determinants of U.S. counties' timing of creation. This table displays the estimates and the adjusted R-squared from regressions of the timing of creation (the date at which a given land area was first politically organized into a U.S. county) on a number of explanatory variables and state fixed effects. Column (1) shows that counties with higher shares of foreign born and out-of-state born settlers were created sooner than others. This explains about 50% of the variation in the timing of counties' creation. In column (2) of Table 3.14, in addition to settlers' population as categorized between foreign born, out-of-state and in-state born migrants, I add a list of county level geographic variables such as latitude, longitude, mean county temperature and rainfall, elevation, distance to lakes and rivers from the county centroid and average potential agricultural yield. The explanatory power of the specification increases to about 58%. In column (3), I further include county level geographic and isolation variables. This explains about 74% of the variation in the timeline of county creation. Lastly, in column (4), I include settlers' culture as proxied by the share of foreign born settlers from countries known to have high FLFP. Results show that settlers' culture has weak explanatory power with an adjusted R-squared almost the same between columns (3) and (4). Results from Table 3.14 thus provide suggestive evidence in support of the weak significance of settlers' culture as a driver of county creation.[16]

Pre-County Creation Population

In this subsection, I describe pre-county creation populations by inferring the time of migration to U.S. counties and restricting to people who plausibly migrated prior to county formation.

Census data is not available for areas before they become politically organized as an administrative entity of the United States. This means that information about the population, population density and composition of inhabitants of geographic areas is available only after the area was formally incorporated into a U.S. county given that our unit of study for this research is a county. The lack of information about the population living in these places before they are formally incorporated might constitute a potential caveat to my construction of the settler population. This is because it might be argued that these areas were settled for a long period of time before they formally became a U.S. county which means that I might not be actually capturing the initial settlers of these counties.

To deal with this, I investigate the timing of migration for a subsample of population (households/families with children) to infer the "time-at-move" of these people to places before they become formally incorporated as U.S. counties. The purpose is to identify the people that plausibly migrated to U.S. counties prior to their formation. The analysis is possible only for families/households with children given that data on individuals' county of birth and migration variables (pre 1940) are missing from the U.S. Census.

[16]The analysis using the shares of settlers from places where women could vote and had property and earning rights also confirm the weak explanatory power of settlers' culture in determining the timing of county creation.

I find that the average "time-at-move" for families with one child only born outside the current state of residence, for residents of counties created in 1860, is about 7 years. For households with more than one child, the average move time is almost 4 years. I repeat this analysis considering the sample of counties created in 1920. I find that families with one child only born outside the current state of residence moved to the area where the county falls about 11 years earlier to county creation. For households with more than one child, the average "time-at-move" to these counties is 3 years prior county creation. While here I have only reported the time at move to counties created in 1860 and 1920, the average time since people moved to other U.S. counties prior to their incorporation is always less than 10 years for households with children.

This provides suggestive evidence that the settler population I observe is not very different from pre-county creation populations, and that these counties were not settled for a long period of time before they formally became relevant U.S. counties. Of note, it is impossible to capture moves within states due to the lack of data from the U.S. census. This means that these settlers might have moved to other counties within the same state before actually moving to their current county of residence, and I will not be able to capture that. Nonetheless, my analysis reflects at least partially on the time of move to these newly created places.

3.8.2 Cultural Correspondence Assumption

In this study, settlers' culture is proxied with values and beliefs from the country/state of birth. The underlying assumption is that when individuals migrate to new places, they carry with them some aspects of their cultural beliefs and values and transmit them to where they move, i.e. that settlers internalize their culture before migrating. This relates to the idea of "cultural continuity" via "portability" (horizontal transmission) of beliefs and values (Alesina, Giuliano, and Nunn, 2013; Antecol, 2000; Fortin, 2005; Fernandez, 2007). A potential caveat that undermines the credibility of this assumption is the following: individuals that migrate might not have beliefs, preferences and values that are representative of the norms of their country/state of birth.

This approach is based on the work of Fernandez and Fogli, 2009 that investigates culture by examining the behavior of second-generation American women using past female labor force participation and fertility rates from women's country of ancestry as proxies for culture in countries of origin. The argument is that female labor force participation in a given country is a function of women's preferences and beliefs, including how she will be treated by others based on her work decision as well as her own perception of the role of women in the household, her perceived impact on children as a result of her work, etc. This is in addition to other economic and institutional factors determining women's work decisions.[17]

The challenge is to examine whether the correspondence assumption between settlers' culture and the dominant culture in their sending country/state is credible for migrants moving to U.S. counties. Table 3.15 provides evidence in support of this in the early stages of the "Age of Mass Migration" in the United States.[18] Using

[17] See Fernandez and Fogli, 2009 for a detailed discussion about the rationale for using countries' female labor force participation to reflect culture.

[18] This individual level result adds to previous findings in the literature showing that individuals internalize their cultural beliefs before moving to new places. One key contribution is the novelty of the setting used in this research to capture horizontal transmission of culture (or "cultural continuity via portability"). Like previous research, I focus on first-generation immigrants, i.e. those who were

Census data from the period of 1860 to 1940 (labor force status is missing for women in 1850) on first generation female immigrants residing in U.S. counties, I show that past female labor force participation in their country of origin at the same time or within a few decades earlier are important determinants of their labor supply. In Table 3.15, my sample is restricted to foreign born women, i.e. women who are born in a country different from the United States. The dependent variable of interest is a binary variable that takes the value 1 if a foreign born woman is in the labor force and 0 otherwise. My analysis is at the individual level, so I include individual level controls that might affect women's likelihood to be in the labor force, such as age, literacy and fertility status. Estimates of the coefficient on my main independent variable of interest, historical female labor force participation in the country of birth of the foreign born woman, are all positive and statistically significant. This is indicative that women are more likely to work if they come from a country with high female labor force participation (FLFP).[19]

3.9 Conclusion

The purpose of this chapter is to examine the relationship between the culture of settlers residing in U.S. counties at early stages of their cultural and institutional development and gender norms historically and in the long run. This research provides an analytical empirical framework that allows for the revisiting of a doctrine in cultural geography proposed by Zelinsky, 1973 and that argues that the characteristics of the first people residing in a given place are crucial for cultural formation.

I focus on counties that were not subdivisioned or partitioned from previously formed ones as those better reflect early stages of development and establishment with regard to the community, society, culture and institutions formed. I document a higher female labor force participation in the short and long run in U.S. counties that were initially occupied by migrants originating from places with liberal gender attitudes. I also document liberal attitudes toward women's roles in societies for individuals currently residing in U.S. counties that historically hosted higher shares of early settlers from places with liberal gender attitudes.

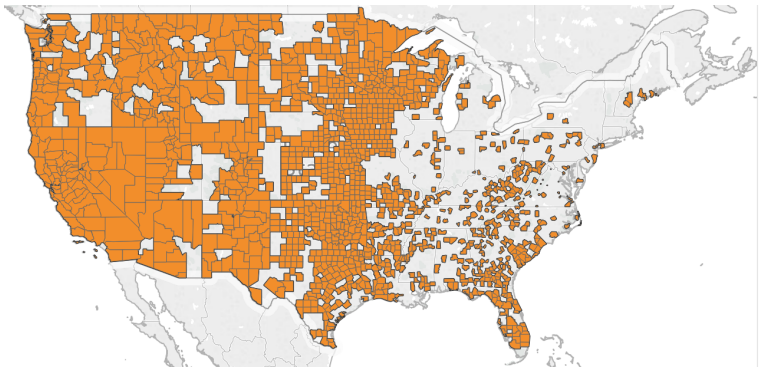
This research sets the stage for future work to look at a host of other cultural traits in the U.S. context. Applying this analysis in other settler societies is also crucial as settlement may have different effects in other countries.

born in foreign countries and then moved to the U.S., which allows the sample of foreign migrants to be exposed to their home country culture before migration.

[19]Of note, while it may be harder to empirically test this for out-of-state born migrants given the lack of data, I adopt the same underlying assumption that individuals born out-of-state internalize the values and beliefs that they were exposed to in their state of origin before migrating to new places.

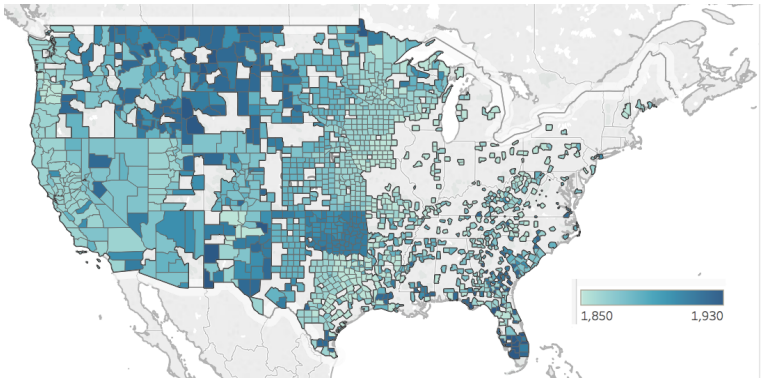
3.10 Figures

FIGURE 3.1: U.S. counties created between 1840 and 1940



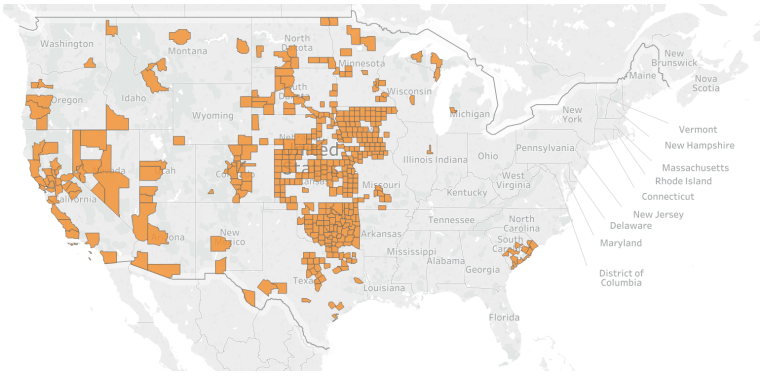
Notes: Sample of U.S. counties created between 1840 and 1940. Orange indicates that a county is included in my analysis while grey areas indicate counties that are excluded from my analysis either because they were created before 1840, between 1880 and 1889 or after 1940. *Source*: Author's compilation based on the ATLAS of Historical County Boundaries data.

FIGURE 3.2: Timing of U.S. counties' creation



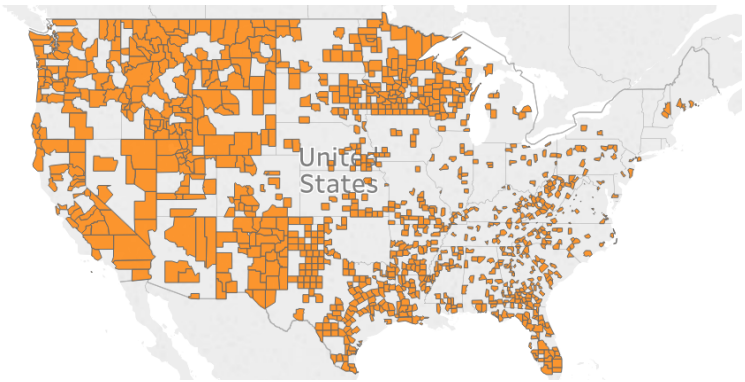
Notes: Chronological timing of 1,494 U.S. counties created between 1840 and 1940. Light blue refers to counties created early on and darker blue refers to counties created later on. Grey areas indicate counties that are excluded from my analysis either because they were created before 1840, between 1880 and 1889 or after 1940. *Source*: Author's compilation based on the ATLAS of Historical County Boundaries data.

FIGURE 3.3: Main sample of "new" counties



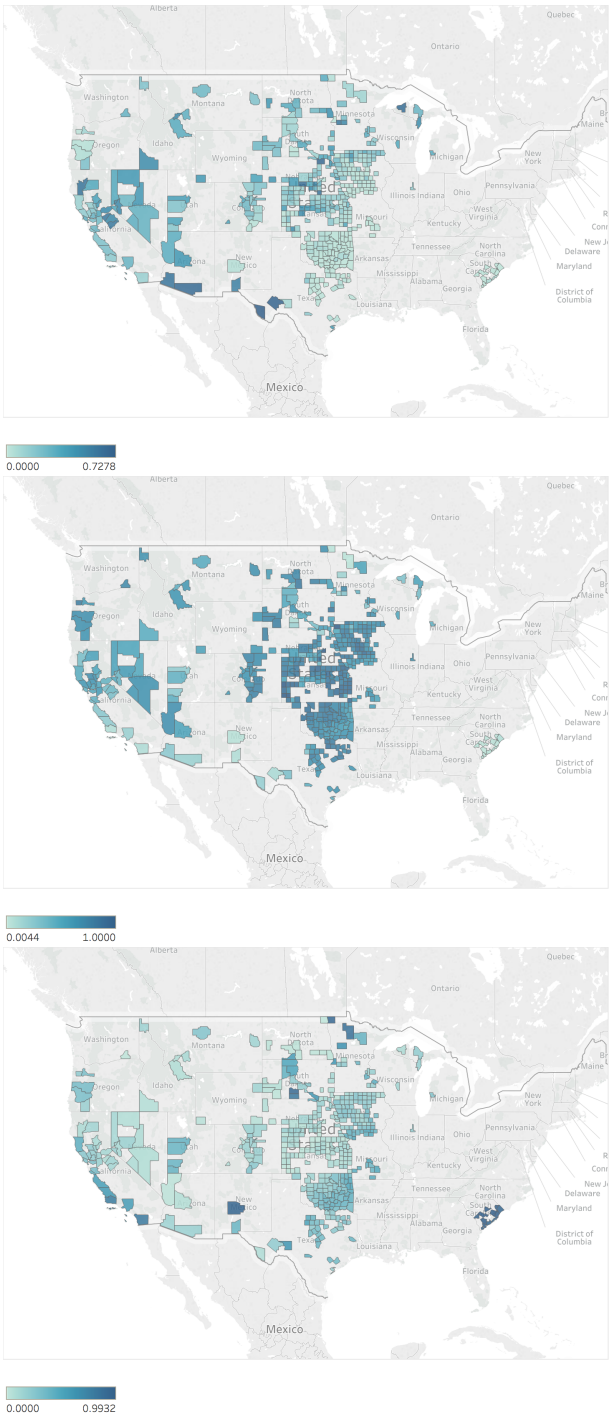
Notes: Main sample of "new" counties which includes counties that were not subdivisioned or partitioned from previously formed counties. Orange indicates that a county is included in my *main* sample because it was created between 1840–1940 and was not subdivisioned or partioned from previously created counties but formed from non-county areas. Grey areas indicate counties that are excluded from my *main* sample. *Source*: Author's compilation based on the ATLAS of Historical County Boundaries data.

FIGURE 3.4: Sample of "partitioned" and "other" counties



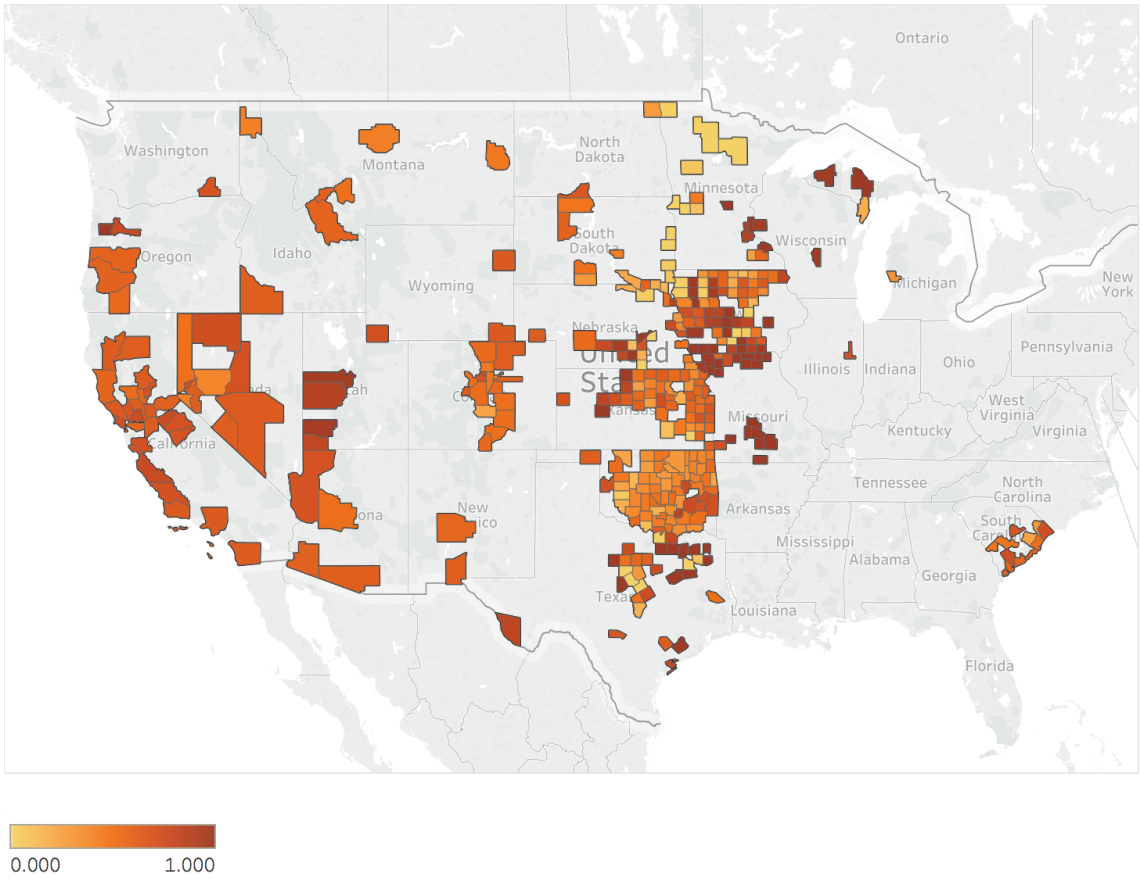
Notes: Alternative sample of "partitioned" and "other" counties which were subdivisioned or partitioned from previously created counties or that were created from a combination of districts and non-county areas, those already created under territorial juridistriction which then changed from an organized incorporated territory to a U.S. state, those created under a given territorial juridistriction which then came under another territorial juridistriction and counties created as a result of the passage of a new constitution converting all judicial districts to counties. Orange indicates that a county is included in my *alternative* sample. Grey areas indicate counties that are excluded from this sample. *Source*: Author's compilation based on the ATLAS of Historical County Boundaries data.

FIGURE 3.5: Settler population by origin: Sample of "new" counties



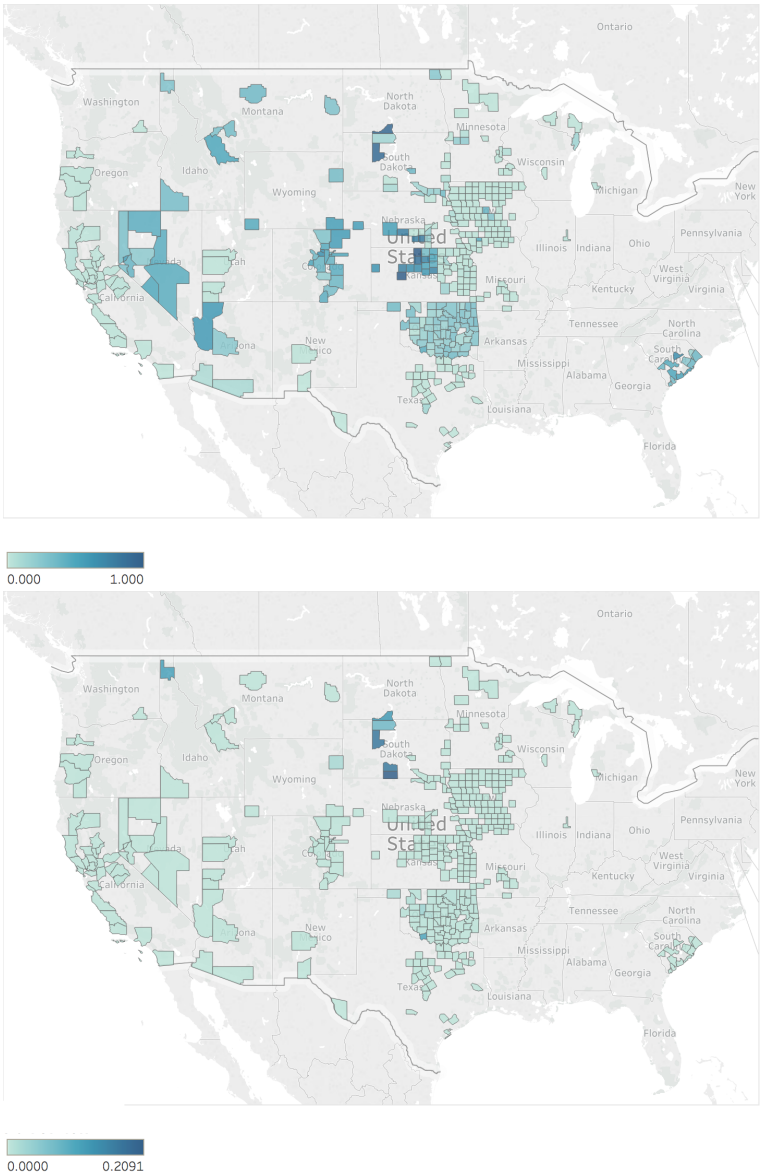
Notes: Shares of foreign born, out-of-state and in-state born individuals out of the total population are displayed respectively. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals are those born in the same state as the one where the household is located. Shares of foreign borns settlers out of the total population range between 0 and 1 with 0 indicating no foreign born settlers in the county and 1 indicating that all individuals living in the county are foreign born. Light blue indicates lower shares and dark blue greater shares. Grey areas indicate counties that are excluded from my sample.

FIGURE 3.6: Map of foreign born settlers from countries with high FLFP: Sample of "new" counties



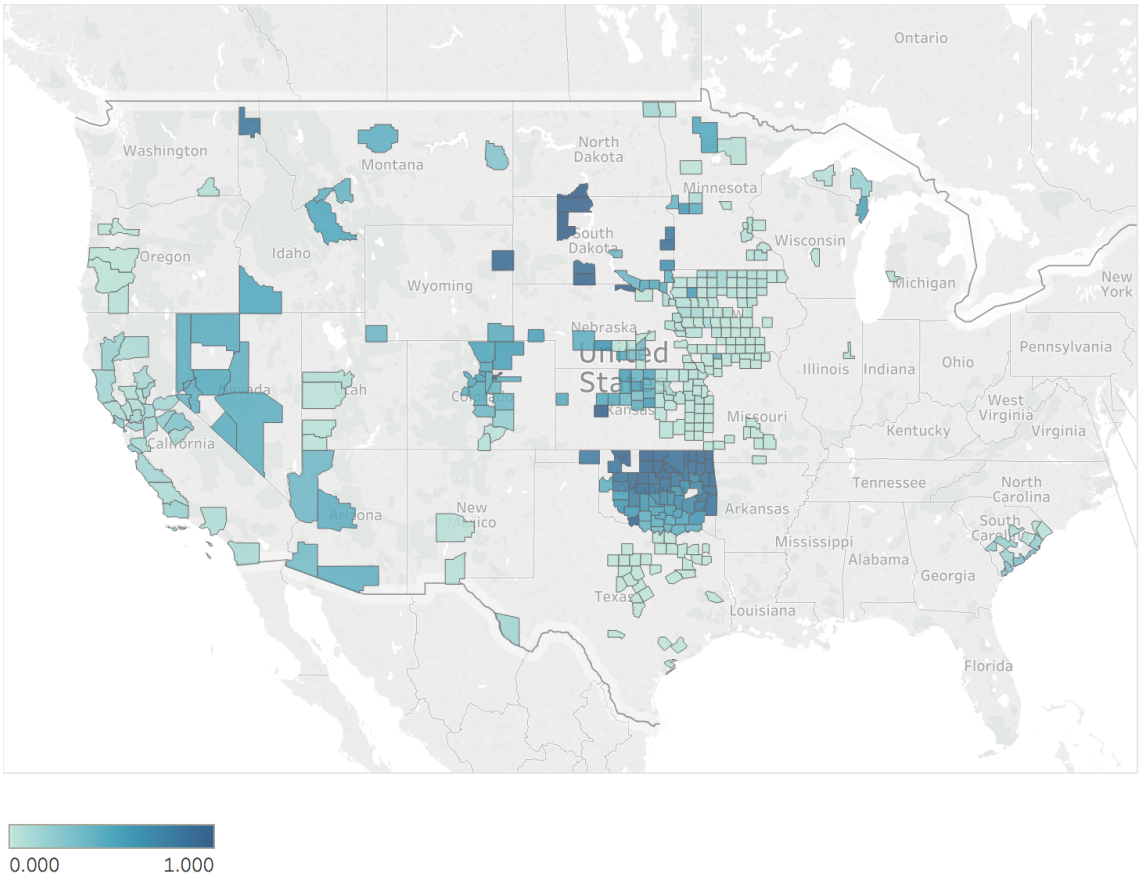
Notes: Map showing the share of foreign born settlers from countries *known* to have *above* median female labor force participation. It is obtained as follows: first median FLFP in a given decade based on data availability on FLFP for sending countries is computed. Then the total number of foreign born settlers coming from countries with above decade specific median FLFP is computed and lastly this total is divided by the total foreign born settler population with *known* FLFP i.e. foreign born settlers from countries with available data on FLFP. Grey areas indicate counties that are excluded from my sample. Light yellow indicates a small share of foreign born settlers coming from counties with *above* median female labor force participation and dark orange indicates a high share. *Source*: Author's compilation.

FIGURE 3.7: Foreign and out-of-state settlers who could vote: Sample of "new" counties



Note: Map showing the share of foreign born settlers out of the total foreign born settler population and out-of-state born settlers out of the total out-of-state settler population coming from countries/U.S. states where partial or full suffrage rights were granted to women anytime before the time when settlers are observed. Grey areas indicate counties that are excluded from my sample. Light blue indicates a small share of foreign born settlers/out-of-state born settlers coming from countries/U.S. states where women could vote and dark blue indicates a high share. *Source*: Author's compilation.

FIGURE 3.8: Out-of-state settlers from countries with women's financial liberation: Sample of "new" counties



Notes: Map showing the share of out-of-state born settlers out of the total out-of-state settler population coming from U.S. states where women's financial liberation was granted anytime before the time when settlers are observed. Grey areas indicate counties that are excluded from my sample. Light blue indicates a small share of out-of-state born settlers coming from U.S. states where women had property and earning rights and blue indicates a high share. *Source*: Author's compilation.

3.11 Tables

TABLE 3.1: Descriptive statistics

		By Gender	
	Entire population	Male population	Female population
	(1)	(2)	(3)
Demographic Characteristics			
Share male population	0.60		
	(0.11)		
Share prime age population	0.57	0.59	0.50
	(0.13)	(0.15)	(0.08)
Share literate population	0.55	0.59	0.47
	(0.20)	(0.20)	(0.19)
Share single population	0.44	0.53	0.24
	(0.18)	(0.18)	(0.10)
Average number of children	1.88		
	(0.54)		
Child women ratio	692.23		
	(226.85)		
Settlers' Population			
Share foreign born	0.15	0.16	0.12
	(0.16)	(0.16)	(0.14)
Share out-of-state born	0.62	0.48	0.60
	(0.19)	(0.19)	(0.21)
Share in-state born	0.22	0.20	0.26
	(0.19)	(0.19)	(0.21)

Notes: Based on complete count Census data from the Integrated Public Use Microdata Series (IPUMS). My sample is restricted to "new" counties. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals are those born in the same state as the one where the household is located. All shares range between 0 and 1. In column (1), shares are displayed out of the total settler population. In columns (2) and (3), summary statistics are displayed by gender. Shares in columns (2) and (3) are displayed out of the male settler population and the female settler population respectively. Prime age refers to ages 15 to 49. The child women ratio is computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000. Standard deviations are reported in parentheses.

TABLE 3.2: Descriptive statistics by settlers' origin and gender

	Foreign born		**Out-of-state born**		**In-state born**	
	Male Population (1)	Female Population (2)	Male Population (3)	Female Population (4)	Male Population (5)	Female Population (6)
Share prime age population	0.77	0.74	0.67	0.59	0.11	0.11
	(0.16)	(0.15)	(0.13)	(0.11)	(0.15)	(0.14)
Share literate population	0.82	0.74	0.68	0.57	0.10	0.09
	(0.17)	(0.21)	(0.18)	(0.20)	(0.15)	(0.12)
Share single population	0.50	0.16	0.52	0.22	0.70	0.44
	(0.25)	(0.15)	(0.19)	(0.11)	(0.24)	(0.27)
Average number of children		2.52		2.00		0.77
		(2.82)		(1.33)		(0.77)
Child women ratio		948.85		738.15		351.70
		(1120.14)		(492.21)		(347.54)

Notes: Based on complete count Census data from the Integrated Public Use Microdata Series (IPUMS). My sample is restricted to "new" counties. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals are those born in the same state as the one where the household is located. All shares range between 0 and 1. The share of male population: *foreign born* is 0.70 (0.13); the share of male population: *out-of-state born* is 0.61 (0.11) and the share of male population: *in-state born* is 0.51 (0.09). In columns (1) and (2) I report statistics for foreign born settlers by gender i.e. shares are displayed out of the total male foreign born settler population and the total female foreign born settler population respectively . In columns (3) and (4) I report statistics for out-of-state born settlers by gender i.e. shares are displayed out of the total male out-of-state born settler population and the total female out-of-state born settler population respectively. In columns (5) and (6) I report statistics for in-state born settlers by gender i.e. shares are displayed out of the total male in-state born settlers' population and the total female in-state born settlers' population respectively. Prime age refers to ages 15 to 49. The child women ratio is computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000. Standard deviations are reported in parentheses.

TABLE 3.3: Settlers' culture: Descriptive statistics

	Mean (1)	SD (2)	(N) (3)
Settlers' culture			
Female labor force participation			
Share foreign born with *known* FLFP	0.77	0.24	426
Share foreign born with *unknown* FLFP	0.23	0.24	426
Share foreign born : *Above* median FLFP	0.54	0.28	436
Share foreign born : *Below* median FLFP	0.46	0.28	436
Suffrage			
Share foreign born: Women suffrage rights	0.16	0.21	431
Share out-of-state born: Women suffrage rights	0.003	0.01	436
Foreign born: suffrage intensity	5.46	9.24	436
Out-of-state born: suffrage intensity	0.07	0.50	436
Financial liberation			
Share out-of-state born: Women's financial liberation	0.33	0.35	436
Out-of-state born: financial liberation intensity	6.70	10.08	436
FLFP in U.S. counties			
FLFP in the short run	0.11	0.11	368
FLFP in the long run	0.28	0.13	334

Notes: My sample is restricted to "new" counties. Data on historical female labor force participation by sending country by decade for foreign born settlers is obtained from Integrated Public Use Microdata Series (IPUMS) International Historical Censuses, (Olivetti, 2014) and (Olivetti and Petrongolo, 2016). Female labor force participation from sending countries is extracted from the same decade, or a decade or two earlier, depending on data availability, from when I observe my foreign born settlers. Share foreign born with *known* FLFP is obtained by dividing the total number of foreign born settlers from countries with *known* FLFP i.e. where data on FLFP is available, over the total foreign born settler population. The share of foreign born settlers from countries *known* to have *above* median female labor force participation is obtained by first computing the median FLFP across countries in a given decade based on data availability on FLFP for sending countries. Then the total number of foreign born settlers coming from countries with above decade specific median FLFP is computed and lastly this total is divided by the total foreign born settler population with *known* FLFP i.e. foreign born settlers from countries with available data on FLFP. Share of foreign born settlers from countries *known* to have *below* median female labor force is computed the same way. Shares of foreign born settlers from countries *known* to have *above* and *below* median FLFP add up to 1. Shares of foreign born settlers from countries with *known* and *unknown* FLFP add up to the entire foreign born settler population. *Share foreign born: Women suffrage rights* and *Share out-of-state born: Women suffrage rights* are the share of foreign born settlers out of the total foreign born settlers population and the share of out-of-state born settlers out of the total out-of-state born settlers population coming from countries/U.S. states where partial or full suffrage rights were granted to women anytime before the time when settlers are observed. Women's suffrage *intensity* is measured as a weighted share of settlers coming from places were partial/full voting rights were granted to women weighted by the number of years between suffrage laws' passage and county creation. *Share out-of-state born: Women's financial liberation* is the share of out-of-state born settlers out of the total out-of-state born settlers population coming from U.S. states where women's financial liberation was granted anytime before the time when settlers are observed. Women's financial liberation *intensity* is measured as a weighted share of out-of-state born settlers coming from U.S. states where property rights and earnings rights were granted to women weighted by the number of years between women's financial liberation laws' passage and county creation.

TABLE 3.4: Analysis in the short run: Settlers' population

	Female labor force participation					
	(1)	(2)	(3)	(4)	(5)	(6)
Share foreign born	0.112*	0.0960	0.111	0.0213	0.00341	0.00546
	(0.060)	(0.071)	(0.067)	(0.035)	(0.046)	(0.049)
Share out-of-state born	0.0265	0.0209	0.0334	0.0408	0.0494	0.0592
	(0.070)	(0.078)	(0.077)	(0.035)	(0.039)	(0.038)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes		Yes	Yes
Additional Controls			Yes			Yes
N	**368**	**368**	**368**	**762**	**762**	**762**

Notes: My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In columns (4)–(6), the sample is restricted to "partitioned" and "other" counties. The dependent variable is female labor force participation in U.S. counties in the short run. Data on labor force participation is based on the first U.S. Census available after county creation. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. State and decade of county creation fixed effects are included in columns (1) to (6). The set of geographic controls include latitude, longitude, mean county temperature and rainfall, elevation, distance to lakes and rivers from the county centroid and average potential agricultural yield. These are included in columns (1)–(6). The set of demographic controls include the share of prime age population, share of literate population, the sex ratio computed as the ratio of the male over the female population, the share of single population and the child women ratio computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000. These are included in columns (2)–(3) and (5)–(6). The set of additional controls includes terrain ruggedness, rainfall risk, the distance to the nearest portage site, the distance to the nearest Indian battle site, the distance to the coast, the number of years that the county has been intersected by railroads since its creation date and the distance to the nearest mineral discovery site. These are included in columns (3) and (6). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 3.5: Settlers' culture: Female labor force participation in sending countries

	Female labor force participation			
	(1)	(2)	(3)	(4)
Settlers' population				
Share foreign born	0.101*	0.105*	0.106	0.00579
	(0.057)	(0.057)	(0.066)	(0.049)
Share out-of-state born	0.0235	0.0287	0.0369	0.0589
	(0.067)	(0.070)	(0.077)	(0.038)
Settlers' culture				
Share foreign born:*Above* median FLFP	0.0549**	0.0527**	0.0596**	0.00698
	(0.026)	(0.026)	(0.027)	(0.022)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes	Yes
Additional Controls			Yes	Yes
N	**368**	**368**	**368**	**762**

Notes: The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the short run. Data on labor force participation is based on the first U.S. Census available after county creation. My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In column (4), the sample is restricted to "partitioned" and "other" counties. Settler population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using female labor force participation from sending countries to reflect gender norms at the place of origin. Data on historical female labor force participation by sending country by decade for foreign born settlers is obtained from Integrated Public Use Microdata Series (IPUMS) International Historical Censuses, (Olivetti, 2014) and (Olivetti and Petrongolo, 2016). Female labor force participation from sending countries is extracted from the same decade, or a decade or two earlier, depending on data availability, from when I observe the foreign born settler population. Share of foreign born settlers from countries *known* to have above median female labor force participation is obtained by first computing the median FLFP in a given decade based on data availability on FLFP for sending countries. Second, the total number of foreign born settlers coming from countries with above decade specific median FLFP is computed and lastly this total is divided by the total foreign born settler population with *known* FLFP. Foreign born individuals from countries with *unknown* FLFP in a given decade are excluded from the foreign born settler population. Share of foreign born settlers from countries *known* to have *below* median female labor force is the omitted category. State and decade of county creation fixed effects are included in columns (1)–(4). The set of geographic controls are included in columns (1)–(4). The set of demographic controls are included in columns (2)–(4). The set of additional controls are included in columns (3) and (4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 3.6: Settlers' culture: Women's suffrage rights

	Female labor force participation			
	(1)	(2)	(3)	(4)
Settlers' population				
Share foreign born	0.105*	0.105*	0.0828	0.00828
	(0.063)	(0.062)	(0.079)	(0.050)
Share out-of-state born	0.0253	0.0264	0.0186	0.0622
	(0.071)	(0.072)	(0.081)	(0.039)
Settlers' culture				
Share foreign born: Suffrage rights	-0.00266	-0.00726	-0.0215	-0.0118
	(0.028)	(0.032)	(0.034)	(0.019)
Share out-of-state born: Suffrage rights	-0.958	-0.384	-0.376	0.0649
	(0.597)	(0.668)	(0.648)	(0.043)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes	Yes
Additional Controls			Yes	Yes
N	**367**	**367**	**367**	**761**

Notes: The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the short run. Data on labor force participation is based on the first U.S. Census available after county creation. My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In column (4), the sample is restricted to "partitioned" and "other" counties. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using women's suffrage rights to reflect gender norms at the place of origin. Data on the variation in the timing of passage of suffrage laws enfranchising women at different points in time across different U.S. states is obtained from Lott and Kenny, 1999 and Miller, 2008. For the timeline of women's suffrage across countries, information is available from Wikipedia using: `https://en.wikipedia.org/wiki/Timeline_of_women's_suffrage`. The country/state of origin is a variable set equal to 1 if partial or full suffrage rights had been granted to women. *Share foreign born: Women suffrage rights* is the share of foreign born settlers out of the total foreign born settler population coming from countries where partial or full suffrage rights were granted to women anytime before the time when settlers are observed. *Share out-of-state born: Women suffrage rights* is the share of out-of-state born settlers out of the total out-of-state born settlers population coming from U.S. states where partial or full suffrage rights were granted to women anytime before the time when settlers are observed. The omitted categories are share out-of-state born: Women *No* suffrage rights and Share foreign born: Women *No* suffrage rights, i.e. settlers from countries/U.S. states were women were not granted neither partial nor full suffrage rights. State and decade of creation fixed effects are included in columns (1)–(4). Geographic controls are included across columns (1)–(4), demographic controls are introduced starting column (2) to (4) and lastly additional county level controls for geography and isolation are introduced in columns (3) and (4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 3.7: Settlers' culture: Women's financial liberation

	Female labor force participation			
	(1)	(2)	(3)	(4)
Settlers' population				
Share foreign born	0.0961	0.101*	0.111	0.00757
	(0.059)	(0.058)	(0.067)	(0.049)
Share out-of-state born	0.0272	0.0309	0.0459	0.0644*
	(0.070)	(0.072)	(0.078)	(0.039)
Settlers' culture				
Share out-of-state born: Financial liberation	0.0751**	0.0672*	0.0683*	0.0358
	(0.034)	(0.034)	(0.038)	(0.030)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes	Yes
Additional Controls			Yes	Yes
N	**368**	**368**	**368**	**762**

Notes: The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the short run. Data on labor force participation is based on the first U.S. Census available after county creation. My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In column (4), the sample is restricted to "partitioned" and "other" counties. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using women's financial liberation to reflect gender norms at the place of origin. Data on the timing of passage and implementation of women's financial liberation granting property rights and earnings rights to women across U.S. sending states is obtained from Geddes and Dean, 2002. A state of origin is a variable set equal to 1 if women's financial liberation has been granted. *Share out-of-state born: Women's financial liberation* is the share of out-of-state born settlers out of the total out-of-state born settler population coming from U.S. states where women's financial liberation was granted anytime before the time when settlers are observed. The omitted category is the share out-of-state born: Women *No* women's financial liberation, i.e. settlers from U.S. states where women were not yet granted financial liberation. State and decade of creation fixed effects are included in columns (1)–(4). Geographic controls are included across columns (1)–(4), demographic controls are introduced starting column (2) to (4) and lastly additional county level controls for geography and isolation are introduced in columns (3) and (4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 3.8: Settlers' culture: Financial liberation intensity

	Female labor force participation			
	(1)	(2)	(3)	(4)
Settlers' population				
Share foreign born	0.105*	0.110*	0.117*	0.00790
	(0.058)	(0.057)	(0.067)	(0.049)
Share out-of-state born	0.0343	0.0391	0.0527	0.0650*
	(0.069)	(0.072)	(0.079)	(0.039)
Settlers' culture				
Out-of-state born: Financial liberation intensity	0.00321**	0.00300**	0.00320**	0.00186*
	(0.001)	(0.001)	(0.002)	(0.001)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes	Yes
Additional Controls			Yes	Yes
N	**368**	**368**	**368**	**762**

Notes: The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the short run. Data on labor force participation is based on the first U.S. Census available after county creation. My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In column (4), the sample is restricted to "partitioned" and "other" counties. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using women's financial liberation *intensity* to reflect gender norms at the place of origin. Women's financial liberation intensity is measured as a weighted share of out-of-state born settlers coming from U.S. states where property rights and earnings rights were granted to women weighted by the number of years between women's financial liberation laws' passage and county creation. If women's financial liberation laws were not yet granted, the negative difference between decade of county creation and year of passage of financial liberation is replaced by *zero*. State and decade of creation fixed effects are included in columns (1)–(4). Geographic controls are included across columns (1)–(4), demographic controls are introduced starting column (2) to (4) and lastly additional county level controls for geography and isolation are introduced in columns (3) and (4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 3.9: Analysis in the long run: Settlers' population

	Female labor force participation					
	(1)	(2)	(3)	(4)	(5)	(6)
Share foreign born	0.0676**	0.0473	0.0274	0.0714***	0.0657**	0.0553*
	(0.028)	(0.035)	(0.034)	(0.026)	(0.028)	(0.029)
Share out-of-state born	0.00106	-0.0182	-0.0292	0.0174	0.0169	0.0233
	(0.028)	(0.038)	(0.037)	(0.018)	(0.020)	(0.020)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes		Yes	Yes
Additional Controls			Yes			Yes
N	**331**	**331**	**331**	**723**	**723**	**723**

Notes: My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In columns (4)–(6), the sample is restricted to "partitioned" and "other" counties. The dependent variable is female labor force participation in U.S. counties in the long run. Data on labor force participation is based on the tenth U.S. Census available after county creation (100 years later). Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. State and decade of county creation fixed effects are included in columns (1) to (6). The set of geographic controls include latitude, longitude, mean county temperature and rainfall, elevation, distance to lakes and rivers from the county centroid and average potential agricultural yield. These are included in columns (1)–(6). The set of demographic controls include the share of prime age population, share of literate population, the sex ratio computed as the ratio of the male over the female population, the share of single population and the child women ratio computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000. These are included in columns (2)–(3) and (5)–(6). The set of additional controls includes terrain ruggedness, rainfall risk, the distance to the nearest portage site, the distance to the nearest Indian battle site, the distance to the coast, the number of years that the county has been intersected by railroads since its creation date and the distance to the nearest mineral discovery site. These are included in columns (3) and (6). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 3.10: Analysis in the long run: Settlers' culture (FLFP)

	Female labor force participation			
	(1)	(2)	(3)	(4)
Settlers' population				
Share foreign born	0.0646**	0.0443*	0.0241	0.0542*
	(0.028)	(0.026)	(0.034)	(0.029)
Share out-of-state born	0.00127	-0.0101	-0.0294	0.0215
	(0.028)	(0.029)	(0.038)	(0.020)
Settlers' culture				
Share foreign born: *Above* median FLFP	0.0179*	0.0163*	0.0185*	0.0141*
	(0.009)	(0.009)	(0.010)	(0.008)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes	Yes
Additional Controls			Yes	Yes
N	**331**	**331**	**331**	**723**

Notes: The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the long run. Data on labor force participation is based on the tenth U.S. Census available after county creation (100 years later). My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In column (4), the sample is restricted to "partitioned" and "other" counties. Settlers' population is based on the first U.S. Census available after county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using female labor force participation from sending countries to reflect gender norms at the place of origin. Data on historical female labor force participation by sending country by decade for foreign born settlers is obtained from Integrated Public Use Microdata Series (IPUMS) International Historical Censuses, (Olivetti, 2014) and (Olivetti and Petrongolo, 2016). Female labor force participation from sending countries is extracted from the same decade, or a decade or two earlier, depending on data availability, from when I observe my foreign born settlers' population. Share of foreign born settlers from countries *known* to have above median female labor force participation is obtained by first computing the median FLFP in a given decade based on data availability on FLFP for sending countries. Second, the total number of foreign born settlers coming from countries with above decade specific median FLFP is computed and lastly this total is divided by the total foreign born settlers population with *known* FLFP. Foreign born individuals from countries with *unknown* FLFP in a given decade are excluded from the foreign born settler population. Share of foreign born settlers from countries *known* to have *below* median female labor force is the omitted category. State and decade of creation fixed effects are included in columns (1)–(4). The set of geographic controls are included in columns (1)–(4). The set of demographic controls are included in columns (2)–(4). The set of additional controls are included in columns (3) and (4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 3.11: Analysis in the long run: Settlers' culture (suffrage)

	Female labor force participation			
	(1)	(2)	(3)	(4)
Settlers' population				
Share foreign born	0.0726**	0.0505*	0.0245	0.0595*
	(0.029)	(0.027)	(0.033)	(0.030)
Share out-of-state born	-0.00496	-0.0178	-0.0450	0.0250
	(0.030)	(0.029)	(0.037)	(0.020)
Settlers' culture				
Share foreign born: Suffrage rights	0.0763*	0.0769*	0.0817*	0.00528
	(0.044)	(0.043)	(0.045)	(0.018)
Share out-of-state born: Suffrage rights	-0.382	-0.281	-0.233	0.0325
	(1.292)	(1.345)	(1.285)	(0.064)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes	Yes
Additional Controls			Yes	Yes
N	**330**	**330**	**330**	**716**

Notes: The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the long run. Data on labor force participation is based on the tenth U.S. Census available after county creation (100 years later). My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In column (4), the sample is restricted to "partitioned" and "other" counties. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using women's suffrage rights to reflect gender norms at the place of origin. Data on the variation in the timing of passage of suffrage laws enfranchising women at different points in time across different U.S. states is obtained from Lott and Kenny, 1999 and Miller, 2008. For the timeline of women's suffrage across countries, information is available from Wikipedia using: `https://en.wikipedia.org/wiki/Timeline_of_women's_suffrage`. The country/state of origin is a variable set equal to 1 if partial or full suffrage rights where granted to women. *Share foreign born: Women suffrage rights* is the share of foreign born settlers out of the total foreign born settler population coming from countries where partial or full suffrage rights were granted to women anytime before the time when settlers are observed. *Share out-of-state born: Women suffrage rights* is the share of out-of-state born settlers out of the total out-of-state born settler population coming from U.S. states where partial or full suffrage rights were granted to women anytime before the time when settlers are observed. The omitted categories are share out-of-state born: Women *No* suffrage rights and Share foreign born: Women *No* suffrage rights, i.e. settlers from countries/U.S. states where women were not granted neither partial nor full suffrage rights. State and decade of creation fixed effects are included in columns (1)–(4). Geographic controls are included across columns (1)–(4), demographic controls are introduced starting column (2) to (4) and lastly additional county level controls for geography and isolation are introduced in columns (3) and (4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

TABLE 3.12: Analysis in the long run: Settlers' culture (financial liberation)

	Female labor force participation			
	(1)	(2)	(3)	(4)
Settlers' population				
Share foreign born	0.0445	0.0283	0.0147	0.0577**
	(0.030)	(0.028)	(0.033)	(0.028)
Share out-of-state born	-0.0115	-0.0179	-0.0314	0.0283
	(0.029)	(0.029)	(0.037)	(0.019)
Settlers' culture				
Share out-of-state born: Financial liberation	0.0994***	0.0935***	0.0909**	0.0318
	(0.032)	(0.033)	(0.036)	(0.023)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes	Yes
Additional Controls			Yes	Yes
N	**331**	**331**	**331**	**723**

Notes: The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the long run. Data on labor force participation is based on the tenth U.S. Census available after county creation (100 years later). My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In column (4), the sample is restricted to "partitioned" and "other" counties. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using women's financial liberation to reflect gender norms at the place of origin. Data on the timing of passage and implementation of laws granting property rights and earnings rights to women across U.S. sending states is obtained from Geddes and Dean, 2002. A state of origin is a variable set equal to 1 if women's financial liberation is granted. *Share out-of-state born: Women's financial liberation* is the share of out-of-state born settlers out of the total out-of-state born settler population coming from U.S. states where women's financial liberation was granted anytime before the time when settlers are observed. The omitted category is the share out-of-state born: Women *No* women's financial liberation, i.e. settlers from U.S. states where women were not yet granted financial liberation. State and decade of creation fixed effects are included in columns (1)–(4). Geographic controls are included across columns (1)–(4), demographic controls are introduced starting column (2) to (4) and lastly additional county level controls for geography and isolation are introduced in columns (3) and (4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 3.13: Attitudes regarding women's roles: GSS

	Women Work	Women Country	Women Work	Women Country	Women Work	Women Country
	(1)	(2)	(3)	(4)	(5)	(6)
Settlers' population						
Share foreign born	-1.137***	-0.0401	0.958***	1.006***	-0.136	-0.00597
	(0.189)	(0.170)	(0.262)	(0.184)	(0.126)	(0.115)
Share out-of-state born	2.885***	1.012***	-5.684***	-2.378**	0.0686	0.190**
	(0.226)	(0.126)	(1.095)	(0.827)	(0.090)	(0.082)
Settlers' culture						
Foreign born: *Above* median FLFP	7.422***	3.238***				
	(1.152)	(0.844)				
Foreign born: Suffrage rights			6.795***	1.228***		
			(0.744)	(0.365)		
Out-of-state born: Suffrage rights			159.6***	57.03***		
			(16.445)	(8.955)		
Out-of-state born: Financial liberation					0.229*	0.156
					(0.130)	(0.118)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey year FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	**959**	**952**	**959**	**952**	**2384**	**2342**

Notes: The unit of observation is a respondent. The period covered from the General Social Survey (GSS) is 1993–1998. Attitudes toward women's roles in societies is assessed through the following two questions: "Do you approve or disapprove of a married woman earning money in business or industry if she has a husband capable of supporting her?" where respondents choices are recoded as follows (1=approve, 0=disapprove) and "Do you agree or disagree with this statement? Women should take care of running their homes and leave running the country up to men?" where respondents choices are recoded as follows (1=disagree, 0=agree). My sample of U.S. counties is restricted to "new" counties. **Settlers' Culture$_{cs}$** is the independent variable of interest. It is proxied for using the share of foreign born settlers from countries *known* to have above median female labor force participation in columns (1) and (2). In columns (3) and (4), **Settlers' Culture$_{cs}$** is measured using the share of foreign born settlers out of the total foreign born settlers population and the share of out-of-state born settlers out of the total out-of-state born settlers population coming from countries/U.S. states where partial or full suffrage rights were granted to women anytime before the time when settlers are observed. Lastly, in columns (5) and (6), **Settlers' Culture$_{cs}$** is proxied for using the share of out-of-state born settlers out of the total out-of-state born settlers population coming from U.S. states where women's financial liberation was granted anytime before the time when settlers are observed. Settlers' population is based on the first U.S. Census available after the county creation date. State, decade of county creation and GSS survey year fixed effects are included in columns (1) to (6). The set of geographic controls include latitude, longitude, mean county temperature and rainfall, elevation, distance to lakes and rivers from the county centroid and average potential agricultural yield. These are included in columns (1)–(6). The set of demographic controls include the share of prime age population, share of literate population, the sex ratio computed as the ratio of the male over the female population, the share of single population and the child women ratio computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000. These are included in columns (1)–(6). The set of additional controls includes terrain ruggedness, rainfall risk, the distance to the nearest portage site, the distance to the nearest Indian battle site, the distance to the coast, the number of years that the county has been intersected by railroads since its creation date and the distance to the nearest mineral discovery site. These are included in columns (1)–(6). Individual's characteristics include the individual's gender, age, age squared, six education dummies, three race dummies and five marital status dummies. These are included in columns (1)–(6). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

TABLE 3.14: Determinants of county creation date

	(1)	(2)	(3)	(4)
Settlers' population				
Share foreign born	-46.19***	-49.75***	-37.90***	-36.97***
	(7.541)	(6.953)	(5.300)	(5.316)
Share out-of-state born	-45.79***	-39.86***	-25.90***	-24.88***
	(8.672)	(7.224)	(6.192)	(6.018)
Geographic controls				
Latitude		1.992**	1.574**	1.455**
		(0.964)	(0.720)	(0.697)
Longitude		1.378***	0.705**	0.616**
		(0.381)	(0.278)	(0.276)
Mean county temperature		2.368**	1.506*	1.511**
		(1.016)	(0.770)	(0.742)
Mean county rainfall		-0.0000861	-0.000417	0.000134
		(0.005)	(0.004)	(0.004)
Elevation		0.00413	0.00605*	0.00661**
		(0.004)	(0.003)	(0.003)
Distance to rivers		0.0000168	0.0000143	0.0000138
		(0.000)	(0.000)	(0.000)
Distance to lakes		-0.00000965	-0.00000957*	-0.00000896
		(0.000)	(0.000)	(0.000)
Average potential agricultural yield		-69.51***	-51.45***	-46.00***
		(11.255)	(10.559)	(10.561)
Geography and Isolation				
Terrain ruggedness			2.956	1.484
			(8.881)	(8.582)
Rainfall risk			11.41	10.48
			(9.594)	(8.970)
Distance to the nearest portage site			-2.237	-1.279
			(3.065)	(2.903)
Distance to the nearest Indian battle site			0.00000589	0.00000783
			(0.000)	(0.000)
Distance to coast			-0.00000221	-0.00000131
			(0.000)	(0.000)
Years connected to railroad since county creation			1.056***	1.042***
			(0.061)	(0.060)
Distance to the nearest mineral discovery site			0.000153***	0.000154***
			(0.000)	(0.000)
Settlers' culture				
Share foreign born: *Above* median FLFP				-7.997***
				(1.572)
State FE	Yes	Yes	Yes	Yes
N	1,362	1,362	1,362	1,362
Adjusted R-squared	0.522	0.575	0.739	0.747

Notes: The dependent variable in columns (1)–(4) is the date at which a given land was first politically organized into a county (county creation date). My sample of U.S. counties includes those created between 1840–1940. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using the share of out-of-state born settlers from countries known to have *above* median FLFP. Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

TABLE 3.15: Foreign born women's likelihood to be in the labor force

	Women's likelihood to be in the labor force		
	1860	*1870*	*1880*
	(1)	(2)	(3)
FLFP: sending countries	0.223***	0.288***	0.198**
	(0.046)	(0.075)	(0.077)
Individual level controls	Yes	Yes	Yes
N	1,435,237	2,042,244	2,512,342

Notes: The sample is restricted to foreign born women i.e. women who are born in a country different from the United States. The dependent variable in columns (1)–(4) is a dummy variable that takes the value 1 if a foreign born woman is in the labor force and 0 otherwise. Data on labor force status is obtained from the complete count Census data extracted from the Integrated Public Use Microdata Series (IPUMS). FLFP in sending countries is the female labor force participation in the country of birth of the foreign born woman and it is extracted from the same decade, or a decade or two earlier, depending on data availability, from when I observe my foreign born women settler population. Data on historical female labor force participation by sending country by decade for foreign born settlers is obtained from Integrated Public Use Microdata Series (IPUMS) International Historical Censuses, (Olivetti, 2014) and (Olivetti and Petrongolo, 2016). Individual level controls include age, age squared, literacy level and the number of children. These are included throughout the analysis. Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

3.12 Appendix

3.12.1 Data on U.S. Counties

ATLAS of Historical County Boundaries provides detailed information about each county event. Events are dated and any change is stated along with the start and end date. Each U.S. county has a unique identifier, however, many versions exist depending on the number of changes to either the size or the shape of the U.S. county. The focus of this chapter is on the first event for each county identifier which is the creation of a given county. Subsequent changes in the size, shape, location or administrative status of counties are disregarded.

3.12.2 Data on FLFP in U.S. Counties

Data on female labor force participation in U.S. counties is obtained from the complete count United States Census data (1860–1940) from the Integrated Public Use Microdata Series (IPUMS). I rely on a labor force status dichotomous variable that indicates whether a person participated in the labor force to compute female labor force participation.

It must be noted that official Census accounts of female labor force participation before 1890 may be subject to under-reporting.[20]

3.12.3 Descriptive Statistics: Entire Sample of Counties Created

In Appendix Table 3.16, I provide summary statistics of the characteristics of settlers living in U.S. counties created between 1840–1940. This sample combines "new", "partitioned" and "other" counties. I present statistics for my entire sample of settlers in column (1) of Appendix Table 3.1. I also report statistics by gender in columns (2) and (3). Men constitute 57% of the settler population, and more than half of the entire settler population are men and women of prime age (15–49). The literacy rate is equal to 52%, 41% of settlers are single and lastly the average number of children is close to 2. Foreign born individuals constitute 13% of settlers, 48% were born out-of-state and the remaining 39% were born in-state. Columns (2) and (3) of Appendix Table 3.16 show that male settlers are more likely to be in their prime age, literate and single in comparison to female settlers. While out-of-state born individuals are equally distributed between men and women, the share of foreign born individuals is higher for male settlers than for female settlers.

Appendix Table 3.17 repeats these descriptive statistics by gender for foreign born, out-of-state and in-state born settlers seperately in columns (1) and (2), (3) and (4), and (5) and (6) respectively. This table shows that among the population of foreign born settlers, close to 70% are men; and that male foreign borns are more likely to be in their prime age, literate and single. Appendix Table 3.17 also shows that male individuals constitute a slightly smaller share out of the out-of-state born settlers' population in comparison to the male share of foreign born individuals. Of those born out-of-state, men are more likely to be in their prime age, literate and single in comparison to women. However, the shares of prime age, literate and single migrants are higher for male foreign born individuals than for male out-of-state born individuals (comparing column (1) to (3)). Comparing statistics for women between those born abroad and out-of-state (i.e. comparing column (2) to (4)), I find that the

[20] See Chiswick and Robinson, 2020 for a discussion of the nineteenth century Census female labor force participation measurement problems.

shares of foreign born women who are in their prime age and literate are significantly larger than for women born out-of-state. Foreign born women are less likely to be single. Columns (5) and (6) of Appendix Table 3.17 report summary statistics for in-state born individuals by gender. Interestingly, in-state born individuals are equally distributed between men and women. The prime age and literacy shares are also the same among the male and female in-state born populations, but men are more likely to be single. Finally, doing an across place of origin comparison reveals that in-state born individuals (both men and women) are less likely to be in their prime age and to be literate, whereas men and women born in-state are more likely to be single in comparison to those not born in-state.

In Appendix Table 3.18, I report summary statistics related to settlers' culture for my entire sample of U.S. counties created between 1840–1940. This sample combines "new", "partitioned" and "other" counties. The share of foreign born settlers with known female labor force participation constitute more than 70% of the entire population of foreign born settlers. This figure is obtained by dividing the total number of foreign born settlers from countries where data on labor force participation for women is available by the total foreign born settler population. Appendix Table 3.3 shows that 56% of foreign born individuals from countries with known female labor force participation are from countries with *above* decade specific median female labor force participation.

Appendix Table 3.18 shows descriptive statistics related to share of settlers coming from countries/states where women could vote. I document that 23% and 4% of the foreign born settler and out-of state settler population respectively came from places where women could vote. Lastly, Appendix Table 3.18 shows that almost 40% of out-of-state born settlers came from U.S. states where women had property and earnings rights.

3.12.4 Settlers' Culture

The share of foreign born settlers from countries known to have *above* decade specific median female labor force participation is obtained by computing first the median female labor force participation in a given decade based on data availability on labor force participation for women for sending countries. The next step is to generate the total number of foreign born migrants coming from countries with above decade specific median FLFP. The last step is to divide this by the total foreign born settler population with known female labor force participation.

To examine the share of settlers coming from places where women could vote, I construct a country/state of origin variable set equal to 1 if partial or full suffrage rights were granted to women. I then calculate the share of foreign born settlers out of the total foreign born settler population and the share of out-of-state born settlers out of the total out-of-state born settler population coming from countries/U.S. states where partial or full suffrage rights were granted to women anytime before the time when settlers are observed by me. Similarly, I compute the share of out-of-state born settlers, out of the total out-of-state born settler population, coming from U.S. states were women's financial liberation was granted anytime before the time when settlers are observed by me.

3.12.5 Settlers' Culture: All Three Measures

In this subsection, I examine the impact of settlers' culture on female labor force participation in the short and long run using all three variables that capture values

and beliefs from their place of origin in one regression. The first measure captures female labor force participation by country of origin of settlers by decade. The two other measures explore the chronological implementation and passage of women's suffrage rights across U.S. sending states and countries and variation in the passage of women's financial rights across U.S. sending states.

I report the results from doing this analysis in Appendix Table 3.20. Columns (1) and (2) display results in the short run and columns (3) and (4) report the long run results. My sample of U.S. counties is restricted to "new" counties. I include state, decade of country creation fixed effects and county level geographic controls throughout columns (1)–(4). I further include controls for geography and isolation in columns (2) and (4).

Results in Appendix Table 3.20 suggest that settlers' culture, as proxied for using the share of out-of-state born settlers coming from states where women's financial rights were granted, is robust to the inclusion of the two other measures of culture both in the short and long run. Estimates remain positive and statistically significant across columns (1)–(4) with magnitudes ranging between 0.06 and 0.09 percentage point.

Conditional on the share of settlers coming from places where women's suffrage and financial rights were implemented, I find that having more foreign born settlers from countries with high FLFP have a positive and statistically significant impact on female labor force participation in the U.S. in the short run with a magnitude of about 0.06 percentage point.

3.12.6 Selective Ex-Post Migration

One possible channel for persistence of gender values and norms is selective ex-post migration of settlers. Early settlers can determine the type of people that populate a given place/area thereafter. Transmission may thus occur by attracting similar people making locational decisions of later migrants, a function of the size of the population of early migrants already in that location.

Appendix Table 3.21 shows evidence in support of selective ex-post migration of foreign born settlers. I compute the share of settlers from a given foreign country for the top sending countries in 1850 out of total foreign born settlers. The dependent variable in columns (1) to (7) is the share of foreign born settlers from a given country of origin o out of total foreign born settlers residing in county c in state s in 1860, 1870, 1880, 1900, 1910, 1920 and 1930 respectively. The independent variable of interest is the share of foreign born settlers from a given country of origin o out of total foreign born settlers residing in county c in state s in 1850. Across columns (1)–(7), I include state fixed effects, geographic county level and additional controls for geography and isolation.

Positive and statistically significant estimates reported in Appendix Table 3.21 document that a higher share of early settlers from a given country of origin residing in U.S. counties in 1850, increases the size of later settlers population from that country of origin in that location thereafter. This thus provides suggestive evidence in support of selective ex-post migration of settlers as one plausible channel of persistence of gender norms.

3.12.7 Pre-County Creation Population

To infer the "time-at-move", I rely on two key variables in the full count Census data. First, I use household identifiers and second, I rely on a variable which indicates the relationship to the head of the household of each member of the household to remove single person households or households without at least one child. I also impose that the country/state of birth of children of households with only one child must be different from the state where the Census enumerator conducted the interview (i.e. the state of residence of the household). Otherwise, I would be capturing some households that never moved/migrated. If the household is composed of more than one child, at least one should be born in a place that is different from the current place of residence.

I follow the method that Bazzi, Fizbein, and Gebresilasse, 2018 adopt in their analysis and that is similar to Collins and Zimran, 2019 to infer the timing of migration as the difference between the current Census year and the child birth year for families with one child born before the move and zero children born after migration, divided by two. For families with one child born in the current state of residence and one child born earlier in a different country/state, I infer the move time as the difference between child birth years, divided by two.

3.12.8 Later Settlers' Culture

In this subsection, I examine the robustness of my findings for the impact of settlers' culture on female labor force participation in the U.S. in the long run by further controlling for settlers' culture in later decades. I modify Equation 3.1 to incorporate measures for later settlers' culture. These are computed as shares of foreign borns out of total foreign settlers coming from countries with high FLFP in 1920, 1930, 1940, 1950, 1980 and 1990. In addition to my main independent variable of interest, the culture of settlers residing in U.S. counties at early stages of cultural and institutional development (i.e., using data from the first census after county creation date), I therefore examine whether the culture of later settlers matters for cultural formation.

To do so, I compute my later settlers' population and culture using female labor force participation by country of origin of settlers in 1920 to 1990 (excluding 1960 and 1970 due to data restrictions) for my sample of U.S. counties created between 1850 and 1930. To fix ideas, this means that for counties created in 1850, I examine whether the culture of settlers residing in these counties in 1920 (1990) for instance, i.e. 70 (140) years later impacts FLFP in the long run. For counties created in 1910, I examine whether settlers' culture about 10 to 80 years later affects FLFP.

Appendix Table 3.22 reports the results from doing this analysis. The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the long run (using 1990 Census data). In columns (1)–(3) the sample is restricted to "new" counties. In column (4), I carry my analysis for the placebo sample of "partitioned" and "other" counties. Throughout columns (1)–(4), I include state, decade of county creation fixed effects and county level geographic controls. In columns (2)–(4), I control for initial settlers' demographic characteristics including the share of prime age population, share of literate population, the sex ratio computed as the ratio of the male over the female population, the share of the single population and the child women to ratio computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000.

Later settlers' culture is measured using the share of foreign born settlers from countries known to have *above* median female labor force participation in 1920 to 1990. The data on county level total number of foreign born individuals by country of birth is extracted from the *National Historical Geographic Information System (NHGIS)*. This data is available for the white foreign born population only.

Results in column (3) of Appendix Table 3.22 document that early settlers' culture remains robust to the inclusion of later settlers' culture. My estimates show about 0.02 percentage point increase in FLFP in U.S. counties in the long run with higher shares of settlers at early stages of cultural and institutional development coming from places with high FLFP. This is robust to controlling for early settlers' population as categorized into shares of foreign, out-of-state and in-state born settlers, early settlers' demographic characteristics, county level geographic controls, state and decade of creation fixed effects and most importantly later settlers' culture.

Estimates in column (4) of Appendix Table 3.22 show that this relationship is not robust to restricting to partitioned and subdivisioned counties and other type of counties that were not created from non-county areas ("empty" territories). The estimate for early settlers' culture is close to zero and not statistically significant.

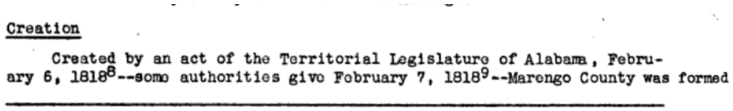
Findings from this analysis provide evidence in support of Zelinsky, 1973's doctrine that the first group of people matter much more for cultural formation than the contribution of new immigrants a few generations later.

Appendix Figures and Tables

FIGURE 3.9: Territorial expansion events

Notes: Map showing the five largest territorial expansion events after the Thirteen original colonies which constituted the United States. These include the Louisiana Purchase (1803); the Adams-Onis Treaty (1819); the Texas Annexation (1845); the Mexican Cession (1848) and the Alaska Purchase (1867). *Source*: https://www.visualcapitalist.com/us-territorial-expansion/

FIGURE 3.10: Sample Act for county creation

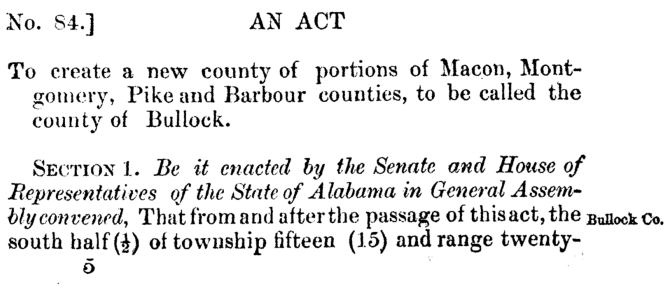


Creation

Created by an act of the Territorial Legislature of Alabama, February 6, 1818[8]--some authorities give February 7, 1818[9]--Marengo County was formed

Notes: Archived information from "Inventory of the County Archives of Alabama", Issue 46 showing the territorial act enacted in 1818 by the Territorial Legislature of Alabama which established Morengo county. *Source*: Inventory of the County Archives.

FIGURE 3.11: Sample Act for county creation as a subdivision



No. 84.] AN ACT

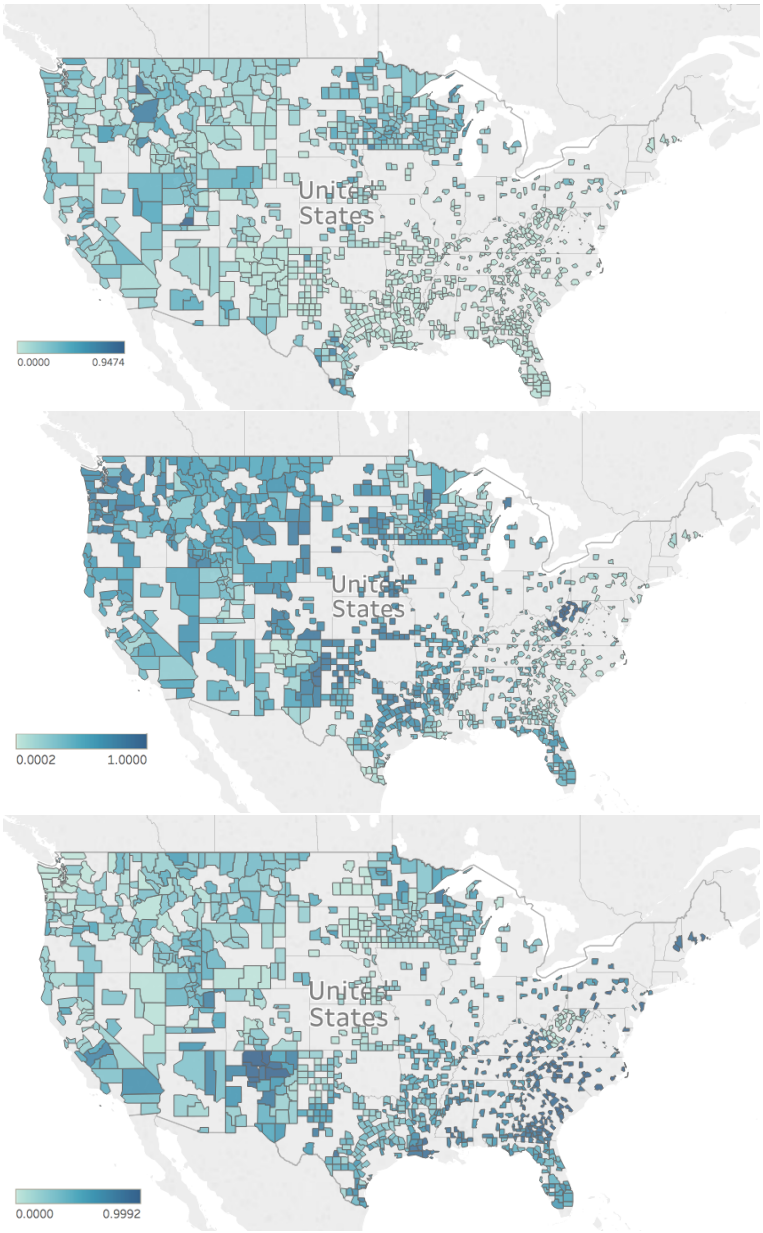
To create a new county of portions of Macon, Montgomery, Pike and Barbour counties, to be called the county of Bullock.

SECTION 1. *Be it enacted by the Senate and House of Representatives of the State of Alabama in General Assembly convened*, That from and after the passage of this act, the south half (½) of township fifteen (15) and range twenty- Bullock Co.

5

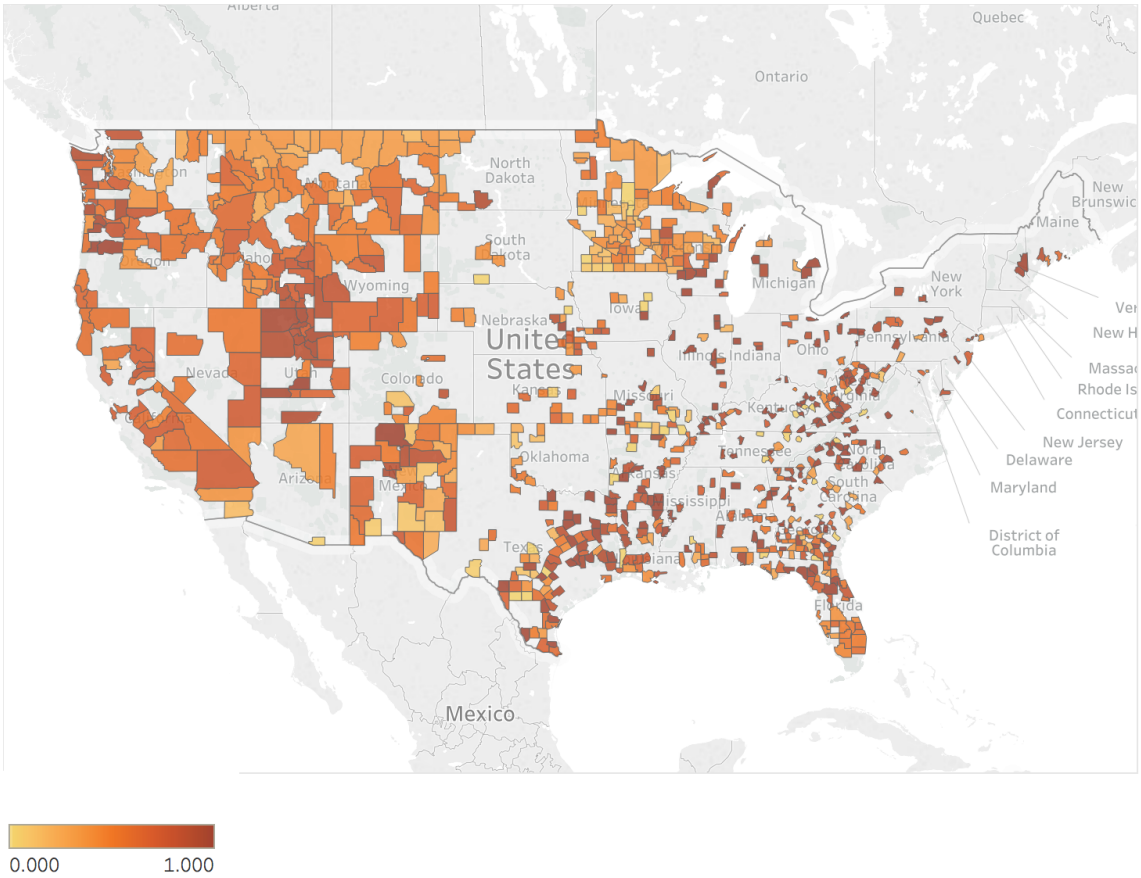
Notes: Extract from the "General Assembly of Alabama", showing an act enabled by Alabama state to establish a new county as a subdivision of previously formed counties. *Source*: U.S. state constitution for Alabama.

FIGURE 3.12: Settler population by origin: Sample of "partitioned" and "other" counties



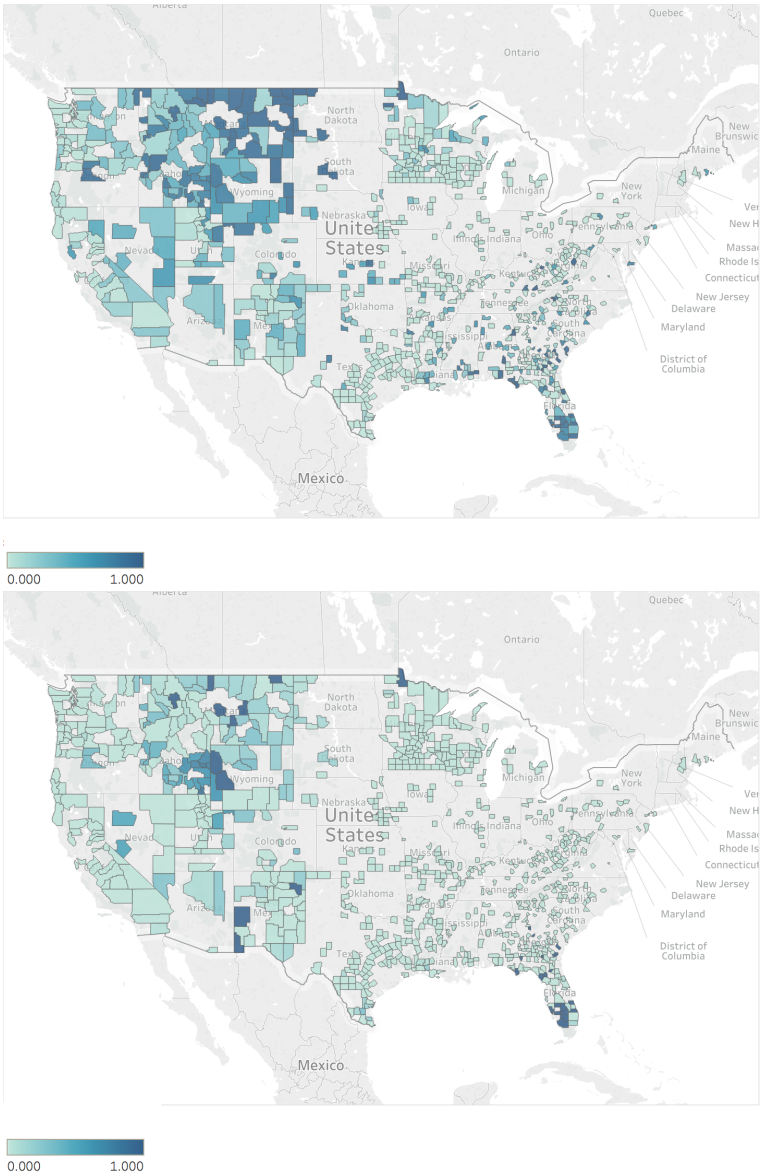
Notes: Shares of foreign born, out-of-state and in-state born individuals out of the total population are displayed respectively. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals are those born in the same state as the one where the household is located. Shares of foreign borns settlers out of the total population range between 0 and 1 with 0 indicating no foreign born settlers in the county and 1 indicating that all individuals living in the county are foreign born. Light blue indicates lower shares and dark blue greater shares. Grey areas indicate counties that are excluded from my sample. *Source*: Author's compilation based on data from the Integrated Public Use Microdata Series (IPUMS).

FIGURE 3.13: Map of foreign born settlers from countries with high FLFP: Sample of "partitioned" and "other" counties



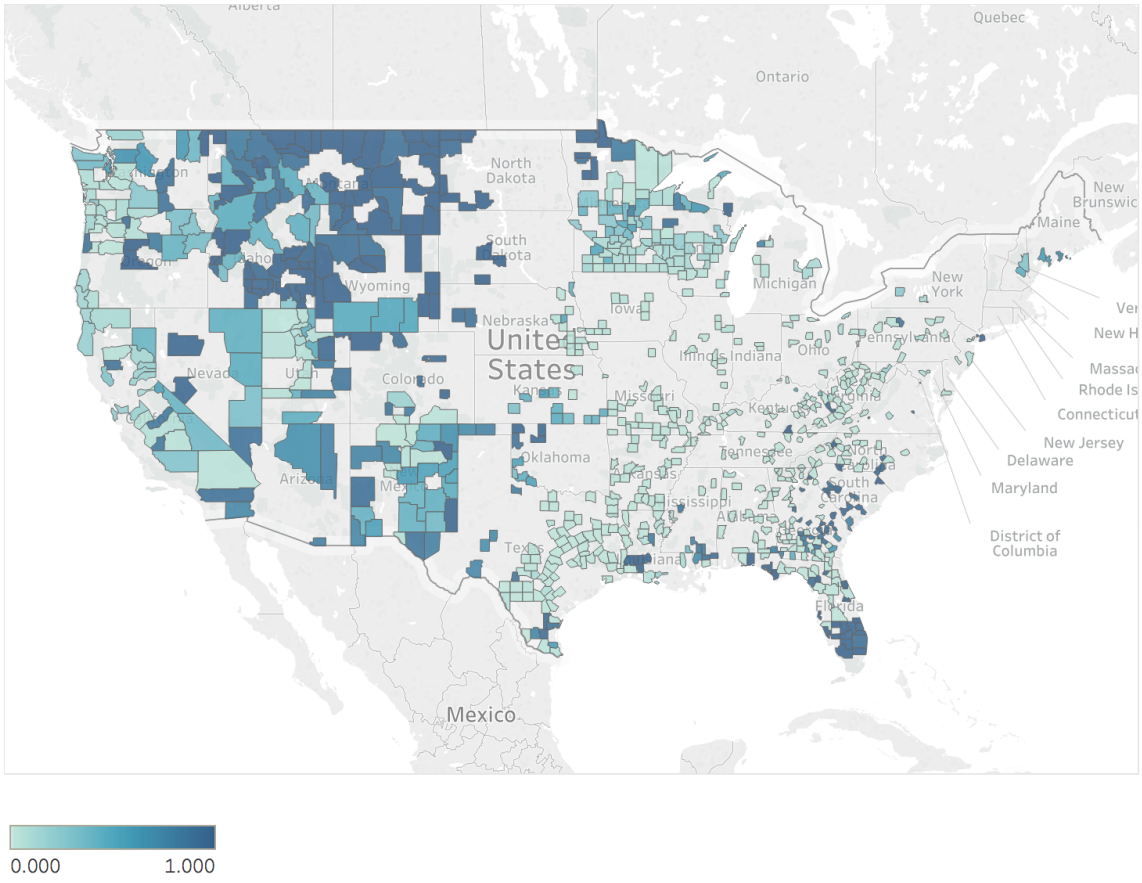
Notes: Map showing the share of foreign born settlers from countries *known* to have *above* median female labor force participation. It is obtained as follows: first median FLFP in a given decade based on data availability on FLFP for sending countries is computed. Then the total number of foreign born settlers coming from countries with above decade specific median FLFP is computed and lastly this total is divided by the total foreign born settler population with *known* FLFP i.e. foreign born settlers from countries with available data on FLFP. Grey areas indicate counties that are excluded from my sample. Light yellow indicates a small share of foreign born settlers coming from counties with *above* median female labor force participation and dark orange indicates a high share. *Source*: Author's compilation.

FIGURE 3.14: Foreign and out-of-state settlers who could vote: Sample of "partitioned" and "other" counties



Note: Map showing the share of foreign born settlers out of the total foreign born settler population and out-of-state born settlers out of the total out-of-state settler population coming from countries/U.S. states where partial or full suffrage rights were granted to women anytime before the time when settlers are observed. Grey areas indicate counties that are excluded from my sample. Light blue indicates a small share of foreign born settlers/out-of-state born settlers coming from countries/U.S. states where women could vote and dark blue indicates a high share. *Source*: Author's compilation.

FIGURE 3.15: Out-of-state settlers from countries with women's financial liberation: Sample of "partitioned" and "other" counties



Notes: Map showing the share of out-of-state born settlers out of the total out-of-state settler population coming from U.S. states where women's financial liberation was granted anytime before the time when settlers are observed. Grey areas indicate counties that are excluded from my sample. Light blue indicates a small share of out-of-state born settlers coming from U.S. states where women had property and earning rights and blue indicates a high share. *Source*: Author's compilation.

TABLE 3.16: Descriptive statistics: Entire sample of counties created between 1840–1940

		By Gender	
	Entire population	Male population	Female population
	(1)	(2)	(3)
Demographic Characteristics			
Share male population	0.57		
	(0.10)		
Share prime age population	0.54	0.56	0.49
	(0.11)	(0.13)	(0.07)
Share literate population	0.52	0.55	0.46
	(0.20)	(0.20)	(0.18)
Share single population	0.41	0.49	0.24
	(0.15)	(0.16)	(0.09)
Average number of children	1.9		
	(0.58)		
Child women ratio	677.02		
	(226.17)		
Settlers' Population			
Share foreign born	0.13	0.15	0.11
	(0.16)	(0.16)	(0.14)
Share out-of-state born	0.48	0.48	0.47
	(0.25)	(0.25)	(0.26)
Share in-state born	0.38	0.36	0.41
	(0.28)	(0.28)	(0.28)

Notes: Based on complete count Census data from the Integrated Public Use Microdata Series (IPUMS). My sample of U.S. counties includes those created between 1840–1940. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals are those born in the same state as the one where the household is located. All shares range between 0 and 1. In column (1), shares are displayed out of the total settler population. In columns (2) and (3), summary statistics are displayed by gender. Shares in columns (2) and (3) are displayed out of the male settler population and the female settler population respectively. Prime age refers to ages 15 to 49. The child women ratio is computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000. Standard deviations are reported in parentheses.

TABLE 3.17: Descriptive statistics by settlers' origin and gender: Entire sample of counties created between 1840–1940

	Foreign born		**Out-of-state born**		**In-state born**	
	Male Population (1)	Female Population (2)	Male Population (3)	Female Population (4)	Male Population (5)	Female Population (6)
Share prime age population	0.74	0.71	0.65	0.60	0.22	0.22
	(0.17)	(0.19)	(0.12)	(0.11)	(0.19)	(0.18)
Share literate population	0.81	0.74	0.68	0.58	0.21	0.19
	(0.18)	(0.23)	(0.19)	(0.21)	(0.21)	(0.19)
Share single population	0.46	0.15	0.45	0.21	0.64	0.41
	(0.46)	(0.15)	(0.19)	(0.09)	(0.22)	(0.20)
Average number of children		2.68		2.09		1.03
		(2.87)		(1.06)		(0.76)
Child women ratio		954.82		733.30		421.21
		(1021.86)		(408.92)		(306.53)

Notes: Based on complete count Census data from the Integrated Public Use Microdata Series (IPUMS). My sample of U.S. counties includes those created between 1840–1940. Settlers' population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals are those born in the same state as the one where the household is located. All shares range between 0 and 1. The share of male population: *foreign born* is 0.69 (0.14); the share of male population: *out-of-state born* is 0.59 (0.10) and the share of male population: *in-state born* is 0.51 (0.09). In columns (1) and (2) I report statistics for foreign born settlers by gender i.e. shares are displayed out of the total male foreign born settler population and the total female foreign born settler population respectively . In columns (3) and (4) I report statistics for out-of-state born settlers by gender i.e. shares are displayed out of the total male out-of-state born settler population and the total female out-of-state born settler population respectively. In columns (5) and (6) I report statistics for in-state born settlers by gender i.e. shares are displayed out of the total male in-state born settlers' population and the total female in-state born settlers' population respectively. Prime age refers to ages 15 to 49. The child women ratio is computed as the ratio of the number of children less than 5 years old over the number of women in their childbearing age times 1000. Standard deviations are reported in parentheses.

TABLE 3.18: Settlers' culture: Descriptive statistics: Entire sample of counties created between 1840–1940

	Mean (1)	SD (2)	(N) (3)
Settlers' culture			
Female labor force participation			
Share foreign born with *known* FLFP	0.76	0.25	1,444
Share foreign born with *unknown* FLFP	0.24	0.25	1,444
Share foreign born : *Above* median FLFP	0.56	0.29	1,486
Share foreign born : *Below* median FLFP	0.44	0.29	1,486
Suffrage			
Share foreign born: Women suffrage rights	0.23	0.28	1,464
Share out-of-state born: Women suffrage rights	0.04	0.16	1,486
Foreign born: suffrage intensity	7.75	13.31	1,486
Out-of-state born: suffrage intensity	0.87	3.6	1,486
Financial liberation			
Share out-of-state born: Women's financial liberation	0.39	0.39	1,486
Out-of-state born: financial liberation intensity	8.75	12.32	1,486

Notes: Data on historical female labor force participation by sending country by decade for foreign born settlers is obtained from Integrated Public Use Microdata Series (IPUMS) International Historical Censuses, (Olivetti, 2014) and (Olivetti and Petrongolo, 2016). Female labor force participation from sending countries is extracted from the same decade, or a decade or two earlier, depending on data availability, from when I observe my foreign born settlers. Share foreign born with *known* FLFP is obtained by dividing the total number of foreign born settlers from countries with *known* FLFP i.e. where data on FLFP is available, over the total foreign born settler population. The share of foreign born settlers from countries *known* to have *above* median female labor force participation is obtained by first computing the median FLFP across countries in a given decade based on data availability on FLFP for sending countries. Then the total number of foreign born settlers coming from countries with above decade specific median FLFP is computed and lastly this total is divided by the total foreign born settler population with *known* FLFP i.e. foreign born settlers from countries with available data on FLFP. Share of foreign born settlers from countries *known* to have *below* median female labor force is computed the same way. Shares of foreign born settlers from countries *known* to have *above* and *below* median FLFP add up to 1. Shares of foreign born settlers from countries with *known* and *unknown* FLFP add up to the entire foreign born settler population. *Share foreign born: Women suffrage rights* and *Share out-of-state born: Women suffrage rights* are the share of foreign born settlers out of the total foreign born settlers population and the share of out-of-state born settlers out of the total out-of-state born settlers population coming from countries/U.S. states where partial or full suffrage rights were granted to women anytime before the time when settlers are observed. Women's suffrage *intensity* is measured as a weighted share of settlers coming from places were partial/full voting rights were granted to women weighted by the number of years between suffrage laws' passage and county creation. *Share out-of-state born: Women's financial liberation* is the share of out-of-state born settlers out of the total out-of-state born settlers population coming from U.S. states where women's financial liberation was granted anytime before the time when settlers are observed. Women's financial liberation *intensity* is measured as a weighted share of out-of-state born settlers coming from U.S. states where property rights and earnings rights were granted to women weighted by the number of years between women's financial liberation laws' passage and county creation.

TABLE 3.19: Settlers' culture: Suffrage intensity

	Female labor force participation			
	(1)	(2)	(3)	(4)
Place of origin				
Share foreign born	0.112*	0.114*	0.111	0.00621
	(0.061)	(0.059)	(0.068)	(0.050)
Share out-of-state born	0.0262	0.0290	0.0333	0.0617
	(0.070)	(0.072)	(0.078)	(0.039)
Norms at the place of origin				
Foreign born: Suffrage intensity	-0.0000349	-0.0000379	-0.0000226	0.000172
	(0.000)	(0.000)	(0.000)	(0.000)
Out-of-state born: Suffrage intensity	-0.0149	-0.00360	-0.00351	0.00108
	(0.026)	(0.025)	(0.025)	(0.001)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes	Yes
Additional Controls			Yes	Yes
N	**368**	**368**	**368**	**762**

Notes: The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the short run. Data on labor force participation is based on the first U.S. Census available after county creation. My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In column (4), the sample is restricted to "partitioned" and "other" counties. Settler population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using women's suffrage rights *intensity* to reflect gender norms at the place of origin. Women's suffrage intensity is measured as a weighted share of settlers coming from places where partial/full voting rights were granted to women weighted by the number of years between suffrage laws' passage and county creation time. If suffrage rights were not yet granted, the negative difference between decade of county creation and year of passage of suffrage rights is replaced by *zero*. State and decade of county creation fixed effects are included in columns (1)–(4). Geographic controls are included across columns (1)–(4), demographic controls are introduced starting column (2) to (4) and lastly additional county level controls for geography and isolation are introduced in columns (3) and (4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

TABLE 3.20: Settlers culture: All three measures

	Female labor force participation			
	Short Run		*Long Run*	
	(1)	(2)	(3)	(4)
Place of origin				
Share foreign born	0.0822	0.0846	0.0483	0.0315
	(0.058)	(0.059)	(0.031)	(0.029)
Share out-of-state born	0.0272	0.0299	-0.0137	-0.0223
	(0.067)	(0.070)	(0.030)	(0.030)
Norms at the place of origin				
Share foreign born: *Above* median FLFP	0.0582**	0.0557*	0.0128	0.0123
	(0.028)	(0.028)	(0.009)	(0.009)
Share foreign born: Suffrage rights	-0.0575*	-0.0548	0.0549	0.0572
	(0.034)	(0.038)	(0.042)	(0.042)
Share out-of-state born: Suffrage rights	-0.936	-0.496	-0.697	-0.617
	(0.655)	(0.706)	(1.251)	(1.303)
Share out-of-state born: Financial liberation	0.0687*	0.0634*	0.0912***	0.0838**
	(0.037)	(0.037)	(0.034)	(0.035)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Additional Controls		Yea		Yes
N	**367**	**367**	**360**	**360**

Notes: The dependent variable in columns (1) and (2) is female labor force participation in U.S. counties in the short run and in columns (3) and (4), the dependent variable is female labor force participation in U.S. counties in the long run. Data on labor force participation is based on the first (tenth) U.S. Census available after county creation for the short (long) run analysis. My sample of U.S. counties is restricted to "new" counties. Settler population is based on the first U.S. Census available after the county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using female labor force participation, women's suffrage rights and women's financial rights to reflect gender norms at the place of origin. State and decade of county creation fixed effects are included in columns (1)–(4). Geographic controls are included across columns (1)–(4), and additional county level controls for geography and isolation are introduced in columns (2) and (4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

TABLE 3.21: Selective ex post migration

	Share foreign born$_{ocsd}$ in U.S. counties						
	1860	1870	1880	1900	1910	1920	1930
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share foreign born$_{ocs1850}$	0.729***	0.646***	0.600***	0.431***	0.337***	0.249***	0.218***
	(0.016)	(0.016)	(0.016)	(0.019)	(0.018)	(0.016)	(0.015)
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	**11018**	**11053**	**11060**	**11067**	**11074**	**11074**	**11074**

Notes: The dependent variable in columns (1) to (7) is the share of foreign born settlers from a given country of origin *o* out of total foreign born settlers residing in county *c* in state *s* in 1860, 1870, 1880, 1900, 1910, 1920 and 1930 respectively. The independent variable of interest is the share of foreign born settlers from a given country of origin *o* out of total foreign born settlers residing in county *c* in state *s* in 1850. State fixed effects are included in columns (1)–(7). Geographic controls and additional county level controls for geography and isolation are included across columns (1)–(7). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

TABLE 3.22: Later settlers' culture: FLFP

	Female labor force participation			
	(1)	(2)	(3)	(4)
Settlers' population				
Share foreign born	0.0829**	0.0525	0.0652	0.0737**
	(0.041)	(0.048)	(0.048)	(0.029)
Share out-of-state born	0.0596	0.0452	0.0469	0.0186
	(0.040)	(0.047)	(0.048)	(0.023)
Settlers' culture				
Share foreign born: *Above* median FLFP	0.0236*	0.0235*	0.0233*	0.00511
	(0.013)	(0.012)	(0.012)	(0.009)
Later settlers' culture				
In 1920: *Above* median FLFP			0.00351	-0.00798
			(0.016)	(0.009)
In 1930: *Above* median FLFP			0.0258	0.0121
			(0.030)	(0.010)
In 1940: *Above* median FLFP			-0.000359	0.0325***
			(0.022)	(0.010)
In 1950: *Above* median FLFP			-0.0349	-0.0119
			(0.031)	(0.030)
In 1980: *Above* median FLFP			-0.0120	0.0119*
			(0.011)	(0.007)
In 1990: *Above* median FLFP			0.00680	0.00833
			(0.015)	(0.011)
State FE	Yes	Yes	Yes	Yes
Decade of creation FE	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Demographic Controls		Yes	Yes	Yes
N	**409**	**409**	**409**	**802**

Notes: The dependent variable in columns (1)–(4) is female labor force participation in U.S. counties in the long run. Data on labor force participation is based on the 1990 U.S. Census. My sample of U.S. counties includes those created between 1840–1940. In columns (1)–(3) the sample is restricted to "new" counties. In column (4), the sample is restricted to "partitioned" and "other" counties. Settlers' population is based on the first U.S. Census available after county creation date. Foreign born individuals are those who are born in a country different from the United States. Out-of-state born individuals are those who are born in a U.S. state that is different from the state in which the household is located when the Census enumerator conducted the interview. In-state born individuals (the omitted category) are those born in the same state as the one where the household is located. All shares range between 0 and 1. Settlers' culture is proxied for using female labor force participation from sending countries to reflect gender norms at the place of origin. Data on historical female labor force participation by sending country by decade for foreign born settlers is obtained from Integrated Public Use Microdata Series (IPUMS) International Historical Censuses, (Olivetti, 2014) and (Olivetti and Petrongolo, 2016). Female labor force participation from sending countries is extracted from the same decade, or a decade or two earlier, depending on data availability, from when I observe my foreign born settlers' population. Share of foreign born settlers from countries *known* to have above median female labor force participation is obtained by first computing the median FLFP in a given decade based on data availability on FLFP for sending countries. Second, the total number of foreign born settlers coming from countries with above decade specific median FLFP is computed and lastly this total is divided by the total foreign born settlers population with *known* FLFP. Foreign born individuals from countries with *unknown* FLFP in a given decade are excluded from the foreign born settler population. Share of foreign born settlers from countries *known* to have *below* median female labor force is the omitted category. Later settlers' culture is measured using the share of foreign born settlers from countries *known* to have above median female labor force participation in 1920 to 1990. Data on county level total number of foreign born individuals by country of birth is extracted from the *National Historical Geographic Information System (NHGIS)*. This data is available for the white foreign born population. Data on FLFP for later decades is extracted from (Olivetti, 2014) and (Olivetti and Petrongolo, 2016). State and decade of creation fixed effects are included in columns (1)–(4). The set of geographic controls are included in columns (1)–(4). The set of demographic controls are included in columns (2)–(4). Standard errors clustered on 60-square-mile grid cells are reported between parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Bibliography

Adamczyk, Amy and Cassady Pitt (2009). "Shaping Attitudes about Homosexuality: The Role of Religion and Cultural Context". In: *Social Science Research* 38.2, pp. 338–351.

Akerlof, George A and Rachel E Kranton (2000). "Economics and identity". In: *Quarterly Journal of Economics* 115.3, pp. 715–753.

Aksoy, Cevat Giray et al. (2019). "Gay Glass Ceilings: Sexual Orientation and Workplace Authority in the UK". In: *Journal of Economic Behavior & Organization* 159, pp. 167–180.

Alesina, Alberto and Paola Giuliano (2010). "The power of the family". In: *Journal of Economic growth* 15.2, pp. 93–125.

— (2015). "Culture and institutions". In: *Journal of Economic Literature* 53.4, pp. 898–944.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn (2013). "On the origins of gender roles: Women and the plough". In: *Quarterly Journal of Economics* 128.2, pp. 469–530.

Algan, Yann and Pierre Cahuc (2006). "Job protection: The macho hypothesis". In: *Oxford Review of Economic Policy* 22.3, pp. 390–410.

Allegretto, Sylvia A and Michelle M Arthur (2001). "An Empirical Analysis of Homosexual/Heterosexual Male Earnings Differentials: Unmarried and Unequal?" In: *ILR Review* 54.3, pp. 631–646.

Altemeyer, Bob (2002). "Changes in Attitudes Toward Homosexuals". In: *Journal of Homosexuality* 42.2, pp. 63–75.

Anderson, Siwan and Debraj Ray (2010). "Missing Women: Age and Disease". In: *Review of Economic Studies* 77.4, pp. 1262–1300.

Angrist, Josh (2002). "How do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation". In: *Quarterly Journal of Economics* 117.3, pp. 997–1038.

Antecol, Heather (2000). "An examination of cross-country differences in the gender gap in labor force participation rates". In: *Labour Economics* 7.4, pp. 409–426.

Asal, Victor, Udi Sommer, and Paul G Harwood (2013). "Original Sin: A Cross-National Study of the Legality of Homosexual Acts". In: *Comparative Political Studies* 46.3, pp. 320–351.

Ashraf, Quamrul and Oded Galor (2011). "Dynamics and stagnation in the Malthusian epoch". In: *American Economic Review* 101.5, pp. 2003–41.

Barro, Robert J and Rachel M McCleary (2003). "Religion and economic growth across countries". In: *American Sociological Review* 68.5, p. 760.

Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilasse (2017). "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States". NBER Working Paper 23997.

Bazzi, Samuel, Martin Fizbein, and Mesay Gebresilasse (2018). ""Frontier Culture: The roots and persistence of "rugged Individualism" in the United States"". Boston University Working Paper.

Becker, Sascha O and Ludger Woessmann (2008). "Luther and the girls: Religious denomination and the female education gap in nineteenth-century Prussia". In: *Scandinavian Journal of Economics* 110.4, pp. 777–805.

Bester, C Alan, Timothy G Conley, and Christian B Hansen (2011). "Inference with dependent data using cluster covariance estimators". In: *Journal of Econometrics* 165.2, pp. 137–151.

Bisin, Alberto and Thierry Verdier (2011). "The economics of cultural transmission and the dynamics of preferences". In: *Journal of Economic Theory* 97.2, pp. 298–319.

— (2017). *On the joint evolution of culture and institutions*. Tech. rep. National Bureau of Economic Research.

Black, Dan et al. (2000). "Demographics of the Gay and Lesbian Population in the United States: Evidence from Available Systematic Data Sources". In: *Demography* 37.2, pp. 139–154.

— (2002). "Why do Gay Men Live in San Francisco?" In: *Journal of Urban Economics* 51.1, pp. 54–76.

Black, Dan A, Seth G Sanders, and Lowell J Taylor (2007). "The Economics of Lesbian and Gay Families". In: *Journal of Economic Perspectives* 21.2, pp. 53–70.

Blau, Francine D (1992). "The fertility of immigrant women: Evidence from high-fertility source countries". In: *Immigration and the work force: Economic consequences for the United States and source areas*. University of Chicago Press, pp. 93–134.

Blau, Francine D and Lawrence M Kahn (2015). "Substitution between individual and source country characteristics: Social capital, culture, and us labor market outcomes among immigrant women". In: *Journal of Human Capital* 9.4, pp. 439–482.

Blau, Francine D, Lawrence M Kahn, and Kerry L Papps (2011). "Gender, source country characteristics, and labor market assimilation among immigrants". In: *Review of Economics and Statistics* 93.1, pp. 43–58.

Blau, Francine D et al. (2020). "Is there still son preference in the United States?" In: *Journal of Population Economics*, pp. 1–42.

Boserup, Ester (1970). *Woman's Role in Economic Development*. George Allen and Universe.

Boyd, Nan Alamilla (2003). *Wide-Open Town: A History of Queer San Francisco to 1965*. University of California Press.

Brodeur, Abel and Joanne Haddad (2018). "Institutions, Attitudes and LGBT: Evidence from the Gold Rush". IZA Discussion Paper 11957.

Buonanno, Paolo et al. (2015). "Poor Institutions, Rich Mines: Resource Curse in the Origins of the Sicilian Mafia". In: *Economic Journal* 125.586, pp. 175–202.

Cagé, J and V Rueda (2017). "Sex and the Mission: The Conflicting Effects of Early Christian Investments on Sub-Saharan Africa's HIV Epidemic". CEPR Working Paper DP12192.

Campa, Pamela and Michel Serafinelli (2019). "Politico-economic regimes and attitudes: Female workers under state socialism". In: *Review of Economics and Statistics* 101.2, pp. 233–248.

Carpenter, Christopher and Gary J Gates (2008). "Gay and Lesbian Partnership: Evidence from California". In: *Demography* 45.3, pp. 573–590.

Chauncey, George (1995). *Gay New York: Gender, Urban Culture, and the Making of the Gay Male World, 1890–1940*. Hachette UK.

Chiswick, Barry R and RaeAnn Halenda Robinson (2020). "Women at Work in the Pre-Civil War United States: An Analysis of Unreported Family Workers". In:

Christafore, David, J Sebastian Leguizamon, and Susane Leguizamon (2013). "Are Black Neighborhoods Less Welcoming to Homosexuals than White Neighborhoods?" In: *Regional Science and Urban Economics* 43.4, pp. 579–589.

Clay, Karen and Randall Jones (2008). "Migrating to Riches? Evidence from the California Gold Rush". In: *Journal of Economic History* 68.4, pp. 997–1027.

Clay, Karen and Werner Troesken (2000). "Squatting and the Settlement of the United States: New Evidence from Post-Gold Rush California". In: *New Frontiers in Agricultural History*. Emerald Group Publishing Limited, pp. 207–233.

Clay, Karen and Gavin Wright (2005). "Order Without Law? Property Rights during the California Gold Rush". In: *Explorations in Economic History* 42.2, pp. 155–183.

Coffman, Katherine B, Lucas C Coffman, and Keith M Marzilli Ericson (2016). "The Size of the LGBT Population and the Magnitude of Antigay Sentiment Are Substantially Underestimated". In: *Management Science* 63.10, pp. 3168–3186.

Collins, William J and Ariell Zimran (2019). "The economic assimilation of Irish Famine migrants to the United States". In: *Explorations in Economic History* 74, p. 101302.

Conley, Timothy G (1999). "GMM Estimation with Cross Sectional Dependence". In: *Journal of Econometrics* 92.1, pp. 1–45.

— (2008). "Spatial Econometrics". In: *The New Palgrave Dictionary of Economics*. Ed. by S N Durlauf and L E Blume. Palgrave Macmillan.

Couttenier, M and M Sangnier (2015). "Living in the Garden of Eden: Mineral Resources and Preferences for Redistribution". In: *Journal of Comparative Economics* 43.2, pp. 243–256.

Couttenier, Mathieu, Pauline Grosjean, and Marc Sangnier (2017). "The Wild West is Wild: The Homicide Resource Curse". In: *Journal of the European Economic Association* 15.3, pp. 558–585.

Cushman, Barry (1998). *Rethinking the new deal court: The structure of a constitutional revolution*. Oxford University Press.

Fernandez, Raquel (2007). "Women, work, and culture". In: *Journal of the European Economic Association* 5.2-3, pp. 305–332.

— (2011). "Does culture matter?" In: *Handbook of social economics*. Vol. 1. Elsevier, pp. 481–510.

Fernandez, Raquel and Alessandra Fogli (2009). "Culture: An empirical investigation of beliefs, work, and fertility". In: *American Economic Journal: Macroeconomics* 1.1, pp. 146–77.

Fernandez, Raquel, Alessandra Fogli, and Claudia Olivetti (2004). "Mothers and sons: Preference formation and female labor force dynamics". In: *Quarterly Journal of Economics* 119.4, pp. 1249–1299.

Fernández, Raquel, Sahar Parsa, and Martina Viarengo (2019). "Coming out in America: AIDS, Politics, and Cultural Change". NBER Working Paper 25697.

Fisher, Robin, Geof Gee, and Adam Looney (2018). "Same-Sex Married Tax Filers After Windsor and Obergefell". In: *Demography* 55.4, pp. 1423–1446.

Fortin, Nicole M (2005). "Gender role attitudes and the labour-market outcomes of women across OECD countries". In: *Oxford review of Economic Policy* 21.3, pp. 416–438.

Gallman, Robert E et al. (1972). "American economic growth: An Economist's history of the United States". In:

Gates, Gary J (2013). "Geography of the LGBT Population". In: *International Handbook on the Demography of Sexuality*. Ed. by A Baumle. Springer, pp. 229–242.

— (2015). "Comparing LGBT Rankings by Metro Area: 1990-2014". Williams Institute, UCLA School of Law.

Geddes, Rick and Lueck Dean (2002). "The gains from self-ownership and the expansion of women's rights". In: *American Economic Review* 92.4, pp. 1079–1092.
Giuliano, Paola (2007). "Living arrangements in western europe: Does cultural origin matter?" In: *Journal of the European Economic Association* 5.5, pp. 927–952.
— (2020). "Gender and culture". In: IZA Discussion Paper 13607.
Giuliano, Paola and Nathan Nunn (2017). "Understanding Cultural Persistence and Change". NBER Working Paper 23617.
Goldin, Claudia (1995). "The U-Shaped Female Labor Force Function in Economic Development and Economic History". In: *Investment in Women's Human Capital and Economic Development*. Ed. by T. P. Schultz. University of Chicago Press, pp. 61–90.
Gonzales, C. et al. (2015). "Fair Play: More Equal Laws Boost Female Labor Force Participation". In: *IMF Staff Discussion Note*.
Goodrich, C A (1859). *The Land we Live in: Travels, Sketches and Adventures in North and South America*. H. M. Rulison Queen City Publishing House.
Gordon, R. J. (2016). *The Rise and Fall of American Growth: The U.S. Standard of Living Since the Civil War*. Princeton University Press.
Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. (2005). "Engines of Liberation". In: *Review of Economic Studies* 72.1, pp. 109–133.
Grosjean, Pauline and Rose Khattar (2019). "It's raining men! Hallelujah? The long-run consequences of male-biased sex ratios". In: *Review of Economic Studies* 86.2, pp. 723–754.
— (Forthcoming). "It's Raining Men! Hallelujah?: The Long-Run Consequences of Male-Biased Sex Ratios". In: *Review of Economic Studies*.
Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2006). "Does Culture Affect Economic Outcomes?" In: *Journal of Economic Perspectives* 20.2, pp. 23–48.
— (2009). "Cultural biases in economic exchange?" In: *Quarterly Journal of Economics* 124.3, pp. 1095–1131.
Hamermesh, D. S. (2019). "Marital Surplus and Sexual Preference: The Importance of the Legal Environment". Mimeo Barnard College.
Hansen, Casper Worm, Peter Sandholt Jensen, and Christian Volmar Skovsgaard (2015). "Modern gender roles and agricultural history: the Neolithic inheritance". In: *Journal of Economic Growth* 20.4, pp. 365–404.
Heath, R. and Seema Jayachandran (2017). " The Causes and Consequences of Increased Female Education and Labor Force Participation in Developing Countries". In: *Oxford Handbook of Women and the Economy*.
Hesketh, Therese and Zhu Wei Xing (2006). "Abnormal Sex Ratios in Human Populations: Causes and Consequences". In: *Proceedings of the National Academy of Sciences* 103.36, pp. 13271–13275.
Holliday, James S (2015). *The World Rushed in: The California Gold Rush Experience*. University of Oklahoma Press.
Hsiang, Solomon M (2010). "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America". In: *Proceedings of the National Academy of Sciences* 107.35, pp. 15367–15372.
Hurtado, Albert L (1999). "Sex, Gender, Culture, and a Great Event: The California Gold Rush". In: *Pacific Historical Review* 68.1, pp. 1–19.
Jayachandran, Seema (2015). "The Roots of Gender Inequality in Developing Countries". In: *Annual Review of Economics* 7, pp. 63–88.
Lipsky, William (2006). *Gay and Lesbian San Francisco*. Arcadia Publishing.
Lott, J. R. and L. W. Kenny (1999). "Did women's suffrage change the size and scope of government?" In: *Journal of Political Economy* 107.6, pp. 1163–1198.

MacKinnon, James G and Matthew D Webb (2017). "Wild Bootstrap Inference for Wildly Different Cluster Sizes". In: *Journal of Applied Econometrics* 32.2, pp. 233–254.

Maffly-Kipp, Laurie F (1994). *Religion and Society in Frontier California*. Yale University Press.

McDowell, Andrea G (2002). "From Commons to Claims: Property Rights in the California Gold Rush". In: *Yale Journal of Law & the Humanities* 14, pp. 1–72.

Megdal, Sharon Bernstein and Michael R. Ransom (1985). "Longitudinal Changes in Salary at a Large Public University: What Response to Equal Pay Legislation?" In: *American Economic Review* 75.2, pp. 271–274. ISSN: 00028282.

Miller, G. (2008). "Women's suffrage, political responsiveness, and child survival in American history". In: *Quarterly Journal of Economics* 123.3, pp. 1287–1327.

Mullen, L (2018). "Divergence in U.S. Sex Ratios by County, 1820-2010". Interactive Map, http://lincolnmullen.com/projects/sex-ratios/.

Munshi, Kaivan (2003). "Networks in the Modern Economy: Mexican Migrants in the US Labor Market". In: *Quarterly Journal of Economics* 118.2, pp. 549–599.

Nugent, Walter (2007). *Into the West: The Story of Its People*. Vintage.

Nunn, Nathan (2010). "Religious Conversion in Colonial Africa". In: *American Economic Review: Papers & Proceedings* 100.2, pp. 147–52.

— (2020). "History as evolution". NBER Working Paper 27706.

Nunn, Nathan et al. (2014). "Gender and missionary influence in colonial Africa". In: *African Development in Historical Perspective*.

Olivetti, Claudia (2014). "The Female labor force and long-run development". In: *Human Capital in History: The American Record*, p. 161.

Olivetti, Claudia and Barbara Petrongolo (2016). "The evolution of gender gaps in industrialized countries". In: *Annual Review of Economics* 8, pp. 405–434.

Posner, Richard A (1994). *Sex and Reason*. Harvard University Press.

Ramey, V. A. (2009). " Time Spent in Home Production in the Twentieth-Century United States: New Estimates from Old Data". In: *Journal of Economic History* 69.1, pp. 1–47.

Roodman, David et al. (2019). "Fast and Wild: Bootstrap Inference in Stata using Boottest". In: *Stata Journal* 19.1, pp. 4–60.

Sansone, Dario (2019). "Pink Work: Same-Sex Marriage, Employment and Discrimination". In: *Journal of Public Economics* 180, p. 104086.

Schacht, Ryan and Karen L Kramer (2016). "Patterns of Family Formation in Response to Sex Ratio Variation". In: *PloS one* 11.8, e0160320.

Shayo, Moses (2009). "A model of social identity with an application to political economy: Nation, class, and redistribution". In: *American Political Science Review*, pp. 147–174.

Shevsky, Eshred and Wendell Bell (1955). *Social Area Analysis: Theory, Illustrative Application, and Computational Procedures*. Stanford University Press.

Smart, Michael J and Andrew H Whittemore (2017). "There Goes the Gaybourhood? Dispersion and Clustering in a Gay and Lesbian Real Estate Market in Dallas TX, 1986–2012". In: *Urban Studies* 54.3, pp. 600–615.

Tabellini, Guido (2008). "Institutions and culture". In: *Journal of the European Economic association* 6.2-3, pp. 255–294.

— (2010). "Culture and institutions: Economic development in the regions of Europe". In: *Journal of the European Economic association* 8.4, pp. 677–716.

Teso, Edoardo (2019). "The long-term effect of demographic shocks on the evolution of gender roles: Evidence from the transatlantic slave trade". In: *Journal of the European Economic Association* 17.2, pp. 497–534.

Vidart, Daniela (2020). "Human Capital, Female Employment, and Electricity: Evidence from the Early 20th Century United States". University of California San Diego Working Paper.

Von Berlepsch, Viola and Andrés Rodríguez-Pose (2019). "The missing ingredient: distance. Internal migration and its long-term economic impact in the United States". In: *Journal of Ethnic and Migration Studies*, pp. 1–20.

Weichselbaumer, Doris (2003). "Sexual Orientation Discrimination in Hiring". In: *Labour Economics* 10.6, pp. 629–642.

William, Benemann, ed. (1999). *A Year of Mud and Gold: San Francisco in Letters and Diaries, 1849-1850*. University of Nebraska Press.

Woloch, Nancy (2015). *A Class by Herself: Protective Laws for Women Workers, 1890s-1990s*. Princeton University Press.

Woodard, Colin (2011). *American nations: A history of the eleven rival regional cultures of North America*. Penguin.

Yue, Audrey (2013). "Lesbian, Gay, Bisexual, Transgender (LGBT) Migration". In: *The Encyclopedia of Global Human Migration*. Ed. by I Ness. Wiley Blackwell.

Zelinsky, W (1973). "The cultural geography of the United States. Englewood Cliffs, NJ: Prentice Hall". In: