NBER WORKING PAPER SERIES

FRONTIER CULTURE:
THE ROOTS AND PERSISTENCE OF "RUGGED INDIVIDUALISM" IN THE UNITED STATES

Samuel Bazzi
Martin Fiszbein
Mesay Gebresilasse

Working Paper 23997
http://www.nber.org/papers/w23997

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2017, Revised June 2018

We thank Romain Wacziarg for a helpful discussion of the paper at the NBER Summer Institute. We also thank Alberto Alesina, Quamrul Ashraf, Jeremy Atack, Michael Clemens,William Collins, Klaus Desmet, Benjamin Enke, Marcel Fafchamps, James Feigenbaum, Ray Fisman, Oded Galor, Camilo Garcia-Jimeno, Paola Giuliano, Bob Margo, Nathan Nunn, Ömer Özak, Daniele Paserman, Nico Voigtlaender, John Wallis, and David Weil, as well as seminar participants at ASREC, Boston University, Bowdoin, Brown, the Deep-Rooted Factors in Comparative Development Conference at Brown, Georgetown, Harvard, the NBER Political Economy Summer Institute, the NBER Development of the American Economy Summer Institute and Spring Meeting, New Economic School Political Economy and Development Workshop, Northwestern, Stanford, UdeSA, University of Houston, UPenn, University of Virginia, Vanderbilt, and the Second Workshop on Macro-Development, for helpful comments. Yeonha Jung, Max McDevitt, and Huiren Tan provided excellent research assistance. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States
Samuel Bazzi, Martin Fiszbein, and Mesay Gebresilasse
NBER Working Paper No. 23997
November 2017, Revised June 2018
JEL No. D72,H2,N31,N91,P16

ABSTRACT

The presence of a westward-moving frontier of settlement shaped early U.S. history. In 1893, the historian Frederick Jackson Turner famously argued that the American frontier fostered individualism. We investigate the Frontier Thesis and identify its long-run implications for culture and politics. We track the frontier throughout the 1790–1890 period and construct a novel, county-level measure of total frontier experience (TFE). Historically, frontier locations had distinctive demographics and greater individualism. Long after the closing of the frontier, counties with greater TFE exhibit more pervasive individualism and opposition to redistribution. This pattern cuts across known divides in the U.S., including urban–rural and north–south. We provide suggestive evidence on the roots of frontier culture: selective migration, an adaptive advantage of self-reliance, and perceived opportunities for upward mobility through effort. Overall, our findings shed new light on the frontier's persistent legacy of rugged individualism.

Samuel Bazzi
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
sbazzi@bu.edu

Martin Fiszbein
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
fiszbein@bu.edu

Mesay Gebresilasse
Boston University
270 Bay State Rd.
Boston, MA 02113
mmelese@bu.edu

1 Introduction

Rapid westward expansion marked the early history of the United States. According to the influential historian Frederick Jackson Turner, the presence of "a continually advancing frontier line" at the "edge of free land" strongly influenced American culture (Turner, 1893). The frontier fostered the development of distinctive cultural traits, including individualism and opposition to government intervention. The combination of these two traits characterizes "rugged individualism," a term popularized by Herbert Hoover in his 1928 presidential campaign.[1]

This paper shows that the American frontier shaped a culture of rugged individualism that persisted throughout time. First, using Census data from the 18th and 19th century, we establish the distinctive demographics and higher levels of individualism that historically characterized frontier locations. Then, using modern survey and Census data, we show that locations exposed to the frontier for a longer period historically exhibit higher contemporary levels of individualism, lower desired and actual levels of redistribution, and stronger opposition to government regulation. Finally, using linked Census records, we document empirical patterns that point to the origins of frontier culture. Frontier individualism is partly but not entirely explained by selective migration. Frontier conditions favored individualism through higher socioeconomic returns, and they created expectations of high income growth through effort, which fueled opposition to government intervention.

To understand the contemporaneous and long-run effects of the frontier, we revisit the classic Frontier Thesis through the lens of modern political economy and social psychology. In our conceptual framework, the significance of the frontier can be explained by three factors. First, frontier locations attracted individualists able to thrive in harsh conditions. Second, frontier conditions—isolation and low population density—further cultivated self-reliance, and they offered favorable prospects for upward mobility through effort, nurturing hostility to redistribution. Finally, frontier conditions shaped local culture at a critical juncture, thus generating persistent effects.

We determine the position of the frontier and track its evolution over time using population data from the Census and applying Geographic Information System (GIS) techniques. Following Turner's classic essay and the *Progress of the Nation* report from the 1890 Census, we define the frontier line as the line at which population density dropped below two people per square mile. We identify the frontier as comprised of counties with low population density in close proximity to the frontier line. This time-varying measure of frontier status is consistent with Turner's view of the frontier as "a form of society" rather than a fixed area. We measure total frontier experience as the time spent on the frontier between 1790 and 1890. This precise and comprehensive measure of frontier history provides, to our knowledge, the first measure of the duration of frontier exposure at a fine geographic level. This measure makes it possible to characterize previously unidentified variation within and across states.

Consistent with historical narratives, we find systematic evidence on the demographic and cultural

[1]The element of anti-statism in rugged individualism makes it different from notions of individualism that point to weak family ties, loose kinship networks, or low levels of communalism (see Alesina and Giuliano, 2014; Enke, 2017, 2018, respectively). The Merriam-Webster Online Dictionary defines rugged individualism as "the practice or advocacy of individualism in social and economic relations emphasizing personal liberty and independence, self-reliance, resourcefulness, self-direction of the individual, and free competition in enterprise." According to Wikipedia, rugged individualism "refers to the idea that individuals should be able to help themselves out and that the government does should not involve itself in the economic lives of people or the nation in general."

distinctiveness of frontier locations. Frontier settlers were disproportionately male, prime-age, and foreign-born. These distinctive traits display strong associations with each of the two defining features of the frontier—sparse population and isolation. Using semiparametric regressions, we identify sharp structural breaks in these traits close to the density cutoff defining the frontier line in historical accounts. Moreover, event study specifications show how these demographics evolve as counties exit the frontier.

Frontier locations also had sharply higher levels of individualism, as reflected in the prevalence of infrequent children's names. This result holds for several alternative ways of measuring infrequency and restricting to children with native-born parents or grandparents. The informational content of given names has been emphasized in economics (e.g., Abramitzky et al., 2016; Bertrand and Mullainathan, 2004; Fryer and Levitt, 2004; Olivetti and Paserman, 2015) as well as psychology and sociology (e.g., Gerrit and Onland, 2011; Gureckis and Goldstone, 2009; Lieberson and Bell, 1992). We borrow our names-based measure of individualism from social psychologists, who note that individualistic types are prone to give their children infrequent names, reflecting a desire to stand out, as opposed to common names, reflecting a desire to fit in (Twenge et al., 2010). Names have the crucial advantage of allowing us to measure individualism historically. Furthermore, name choices are particularly useful for studying cultural persistence as they represent a primordial act of cultural transmission by parents, which can have lasting effects on children's identity and behavior (Nelson and Simmons, 2007; Yadin, 2016).

We investigate the long-run effects of frontier exposure on culture using our novel, county-level measure of total frontier experience (TFE). First, we show that TFE positively correlates with infrequent naming patterns several generations after the closing of the frontier. Second, we find a robust association of TFE with opposition to redistribution and public spending based on contemporary surveys capturing different notions of government intervention. TFE is also associated with actual policy differences such as lower property tax rates. The results are robust to state fixed effects as well as geoclimatic controls including area, latitude and longitude, rainfall and temperature, distance to waterways, and potential agricultural productivity.

These long-run differences in preferences have translated into stronger contemporary support for the Republican Party. Each decade of TFE is associated with 3.5 percent more votes for Republican candidates in presidential elections since 2000. This association ratchets up over the 2000s as each election exhibits a significantly larger effect of TFE. Frontier exposure explains a significant part of the increase in Republican vote shares in the American heartland from 2000 to 2016, a period of rising political polarization. The effect of TFE remains significant when compared to the effect of Chinese import competition, a leading proximate explanation for the regional Republican shift (Autor et al., 2016).

We provide insights into why Republican Party support is stronger in areas with greater TFE by considering preferences over a set of contentious policy issues: the Affordable Care Act, increases in the minimum wage, the ban on assault rifles, and the regulation of CO_2 emissions. Republican Party positions on these issues can be linked to salient aspects of frontier culture described in the historical literature, including opposition to state intervention, strong belief in effort versus luck in reward, necessity of self-defense, and notions of "manifest destiny." We show that locations with greater TFE exhibit stronger opposition to each of these policies.

These results survive a battery of robustness checks. Our findings hold for alternative proximity and low density cutoffs defining TFE. They are robust to including regions exposed to frontier conditions

after 1890, the year in which the Census declared the frontier closed. The same is true when adding the West Coast, including California, which experienced its own frontier expansion eastward from the Pacific Coast in the second half of the 19th century. In fact, we find similar effects of TFE across Census regions of the country despite large average differences in preferences across those regions.

We address a host of potential confounders of frontier experience and contemporary culture. We show that the effects of TFE are not explained by low population density today nor by a history of low density.[2] Rather, they reflect the history of frontier settlement, which entailed not only low density but also opportunity afforded by relative resource abundance in remote areas. We account for other potential confounders, including mineral resource abundance, rainfall risk, access to railroads, slavery, and birthplace diversity. We also show that African Americans' preferences today are unaffected by TFE, consistent with the fact that mechanisms linking frontier experience to rugged individualism were of limited historical relevance for blacks, especially in the antebellum period. In addition, racial resentment toward blacks does not seem to account for the link between TFE and opposition to redistribution among the white population. Finally, we report results for placebo outcomes that are less plausibly connected with frontier culture.

We clarify the long-run identifying variation using an instrumental variables (IV) strategy. The IV exploits time series variation in immigrant inflows to the United States. For each location, we consider the intensity of inflows in the period starting just before the onset of local frontier settlement. The ups and downs of international immigration from 1790 to 1890 affected the speed of westward expansion and hence the length of time it took for frontier locations to become established settlements. We find similar results when using an alternative instrument based on predicted outflows from Europe due to climatic shocks, isolating push factors abroad as in Nunn et al. (2017).

Next, we explore the roots of frontier culture. We first examine the selective migration of individualists to the frontier. Using a linked sample of households from the 1870 and 1880 Censuses, we track people across locations and decompose county-level differences in individualism into components coming from early versus later settlers. We find that selective migration was significant, though it does not seem to fully explain the prevalence of rugged individualism on the frontier historically.

We identify two factors that may have complemented and reinforced selective migration. First, individualism thrived on the frontier due to its adaptive advantage in a setting of wilderness and isolation, where self-reliance was beneficial for survival and success. We find that infrequent names are associated with greater socioeconomic status on the frontier than elsewhere and also with lower rates of return migration from the frontier. Second, frontier conditions presented opportunities for upward mobility through effort, which would hone opposition to redistribution. This is consistent with existing evidence of widespread access to land and high rates of wealth accumulation on the frontier.[3]

Frontier culture may have persisted through several mechanisms, even after frontier conditions were long gone. Cultural traits established at early stages can maintain and even increase their prevalence through various forms of intergenerational transmission. As frontier experience started by definition at

[2]One important robustness check shows that the effects of TFE hold when comparing across counties within the same decile of historical or contemporary population density within the same state.

[3]We also consider a competing, disease-based explanation for the origins of individualism rooted in biology and known as the parasite-stress theory of values (Fincher and Thornhill, 2012). However, using data on disease and illness in the 1880 Census, we do not find evidence in support of this mechanism as a factor explaining frontier culture.

the earliest stages of settlement, it was bound to influence the formation of local institutions and social identity, which probably affected subsequent cultural evolution.

This paper contributes to a growing literature on the historical origins and persistence of cultural traits (e.g., Alesina, Giuliano and Nunn, 2013; Fernández, 2010; Guiso, Sapienza and Zingales, 2016; Nunn and Wantchekon, 2011; Spolaore and Wacziarg, 2013; Voigtländer and Voth, 2012).[4] The deep roots of rugged individualism shed new light on a puzzle in American political economy, namely the relative stability of preferences for redistribution over the last 40 years despite significant increases in inequality (see Ashok et al., 2015).[5] Our findings make a distinct addition to the economics literature on individualism (e.g., Greif, 1994; Gorodnichenko and Roland, 2016; Olsson and Paik, 2016) and preferences for redistribution (see Alesina and Giuliano, 2010, for a survey). We use a wealth of subnational data spanning over two centuries to link the American frontier with rugged individualism, drawing on core political economy theories in our analysis of mechanisms.

Our analysis relies on subnational variation and does not directly address the roots of cultural traits in the U.S. as a whole or cultural variation across countries. However, our results may have broad implications. First, the mere reshuffling of individualists through selective migration may have increased the long-run prevalence of individualism in the aggregate. Most U.S. locations experienced frontier conditions for at least a few years in their early history. During this critical juncture, individualists could exert decisive impacts on local cultural formation, whereas their influence in settled areas outside the frontier would be limited. Second, our findings suggest frontier conditions not only attracted individualists but also fostered further individualism. Thus, over the process of westward expansion, the frontier imbued a culture of rugged individualism throughout the U.S. (with varying intensity across locations based on length of exposure).

Differences in rugged individualism across the U.S. may help to explain cultural differences between the U.S. and Europe, insofar as there are common underlying determinants. The mechanisms we explore—selective migration, an adaptive advantage of individualism, and prospects of upward mobility through effort—were arguably important in shaping American culture as distinct from European culture. Indeed, according to Turner, "the Atlantic coast. . . was the frontier of Europe." In an early contribution on differences between the U.S. and Europe, Alesina et al. (2001) conjecture that "American anti-statism" may be partly linked to the frontier, which "strengthened individualistic feelings and beliefs in equality of opportunities rather than equality of outcomes." We provide the first systematic evidence on this hypothesis.

We make a novel contribution to the large social science literature animated by Turner's ideas. Many studies in history and sociology describe the demographic characteristics of the frontier. We provide a comprehensive analysis of its distinctive features, measuring the local prevalence of individualism

[4]This includes recent work in the U.S. context exploring the historical legacy of Scots-Irish honor culture (Grosjean, 2014), slavery (Acharya et al., 2016), mining (Couttenier et al., 2017; Couttenier and Sangnier, 2015; Glaeser et al., 2015), and agricultural risk (Ager and Ciccone, forthcoming; Davis, 2016).

[5]In a *New York Times* article from July 4th, 2017, titled "What's the Matter with Republicans?", journalist David Brooks wonders about voters' proclivities for supporting policies that are seemingly against their economic self-interest, and conjectures: "My stab at an answer would begin in the 18th and 19th centuries. Many Trump supporters live in places that once were on the edge of the American frontier. Life on that frontier was fragile, perilous, lonely and remorseless. [. . .] discipline and self-reliance were essential. [. . .] In their view, government doesn't reinforce the vigorous virtues. On the contrary, it undermines them [. . .] I'd say they believe that big government support would provide short-term assistance, but that it would be a long-term poison to the values that are at the core of prosperity."

for the first time. The only study of the Frontier Thesis in economics is García-Jimeno and Robinson (2011), which links variation in the quality of democratic institutions across countries in the Americas to variation in the historical importance of frontier land. Our paper examines subnational variation in individualism and preferences for redistribution in the U.S., and relies on a novel measure of historical frontier experience. Some studies in social psychology use state-level data to study variation in contemporary individualism, comparing demographic features (Vandello and Cohen, 1999) or infrequent names (Varnum and Kitayama, 2011) between the western U.S. and the rest of the country. Our empirical analysis goes beyond these broad geographic correlations to provide a precise definition of the frontier and to rule out potential confounders of settlement history. Furthermore, we link the frontier with both historical and modern individualism and identify several possible explanations for the origins of frontier culture.

Turner's work has attracted immense attention as well as vast criticism (see, e.g., Larson, 1993, and other articles in the same journal issue). His narratives contain departures from the historical record, overblown statements, and mythological elements. They paint an idealized portrait of the frontiersman and leave women and minorities out of the picture. The term "free land" appears repeatedly when, in fact, land was violently taken from Native Americans and, in many areas, westward expansion was more about "conquest" than "settlement" (Limerick, 1988). Our research provides empirical support for some important elements of the Frontier thesis, but it is not a general assessment of Turner's work nor an endorsement of its ideological overtones. Ultimately, frontier culture is likely rooted in both facts and myths. Rugged individualism may be stronger in areas with greater TFE partly due to pervasive subjective beliefs about frontier history.

The paper is organized as follows. Section 2 provides a general discussion of individualism and opposition to redistribution as well as economic theories about their origins and consequences. We also link these theories to the Frontier Thesis and offer a simple conceptual framework to understand its significance. Section 3 explains how we locate the frontier and measure total frontier experience. Section 4 documents the distinctive features of frontier populations. Section 5 provides estimates of the long-run effects of frontier experience on culture. Section 6 then offers evidence for why the frontier may have favored individualism and opposition to redistribution. Section 7 concludes with key lessons and caveats about extrapolating to other countries or even the U.S. as a whole.

2 A Modern Reading of The Frontier Thesis

This section provides a conceptual background that connects Turner's ideas about the American Frontier with contemporary political economy, cultural economics, and social psychology. We start by discussing contributions in these fields that provide insight into outcomes of interest. Then, we restate the Frontier Thesis, spelling out the potential channels for initial influence and subsequent persistence.

2.1 Individualism and Preferences for Redistribution

A large literature in social psychology portrays individualism as the most important dimension of cross-country variation in culture (e.g., Heine, 2010). Following Hofstede (1980, 1991) and Triandis (1988, 1995, 2001), we think of individualism (in contrast to collectivism) as comprising several related traits: a

view of the self as independent rather than interdependent, the emphasis on self-reliance, the primacy of self-interest, and the regulation of behavior by personal attitudes rather than social norms.

Empirical measures of individualism illustrate the concept more concretely. Some studies use Hofstede's survey-based index while others propose coarse proxies like divorce rates or the percentage of people living alone. Social psychologists propose other creative indicators. Kashima and Kashima (1998) show that in individualistic cultures, "I" and "you" are never dropped, as that would de-emphasize the individual. Twenge et al. (2010) argue that infrequent (common) children names reflect parents' desires to stand out (fit in). In this respect, name choices echo the behavior characterized by the Kim and Markus (1999) study contrasting preference for uniqueness in American culture with preference for conformity in East Asia. Given the choice among a set of colored pens, Americans chose the minority color while East Asians chose the majority color.

In economics, a small set of contributions has focused on individualism, starting with the work of Greif (1994) on how individualistic and collectivistic cultures shaped different trade institutions in the Middle Ages. The recent contributions of Gorodnichenko and Roland (2011, 2015, 2016) show that individualistic countries have higher levels of income, productivity, and innovation, as well as more democratic institutions. Gorodnichenko and Roland (2016) explain some of these effects through an endogenous growth model in which individualism fosters innovation by creating incentives to stand out.

Preferences for redistribution are distinct but closely related to individualism. Paul Samuelson (1965) once noted that "to an economist the word 'individualism' is tied up with *laissez faire*." In fact, Alesina and Giuliano (2010) measure preferences for redistribution using the same question from the General Social Survey that Di Tella et al. (2008) use to measure individualism. Using data from the European Social Survey, Quattrociocchi (2014) shows that immigrants who were born in countries with a more individualistic culture tend to have weaker preferences for redistribution in their country of residence. Intuitively, the defining characteristics of individualism—self-interest and inclination toward self-reliance—may be associated with opposition to redistribution and other forms of government intervention.[6] The connection is explicit in the ideology of "rugged individualism," which promotes self-reliance and opposes state intervention through taxes or regulations.

A rich literature on preferences for redistribution offers several useful insights for understanding the roots and persistence of frontier culture (see, e.g., Alesina and Angeletos, 2005; Alesina and La Ferrara, 2005; Bénabou and Ok, 2001; Bénabou and Tirole, 2006; Piketty, 1995). First, expectations about future income are central, as favorable prospects of upward mobility tend to generate opposition to redistribution. Second, the importance of effort (relative to luck) in the income-generation process may lead to lower desired tax rates. The greater the importance of effort, the larger the efficiency costs of taxes due to adverse incentives, and the larger the perception that they are unfair. Finally, these studies offer models with multiple equilibria that can shed light on the persistent nature of cultural traits.

[6] Gorodnichenko and Roland (2012) note that "The individualist view of government would tend to be wary of possible infringements of government on the individual's drive to self-achievement." In the sociology literature, Celinska (2007) notes that an aspect of "utilitarian individualism and the consequence of a strong belief in self-reliance" is the "opposition toward governmental efforts to equalize citizens' economic position, to limit private business, and to build strong social programs that provide assistance to the most disadvantaged."

2.2 The Frontier Thesis: A Restatement

According to the classic thesis advanced by F. J. Turner, the presence of a frontier separating established settlements from vast tracts of unsettled land during a formative period shaped the distinctive aspects of American culture. According to his thesis, "these free lands promoted individualism, economic equality, freedom to rise, democracy" (Turner, 1893). He also observed that, on the frontier, the "tax-gatherer is viewed as a representative of oppression," since the environment "produces antipathy to control."

The conceptual framework guiding our analysis combines some of Turner's ideas with insights from economics and social psychology. We think of the contemporaneous and long-run effects of the frontier as the result of three main forces. First, frontier locations attracted people with distinctive characteristics, both in terms of demographics and the prevalence of individualism. Second, the frontier experience, characterized by isolation and wilderness, fostered the development of self-reliance and related cultural traits. Finally, the distinctive features of frontier populations affected preferences and social norms at a critical juncture of institutional formation and thus left a persistent imprint on local culture. While these mechanisms may be relevant beyond the context of our study, it is of course possible that certain preconditions were specific to American history. We discuss the three mechanisms below and revisit external validity in the conclusion.

Selective Migration. Traditional narratives characterize the frontier by the prevalence of young single men, mostly of low socioeconomic status. Harsh living conditions on the frontier were particularly hostile to women and the elderly, which helps explain the skewed sex ratio and age distribution in most accounts. In addition, the frontier attracted workers from the lower end of the urban skill distribution in search of opportunity, as suggested by the theory of the "safety valve" (see Ferrie, 1997; Goodrich and Davison, 1935, 1936; Steckel, 1989; Stewart, 2006; Turner, 1893).

Frontier residents also tended to exhibit a high degree of individualism. Migrants generally have independent mindsets. This trait may be stronger among those moving to the frontier, giving up their social environment to settle in remote and isolated contexts (see Beck-Knudsen, 2017; Hoang-Anh et al., 2018; Jokela, 2009; Kitayama et al., 2006, 2010; Olsson and Paik, 2016). Moreover, as discussed next, the adaptive advantage of self-reliance in such conditions might further amplify the self-selection of individualist types to the frontier.

Effects of Frontier Conditions. While frontier locations attracted people with specific traits, the frontier's unique natural and social conditions, in turn, influenced the settlers' values, beliefs, and behavior. In Turner's words, "a modification of the original stock occurred." Remoteness and isolation implied a particular set of opportunities and challenges. The abundance of land and other natural resources offered ample profit opportunities, insofar as they were deftly exploited. On the other hand, as Overmeyer (1944) argues, "life was rough, crude, hard, and dangerous." Frontier settlers often faced harsh climatic conditions and multiple types of danger, such as plagues, droughts, blizzards, and crop failure, as well as attacks from wild animals, Native Americans, and other settlers. Violence was commonplace, and social infrastructure providing protection and care was limited or nonexistent.[7]

[7]Rampant violence, noted in many historical narratives of the frontier, is a common characteristic of contexts with low population density, high population mobility, lack of well-defined property rights, and absence of clear mechanisms for law enforcement (Couttenier et al., 2017; Grosjean, 2014; Nisbett and Cohen, 1996; Restrepo, 2015).

These opportunities and threats on the frontier may have favored individualism through an adaptive mechanism. In the frontier context, people often had to rely on themselves for protection and prevention, and to improve their living conditions.[8] Moreover, the resourcefulness associated with individualism would prove useful in a context characterized by novel and uncertain conditions. Thus, individualistic traits had an adaptive value: beliefs and behavior based on independence and self-reliance made people better suited to cope with the frontier environment (Kitayama et al., 2010; Plaut et al., 2002). In turn, the adaptive advantage of individualism may have increased its prevalence in the population through differential reproductive success, learning, or both (see Galor and Özak, 2016).

Moreover, land abundance and remoteness also offered favorable prospects of upward mobility and a large perceived importance of effort in income generation. Based on the political economy theories mentioned in 2.1, and as conjectured by Alesina et al. (2001), these conditions would naturally foster opposition to government intervention. This resonates with historical narratives. Billington (1974), a noted Turnerian, argued that on the frontier "every man was a self-dependent individual, capable of caring for himself without the fostering care of society," which "seemed just in a land that provided equal opportunity for all to ascend the social ladder."

Frontier mentality may also be shaped by "motivated beliefs," i.e. departures from objective cognition that fulfill psychological and functional needs (Bénabou and Tirole, 2016). Perhaps the inclination toward self-reliance served to cope with the frontier's loneliness. Perhaps the importance of effort and the prospects of upward mobility were exaggerated as this helped to bear the exertion required by frontier life. Opposition to redistribution may partly reflect a psychological incentive to consider government intervention inefficient as the state had more limited presence on the frontier (Bénabou, 2008).

Cultural Persistence. Culture can be remarkably persistent but can also change rapidly (Gershman, 2017; Giuliano and Nunn, 2017). In our view, the persistence of frontier culture long after frontier conditions abated can be linked to the distinctive traits of early settlers at critical junctures of institutional development. A seminal theory in cultural geography due to Zelinsky (1973) captures this potential channel. The "doctrine of first effective settlement" argues that when "an empty territory undergoes settlement [...] the specific characteristics of the first group able to effect a viable, self-perpetuating society are of crucial significance for the later social and cultural geography of the area, no matter how tiny the initial band of settlers may have been."

The economics literature on culture offers several mechanisms by which the distribution of cultural traits in the population at a point in time can influence its subsequent evolution (Bisin and Verdier, 2010). First, it affects the likelihood that new generations adopt these traits through horizontal transmission. Second, it can influence vertical transmission, for example, by affecting beliefs about the behavior of other members of society and thus the expected rewards for different traits (Guiso et al., 2008). Various models of intergenerational transmission imply that initial conditions determine the long-run cultural equilibrium. The initial conditions in our setting were precisely those distinctive traits of frontier populations during the inception of settlement.

[8] Some critics of Turner emphasize the importance of cooperation on the frontier (e.g., Boatright, 1941), but his supporters have argued that cooperation was not inconsistent with individualism. For instance, according to Billington (1974), the frontiersman "spoke for individualism ... even though he was equally willing to find haven in cooperation when danger threatened or need decreed." While returns to cooperation may have been high at times, maintaining reciprocity would have been difficult in frontier settings with such high population mobility, as noted in the literature on social capital (see Munshi, 2014).

Moreover, the initial distribution of traits can influence the long-run cultural equilibrium by shaping early institutions (Alesina and Giuliano, 2015; Bisin and Verdier, 2017; Tabellini, 2008a,b). Several political economy models suggest that initial preferences for redistribution lead to different institutional outcomes, which in turn affect the evolution of preferences (see Alesina and Angeletos, 2005; Alesina et al., 2012; Bénabou and Tirole, 2006; Bisin and Verdier, 2005). Consistent with notions of persistence and path dependence, Turner (1893) noted that "traits [of frontier society] have, while softening down, still persisted as survivals in the place of their origin, even when a higher social organization succeeded."

Motivated beliefs may also contribute to the diffusion of frontier culture and its persistence after frontier conditions are gone. Intergenerational transmission of rugged individualism may be stronger in areas with with longer frontier experience because narratives about local frontier history—whether myth or reality—are more pervasive. And once the beliefs of frontier culture become part of social identity, they are likely to stick even if contradicted by facts.

3 Mapping the History of the Frontier

This section presents our method for mapping the history of the frontier. After providing historical background, we explain how to use U.S. Census data and GIS techniques to determine the position of the frontier line. We then define frontier counties and our new measure of total frontier experience.

From colonial times until the late 19th century, America underwent rapid population growth and a massive westward expansion. Historical sources document this process, and the noteworthy 1890 Census report on the *Progress of the Nation* (Porter et al., 1890) provides a key source of inspiration for Turner's classic 1893 essay. The authors observe that the Thirteen Colonies, already settled communities by 1790, were "the sources of supply for a great westward migration," as people "swarmed from the Atlantic coast to the prairies, plains, mountains, and deserts by millions during the last century." The report describes in great detail the decade-by-decade push westward.[9] From 1790 to 1890, as the nation's total population increased from 3.9 million to 62.6 million, the extent of settled area went from under 240,000 square miles to almost 2,000,000. In the same period, the mean center of population[10] shifted westward over 500 miles, from just east of Washington D.C. to Decatur, Indiana.

The Porter et al. (1890) Census report considered the process of westward expansion complete, and the frontier closed, by 1890. In a passage famously quoted in Turner's essay, it stated that "up to and including 1890 the country had a frontier of settlement, but at present the unsettled area has been so broken into by isolated bodies of settlement that there can hardly be said to be a frontier line." As one of the authors of the Census report put it elsewhere, "the frontier line has disappeared . . . the settled area has become the rule and the unoccupied places the exception" (Gannett, 1893).

3.1 Locating the Frontier and Tracking its Movements

Prior research on the American frontier adopted simplifying definitions. In a study of westward migrants in 1850 and 1860, Steckel (1989) identifies the frontier with the states of Minnesota, Iowa, Kansas,

[9] The report included detailed maps of population density, which closely resemble our maps (see Appendix Figure A.1).

[10] This is the point at which weights of equal magnitude corresponding to the location of each person in an imaginary flat surface representing the U.S. would balance out.

Texas, and those farther west. Ferrie (1997) studies migration to the frontier between 1850 and 1870 and defines 90° west longitude as the frontier's eastern boundary. Kitayama et al. (2010) associate the frontier with the Western United States.

We take a different approach. Following Porter et al. (1890) and Turner's classic essay, we define the frontier line as the line dividing settlements with population density of two or more per square mile from those with less.[11] We therefore define frontier counties as those (i) in close proximity to the frontier line (100 kilometers in our baseline) so as to capture Turner's notion of the "frontier belt",[12] and (ii) with population density below six people per square mile, a cutoff suggested by Porter et al..[13] While these cutoffs are necessarily arbitrary, we offer empirical support for these definitions in Section 4.2, and our primary results in Section 5 are largely unchanged using different distance and density thresholds.

This definition allows us to precisely locate the frontier at any point in time and also to trace its evolution. As Turner noted, the frontier was "a form of society, rather than an area." Its two defining features implied distinct but mutually reinforcing forms of isolation, the combination of which was conducive to producing frontier culture according to Turner's thesis. Low density entailed isolation from other people within a given location. Proximity to the frontier line entailed isolation from other population centers, and in many cases limited interaction with agents of the federal government. Both dimensions entailed a lack of social infrastructure as well as relative resource abundance. However, while low density locations could be found amidst densely settled areas in the east, such areas arguably offered much more limited prospects for upward mobility than similarly low density areas close to the frontier line, which attracted pioneering settlers in search of opportunity.

For each Census year beginning in 1790, we calculate county-level population density per square mile. For intercensal years, we interpolate county-level population density by assuming a constant annual population growth rate that matches the decadal growth rate. We maintain consistent units of observation over time by harmonizing all data to the 2010 boundaries using an approach suggested in Hornbeck (2010) and detailed in Appendix C.[14]

Using annual county-level population densities, we locate the frontier line for each year by drawing contour lines that divide counties with population densities above and below two people per square mile. Figure 1 plots the resulting lines for 1790, 1820, 1850, and 1890, and full details on the underlying GIS procedure can be found in Appendix A. In order to closely approximate historical notions of the frontier developed above, we discard all line segments less than 500 km, as well as isolated pockets of low density counties within the main area of settled territory (i.e., to the east of the main frontier line).[15] Figure 2 shows the evolution of the resulting, main frontier lines in red for 1790–1890.

[11] Turner (1893) notes, "The most significant thing about the American frontier is, that it lies at the hither edge of free land. In the census reports it is treated as the margin of that settlement which has a density of two or more to the square mile. The term is an elastic one, and for our purposes does not need sharp definition. We shall consider the whole frontier belt including the Indian country and the outer margin of the "settled area' of the census reports."

[12] Traveling 100 kilometers in a covered wagon would have required about 2–3 days in the early 19th century.

[13] The Porter et al. report defined locations with > 6 people/mi^2 as established, post-frontier settlements.

[14] We show in Appendix Figure A.2 that the use of harmonized 2010 county boundaries has little effect on the location of the frontier lines relative to an approach based on contemporaneous county boundaries historically.

[15] Our results are qualitatively similar when discarding isolated pockets of high density settlement to the west of the main frontier line. The 500 km cutoff discards many contour lines but retains some large unconnected lines off of the main east-to-west frontier line, e.g., the ones spanning Maine in 1820 and Michigan in 1850. Like other cutoffs we are forced to specify, this one is arbitrary but also robust to other rules, including having no cutoff at all. We detail robustness to many alternative specifications of frontier counties in Section B.7.

A second major frontier line emerged on the West Coast, starting in California, in the mid-19th century. Initiated by the Gold Rush, this process of settlement constituted a large, discontinuous leap in the advance of east-to-west expansion. While Eastern cities were very far away, proximity to the ocean reduced transportation costs. This facilitated the arrival of settlers as well as international trade, and implied different types of isolation compared to frontier locations in the heartland. We leave this secondary frontier out of the baseline analysis but show in Section 5.5 that, in fact, frontier experience has similar long-run effects in the West as elsewhere.

3.2 Total Frontier Experience

The westward movement of the frontier was fast at times, slow at others. Thus, some locations spent little time in frontier conditions, while others remained on the frontier for decades. This variation in the duration of frontier exposure is key for our investigation of long-run persistence.

To measure the intensity of historical frontier experience for each location, we calculate the number of years spent within the frontier belt from 1790 to 1890. For each year, we assign each county a dummy variable equal to one if it is on the frontier according to the abovementioned definitions of proximity to the frontier line and low density. Then, total frontier experience (TFE) for each county is the sum of indicators of frontier status from 1790 to 1890.

We set 1890 as the endpoint for measuring TFE following the *Progress of the Nation* report and Turner. While many places in the U.S. remained sparsely populated long after 1890, the other defining feature of the frontier—isolation—did not persist with the same intensity. Transcontinental railroads were built in the early 1880s, which was also when the Indian wars came to an end. Large-scale federal irrigation efforts came soon after. As a robustness check, in Section 5.5 we consider a longer time frame for the measurement of TFE, changing the endpoint to 1950.

Our analysis excludes counties to the east of the 1790 east-to-west main frontier line for which we do not observe *total* frontier experience given the available data. To be clear, we do not claim that counties close to the East Coast were never on the frontier. We simply cannot measure their frontier experience.

Figure 3 shows the spatial distribution of TFE, measured in years and using the 100 km frontier cutoff, for the counties in our main analysis. Total frontier experience ranges from 0 to 63 years with a mean of 18.2 years and a standard deviation of 11.2 years. TFE exhibits considerable variation both across and within states and bands of latitude and longitude more generally. Moreover, the variation in TFE goes well beyond simply capturing differences in contemporary population density, as seen in the maps in Appendix B.1 and shown empirically in Section 5.5. Within Illinois, for instance, Cass County has TFE of 10 years, while Johnson County stayed on the frontier for 32 years. Today, however, the two counties have nearly identical population density of 36 people/mi^2. While some of the greatest TFE is found in the South, it is important to note that we find similar long-run effects across different Census regions (see Section 5.5).

4 The Distinctive Features of the Frontier

This section documents the unique demographic features of the frontier and its higher levels of individualism. Historians and sociologists have devoted considerable effort to analyzing the demographics of

frontier locations.[16] However, these studies usually focus on a specific place at a particular time, making it difficult to establish empirical regularities. In contrast, we characterize the demographics of the frontier using data from all Census rounds from 1790 to 1890. Moreover, we provide the first empirical validation of differential individualism on the frontier.

We document the distinctive features of the frontier using three complementary strategies. Section 4.1 establishes basic cross-sectional differences between frontier and non-frontier counties (east of the frontier line), and it shows that both remoteness and sparsity of frontier counties matter in explaining these differences. Section 4.2 validates these two defining features of frontier counties by identifying nonlinear relationships between these population traits and both density and distance. Finally, Section 4.3 exploits time-series variation comparing counties before and after exiting the frontier.

We focus on a set of demographic characteristics associated with the frontier in historical accounts. These include sex ratios, age distributions, foreign-born population shares, and literacy rates, all of which we draw from historical Census data in Haines and ICPSR (2010). With the exception of immigrant shares, we calculate all variables over the white population as this helps with consistency across time periods and ensures that results are not driven by racial composition.[17]

We measure individualism based on children's names. As suggested by social psychologists, the share of children with infrequent names reflects the prevalence of individualism, correlating strongly with other proxies, as shown by Varnum and Kitayama (2011), Ogihara et al. (2015), and Beck-Knudsen (2017).[18] We measure the share of children in a given county under 10 years of age with infrequent names using full-count, historical Census data from several decades beginning in 1850 and made available by `ancestry.com`. Appendix C provides a list of common names for selected years (e.g., John and Sarah) as well as a random sample of infrequent names (e.g., Luke and Lucinda).[19] Our findings are robust to defining infrequent names in various ways, (i) focusing on those outside the top 10, 25, or 100, (ii) considering the top names in the nation, Census division, state, or county, and (iii) restricting the sample to children born in the U.S., with U.S.-born parents, or with U.S.-born grandparents.[20]

Overall, the results below provide new, systematic empirical support for historical narratives about the individualists settling the American frontier. The three approaches offer a stark and consistent picture of the frontier being a qualitatively different type of society.

[16] See, for example, Bowen (1978), Coombs (1993), Demos (1968), Easterlin et al. (1978), Eblen (1965), Modell (1971), Moller (1945), Schaefer (1985), and Smith (1928).

[17] The average county during the 1790–1890 period is 83 percent white with little difference between frontier and non-frontier locations; using specification (1) below, frontier counties have a 2 percentage point higher white population share than non-frontier counties. In Section 5.5, we consider the role of race in understanding the legacy of frontier experience.

[18] Varnum and Kitayama (2011) shows a strong cross-country correlation between infrequent names and Hofstede's widely used index of individualism. Beck-Knudsen (2017) shows correlations that the names-based measure is strongly correlated with Hofstede's index as well as with the use of first- and second-person singular pronouns across 44 countries (and across regions within five countries). In Japan, Ogihara et al. (2015) shows a strong time-series correlation between the share of common name pronunciations and an index of individualism similar to the one proposed by Vandello and Cohen (1999), which includes divorce rates and the percentage of people living alone.

[19] As an example of infrequent names on the frontier from popular historical fiction, consider the Luckett family at the center of the celebrated trilogy, *The Awakening Land*, by novelist Conrad Richter. Members of this white, native-born family on the Ohio frontier in 1795 had first names that included Chancey, Wyitt, and Worth for boys and Ascha, Sayward, Sulie for girls.

[20] As our baseline measure we choose the share of names outside the top 10 following the social psychology literature (Varnum and Kitayama, 2011). With this measure, the majority of children have infrequent names (e.g., 57 percent of boys and 60 percent of girls in 1850). Hence, an "infrequent name" may not be a very unusual name but simply one that is uncommon. In any case, all these measures are highly correlated, and results are qualitatively similar across all of them (see Appendix B.4).

4.1 Demographics and Individualism on the Frontier: Basic Patterns

We begin by documenting the basic differences in demographics and individualism on the frontier with the following estimating equation:

$$x_{cdt} = \alpha + \beta \text{ frontier}_{ct} + \theta_d + \theta_t + \varepsilon_{cdt}, \tag{1}$$

where x_{cdt} is one of the population traits of interest in county c in Census division d at time t, frontier_{ct} is time-varying frontier status, and θ_d and θ_t are Census division and year fixed effects, respectively. Panel A of Table 1 reports estimates of β, the frontier differential, for each of six x outcomes.[21]

Across columns, we find that frontier populations tend to have significantly more (1) males, (2) prime-age adults, and (3) foreign-born. Frontier counties have 0.15 additional males for every female relative to non-frontier counties where the average sex ratio is 1.09. The share of prime-age adults (15–49 years old) in the population is 2.6 percentage points (p.p.) higher than in non-frontier counties, for which that share is around 46 percent. Additionally, frontier counties have nearly 6 p.p. higher foreign-born population shares than the average non-frontier county where 7 percent of residents are immigrants. Finally, in column (4), literacy rates are not significantly different on the frontier. However, literacy may simply be a noisy measure of skill.

Columns 5 and 6 show that individualism is more pervasive in frontier counties as reflected in the share of children with infrequent names. In frontier counties, around 2 p.p. more girls (boys) have infrequent names relative to the average non-frontier county with around 66 (58) percent having infrequent names. These measures capture the share of children aged 0–10 with names that are outside the top 10 most popular names in that decade's birth cohort within the Census division.[22] We restrict here to white children with native-born parents, but results are similar using the other ancestry restrictions and other measures of popularity noted above.

We further clarify the frontier differential by distinguishing the two attributes of frontier locations: (i) proximity to the frontier line and (ii) low population density. Panel B of Table 1 estimates

$$x_{cdt} = \alpha + \beta_1 \text{ near frontier line}_{ct} + \beta_2 \text{ low population density}_{ct} + \theta_d + \theta_t + \varepsilon_{cdt}, \tag{2}$$

where *near frontier line*$_{ct}$ is an indicator for having a centroid within 100 km of the frontier line at time t, and *low population density*$_{ct}$ is an indicator for population density below six people per square mile.

Overall, the results in Panel B suggest that both remoteness and sparsity contribute to the distinctive demographics and higher rates of individualism in frontier counties.[23] Frontier differentials are not

[21] These outcomes are all measured at the 2010 county-level but are not all available in every Census round, which explains why the sample varies across columns. Non-frontier counties include those to the east of the 100 km belt around the frontier line in the given decade but west of the 1790 frontier line, for consistency with the sample used in the long-run analysis. Including counties to the east of that line leaves results unchanged. Treating counties to the west of the frontier belt in the given decade as frontier amplifies the differential β, as those living in wilderness conditions tend to be even more distinctive.

[22] A potentially confounding naming practice lies in the passing on of parental first names to children. Using data discussed in Section 6.1, we find that while only around three (five) percent of girls (boys) have such matronymics (patronymics), this practice is less common on the frontier and significantly so for boys. Choosing novel names for one's children rather than passing on one's own arguably reflects a desire to instill independence. As such, this finding is consistent with our broader claim that the higher prevalence of infrequent names on the frontier reflects greater individualism.

[23] Column 4 shows that the null for illiteracy in Panel A is due to offsetting positive effects of low density and negative effects of proximity. This pattern does not arise for other outcomes and suggests scope for further work on the safety valve hypothesis.

merely an artifact of their low population density. There is something distinct about people living close to the frontier line. Nor are these differences in frontier counties due to arbitrary density or proximity cutoffs, as we show next.

4.2 The Frontier Is Qualitatively Different: Semiparametric Evidence

In this section, we validate the notion that the frontier was a structurally distinct type of settlement. We use the following semiparameric specification:

$$x_{cdt} = \alpha + g(\text{population density}_{cdt}) + \theta_d + \theta_t + \varepsilon_{cdt}, \tag{3}$$

where $g(\cdot)$ is a nonlinear function recovered using the partially linear (Robinson, 1988) estimator. While we estimate $g(\cdot)$ across all counties in the sample, we restrict the graphs presented in this section to counties with less than 50 people/mi^2 in order to focus on changes close to the assumed frontier threshold. In Appendix B.2, we conduct an analogous exercise for proximity to the frontier line instead of density.

Figure 4 provides a stark illustration of the qualitative differences in demographics and individualism in low density areas. Each graph shows the local linear regression function and 95 percent confidence interval around $g(\cdot)$. In graph (a), the sex ratio displays levels around 1.6 in the most sparsely populated counties and declines sharply as population density rises to 3–4 people/mi^2. The slope of $g(\cdot)$ then abruptly flattens out as the sex ratio stabilizes at around 1.05–1.1 males for every female. In graph (b), the prime-age adult share declines sharply as we move towards density levels of 2–3 people/mi^2 and then levels off. Graphs (c) and (d) show similar downward-sloping albeit less sharply nonlinear shapes. However, graphs (e) and (f) show stark nonlinear shapes for infrequent child name shares. Appendix Figure B.2 shows evidence of similar nonlinear patterns, though less stark, when we consider proximity to the frontier instead of density.

Overall, the results in Figure 4 point to a structural break in demographics and individualism at levels of population density consistent with the seemingly arbitrary cutoffs in the historical literature. In fact, using the Chow (1960) test, we can easily reject the null hypothesis of a constant effect of population density above and below 6 people/mi^2 (the upper bound of frontier settlement according to Porter et al., 1890), or above and below any cutoff in the 2–6 range. We can also be agnostic about the relevant cutoff, using the Zivot and Andrews (2002) test to identify unknown structural break points in each decade. In 1850, for example, we find a break in the sex ratio at 2.7 people/mi^2, the prime-age adult share at 2.0, and infrequent names for boys (girls) at 3.2 (2.6).

4.3 Frontier Transitions: Event-Study Evidence

Building on the prior cross-sectional results, we now use time-series variation as counties transition from frontier to established settlement. We estimate an event-study analogue to equation (1):

$$x_{cdt} = \alpha + \sum_{j=-20}^{40} \gamma_j \mathbf{1}(\text{years since exiting frontier = j}) + \theta_d + \theta_t + \varepsilon_{cdt}, \tag{4}$$

where the γ_j coefficients identify the average x for counties that have exited or will exit the frontier j years prior or in the future, respectively. We plot 95 percent confidence intervals for the γ terms, each of which are estimated with reference to the decade in which the county transitioned out of the frontier.

The estimates in Figure 5 provide additional insight into the process of demographic and cultural change along the frontier. Panel (a) reveals an abrupt shift in the sex ratio as counties exit the frontier. On average, counties have 0.25 higher sex ratios in the two decades prior to exiting the frontier whereas those decades thereafter exhibit lower ratios that stabilize by the second decade. Panel (b) provides similar evidence of convergence towards a lower prime-age adult share as counties exit the frontier. Panel (c) shows that the foreign-born population share exhibits a steady and roughly linear decline along the frontier transition path. Results are noisier for illiteracy rates in panel (d).

Panels (e) and (f) demonstrate the declining prevalence of infrequent children's names as counties approach the decade in which they exit the frontier. Thereafter, we see naming patterns stabilize around a less individualistic equilibrium in which popular names becomes more common at the local level. However, as we show next, the length of exposure to frontier culture in these early stages of settlement has implications for differences in the prevalence of rugged individualism across counties today.

5 Long-Run Effects of Frontier Experience on Culture

In this section, we examine the long-run effects of frontier experience on culture and discuss their implications for modern political economy debates. We present our empirical framework, discuss key data sources, and then move to our main cross-county results

Our motivation stems from the theories of cultural persistence discussed in Section 2.2. While the high levels of individualism on the frontier historically could have dissipated, it is also possible that frontier experience shaped the long-run evolution of local culture. The duration of exposure to frontier conditions determined the scope for the mechanisms through which rugged individualism came to thrive on the frontier, which we analyze in Section 6. In the presence of multiple equilibria and path dependence, these early stages of cultural formation would represent a critical juncture, and frontier experience could have a lasting legacy.

5.1 Estimating Equation

We relate historical frontier exposure to modern measures of individualism and preferences for individuals in county c. In particular, our main county-level, cross-sectional estimating equation is given by:

$$y_c = \alpha + \beta \text{ total frontier experience}_c + \mathbf{x}_c'\boldsymbol{\gamma} + \theta_{FE(c)} + \varepsilon_c, \tag{5}$$

where y_c is some long-run outcome capturing cultural traits (e.g., individualism or preferences for redistribution). Total frontier experience (TFE) is the amount of time *in decades* a given county remained on the frontier according to our baseline definition in Section 3.2. Our baseline sample, seen in Figure 3, is restricted to those counties for which the 1790–1890 period contains the whole extent of frontier experience as discussed in Section 3.2. In baseline specifications, the vector $\mathbf{x}_c$ comprises predetermined or fixed county-level covariates including latitude, longitude, county area, average rainfall and temper-

ature, elevation, potential agricultural yield, and distance to rivers, lakes, and the coast (see Appendix C for details). The $\theta_{FE(c)}$ term includes state fixed effects (FE) in the preferred specification. The coefficient β therefore identifies a local effect of TFE after accounting for geoclimatic factors that shape culture and also determined the duration of frontier experience historically. Following the approach suggested by Bester et al. (2011), standard errors in all specifications are clustered on 60-square-mile grid cells that completely cover counties in our sample.[24] When considering several correlated outcomes, we also estimate mean effects based on the Kling et al. (2007) approach.[25]

The main threat to causal identification of β lies in omitted variables correlated with both contemporary culture and TFE. We take many steps to address this concern with the goal of isolating variation in TFE that influences modern culture through the channels outlined in Section 2.2. First, in Section 5.5, we add a richer set of controls to $\mathbf{x}_c$ aimed at removing variation that may be associated with confounding factors such as current population density. Second, we use the Oster (2016) approach to show that unobservables are unlikely to drive our results. Third, in Section 5.6, we pursue an instrumental variables (IV) strategy that exploits variation in the speed of the frontier's westward movement induced by changes in the intensity of national immigration flows over time.

In Appendix B.3, we revisit the Illinois counties of Cass and Johnson noted in Section 3.2 to illustrate the link from historical TFE to contemporary political economy. This case study also clarifies how our empirical strategy, robustness checks, and IV approach help to isolate variation in TFE that is less confounded with other drivers of long-run cultural differences.

5.2 Data on Contemporary Culture and Political Economy

We measure contemporary culture and policy outcomes using several data sources. We draw upon three nationally representative surveys: the Cooperative Congressional Election Study (CCES), the General Social Survey (GSS), and the American National Election Study (ANES). These surveys are staples in the social science literature on political preferences and social norms, often asking different questions about similar underlying preferences. Their geographic coverage differs and is quite narrow for the GSS and ANES. Appendix C describes each survey, discusses advantages and disadvantages, and also provides definitions and sample coverage maps.

Additionally, for the full sample of counties, we observe infrequent names from post-1890 Census rounds and two salient policy outcomes. First, we measure the Republican vote share in recent presidential elections using data from Leip's Atlas. Second, we take estimates of property tax rates from the American Community Survey in 2010 as prepared by the National Association of Home Builders. Together, these local voting records and survey data allow us to paint a rich picture of the persistent culture of rugged individualism in areas exposed to the frontier for a longer time.

[24] Inference remains unchanged when using the computationally more intensive Conley (1999) spatial HAC estimator with a bandwidth of 150–300 km. We retain the arbitrary grid-cell approach as it is considerably easier to implement and less prone to instabilities, which becomes important with the sparser geographic coverage in some of the survey data.

[25] This takes a weighted average of the estimates of β for each of K related outcome variables (on the same scale) with the weights equal to the inverse sample standard deviation of that variable for a suitable control group. The choice of that control group affects the mean effect size but not significance. Results are very similar across three alternative approaches, including the few counties with zero frontier experience, those with less than a decade, and those with below the median.

5.3 Total Frontier Experience and Persistent Individualism

We begin by documenting a long-run association between total frontier experience (TFE) and contemporary individualism. Nearly five decades after the closing of the frontier, infrequent children's names are more pervasive in counties with greater TFE. In Table 2, we report the effect of TFE on the share of white boys (Panel A) and girls (Panel B) age 0–10 given infrequent names in the 1930s. The data come from the full count 1940 Census and capture naming choices multiple generations after counties exited the frontier.[26] Our baseline measure of infrequent names considers those outside the top 10 within the county's Census division. In the average county, 72 percent of boys and 79 percent of girls have infrequent names with standard deviations of 0.07 and 0.04, respectively. We normalize these variables so that standard deviation effect sizes can be read directly from the coefficients.

The most demanding OLS specification in column 4 of Table 2 suggests that each additional decade of TFE is associated with 0.11 (0.16) standard deviations higher share of infrequent names for boys and girls, respectively. Comparing counties across the interquartile range of TFE (11 vs. 24 years) implies 1.4 (1.1) percent more boys (girls) with infrequent names. We build up to this result by expanding the set of control variables, starting in column 1 with no controls. Columns 2 and 3 add Census division and state fixed effects, respectively, to rule out broad regional differences in TFE and culture. Our main specification in column 4 includes the full set of baseline geoclimatic controls. Comparing across columns 1–4, the coefficient remains relatively stable despite large changes in the R^2. This pattern is consistent with limited selection-on-unobservables according to the parameter δ reported in the table; Oster (2016) suggests that $|\delta| > 1$ implies limited scope for unobservables to explain observational results.

Furthermore, these results are not sensitive to the particular measure of infrequent names or the national background of the parents assigning names. We document this robustness in Appendix B.4. After replicating the baseline result in column 1, Appendix Table B.2 restricts columns 2–4, respectively, to children with native-born fathers, native-born parents, native-born grandparents. Together, these help address concerns about immigrants having infrequent name preferences and being more likely to settle in frontier areas historically.[27] Column 5 defines infrequent names based on the top 10 names nationally while columns 6 and 7 do so at the state and county level, respectively. Columns 8 and 9 define infrequent names as those outside the top 25 and top 100 names, respectively. Finally, column 10 restricts to non-biblical names to account for the fact that religiosity may be confounded with TFE and naming choices. Across measures, we see a similar effect of TFE.

Together with the findings in Section 4, these results suggest that infrequent name choices were not only more common in frontier areas historically but are also more prevalent in the long run in areas with greater TFE. Indeed, the effect of TFE on infrequent name choices can be seen in the early 1900s with little change thereafter (see Appendix Table B.4). This points to the persistence of the early frontier culture of individualism long after frontier conditions abated.

In Appendix B.4, we further validate the link between TFE and individualism today using a well-

[26] Ideally, we could carry these results through to the contemporary period, but, unfortunately, the 1940 Census is the latest round that provides information on names. Although the Social Security Administration releases baby name counts by state, it does not do so at the county level as required for our empirical strategy.

[27] The robustness to these alternative measures of ancestry is consistent with the rapid speed of assimilation to American name choices reported in Abramitzky et al. (2016). They show that the immigrant–native gap in Americanized name choice is halved within 20 years after parents arrive in the U.S.

suited measure from the ANES data in 1990. In particular, greater TFE is associated with respondents identifying more strongly with self-reliant as opposed to cooperative behaviors. We turn now to identify the closely related link between frontier experience and opposition to government intervention.

5.4 Total Frontier Experience and Opposition to Redistribution and Regulation

This section identifies a long-run effect of total frontier experience (TFE) on contemporary political preferences. First, greater TFE is associated with opposition to redistribution, preferences for limited government, and low levels of local taxation. Second, these differences in preferences translate into stronger Republican Party support today. Finally, we identify a link between TFE and opposition to government regulations surrounding issues that were salient in frontier culture historically. We view all of these outcomes as closely connected measures capturing the same underlying opposition to government intervention. In all cases, we report estimates of equation (5) controlling for the geoclimatic characteristics used in column 4 of Table 2 as well as individual demographics (age, age squared, gender, and race dummies) and survey wave fixed effects where relevant. We continue finding supportive Oster (2016) tests for selection-on-unobservables while other robustness checks and instrumental variables results are discussed in Sections 5.5 and 5.6.

Redistribution and Limited Government. Table 3 shows that greater TFE is associated with stronger opposition to income redistribution today. In column 1, we use ANES data from 1992 and 1996, which asks respondents whether they would like to see "federal spending on poor people be increased, decreased (or cut entirely) or kept about the same." Around nine percent of individuals would like to see such redistributive spending decreased. Each additional decade of TFE is associated with one additional p.p. increase in support of cuts. Column 2 provides complementary evidence, showing that each decade of TFE is associated with 0.7 p.p. higher support for cutting state spending on welfare as reported in the CCES. Following Alesina and La Ferrara (2005), column 3 uses a measure from the GSS indicating the intensity of preferences for redistribution on a scale from 1 to 7 (with 1 being that the government should not be engaged in redistribution and 7 being that the government should reduce income differences through redistribution). Each additional decade of TFE is associated with around 0.02 standard deviations lower support for redistribution. These effect sizes are akin to a 5–10 year age gap in preferences among respondents, with older respondents more in favor of welfare spending cuts, reflecting well-known cohort differences.

Turning to other measures, columns 4 and 5 show that residents of areas with greater TFE exhibit stronger fiscal conservatism. Column 4 uses a CCES question on whether individuals would prefer to cut domestic spending or to raise taxes to balance the federal budget. Column 5 uses an index based on the principal components of a set of questions from the GSS on whether the government spends too much on an array of public goods and social transfers. In both cases, we find that individuals are significantly more opposed to high levels of government spending in areas with greater TFE. The Kling et al. (2007) mean effects analysis yields similar insights. For example, combining the CCES measures in columns 2 and 4 into a single index yields a statistically significant effect of around 0.02.

Finally, column 6 of Table 3 shows that these reported preferences line up with actual policy differences across counties. In particular, each decade of TFE is associated with around 3.4 percent lower

reported property tax rates, which range from 0.1 to 2.9 percent across counties in our study. Given that much of the variation in tax rates lies across rather than within states, this is not a small effect. In fact, it is roughly akin to the within-state difference between counties that are 10 percent more versus less aligned with the Republican Party, a policy outcome we consider next.

Party Identification. We show in Table 4 that the persistent effects of TFE have strong implications for the growing strength of the Republican Party. Republican Party platforms have been increasingly associated with broad opposition to government intervention and aversion to redistribution. As a baseline, we consider the average vote share across the five elections since 2000. Column 1 shows that each decade of TFE is associated with around a 2 p.p. greater Republican vote share relative to the mean of 60 percent. This effect size is plausible and in line with individual-level regressions using degree of stated support for the Republican Party in the CCES.[28] For perspective, the 2 p.p. effect is roughly the difference in population-weighted, average county-level vote shares in Iowa (48.4 percent) and Wisconsin (46.3 percent) over these five elections. This average effect across the 2000s masks an interesting ratcheting up over time as seen in columns 2–6. An additional decade of TFE is associated with a significantly higher Republican vote share in each subsequent election, based on cross-equation tests of relative effect sizes.

Columns 7 and 8 then provide marked evidence of the relatively larger shift towards the Republican Party in areas with greater TFE. The average heartland county in our long-run analysis exhibits a 9 p.p. shift towards Republican candidates from 2000 to 2016. Each decade of TFE is associated with an additional 1.6 p.p. relative to that mean. Alternatively, comparing a county at the 25th percentile of TFE (11 years) to a county at the 75th percentile of TFE (24 years), implies an additional 2.2 p.p. Republican Party shift. As a benchmark, Autor et al. (2016) find that an interquartile shift in exposure to import competition from China implies a 1.7 p.p. Republican shift relative to a mean of -0.6 p.p. over the same period in their full-country sample of commuting zones.[29] A similarly large shift can be seen in column 8, which shows the frontier effect on the differential between 2012 and 2016. These findings offer suggestive evidence of a potential link between frontier culture and the growing strength of the Republican Party in certain areas of the American heartland.

In sum, Table 4 suggests relatively more conservative voting preferences in areas that were part of the frontier for a longer duration in the 19th century. While people vote Republican for many reasons, one recurring theme in recent years is the view that government should not be too heavily relied upon and hence government should be small. Some of these views bear an interesting similarity to the individualistic norms described in historical narratives of the frontier. It is in this respect that we view these voting outcomes as reflecting preferences shaped by frontier culture whether by way of motivated beliefs or other mechanisms discussed in Section 2.2.

Using the CCES, we provide further insight into why TFE may be associated with increasing Republican Party support today. In Table 5, we relate TFE to measures of opposition to (1) the Affordable Care

[28] Using the CCES 2007, 2012, and 2014 survey rounds, we construct an indicator equal to one if the respondent identifies as a "strong Republican" on a seven point scale ranging from "strong Democrat" to "strong Republican" with around 17 percent of individual–years reporting the latter. The estimates imply that an additional decade of TFE is associated with around 4.5 percent greater intensity of strong Republican support. As a benchmark, consider that with each additional year of age, individuals are around 2 percent more likely to report strong Republican support.

[29] In a more direct comparison, we use the original data from Autor et al. (2013) and map the China shock to our sample of counties. Estimating a single equation with both measures, we find that the TFE effect is around one-quarter as large as the effect of the more proximate China shock, though both remain statistically and economically significant.

Act (ACA or "Obamacare"), (2) increases in the minimum wage, (3) the ban on assault rifles, and (4) Environmental Protection Agency (EPA) regulations on pollution. These policy issues have been sharply contentious, with the two parties adopting increasingly polarized positions. Moreover, they can be connected to norms and beliefs pervasive on the frontier in terms of the link between effort and reward (ACA and minimum wage), the right to bear arms (ban on assault rifles), and the salience of manifest destiny (EPA regulations). The results in Table 5 show that places with greater TFE display significantly stronger opposition to each of these government regulations. Combining all estimates into a single index implies a mean effect size of around 0.04 that is significant at the 1 percent level. Moreover, as shown in Appendix B.5, these results are also robust to controlling for individual-level education, family income, and reported strength of party identification.

Summary of Results. Overall, the findings in Tables 3–5 paint a rich picture of the cultural and political legacy of historical frontier exposure. It is plausible that the early settlers left a lasting imprint on frontier locations and that the degree of that imprint increased with duration of exposure. As a summary takeaway, the Kling et al. (2007) mean effect on individualism (infrequent name share), conservative political preferences (Republican vote share) and policy (property tax rate) suggests that each decade of TFE is associated with roughly 0.15 standard deviations more frontier culture today.

Moreover, in connecting these outcomes to define a culture of "rugged individualism," it is important to note that infrequent name shares (in 1940) are strongly associated with greater Republican vote shares and lower property tax rates today.[30] In other words, the estimated effects in Table 2 are identified off of a similar set of counties as in Tables 3 and 4, again pointing to the close connection between individualism and opposition to redistribution.

5.5 Further Analysis and Robustness Checks

We report here several results that demonstrate the robustness of our key findings and interpretation. In Table 6, we expand our sample to include the secondary frontier on the West Coast, we measure TFE through 1950, and we show results separately by major Census region.[31] In Tables 7 and Table 8, we account for a battery of potential confounders, including present-day population density, mineral resource abundance, historical access to railroads, population diversity and slavery. Tables 6-8 focus on the key outcomes of infrequent names, property taxes, and Republican voting. Appendix B.8 reports these checks for the other survey-based outcomes. Although we include controls that may be outcomes of frontier experience and hence "bad controls," robustness to their inclusion rules out some long-run channels. In Table 9, we analyze racial differences to help sharpen our interpretation of the long-run effects of TFE. We conclude by considering placebo outcomes.

Adding West Coast, Extended Time Frame, Regional Heterogeneity. We start in Table 6 by adding West Coast frontier counties to our sample. These 105 counties were settled starting in the mid-19th century and were located to the west of the major frontier line on the West Coast in 1890 (the year

[30] Conditional on state fixed effects, a one standard deviation increase in (average male and female) infrequent name shares in 1940 is associated with 0.3 greater (0.23 lower) standard deviations of Republican vote shares (property tax rates).

[31] Appendix B.7 shows robustness to a battery of alternative ways of measuring TFE by varying the density and proximity cutoffs as well as the treatment of small pockets of dense (sparse) settlement beyond (inside) the main frontier line.

in which the Census declared the frontier closed). As shown in column 1, for all key outcomes, the estimated effects of TFE remain significant, with very similar magnitudes.

Then, we split the sample by Census region and show that the effects of TFE hold separately in the Midwest (column 2), the South (column 3), and the West (column 4).[32] The coefficient estimates are generally largest in the Midwest and smallest in the West. The relatively noisier estimates for the West are due to the small sample size (152 counties) as seen in subsequent columns 6–8 that extend the frontier time period through 1950, incorporating counties that experienced frontier conditions beyond 1890. Here, the effects of TFE are economically and statistically significant across all regions.

At first glance, the stability across regions may seem puzzling insomuch as the West region on average exhibits more collectivism and less opposition to redistribution.[33] However, our empirical strategy isolates the effects of TFE within-region and, moreover, within-state. That is, counties within California with greater TFE exhibit more prevalent individualism and opposition to redistribution compared to counties within California with lower TFE. If anything, the similar results across regions is reassuring and suggests that our findings capture a specific cultural legacy of settlement history, a point we further substantiate in subsequent tables.

Finally, note in columns 5–8 that the extended 1950 sample delivers estimated effects of TFE that are somewhat smaller in magnitude than the baseline 1890 sample. This could reflect that in the 20th century, transcontinental railroads and improved communications made frontier locations less isolated than were historically. According to Lang et al. (1995), "the modern-day [post-1890] frontier is not the nineteenth-century one. It is smaller, more law-abiding and regulated, less isolated, less rugged, and less dangerous." Moreover, these authors note that "the frontier has not for generations been the dream of those who seek a fortune or a new life."[34] Together, these factors may imply more limited scope for the mechanisms generating and amplifying frontier culture during early stages of settlement.

Disentangling Population Density. Contemporary population density is a key potential confounder of the effects of TFE. Population density can be very persistent (Bleakley and Lin, 2012). Places with low TFE may display different cultural and political attributes simply because they are sparsely populated today. For example, as is well-known, there is a large rural–urban divide in Republican vote shares.

We take several steps in Table 7 to disentangle the effects of frontier settlement history from those of present-day population density. Column 1 displays the baseline estimates for each of our key outcomes. Column 2 controls for population density in 1890, the final year of the frontier era. Column 3 implements a flexible specification including dummies for each decile of *within-state* population density in 1890. Column 4 controls for population density measured contemporaneously with the outcome variable. Column 5 repeats the decile-specific fixed effects approach for contemporary density. All specifications include state fixed effects and our baseline geographic and agroclimatic controls. The flexible specifications of density are very demanding, leaving limited variation to identify our coefficient of in-

[32] There are four Census regions: Northeast, South, Midwest, and West. We do not consider the Northeast as there are too few counties in this region (66 east of the 1790 frontier line) for credible within-state regression analysis. Note that our baseline sample includes 47 counties in Colorado, New Mexico, and Wyoming in the West region (see Figure 3) but not on the West Coast frontier as used in column 1 of Table 6.

[33] The large effects of TFE in the South may also seem surprising given the legacy of slavery. We address this issue below.

[34] Going further, the "nineteenth-century frontier was a dynamic settlement process" whereas the "twentieth century frontier is a more static place." It is this dynamic process that most concerned Turner and which forms the core of our analysis.

terest. Nevertheless, statistical and economic significance remain for all outcomes. The same holds for alternative specifications relying on polynomials or splines of population density. This suggests that modern preferences were shaped by the history of frontier settlement through channels other than simply persistently low population density.[35]

The remaining columns of Table 7 further clarify that the long-run effects of TFE are driven by the history of frontier settlement rather than the history and persistence of low or high density. Columns 6 and 7 show that TFE has similar effects in urban and rural areas, splitting the sample into counties above and below the 90th percentile of urban population shares, respectively. Finally, column 8 separates out the history of low density—the number of decades with density below 6 people/mi^2—from the measure of total frontier experience. Recall that this was one of the defining features of frontier locations, proximity to the frontier line being the other. The coefficient on TFE remains significant, indicating that both dimensions of the history of frontier experience are important. In sum, the effects of TFE go above and beyond the correlated effects of present-day and historical low population density.

Additional Controls. While our baseline regressions include spatial fixed effects and an array of geoclimatic controls, there are of course other factors that may be correlated with TFE and contemporary culture. We address this concern in part by including a number of additional controls in Table 8.[36]

Column 1 reports once again our baseline estimates. Columns 2–5 add, one at a time, four geoclimatic features whose importance has been emphasized in previous work: ruggedness (Nunn and Puga, 2012a), rainfall risk (Ager and Ciccone, forthcoming; Davis, 2016), distance to portage sites (Bleakley and Lin, 2012), and distance to mines. The latter seems likely to generate a downward bias in our estimates, as Couttenier and Sangnier (2015) show that across U.S. states opposition to redistribution is associated with mineral abundance (which in turn is likely to reduce TFE by leading to more rapid settlement). In all cases, though, the point estimates remain nearly identical.

We also consider a number of other controls. Column 6 adds distance to the nearest Indian battle site. Conflict with Native Americans probably increased TFE and plausibly affected preferences. Columns 7 adds the prevalence of slavery (in 1860, just before the Civil War), another potential confounder. Column 8 adds the sex ratio, which was systematically higher in the frontier and could by itself be a force shaping long-run cultural outcomes (see Grosjean and Khattar, forthcoming).

Columns 9–11 add additional demographic variables: the share of migrants in the population, the share of Scottish and Irish migrants (which have been associated with higher levels of violence, as shown by Grosjean, 2014), and a measure of birthplace diversity. We measure all of these in 1890, the endpoint of the frontier period.[37]

Column 12 controls for the number of years that each location was connected to the railroad network (until 1890), which is likely to reduce TFE and may also affect attitudes toward government intervention. Column 13 adds the share of the population employed in manufacturing (in 1890), a basic measure of economic development. In column 14, we include all the additional controls in the same regression.

[35] While TFE and modern density are indeed highly correlated, the latter may have independent effects or capture a mechanism through which TFE affects outcomes today. For example, high TFE locations with greater individualism and less redistribution (and thus low provision of public goods) may have been unattractive and thus grew slowly over the last century. This might have discouraged outsiders with different cultural traits from moving into such areas in the post-frontier period.

[36] See Appendix Table B.7 for the full elaboration of coefficients.

[37] We also tried analogous controls usign the average values over the 1820–1890 period and obtained similar results.

Throughout Table 8, the estimated effects of TFE remain significant and relatively stable despite the substantial added explanatory power of the additional controls. Finally, note that adding controls for population density (as in Table 7) leaves the results in column 14 largely unchanged.

Racial Differences. To shed further light on the link between historical frontier experience and modern preferences, we show that there are stark differences in the long-run effects of TFE by race. We saw in Table 6 that TFE has large effects on Republican vote shares in the South (Census region), but in Table 8, controlling for the slave share in the 1860 county population reduced the average long-run effect for the full sample. Table 9 offers some insight into these results. We find precise null effects of TFE for African American respondents across the six measures of opposition to redistribution and regulation in the CCES. Moreover, while non-whites (i.e., blacks and others) generally exhibit weaker opposition to government intervention, it is only the black population for whom we find null effects.

These results lend support to our interpretation of the long-lasting effects of frontier experience on culture and political preferences. Most of today's black population in the U.S. can trace their familial roots to slavery. Given the extreme barriers to geographic and socioeconomic mobility faced by slaves and their postbellum descendants, the mechanisms linking frontier experience to modern outcomes (see Section 2.2) would have been largely irrelevant to blacks living in high TFE regions.[38]

These racial differences in the effects of TFE are connected to the possibility that racial resentment by whites may explain some of our findings on opposition to redistribution and shifts in voting patterns across the United States. Indeed, part of that resentment may be linked to beliefs—accurate or not—about the role of effort versus luck in generating income (e.g., pertaining to views of affirmative action and welfare programs). Using measures from the CCES, we find a significant association between this type of racial resentment and TFE, but it does not survive controlling for contemporary population density, thus pointing to an urban–rural divide rather than a high TFE–low TFE divide.[39] This stands in contrast to the robustness of the association of TFE with our key outcomes of interest.

Placebo Outcomes. While frontier experience affects many cultural traits, preferences over certain policies do not have a clear connection with historical frontier experience. For illustration, we consider preferences over a few foreign policy issues as placebo outcomes: support for U.S. military intervention abroad in the case of genocide or civil war (35 percent in the CCES), opposition to the Iran sanctions regime (20 percent), and opposition to the U.S.-Korea Free Trade Agreement (45 percent). Estimating our baseline specification, we find relatively precise null effects of TFE on these three measures (-0.004, -0.003, and 0.003, respectively).

[38] The results in Table 9 are driven largely by the South. When splitting the sample into the South and non-South Census regions, we find more muted and in some cases no differences in the effects of TFE between black and non-black respondents. This may be due in part to the selective migration of blacks out of the South and into frontier areas in the late 1800s. While still subject to different barriers to upward mobility than whites, such self-selected black migrants were arguably more exposed to the influence of frontier conditions than those remaining in the postbellum South. See Billington and Hardaway (1998) for a rich exploration of African Americans on the frontier.

[39] The 2010, 2012, and 2014 rounds of the CCES make two statements about racial resentment and ask respondents to state their degree of agreement on a scale from strongly agree to strongly disagree: (i) "The Irish, Italians, Jews and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors." (64 percent somewhat or strongly agree), and (ii) "Generations of slavery and discrimination have created conditions that make it difficult for Blacks to work their way out of the lower class." (51 percent somewhat or strongly disagree). While TFE exhibits a significant positive association with both measures in our baseline regression (0.010** and 0.012***, respectively), controlling for 2010 population density, as in column 3 of Table 7, renders the estimates null and insignificant (0.001 and 0.004, respectively).

5.6 Instrumental Variable Strategy

Despite the battery of robustness checks, omitted variables remain a concern. This section introduces an instrumental variables (IV) strategy that isolates external variation in TFE, beyond what can be explained by geoclimatic controls. We exploit time series variation in the intensity of immigration inflows to the U.S. Our approach ensures that key results are driven by historical variation in *local* TFE associated with *national* population shocks.

Immigrants contributed to westward expansion by exerting population pressure on the Eastern seaboard and by going west themselves. Appendix B.9 documents the connection between the intensity of migrant inflows and the speed of westward expansion. For a given location, TFE partly reflects the speed of westward movement at the national level during the relevant time frame. To construct the instrument, we determine the first year in which each county is within 110 km of the frontier line. At this time, the county's local conditions do not affect the contemporaneous process of westward expansion, but the moving frontier is getting close. We then consider the average annual immigrant inflow (in logs) in the next 30 years.[40]

An important identifying assumption is that the intensity of immigrant flows to the U.S. is unaffected by the conditions of any given frontier county. This would not hold if, for example, Europeans' migration decisions in a given period were influenced by knowledge of frontier locations (e.g., their levels of land productivity). To address this concern, we aim to eliminate potential pull factors and isolate push factors unrelated to conditions on the frontier. Following Nunn et al. (2017), we predict migrant outflows from Europe based on climate shocks, and use these predictions to construct an alternative version of the instrument. Appendix B.9 provides full details.

Table 10 presents IV estimates for the same four primary outcomes as in Table 6.[41] In Panel A, we find large and statistically significant effects of TFE that are slightly larger but generally indistinguishable from the OLS estimates.[42] Panel B shows similar results when using predicted rather than actual migrant flows in the IV construction. Both instruments are quite strong (see Appendix Table B.11 for the first stage). Overall, the IV exercises strengthen our understanding of the identifying variation linking frontier experience to modern culture. While the shocks underlying the IV may not be perfectly excludable, they do generate external variation in TFE and move us closer to a causal interpretation.

6 The Roots of Frontier Culture

This section explores mechanisms through which the frontier shaped a culture of individualism and opposition to government intervention. Section 6.1 shows that there was significant selective migration of individualistic types to the frontier, though not all of the differential individualism on the frontier can be explained by selection.[43] We then document empirical patterns consistent with frontier conditions

[40] We choose this window as a baseline because nearly 85 percent of counties exit the frontier within that time. Results are similar for other time windows.

[41] The alternative measures of infrequent names are also robust to the IV approach (see Appendix Table B.3).

[42] The slightly larger effect sizes in the IV could be due to measurement error or a local average treatment effect arising from the instrument isolating variation in TFE that is more closely linked to the mechanisms we explore in the following section.

[43] While emphasizing the implications of selective migration of individualists for frontier culture, selection on other attributes might have been important as well. For example, selective migration of men, reflected in the high sex ratios seen in Section 4, might have also contributed to rugged individualism as men may be less inclined toward cooperation and interdependence

shaping people's values and behavior. Section 6.2 shows that individualism was differentially rewarded on the frontier, which may have fostered the prevalence of this trait over time. Section 6.3 shows that frontier conditions implied favorable prospects of upward mobility and a large perceived importance of effort in income generation, which would hone opposition to redistribution.[44]

The results presented here are meant to be suggestive rather than conclusive. We do not try to disentangle the quantitative importance of the different mechanisms, which we view as complementary and mutually reinforcing. For instance, a greater adaptive advantage of individualism on the frontier would induce more (less) selective migration (outmigration) of individualists. And conversely, selective migration of individualists to the frontier would likely increase the advantage of this trait, insofar as conformity to group norms would be of limited value in a society of individualists. In addition, the greater the advantage of individualism in the frontier, the more favorable their upward mobility prospects, which would feed into opposition to redistribution. And if this shaped local institutions and the provision of public goods, it could further reinforce selective migration of individualists over time.

Once frontier culture put down roots, it may have persisted through various mechanisms, even if the distinctive features of frontier settlement were long gone. Initial conditions can determine the long-run equilibrium through the dynamics of intergenerational cultural transmission. Moreover, since the frontier shaped culture at the earliest stages, it was bound to influence the formation of local institutions and social identity, which likely affected the subsequent evolution of cultural traits.

6.1 Selective Migration

This section investigates the role of selective migration in explaining the pervasiveness of individualism on the frontier. Our basic strategy is to distinguish the relative contributions of early versus later frontier settlers to the overall differential in individualistic naming patterns. The key intuition is that because the latter have lived in the given location for a longer period of time, local conditions have a greater scope for affecting their preferences by the time we observe them. To estimate selection patterns, we need to track households across time. For this purpose, we use full count data from the 1870 and 1880 Censuses provided by `ancestry.com` and the North American Population Project or NAPP (Sobek et al., 2017), respectively. We focus on the latest consecutive rounds available within the frontier period to ensure a large sample. The data include location, names, and demographics. We link individuals across rounds using an algorithm developed by Feigenbaum (2016).[45]

(Cross and Madson, 1997; Gabriel and Gardner, 1999) and more opposed to redistribution (Ashok et al., 2015). In a recent contribution, Grosjean and Khattar (forthcoming) establish the long-run effects of historical male-biased sex ratios on gender norms and female labor force participation in Australia.

[44] In Appendix B.10, we consider another potential mechanism, the prevalence of infectious diseases, and show that the evidence does not support its relevance in explaining differential individualism on the frontier.

[45] The base sample in 1880 is restricted to male household heads, native-born, aged 30–50, white, and who have at least one (biological) child aged 0–10. The target year is 1870. The set of potential matches for these men are first identified based on first and last name, birth state and birth year. A random training sample is then drawn from among the potential matches and manually trained. The importance of each match feature is quantified using a probit model, and used to estimate a probability score for each link. A true match is defined as one with a sufficiently high score both in absolute and relative terms. The match rate was 25 percent, which is comparable with the rates achieved by recent studies linking records with broadly comparable data albeit different target populations (e.g., 29 percent in Abramitzky, Boustan and Eriksson, 2012; 26 percent in Collins and Wanamaker, 2017; and 22 percent in Long and Ferrie, 2013). Although matching on names leaves scope for sample selection, our core results in Tables 11 and 12 look similar when reweighting using the inverse probability of being linked across Census rounds (following Bailey et al., 2017). We estimate these probabilities using the same characteristics used for linking as well

Table 11 reports estimates of the frontier differential in infrequent naming patterns based on versions of the following equation for different sub-populations of movers and stayers:

$$\text{child has infrequent name}_{ic,1880} = \alpha + \beta\ \text{frontier}_{c,1880} + \mathbf{x}_{ic}'\boldsymbol{\zeta} + \varepsilon_{ic,1880}, \tag{6}$$

where the binary dependent variable equals one if child i residing in county c in 1880 has a name that falls outside the top 10 nationally in that decade, and the frontier indicator equals one if county c lies on the frontier according to our baseline definition. We restrict attention to white children aged 0–10 with native-born parents and cluster standard errors at the county level. The $\mathbf{x}_{ic}$ vector includes age×gender and birth order fixed effects as well as indicators for whether the parents have infrequent names, but results are identical without these controls.

Column 1 of Table 11 identifies the significance of selective migration. Children in households that migrated to the frontier between 1870 and 1880 are 4.2 p.p. more likely to have infrequent names than those remaining in non-frontier areas during that period, 71 percent of whom have infrequent names. While we do not observe whether these children were born before or after arriving on the frontier, this differential points to the self-selection of individualist types.

Column 2 captures the overall frontier differential in individualism. Children in frontier counties in 1880 are 7.5 p.p. more likely to have an infrequent name relative to children in non-frontier locations. Next, we show that the longer-term frontier residents (stayers) exhibit stronger individualism than recent arrivals from other counties. Column 3 decomposes the 7.5 p.p. differential into differences coming from early versus later frontier settlers. Early settlers in frontier counties are nearly three times more likely to give their children infrequent names than those that arrived more recently during the 1870s. Column 4 corroborates this differential, restricting the sample to those living in frontier counties in 1880. These results suggest that greater time on the frontier is associated with more individualistic naming patterns.

Overall, the findings in Table 11 provide suggestive evidence that selection was significant does not fully explain the frontier differential in individualism. It is of course still possible that selective migration before 1870, which cannot be observed in this data, helps explain some of the differential.[46] For example, pre-1870 frontier migrants may be more individualistic than post-1870 frontier migrants. However, for this to fully explain the differences, the degree of differential selection would have to be nearly three times as large, which seems unlikely given that both groups of individuals migrated when the county was characterized by frontier conditions.

If the results instead suggested that selective migration explained all of the frontier differential, it might seem that the presence of a frontier had no aggregate implications for American culture. While the reallocation of people across the country cannot change the national prevalence of individualism at a given point in time, it may do so in the long-run because individualists arrived in frontier locations at a time when local culture and institutions were taking shape. By comparison, individualists' emigration from non-frontier regions arguably had more limited countervailing effects since those areas had already reached more advanced stages of settlement.

If selective migration does not fully explain the frontier differential, as our results suggest, the impli-

as an interaction of infrequent name status and frontier location in 1880. These interactions re-balance the linked sample to account for differential missing-ness along our key variables of interest.

[46] Abramitzky et al. (2014) make a similar selection argument about early versus later immigrants to the U.S.

cation is that frontier life reinforced and amplified the already individualistic tendencies of settlers. We now explore one potential explanation for why: individualism had differential returns on the frontier.

6.2 The Adaptive Advantage of Individualism

The opportunities and threats faced by frontier settlers may have favored individualism through an adaptive mechanism. Because people on the frontier primarily had to rely on themselves for protection and material progress, the independent, self-reliant types would arguably have fared better (Kitayama et al., 2010).[47] Moreover, frontier settlers faced novel agroclimatic conditions, and there was little local knowledge about how best to approach the harsh and unfamiliar setting (see Baltensperger, 1979; Libecap and Hansen, 2002; Shannon, 1977). Adherence to old traditions and norms was less suited to the environment than non-conformism and innovation, two traits associated with individualism.[48]

This section presents evidence consistent with an adaptive advantage of individualism in frontier conditions. Using data from the linked Census sample, we show that households exhibiting greater individualism were more successful economically and more likely to stay in frontier locations.

First, we estimate the relationship between father i's economic status in county c in 1880, $y_{ic,1880}$, and infrequent names according to the following difference-in-difference type specification:

$$\begin{aligned} y_{ic,1880} \;=\; & \alpha + \beta \text{ own infrequent name}_i + \eta(\text{own infrequent name}_i \times \text{frontier}_{c,1880}) \\ & + \delta \text{ children infrequent name}_i + \zeta(\text{children infrequent name}_i \times \text{frontier}_{c,1880}) + \theta_c + \varepsilon_{ic}, \end{aligned} \tag{7}$$

where β captures the hedonic returns to the father's own infrequent name outside the frontier and η the differential effect on the frontier. At the same time, δ captures the association of infrequent name choices for children born during the 1870s and the father's economic well-being outside the frontier, and ζ the frontier differential. We restrict attention to white, native-born fathers that did not move between counties from 1870 to 1880 and had at least one child in 1880. Again, we define infrequent names as those outside the top 10 nationally, but other definitions yield similar results.

We measure economic status y_{ic} using data on occupation from the 1880 Census recorded in the linked sample. We consider the Duncan (1961) socioeconomic index (*sei*) and the occupational score (*occscore*) provided by the NAPP. Both measures range from 0 to 100 and capture the income returns associated with occupations in the 1950 Census, and *sei* additionally captures education and occupational prestige. These measures are widely used in the economic history literature and capture broad differences in economic status across individuals (see Olivetti and Paserman, 2015, for a discussion). Finally, we cluster standard errors at the county level, and the county fixed effects, θ_c, account for all differences in outcomes common across individuals within the same county.

The estimates in Panel A of Table 12 provide evidence of differential returns to individualism on the frontier. Across all specifications, fathers with infrequent names outside the frontier exhibit socioeco-

[47] This view can be framed within a notion of culture as decision-making rules-of-thumb used in uncertain environments, as proposed by evolutionary anthropologists (Boyd and Richerson, 1985, 2005). In their models, a process of natural selection governed by the payoffs from different rules-of-thumb determines which rule prevails.

[48] The connection between innovation and individualistic culture is discussed at length in Gorodnichenko and Roland (2012). In characterizing the traits of frontier populations, Turner (1893) himself mentions individualism along with the "coarseness and strength combined with acuteness and inquisitiveness" and the "practical, inventive turn of mind, quick to find expedients."

nomic status that is nearly 0.05 standard deviations lower than fathers with more common names. This is roughly the typical difference between a farmer and a blacksmith or a blacksmith and a carpenter.[49] This apparent economic penalty might be due to various types of discrimination or other mechanisms favoring conformity. However, this penalty is more than offset on the frontier where infrequent names exhibit a differentially positive association with economic status. We find a similar differential for infrequent names of children, which exhibit a positive correlation with father's status outside the frontier and an even stronger positive correlation on the frontier. These results, which hold for both *sei* and *occscore*, suggest that individualists are relatively better off on the frontier.

In Appendix Table B.13, we show that these differential hedonic returns arise not only for levels but also for changes in socioeconomic status. The NAPP linked sample for 1870–1880 allows us to investigate changes in occupational standing for 1 percent of the entire population. We are constrained to this small subset of all individuals in the prior analyses because the full county data for 1870 provided by `ancestry.com` does not include occupational or socioeconomic status measures. The results show that fathers with infrequent names exhibit significantly faster growth in *sei* and *occscore* on the frontier but not outside the frontier. These results and those in Table 12 are robust to allowing those in farming occupations to have a different intercept (see Appendix Table B.14).

In Panel B of Table 12, we provide a second piece of evidence consistent with an individualist advantage on the frontier. We estimate the following equation relating infrequent names to migration choices for household h living in frontier county c in 1870:

$$\begin{aligned} \text{outmigrate}_{hc} \quad = \quad & \alpha + \beta_f \text{ father has infrequent name}_h + \beta_m \text{ mother has infrequent name}_h \\ & + \eta \text{ any children with infrequent name}_h + \theta_c + \varepsilon_{hc}, \end{aligned} \tag{8}$$

where outmigrate_{hc} is a binary outcome indicating whether the household moved from a frontier county in 1870 to a non-frontier county by 1880. The key explanatory variables are defined as above, with the mother's infrequent name status defined similarly. The results suggest that, within a given frontier county, households in which fathers have infrequent names are around 4 percentage points less likely to leave the frontier by 1880. This is a sizable magnitude given that 40 percent of linked households in our sample left the frontier during this period.[50] We observe little relationship to mother's names, but households with children with infrequent names are also significantly less likely to leave the frontier.[51]

Overall, the findings in Table 12 suggest that inherited and revealed individualism are associated with a higher likelihood of socioeconomic success on the frontier. This may explain the self-selection of individualists to the frontier as well as the diffusion of individualistic traits after arrival. The adaptive value of individualism probably favored its prevalence in frontier settlements not only through selective immigration and emigration but also through other mechanisms including differential fertility and survival, and various forms of cultural transmission.

[49] A list of the top 10 occupations in frontier versus non-frontier counties is in Appendix Table B.1.

[50] Appendix Table B.15 shows that this migration accounts for most departures from frontier counties. The other direction of migration—from frontier counties to other frontier counties—is not associated with infrequent names.

[51] Of course, this latter result is less straightforward to interpret as we do not observe the timing of migration within the decade.

6.3 Effort as the Road to Riches

This final section argues that the opportunities and challenges on the frontier contributed to a culture of opposition to government intervention. The frontier's favorable prospects of upward mobility and a large perceived importance of effort in income generation may have fostered opposition to tax redistribution, as suggested by the literature discussed in Section 2.2. This connection between the American frontier and theories of preferences for redistribution, hinted at by Alesina et al. (2001), echoes Billington (1974), who argued that the frontiersman "wanted not government interference with his freedom as he followed the road to riches."

In his reading of the Turner thesis, Billington (1974) emphasizes the implications of the frontier's land abundance and "widespread property holdings." In these conditions, "a man's capacities, not his ancestry, determined his eventual place in the hierarchy, to a greater degree than in older societies." The frontiersman believed that "his own abilities would assure him a prosperous future as he exploited the natural resources about him." Access to land offered profit of opportunities, even for settlers with low initial wealth. Class distinctions were also weakened by the ubiquity of threats characterizing frontier life. As Overmeyer (1944) argues, since everyone "had to face the same hardships and dangers," the frontier was a "great leveling institution."

Numerous historical studies present stylized facts consistent with the frontier presenting both prospects for upward mobility and a large perceived importance of effort.[52] As summarized by Stewart (2006), the frontier was "a place of economic opportunity," where settlers had low levels of initial wealth, but land-holding was widespread and rates of wealth accumulation were high, especially for early settlers and those that were able to endure.

Indeed, as shown in Appendix Figure B.2, historical Census data on landholdings is consistent with the idea that frontier locations offered a more level playing field. Land inequality, captured by the Gini coefficient, was significantly lower on the frontier, with a pattern resembling what we documented for key demographics and individualism in Section 4. Moreover, this difference dissipated over time as counties exited the frontier and the usual forces giving rise to inequality took hold.

In sum, the stylized facts summarized above suggest a relatively limited role for inherited social class as a key determinant of income and wealth generation in the frontier economy. This implied a level playing field offering equality of opportunity, and a relatively high importance of effort as opposed to luck (of being born into a given class). Together with the selection and cultivation of individualism, these conditions plausibly contributed to the origins and persistence of frontier culture.

7 Discussion

This paper provides new evidence on the historical and long-run effects of the American frontier on culture at the subnational level. Historically, frontier locations exhibited starkly different demographics and a higher prevalence of individualism as reflected in name choices for children. Today, counties that remained on the frontier for a longer period historically exhibit stronger opposition to government intervention in the form of redistribution, taxation and various regulations. Guided by the historical

[52] See Curti (1959), Galenson and Pope (1989), Gregson (1996), Kearl et al. (1980), and Schaefer (1987).

record and insights from social psychology and political economy, we offer empirical evidence on the origins of frontier culture, identifying the importance of selective migration, the adaptive advantage of self-reliance, and expectations of high income growth through effort.

The results shed new light on the roots and persistence of rugged individualism in the United States. We provide some of the first systematic evidence on a prominent theme in American history and lend credence to some elements of Turner's famous thesis. Our method for locating and tracking the frontier historically should prove useful in other attempts to understand the legacy of the frontier.

Our findings have suggestive implications about the sharp contrast between the U.S. and Europe in terms of preferences for redistribution and redistributive policies, a recurring topic in the literature (e.g., Alesina et al., 2001; Alesina and Glaeser, 2004; Alesina et al., 2012; Bénabou and Tirole, 2006). According to Turner, initially "the Atlantic coast ... was the frontier of Europe," but subsequently "the advance of the frontier ... meant a steady movement away from the influence of Europe," as "moving westward, the frontier became more and more American." Intuitively, as settlers of European origin shed their former culture and embraced rugged individualism across the U.S., America as a whole became more and more different from Europe.

The results also offer new perspective on contemporary political debates. The deep roots of opposition to redistribution in the United States may explain why their levels remain persistently high even in the face of sharply rising inequality. Our findings suggest that expressions of stark opposition to government intervention amidst growing political polarization may reflect not only a reaction to current events but also a rekindling of long-standing elements of American culture. Admittedly, since our study is based on subnational variation, extrapolation to the national level is speculative.

The persistence of rugged individualism points to the relevance of critical junctures. Early stages of settlement were formative for local culture and institutions, so the frontier was bound to leave a long-lasting imprint. As the frontier moved with time-varying speed, each location experienced frontier conditions in its early history for a different amount of time, and in most cases for more than a few years. Thus, the process of westward expansion created cross-sectional variation in rugged individualism, and may have amplified its prevalence at the national level.

Frontier settlement may have different effects in other countries. For instance, Argentina and Russia also underwent massive territorial expansion in their early history, but were ruled by elites that built very different institutions. In their analysis of how historical frontiers affected later democratic quality across countries in the Americas, García-Jimeno and Robinson (2011) show that the positive effects of frontiers depend on the quality of initial institutions. The national institutions of the United States, which favored relatively high levels of geographic mobility, access to land, and security of property rights, may have been preconditions for the operation of the mechanisms we emphasize.

References

Abramitzky, R., L. P. Boustan, and K. Eriksson, "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration," *American Economic Review*, 2012, *102* (5), 1832–1856.

_ , _ , and _ , "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration," *Journal of Political Economy*, 2014, *122* (3), 467–506.

_ , _ , and _ , "Cultural Assimilation during the Age of Mass Migration," *NBER Working Paper 22381*, 2016, (22381).

Acharya, A., M. Blackwell, and M. Sen, "The Political Legacy of American Slavery," *The Journal of Politics*, 2016, *78* (3), 621–641.

Ager, P. and A. Ciccone, "Agricultural risk and the spread of religious communities," *Journal of the European Economic Association*, forthcoming.

Alesina, A. and E. L. Glaeser, *Fighting Poverty in the US and Europe: A World of Difference*, Oxford, U.K.: Oxford University Press, 2004.

_ and E. La Ferrara, "Preferences for Redistribution in the Land of Opportunities," *Journal of public Economics*, 2005, *89* (5), 897–931.

_ and G. Angeletos, "Fairness and Redistribution," *The American Economic Review*, 2005, *95* (4), 960–980.

_ and P. Giuliano, *Preferences for Redistribution*, Vol. 1, North Holland,

_ and _ , "Family Ties," *Handbook of Economic Growth*, 2014, *2*, 177–215.

_ and _ , "Culture and Institutions," *Journal of Economic Literature*, 2015, *53* (4), 898–944.

_ , E. L. Glaeser, and B. Sacerdote, "Why Doesn't the United States Have a European-Style Welfare State?," *Brookings Papers on Economic Activity*, 2001, *2001* (2), 187–277.

_ , G. Cozzi, and N. Mantovan, "The Evolution of Ideology, Fairness and Redistribution," *The Economic Journal*, 2012, *122* (565), 1244–1261.

_ , P. Giuliano, and N. Nunn, "On the Origins of Gender Roles: Women and the Plough," *The Quarterly Journal of Economics*, 2013, *128* (2), 469–530.

Ansolabehere, S. and B. F. Schaffner, "General Social Surveys, 1972-2014," *CCES Common Content, 2006-2016*, 2017.

Ashok, V., I. Kuziemko, and E. Washington, "Support for Redistribution in an Age of Rising Inequality: New Stylized Facts and Some Tentative Explanations," *NBER Working Paper 21529*, 2015.

Atack, J., F. Bateman, M. Haines, and R. A. Margo, "Did Railroads Induce or Follow Economic Growth?," *Social Science History*, 2010, *34* (2), 171–197.

Autor, D., D. Dorn, and G. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *The American Economic Review*, 2013, *103* (6), 2121–2168.

_ , _ , _ , and K. Majlesi, "A Note on the Effect of Rising Trade Exposure on the 2016 Presidential Election," 2016. Mimeo, MIT.

Bailey, M., C. Cole, and C. Massey, "Representativeness and False Links in the 1850-1930 IPUMS Linked Representative Historical Samples," Technical Report, Working Paper 2017.

Baltensperger, B. H, "Agricultural Adjustments to Great Plains Drought: The Republican River Valley, 1870-1900," in B. W. Blouet and F. C. Luebke, eds., *The Great Plains: Environment and Culture*, Lincoln: University of Nebraska Press, 1979, pp. 43–59.

Beck-Knudsen, A. S., "Historical Individualism: Selective Migration and Cultural Persistence," 2017.

Bénabou, R., "Ideology," *Journal of the European Economic Association*, 2008, *6* (2-3), 321–352.

_ and E. A. Ok, "Social Mobility and the Demand for Redistribution: The Poum Hypothesis," *The Quarterly Journal of Economics*, 2001, *116* (2), 447–487.

_ and J. Tirole, "Belief in a Just World and Redistributive Politics," *The Quarterly Journal of Economics*, 2006, *121* (2), 699–746.

_ and _ , "Mindful Economics: The Production, Consumption, and Value of Beliefs," *Journal of Economic Perspectives*, September 2016, *30* (3), 141–64.

Bertrand, M. and S. Mullainathan, "Are Emily and Greg More Employable than Lakisha and Jamal? A

Field Experiment on Labor Market Discrimination," *American Economic Review*, 2004, *94* (4), 991–1013.

Bester, C. A., T. G. Conley, and C. B. Hansen, "Inference with dependent data using cluster covariance estimators," *Journal of Econometrics*, 2011, *165* (2), 137–151.

Billington, M. L. and R. D. Hardaway, *African Americans on the Western Frontier*, University Press of Colorado, 1998.

Billington, R. A., *America's Frontier Heritage*, University of New Mexico Press, 1974.

Bisin, A. and T. Verdier, "Work Ethic and Redistribution: A Cultural Transmission Model of the Welfare State," *Unpublished Manuscript, New York University*, 2005.

_ **and** _ , "The Economics of Cultural Transmission and Socialization," in J. Benhabib, A. Bisin, and M. O. Jackson, eds., *Handbook of Social Economics*, Vol. 1, North-Holland, 2010, pp. 339–416.

_ **and** _ , "On the Joint Evolution of Culture and Institutions," Technical Report, National Bureau of Economic Research 2017.

Bleakley, H. and J. Lin, "Portage and path dependence," *Quarterly Journal of Economics*, 2012, *127* (2), 587–644.

Boatright, M. C., "The Myth of Frontier Individualism," *The Southwestern Social Science Quarterly*, 1941, *22* (1), 14–32.

Bowen, W. A., *The Willamette Valley: Migration and Settlement on the Oregon Frontier*, Seattle: University of Washington Press, 1978.

Boyd, R. and P. J. Richerson, *Culture and the Evolutionary Process*, Chicago: University of Chicago Press, 1985.

_ **and P. J Richerson**, *The Origin and Evolution of Cultures*, New York: Oxford University Press, 2005.

Celinska, K., "Individualism and Collectivism in America: The Case of Gun Ownership and Attitudes Toward Gun Control," *Sociological Perspectives*, 2007, *50* (2), 229–247.

Chow, G. C., "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 1960, *28* (3), 591–605.

Collins, W. J. and M. H. Wanamaker, "Up from Slavery? African American Intergenerational Economic Mobility Since 1880," *NBER Working Paper 23395*, 2017.

Conley, T. G., "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Coombs, J., "Frontier Patterns of Marriage, Family, and Ethnicity: Central Wisconsin in the 1880s," *Journal of Family History*, 1993, *18* (3), 265–282.

Couttenier, M. and M. Sangnier, "Living in the Garden of Eden: Mineral resources and preferences for redistribution," *Journal of Comparative Economics*, 2015, *43* (2), 243–256.

_ **, P. Grosjean, and M. Sangnier**, "The Wild West is Wild: The Homicide Resource Curse," *Journal of the European Economic Association*, 2017, *15* (3), 558–585.

Cross, S. E. and L. Madson, "Models of the Self: Self-Construals and Gender," *Psychological Bulletin*, 1997, *122* (1), 5–37.

Curti, M. E., *The Making of an American Community: A Case Study of Democracy in a Frontier County*, Stanford, CA: Stanford University Press, 1959.

Davis, Lewis, "Individual responsibility and economic development: Evidence from rainfall data," *Kyklos*, 2016, *69* (3), 426–470.

Demos, J., "Families in Colonial Bristol, Rhode Island: An Exercise in Historical Demography," *The William and Mary Quarterly: A Magazine of Early American History*, 1968, *25* (1), 40–57.

Duncan, O. D., "A Socioeconomic Index for All Occupations," *Class: Critical Concepts*, 1961, *1*, 388–426.

Easterlin, R. A., G. Alter, and G. A. Condran, "Farms and Farm Families in Old and New Areas: The Northern States in 1860," *Quantitative Studies in History*, 1978, pp. 22–84.

Eblen, J. E., "An Analysis of Nineteenth-Century Frontier Populations," *Demography*, 1965, *2* (1), 399–413.

Enke, B., "Kinship Systems, Cooperation and the Evolution of Culture," *NBER Working Paper 23499*, 2017.

_ , "Moral Values and Voting: Trump and Beyond," *NBER Working Paper 24268*, 2018.

Feigenbaum, J. J., "Automated Census Record Linking: A Machine Learning Approach," *Unpublished Manuscript*, 2016.

Fernández, R., "Does Culture Matter?," in J. Benhabib, A. Bisin, and M. O. Jackson, eds., *Handbook of social economics*, Vol. 1, North-Holland, 2010, pp. 481–510.

Ferrie, J. P., "Migration to the Frontier in Mid-Nineteenth Century America: A Re-Examination of Turner's 'Safety-Valve'," *Department of Economics, Northwestern University Manuscript*, 1997.

Fincher, C. L. and R. Thornhill, "Parasite-Stress Promotes In-Group Assortative Sociality: The Cases of Strong Family Ties and Heightened Religiosity," *Behavioral and Brain Sciences*, 2012, *35* (2), 61–79.

Fryer, R. G. and S. D. Levitt, "The Causes and Consequences of Distinctively Black Names," *The Quarterly Journal of Economics*, 2004, *119* (3), 767–805.

Gabriel, S. and W. L. Gardner, "Are There" His" and" Hers" Types of Interdependence? The Implications of Gender Differences in Collective versus Relational Interdependence for Affect, Behavior, and Cognition.," *Journal of Personality and Social Psychology*, 1999, *77* (3), 642–655.

Galenson, D. W. and C. L. Pope, "Economic and Geographic Mobility on the Farming Frontier: Evidence from Appanoose County, Iowa, 1850–1870," *The Journal of Economic History*, 1989, *49* (3), 635–655.

Galor, O. and Ö. Özak, "The Agricultural Origins of Time Preference," *American Economic Review*, 2016, *106* (10), 3064–3103.

Gannett, Henry, "The Movements of Our population," *National Geographic Magazine*, 1893, (5).

García-Jimeno, C. and J. A. Robinson, "The Myth of the Frontier," in D. L. Costa and N. R. Lamoreaux, eds., *Understanding Long-Run Economic Growth: Geography, Institutions, and the Knowledge Economy*, Chicago: University of Chicago Press, 2011, pp. 49–89.

Gerrit, B. and D. Onland, "Socioeconomic Determinants of First Names," *Names*, 2011, *59* (1), 25–41.

Gershman, B., "Long-Run Development and the New Cultural Economics," in M. Cervelatti and U. Sunde, eds., *Demographic Change and Long-Run Development*, Cambridge, MA: MIT Press, 2017, pp. 221–261.

Giuliano, P. and N. Nunn, "Understanding Cultural Persistence and Change," Technical Report, Harvard University, mimeo 2017.

Glaeser, E. L., S. P. Kerr, and W. R. Kerr, "Entrepreneurship and urban growth: An empirical assessment with historical mines," *Review of Economics and Statistics*, 2015, *97* (2), 498–520.

Goodrich, C. and S. Davison, "The Wage-Earner in the Westward Movement I," *Political Science Quarterly*, 1935, *50* (2), 161–185.

_ **and** _ , "The Wage-Earner in the Westward Movement II," *Political Science Quarterly*, 1936, *51* (1), 61–116.

Gorodnichenko, Y. and G. Roland, "Which Dimensions of Culture Matter for Long-Run Growth?," *American Economic Review: Papers & Proceedings*, 2011, *101* (3), 492–498.

_ **and** _ , "Understanding the Individualism-Collectivism Cleavage and its Effects: Lessons from Cultural Psychology," in M. Aoki, T. Kuran, and G. Roland, eds., *Institutions and Comparative Economic Development*, London: Palgrave Macmillan, 2012, pp. 213–236.

_ **and** _ , "Culture, Institutions and Democratization," *NBER Working Paper 21117*, 2015.

_ **and** _ , "Culture, Institutions and the Wealth of Nations," *Review of Economics and Statistics*, 2016, *99* (3), 402–416.

Gregson, M. E., "Wealth Accumulation and Distribution in the Midwest in the Late Nineteenth Century," *Explorations in Economic History*, 1996, *33* (4), 524–538.

Greif, A., "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies," *Journal of Political Economy*, 1994, *102* (5), 912–950.

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

_ **and R. Khattar**, "Its Raining Men! Hallelujah? The Long-Run Consequences of Male-Biased Sex Ratios"," *Review of Economic Studies*, forthcoming.

Guiso, L., P. Sapienza, and L. Zingales, "Social Capital as Good Culture," *Journal of the European Economic Association*, 2008, *6* (2-3), 295–320.

Guiso, L, P. Sapienza, and L. Zingales, "Long-term Persistence," *Journal of the European Economic Association*, 2016, *14* (6), 1401–1436.

Gureckis, T. M. and R. L. Goldstone, "How You Named Your Child: Understanding the Relationship Between Individual Decision Making and Collective Outcomes," *Topics in Cognitive Science*, 2009, *1* (4), 651–674.

Hahn, M. W. and R. A. Bentley, "Drift as a Mechanism for Cultural Change: An Example from Baby Names," *Proceedings of the Royal Society of London B: Biological Sciences*, 2003, *270* (Suppl 1), S120–S123.

Haines, M. R. and ICPSR, *Historical, Demographic, Economic, and Social Data: The United States, 1790-2002*, Inter-University Consortium for Political and Social Research, 2010.

Heine, S. J., "Cultural Psychology," in S. T. Fiske, D. T. Gilbert, and G. Lindzey, eds., *Handbook of Social Psychology*, 2010, pp. 254–266.

Hoang-Anh, H., P. Martinsson, and O. Olsson, "The Origins of Cultural Divergence: Evidence from a Developing Country," *Unpublished Manuscript*, 2018.

Hofstede, G., *Culture's Consequences: International Differences in Work-Related Values*, London: Sage Publications, 1980.

_ , *Cultures and Organizations: Software of the Mind*, McGraw-Hill, 1991.

Hornbeck, R., "Barbed wire: Property rights and agricultural development," *The Quarterly Journal of Economics*, 2010, *125* (2), 767–810.

IIASA/FAO, "Global Agro-ecological Zones (GAEZ v3.0)," *Unpublished Manuscript*, 2012.

Jokela, M., "Personality Predicts Migration Within and Between US states," *Journal of Research in Personality*, 2009, *43* (1), 79–83.

Kashima, E. S. and Y. Kashima, "Culture and Language: The Case of Cultural Dimensions and Personal Pronoun Use," *Journal of Cross-Cultural Psychology*, 1998, *29* (3), 461–486.

Kearl, J. R., C. L. Pope, and L. T. Wimmer, "Household Wealth in a Settlement Economy: Utah, 1850–1870," *The Journal of Economic History*, 1980, *40* (3), 477–496.

Kim, H. and H. R. Markus, "Deviance or Uniqueness, Harmony or Conformity? A Cultural Analysis," *Journal of Personality and Social Psychology*, 1999, *77* (4), 785–800.

Kitayama, S., K Ishii, T. Imada, K. Takemura, and J. Ramaswamy, "Voluntary Settlement and the Spirit of Independence: Evidence from Japan's "Northern Frontier"," *Journal of personality and social psychology*, 2006, *91* (3), 369–384.

_ **, L. G. Conway III, P. R. Pietromonaco, H. Park, and V. C. Plaut**, "Ethos of independence across regions in the United States: the production-adoption model of cultural change," *American Psychologist*, 2010, *65* (6), 559–574.

Kling, J. R., J. B. Liebman, and L. F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Lang, R. E., D. E. Popper, and F. J. Popper, "'Progress of the Nation': The Settlement History of the Enduring American Frontier," *The Western Historical Quarterly*, 1995, *26* (3), 289–307.

Larson, J. L., "Grasping for the significance of the Turner legacy: An afterword," *Journal of the Early Republic*, 1993, *13* (2), 241–249.

Libecap, G. D. and Z. K. Hansen, ""Rain Follows the Plow" and Dryfarming Doctrine: The Climate Information Problem and Homestead Failure in the Upper Great Plains, 1890–1925," *The Journal of Economic History*, 2002, *62* (1), 86–120.

Lieberson, S. and E. O. Bell, "Children's First Names: An Empirical Study Of Social Taste," *American Journal of Sociology*, 1992, *98* (3), 511–554.

Limerick, P. N., *The legacy of conquest: The unbroken past of the American West*, WW Norton & Company, 1988.

Long, J. and J. Ferrie, "Intergenerational Occupational Mobility in Britain and the US since 1850," *American Economic Review*, 2013, *103* (4), 1109–1137.

Markus, G. B., "American Individualism Reconsidered," in J. H. Kuklinski, ed., *Citizens and Politics: Perspectives from Political Psychology*, New York: Cambridge University Press, 2001, pp. 401–32.

McFaul, EJ, GT Mason, WB Ferguson, and BR Lipin, "US Geological Survey mineral databases; MRDS and MAS/MILS," Technical Report, US Geological Survey, 2000.

Migration Policy Institute, "Tabulations of U.S. Department of Homeland Security, Office of Immigration Statistics, Yearbook of Immigration Statistics," 2016.

Minnesota Population Center, "National Historical Geographic Information System: Version 2.0.," *Minneapolis, MN: University of Minnesota*, 2011.

Modell, J., "Family and Fertility on the Indiana Frontier, 1820," *American Quarterly*, 1971, *23* (5), 615–634.

Moller, H., "Sex Composition and Correlated Culture Patterns of Colonial America," *The William and Mary Quarterly: A Magazine of Early American History, and Culture*, 1945, *2* (2).

Munshi, K., "Community networks and the process of development," *The Journal of Economic Perspectives*, 2014, *28* (4), 49–76.

Nelson, L. D. and J. P. Simmons, "Moniker maladies: When names sabotage success," *Psychological Science*, 2007, *18* (12), 1106–1112.

Nisbett, R. E. and D. Cohen, "Culture of Honor: The Psychology of Violence in the South," 1996.

Nunn, N. and L. Wantchekon, "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 2011, *101* (7), 3221–3252.

_ , N. Qian, and S. Sequeira, "Migrants and the Making of America," *Unpublished Manuscript*, 2017.

Nunn, Nathan and Diego Puga, "Ruggedness: The blessing of bad geography in Africa," *Review of Economics and Statistics*, 2012, *94* (1), 20–36.

_ and _ , "Ruggedness: The blessing of bad geography in Africa," *Review of Economics and Statistics*, 2012, *94* (1), 20–36.

Ogihara, Y., H. Fujita, H. Tominaga, S. Ishigaki, T. Kashimoto, A. Takahashi, K. Toyohara, and Y. Uchida, "Are Common Names Becoming Less Common? The Rise in Uniqueness and Individualism in Japan," *Frontiers in psychology*, 2015, *6*.

Olivetti, C. and M. D. Paserman, "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850–1940," *American Economic Review*, 2015, *105* (8), 2695–2724.

Olsson, O. and C. Paik, "Long-run cultural divergence: Evidence from the neolithic revolution," *Journal of Development Economics*, 2016, *122*, 197–213.

Oregon State University, "PRISM Climate Group," 2018.

Oster, E., "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business and Economic Statistics*, 2016, *0* (0), 1–18.

Overmeyer, P. H., "Westward Expansion Before the Homestead Act," in H. F. Williamson, ed., *The Growth of the American Economy: An Introduction to the Economic History of the United States*, New York: Prentice-Hall, Inc, 1944, pp. 82–112.

Piketty, T., "Social Mobility and Redistributive Politics," *Quarterly Journal of Economics*, 1995, *110* (3), 551–584.

Plaut, V. C., H. R. Markus, and M. E. Lachman, "Place Matters: Consensual Features and Regional Variation in American Well-Being and Self," *Journal of Personality and Social Psychology*, 2002, *83* (1), 160–184.

Porter, R., H. Gannett, and W. Hunt, "Progress of the Nation, including the Map of the Population of 1870," *Report on Population of the United States at the Eleventh Census*, 1890, pp. 13–30.

Quattrociocchi, Jeff, "Culture and Redistribution," *PhD Dissertation, York University*, 2014.

Restrepo, P., "The Mounties and the Origins of Peace in the Canadian Prairies," 2015.

Robinson, P. M., "Root-N-Consistent Semiparametric Regression," *Econometrica*, 1988, *56* (4), 931–954.

Samuelson, P., "Modern Economic Realities and Individualism," in G. H. Mills, ed., *Innocence and Power: Individualism in Twentieth-century America*, Austin: University of Texas Press, 1965, pp. 51–71.

Schaefer, D. F., "A Statistical Profile of Frontier and New South Migration: 1850-1860," *Agricultural History*, 1985, *59* (4), 563–578.

_ , "A Model of Migration and Wealth Accumulation: Farmers at the Antebellum Southern Frontier," *Explorations in Economic History*, 1987, *24* (2), 130–157.

Shannon, F. A., *The Farmer's Last Frontier: Agriculture, 1860-1897*, Vol. 5, ME Sharpe, 1977.

Smith, G.H., "The Populating of Wisconsin," *Geographical Review*, 1928, *18* (3), 402–421.

Smith, T. W., P. Marsden, M. Hout, and J. Kim, "General Social Surveys, 1972-2014," 2015.

Sobek, M., S. Ruggles, A. J. Trent, K. Genadek, R. Goeken, and M. B. Schroeder, "Integrated Public Use Microdata Series: Version 7.0 [dataset]," *University of Minnesota, Minneapolis*, 2017.

Spolaore, E. and R. Wacziarg, "How Deep are the Roots of Economic Development?," *Journal of Economic Literature*, 2013, *51* (2), 325–369.

Steckel, R. H., "Household Migration and Rural Settlement in the United States, 1850–1860," *Explorations in Economic History*, 1989, *26* (2), 190–218.

Stewart, J. I., "Migration to the Agricultural Frontier and Wealth Accumulation, 1860–1870," *Explorations in Economic History*, 2006, *43* (4), 547–577.

Tabellini, G., "Institutions and Culture," *Journal of the European Economic Association*, 2008, *6* (2-3), 255–294.

_ , "The Scope of Cooperation: Values and Incentives," *The Quarterly Journal of Economics*, 2008, *123* (3), 905–950.

Tella, R. Di, J. Dubra, and R. MacCulloch, "A Resource Belief-Curse? Oil and Individualism," *NBER Working Paper 14556*, 2008.

The American National Election Studies, *THE ANES GUIDE TO PUBLIC OPINION AND ELECTORAL BEHAVIOR*, University of Michigan, Center for Political Studies: Ann Arbor, MI.

Thornhill, R. and C. L. Fincher, *The Parasite-Stress Theory of Values and Sociality: Infectious Disease, History and Human Values Worldwide*, New York: Springer, 2014.

Triandis, H., "Collectivism v. Individualism: A Reconceptualisation of a Basic Concept in Cross-Cultural Social Psychology," in G. K. Verma and C. Bagley, eds., *Cross-Cultural Studies of Personality, Attitudes and Cognition*, London: Palgrave Macmillan UK, 1988, pp. 60–95.

_ , *Individualism & Collectivism*, Boulder, CO: Westview Press, 1995.

_ , "Individualism and Collectivism: Past, Present, and Future," 2001, pp. 35–50.

Tucker, G., "Progress of Population and Wealth in the United States, in Fifty Years, as Exhibited by the Decennial Census Taken in that Period," *The Merchants' Magazine and Commercial Review*, 1843, *9*, 136–44.

Turner, F. J., "The Significance of the Frontier in American History," *Proceedings of the State Historical Society of Wisconsin*, 1893.

Twenge, J. M., E. M. Abebe, and W. K. Campbell, "Fitting In or Standing Out: Trends in American Parents' Choices for Children's Names, 1880–2007," *Social Psychological and Personality Science*, 2010, *1* (1), 19–25.

Vandello, J. A. and D. Cohen, "Patterns of Individualism and Collectivism Across the United States," *Journal of Personality and Social Psychology*, 1999, *77* (2), 272–292.

Varnum, M E. W. and S. Kitayama, "What's in a name? Popular Names are Less Common on Frontiers," *Psychological Science*, 2011, *22* (2), 176–183.

Voigtländer, N. and H.J. Voth, "Persecution Perpetuated: the Medieval Origins of Anti-Semitic Violence in Nazi Germany," *The Quarterly Journal of Economics*, 2012, *127* (3), 1339–1392.

Willcox, W. F., "International Migrations, Volume I: Statistics," *National Bureau of Economic Research*, 1929.

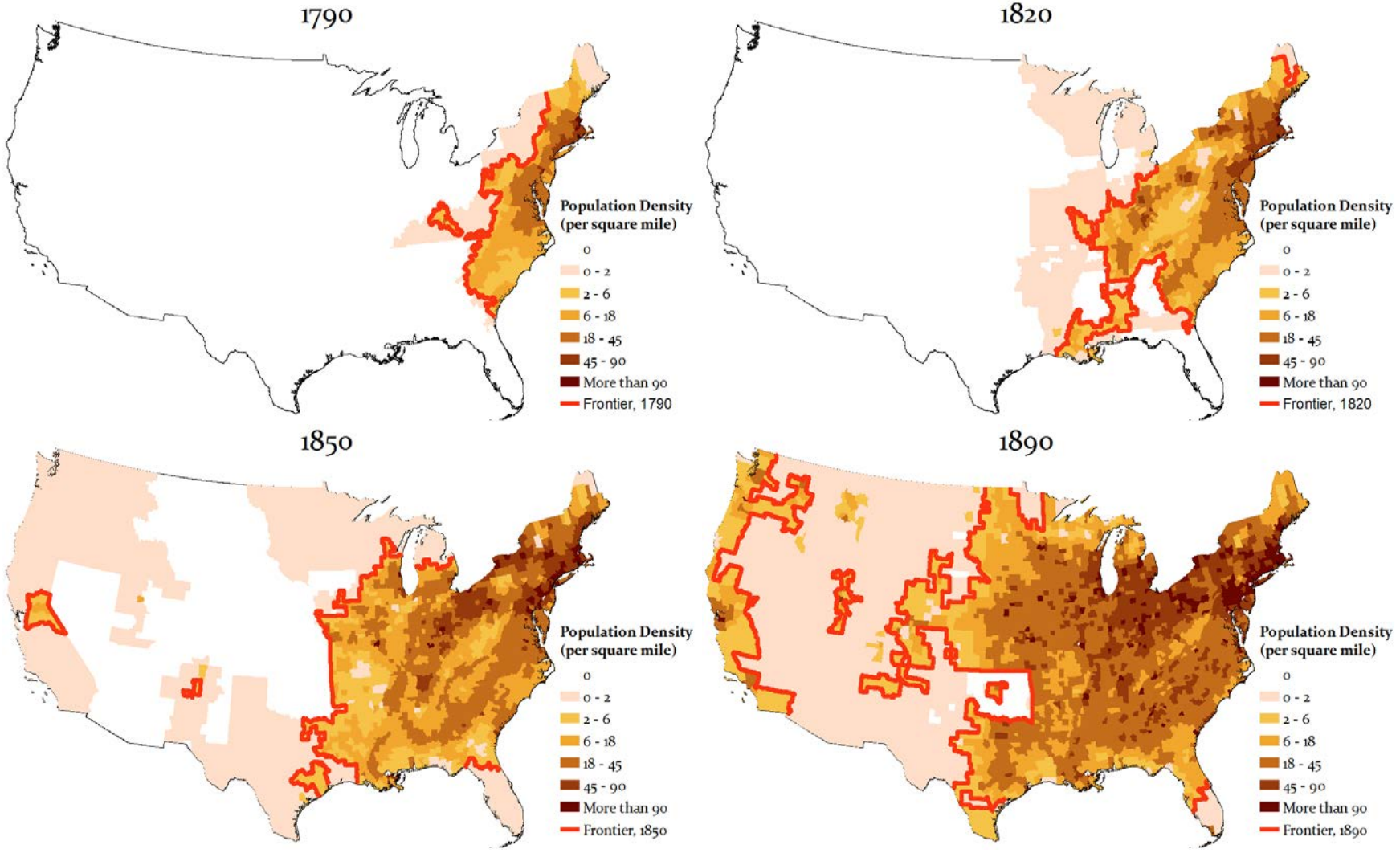
Yadin, Z. S., "Analyzing the Patient's First Name in the Search for Identity," *Contemporary Psychoanalysis*, 2016, *52* (4), 547–577.

Zelinsky, W., *The Cultural Geography of the United States*, Englewood Cliffs, N.J.: Pearson Education, Limited, 1973.

Zivot, E. and D. W. K. Andrews, "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis," *Journal of Business & Economic Statistics*, 2002, *20* (1), 25–44.

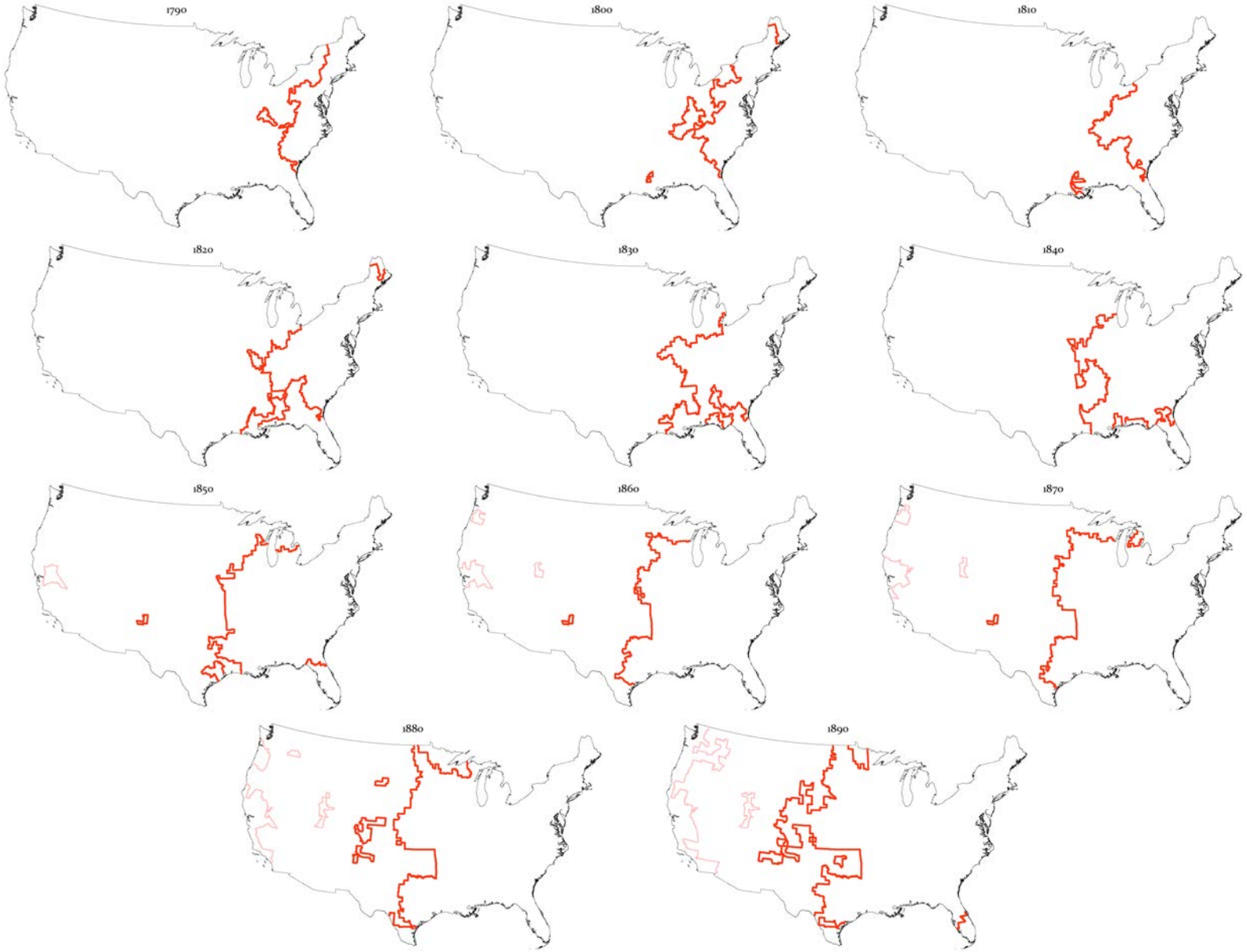
Figures

Figure 1: Population Density and the Frontier for Selected Years



Notes: Based on county-level data from National Historical Geographic Information System: Version 11.0. Population is allocated across years and counties based on the procedure described in Section 3.1, which builds upon Hornbeck (2010). The red frontier line is based on the algorithm described in Section 3.1 and Appendix A.

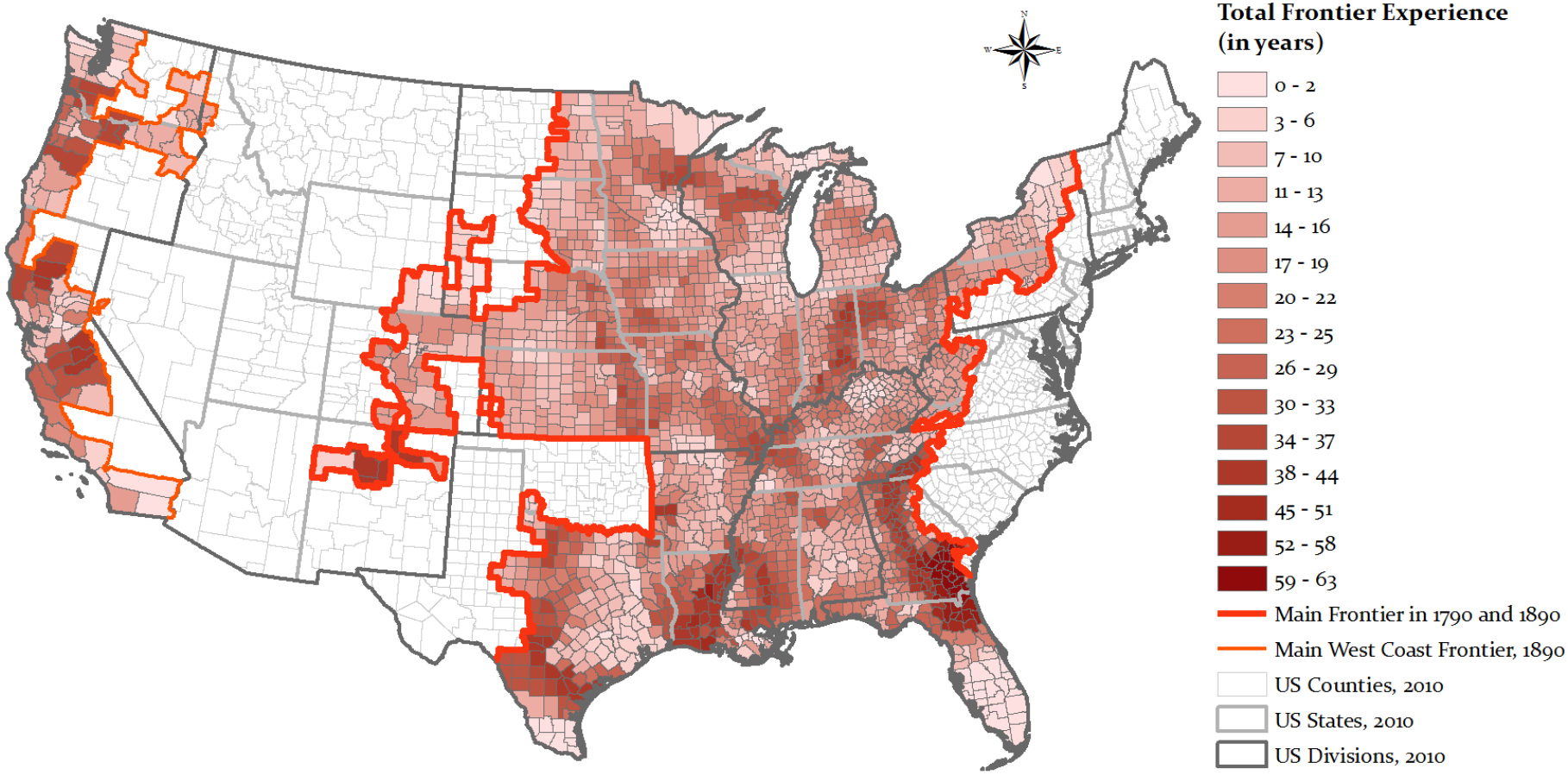
Figure 2: The Evolution of the Frontier, 1790 to 1890



Notes: Based on county-level data from National Historical Geographic Information System: Version 11.0. The frontier lines demarcate the contour of counties with population density below and above 2 people per square mile. The dark red lines correspond to the main frontier lines emerging form east-to-west expansions (our baseline analysis). The light red lines correspond to the frontiers resulting from west-to-east expansions from the West Coast, which we examine for robustness. In both cases, we exclude smaller "island frontiers" lines in the interior. Full details on the frontier line algorithm can be found in Appendix A.

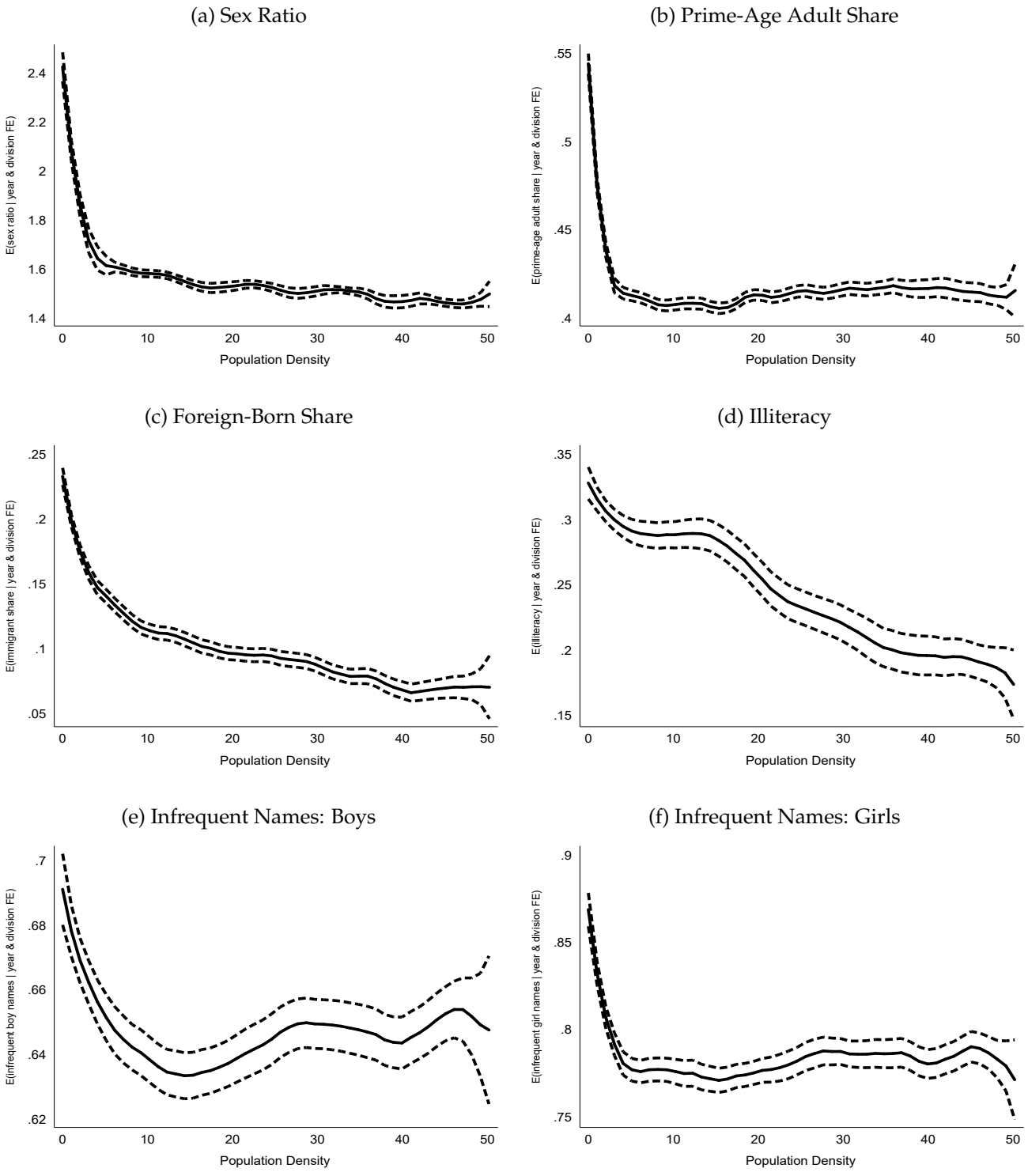
Figure 3: Total Frontier Experience, 1790 to 1890

(baseline sample lies between 1790 and 1890 main frontier lines, see Section 3.2 and notes below)



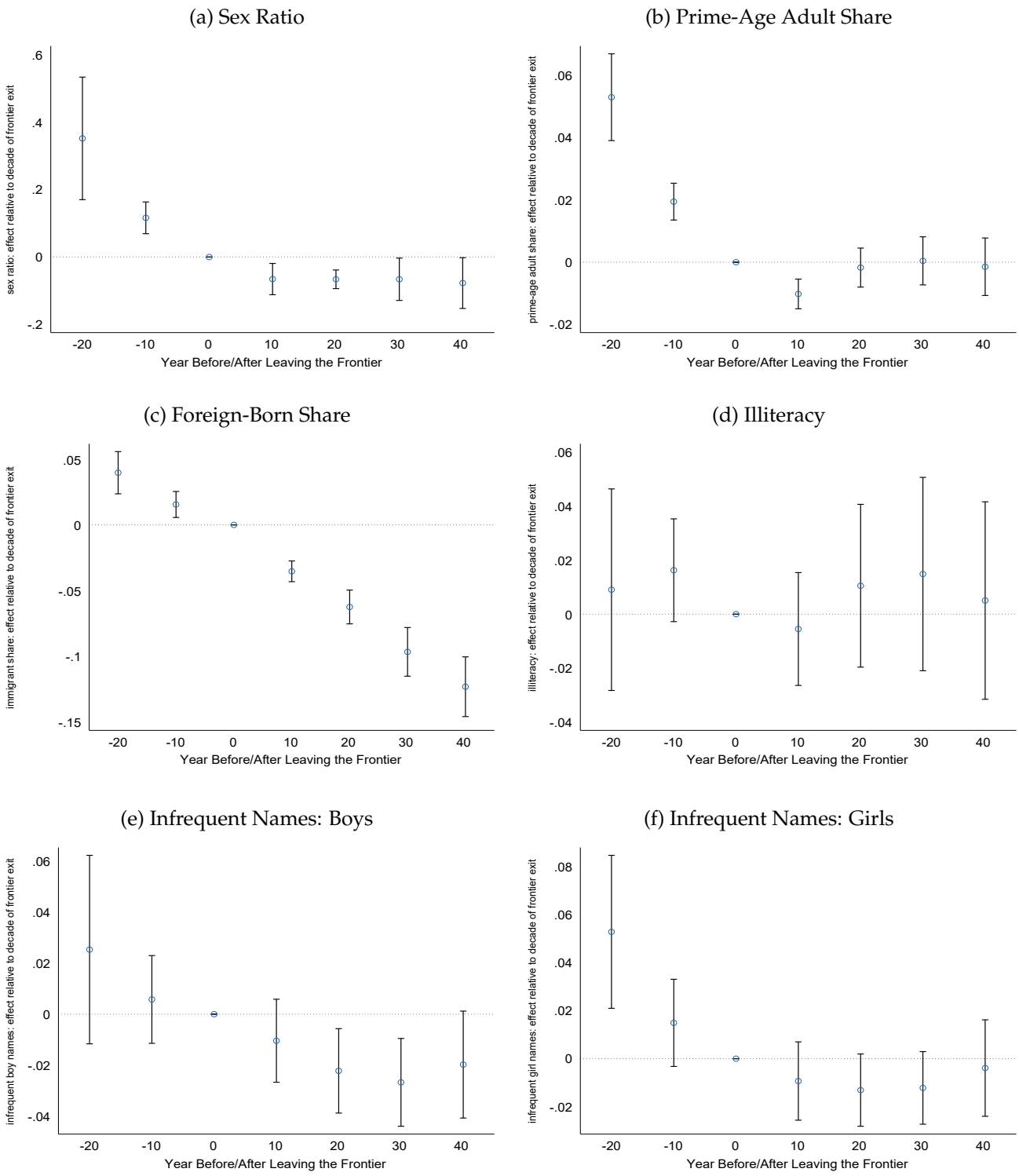
Notes: Based on county-level data from National Historical Geographic Information System: Version 11.0. Total frontier experience is the total number of years the county was within 100 km of the frontier line and its population density was below 6 people per square mile, between 1790–1890. The white areas to the east of the 1790 main frontier line are counties for which we do not know frontier history given the lack of Population Census data before 1790. The white areas to the west are beyond the 1890 frontier line and hence not included in our baseline sample, which is confined to the frontier era as defined by Porter et al. (1890) in the Census *Progress of the Nation* report. We include many of those counties to the west when extending the frontier era through 1950 for robustness.

Figure 4: Demographics and Individualism by Population Density, 1790 to 1890



Notes: These figures plot semiparametric estimates of equation (3) relating population density to demographic characteristics prominent in historical accounts of the frontier (a-d) and proxies for individualism (e-f). We estimate these curves $g(\cdot)$ based on the Robinson (1988) partially linear approach, pooling across all available years 1790–1890 for each county c. The specification includes Census division and year fixed effects, which are partialled out before estimating these shapes, and are based on an Epanechnikov kernel and rule-of-thumb bandwidth. The dashed lines are 95 percent confidence intervals. The estimates are recovered over all counties, but the figure zooms in on those with less than 50 people/mi^2 for presentational purposes. (a) *Sex Ratio* for whites is the ratio of the number of white males over white females. (b) *Prime-Age Adult Share* is the fraction of whites aged 15–49 over the total number of whites. (c) *Foreign-Born Share* is the ratio of foreign-born persons over total population. (d) *Illiteracy* is the illiteracy rate for whites aged 20 or older. (e) and (f) *Infrequent Names* are the share of boys and girls, respectively, with names outside of the top 10 most popular names in their Census division with the sample restricted to children aged 0–10 with native-born parents.

Figure 5: Demographics and Individualism Along the Transition out of the Frontier
Event Study Estimates with Respect to Year of Exiting the Frontier



Notes: This figure plots coefficients from the event study regressions in equation (4) for each of the outcomes in the semiparametric regressions presented in Figure 4. The decade-specific point estimates and 95 percent confidence intervals are each with reference to the county-specific decade of exiting the frontier. All regressions include division and year fixed effects. Standard errors are clustered using the grid cell approach of Bester, Conley and Hansen (2011) as described in Section 5.1.

Tables

Table 1: Demographics and Individualism on the Frontier

Dependent Variable:	Male/Female Ratio	Prime-Age Adult Share	Foreign-Born Share	Illiterate Share	Share of Infrequent Child Girl Names	Share of Infrequent Child Boy Names
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Baseline Frontier Definition: Low Density and Proximity to Frontier Line					
frontier county	0.144***	0.026***	0.060***	-0.007	0.022***	0.021***
	(0.019)	(0.004)	(0.008)	(0.011)	(0.007)	(0.007)
Mean Dep. Var. in Non-Frontier Counties	1.09	0.46	0.07	0.18	0.66	0.58
Number of County-Years	9,628	5,508	10,826	2,779	6,873	6,874
R^2	0.06	0.20	0.34	0.17	0.49	0.13
	Panel B: Distinguishing Low Density and Proximity to Frontier Line					
near frontier line	0.087***	0.022***	0.055***	-0.058***	0.018***	0.014*
	(0.013)	(0.003)	(0.009)	(0.013)	(0.006)	(0.007)
low population density	0.097***	0.005	0.032***	0.055***	0.005	0.010
	(0.011)	(0.003)	(0.008)	(0.011)	(0.006)	(0.007)
Mean Dep. Var. in Non-Frontier Counties	1.09	0.46	0.07	0.18	0.66	0.58
Number of County-Years	9,628	5,508	10,826	2,779	6,873	6,874
R^2	0.07	0.19	0.36	0.19	0.49	0.13

Notes: This table reports OLS estimates of equations (1) and (2) in Panels A and B, respectively. The dependent variables and sample are the same as in Figures 4 and 5. The sample size varies across columns depending on availability in the given Census round. All variables, except foreign-born share, are defined over the white population. Infrequent names capture the share of boys and girls, respectively, with names outside of the top 10 most popular names in their Census division with the sample restricted to children aged 0–10 with native-born parents. *Low population density* equals one if the county has density less than 6 people per square mile, and *near frontier line* equals one if the county is within 100 km of the frontier line in the given year. The sample excludes counties to the east of the 1790 frontier line and west of the main 1890 frontier line in keeping with our baseline long-run sample restrictions. All regressions include year and Census division FE. Standard errors are clustered using the grid cell approach of Bester, Conley and Hansen (2011) as described in Section 5.1.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 2: Total Frontier Experience and 20th Century Individualism

	Dep. Var.: Infrequent Names Among White Children Aged 0-10 with Native-Born Parents, 1940 Census			
	(1)	(2)	(3)	(4)
	Panel A: Boys with Infrequent Names (normalized share)			
total frontier experience	0.220***	0.134***	0.112***	0.112***
	(0.034)	(0.025)	(0.026)	(0.022)
Oster δ for $\beta = 0$		3.36	1.75	1.70
Number of Counties	2,036	2,036	2,036	2,036
R^2	0.06	0.48	0.54	0.61
	Panel B: Girls with Infrequent Names (normalized share)			
total frontier experience	0.202***	0.157***	0.161***	0.161***
	(0.033)	(0.028)	(0.030)	(0.024)
Oster δ for $\beta = 0$		5.12	3.35	3.42
Number of Counties	2,036	2,036	2,036	2,036
R^2	0.05	0.28	0.33	0.42
Division Fixed Effects	No	Yes	Yes	Yes
State Fixed Effects	No	No	Yes	Yes
Geographic/Agroclimatic Controls	No	No	No	Yes

Notes: This table reports estimates of equation (5) for our leading proxy of individualism in the 20th century, the share of boys and girls age 0–10 with infrequent names in the 1940 Census. The dependent variable is normalized so that the coefficient indicates the standard deviation effect of each additional decade of frontier exposure historically. This baseline sample is based only on counties inside the 1790–1890 east-to-west frontier. The baseline measure of infrequent names is given by the share of children with native-born parents in county c with a name that falls outside the top 10 names for children with native-born parents born in the same Census division within the given decade. Other measures of infrequent names are considered in Appendix Table B.2. Frontier experience is expressed in decades. Column 1 is the simple bivariate regression. Columns 2 and 3 add Census division and state fixed effects, respectively. Column 4 adds the following controls: county area; county centroid latitude and longitude; distance to oceans, lakes and rivers from county centroid; mean county temperature and rainfall; elevation; and average potential agricultural yield. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. The Oster (2016) tests in columns 2–4 are each with reference to the baseline specification in column 1 with no controls.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 3: Total Frontier Experience and Opposition to Government Intervention and Redistribution

Dependent Variable:	Prefers Cut Public Spending on Poor	Prefers Cut Public Spending on Welfare	Believes Gov't Should Redistribute	Prefers Reduce Debt by Spending Cuts	Index of Preferences for Spending Cuts	County Property Tax Rate, 2010
Scale:	binary	binary	normalized	binary	normalized	[0, 100]
Data Source:	ANES	CCES	GSS	CCES	GSS	ACS
	(1)	(2)	(3)	(4)	(5)	(6)
total frontier experience	0.010***	0.007**	-0.022*	0.014***	0.028**	-0.034***
	(0.004)	(0.003)	(0.012)	(0.002)	(0.011)	(0.007)
Oster δ for $\beta = 0$	5.59	6.86	5.79	2.40	2.28	1.67
Mean of Dependent Variable	0.09	0.40	0.00	0.41	-0.00	1.02
Number of Individuals	2,322	53,472	9,085	111,853	5,739	2,029
Number of Counties	95	1,863	255	1,963	253	2,029
R^2	0.04	0.04	0.06	0.04	0.07	0.82
Survey Wave Fixed Effects	Yes	Yes	Yes	Yes	Yes	–
Individual Demographic Controls	Yes	Yes	Yes	Yes	Yes	–
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of equation (5) for several measures capturing preferences for redistribution and state spending as well as actual property tax rates. Total frontier experience is expressed in decades. Full details on the outcomes can be found in Appendix C. We use all available survey rounds with the given outcome, and in all cases, we restrict to those counties in our baseline sample as described in the notes to Table 2. All columns are based on the specification in column 4 of Table 2 with additional individual-level controls for age, age squared, gender, and race in columns 1–5. The ANES measure in column 1 equals one if the respondent prefers that federal government spending on poor people be cut. The CCES measure in column 2 equals one if the respondent would prefer to cut public spending on welfare programs. The GSS measure in column 3 is a normalized measure of intensity of support on a 7 point scale of the statement that the government should reduce income differences in society through redistribution. The CCES question in column 4 equals one if the household would prefer that the state budget be balanced through spending cuts rather than tax increases. The GSS measure in column 5 is a normalized first principal component analysis (PCA) index based on a series of questions about whether the government spends too much on different public goods and transfer programs. The measure of county-level property tax rates in column 6 is estimated from American Community Survey data from 2010. Combining estimates from different columns and related outcomes across subsequent tables yields mean effects estimates based on the Kling, Liebman and Katz (2007) approach that retain statistical and economic significance as discussed in Section 5.4. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. The Oster (2016) tests are with reference to a baseline specification with no controls.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 4: Total Frontier Experience and Republican Vote Share

Dependent Variable:	Republican Vote Share in Recent Presidential Election							
	2000–16 (Avg.)	2000	2004	2008	2012	2016	Δ '16–'00	Δ '16–'12
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
total frontier experience	2.055***	1.215***	1.580***	1.979***	2.329***	3.171***	1.956***	0.842***
	(0.349)	(0.312)	(0.327)	(0.364)	(0.390)	(0.416)	(0.265)	(0.134)
Oster δ for $\beta = 0$	13.01	10.47	6.44	6.12	18.68	-24.08	-7.38	-3.65
Number of Counties	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036
Mean of Dependent Variable	60.0	56.6	60.3	57.4	60.6	65.4	8.9	4.9
R^2	0.33	0.32	0.33	0.38	0.37	0.32	0.35	0.33
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates of equation (5) for measures of the county-level Republican vote share in the last five presidential elections with data from the Leip Atlas. Total frontier experience is expressed in decades. Column 1 averages across all five elections. Columns 2–6 report year-specific effects. The sample and measure of frontier experience are as described in the notes to Table 2, and all estimates are based on the specification in column 4 from that table. Cross-equation tests reveal that the effect sizes are statistically different in each subsequent year and each year is statistically different from 2016. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. The Oster (2016) tests are with reference to a baseline specification with no controls.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 5: Total Frontier Experience and Preferences Over Contentious Policy Issues

Dependent Variable:	Opposes Affordable Care Act	Opposes Increasing Minimum Wage	Opposes Banning Assault Rifles	Opposes Regulation of CO_2 Emissions
	(1)	(2)	(3)	(4)
total frontier experience	0.022***	0.023***	0.015***	0.016***
	(0.004)	(0.008)	(0.004)	(0.004)
Oster δ for $\beta = 0$	2.96	3.05	2.46	2.22
Number of Individuals	29,446	5,134	29,404	29,215
Number of Counties	1,728	1,066	1,723	1,718
Mean of Dependent Variable	0.53	0.31	0.37	0.32
R^2	0.06	0.06	0.09	0.08
Survey Wave Fixed Effects	Yes	Yes	Yes	Yes
Individual Demographic Controls	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes

Notes: This table reports estimates of equation (5) for four measures of support for conservative issues that are particularly relevant to the frontier setting in historical accounts. Total frontier experience is measured in decades. The dependent variables are all binary indicators based on questions in the CCES across different years. The measure in Column 1 equals one if the individual in 2014 believes that the Affordable Care Act (ACA) should be repealed, in Column 2 equals one if the individual in 2007 opposes an increase in the minimum wage, in Column 3 equals one if the individual in 2014 opposes a ban on assault rifles, and in Column 4 equals one if the individual in 2014 opposes regulation of pollution by the Environmental Protection Agency (EPA). The set of specifications are otherwise the same as in Table 3; see the notes therein for details. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. The Oster (2016) tests are with reference to a baseline specification with no controls.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 6: Robustness (I): Adding West Coast, Extended Time Frame, Regional Heterogeneity

Frontier Time Frame:	Baseline (1790–1890)				Extended (1790–1950)			
Regional Sample Restriction:	Plus West Coast	Only Midwest	Only South	Only West	Extended Sample	Only Midwest	Only South	Only West
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Infrequent Name Share in 1940 (normalized), Boys							
total frontier experience	0.111***	0.236***	0.111***	0.113*	0.074***	0.125***	0.076***	0.056***
	(0.021)	(0.044)	(0.030)	(0.059)	(0.013)	(0.035)	(0.025)	(0.018)
Mean of Dependent Variable	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00
	Panel B: Infrequent Name Share in 1940 (normalized), Girls							
total frontier experience	0.156***	0.220***	0.162***	0.108	0.092***	0.108***	0.096***	0.083***
	(0.023)	(0.049)	(0.029)	(0.080)	(0.016)	(0.038)	(0.027)	(0.019)
Mean of Dependent Variable	-0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00
	Panel C: County Property Tax Rate in 2010							
total frontier experience	-0.031***	-0.051***	-0.027***	-0.006	-0.025***	-0.042***	-0.031***	-0.009**
	(0.006)	(0.014)	(0.007)	(0.013)	(0.004)	(0.012)	(0.006)	(0.004)
Mean of Dependent Variable	1.01	1.24	0.75	0.76	0.98	1.23	0.78	0.72
	Panel D: Average Republican Vote Share over 2000-2016							
total frontier experience	2.070***	1.882***	2.458***	1.459	1.302***	1.515***	1.429***	1.197***
	(0.332)	(0.414)	(0.396)	(0.890)	(0.256)	(0.350)	(0.422)	(0.274)
Mean of Dependent Variable	59.43	59.15	61.78	48.81	60.49	59.43	63.18	56.10
Number of Counties	2,141	987	936	152	2,500	1,038	1,074	322
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Focusing on four key outcomes across Panels A–D, this table extends our baseline sample of counties and examines region-by-region sample splits. Column 1 adds 105 counties along the secondary West Coast frontier (see Figure 3). Column 2 restricts to counties in the Midwest Census region, column 3 restricts to the South region, and column 4 restricts to the West, which includes the 105 counties added in column 1 plus 47 others in states in the West region but falling inside the 1890 main east-to-west frontier line. Column 5 expands the column 1 sample to include counties beyond the (main and secondary) 1890 frontier lines but inside the eventual frontier line realized by 1950. Columns 6–8 then proceed with the same region-by-region sample splits. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 7: Robustness (II): Population Density and Urbanization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population Density, 1890		✓						
Population Density Decile Within-State, 1890			✓					
Population Density, 1940/2010				✓				
Population Density Decile Within-State, 1940/2010					✓			
Sample Restriction	None	None	None	None	None	$> 90th$ percentile urban pop. share, 1940/2010	$\leq 90th$	None
	Panel A: Infrequent Boy Name Share in 1940 (normalized)							
total frontier experience	0.112***	0.103***	0.044*	0.101***	0.063***	0.099***	0.072***	0.069***
	(0.022)	(0.022)	(0.023)	(0.022)	(0.021)	(0.032)	(0.024)	(0.025)
total low density experience								0.067***
								(0.018)
Oster δ for $\beta = 0$	1.70	1.52	0.53	1.47	0.84	3.13	1.12	0.68
Number of Counties	2,036	2,036	2,021	2,036	2,021	242	1,794	2,036
Mean of Dependent Variable	-0.00	-0.00	-0.01	-0.00	-0.01	-0.00	-0.00	-0.00
R^2	0.61	0.62	0.67	0.62	0.67	0.85	0.60	0.62
	Panel B: Infrequent Girl Name Share in 1940 (normalized)							
total frontier experience	0.161***	0.154***	0.078***	0.151***	0.103***	0.134***	0.123***	0.090***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.023)	(0.039)	(0.026)	(0.028)
total low density experience								0.113***
								(0.022)
Oster δ for $\beta = 0$	3.42	3.09	1.20	3.02	1.82	2.49	2.95	1.03
Number of Counties	2,036	2,036	2,021	2,036	2,021	242	1,794	2,036
Mean of Dependent Variable	0.00	0.00	-0.01	0.00	-0.01	0.00	-0.00	0.00
R^2	0.42	0.43	0.50	0.43	0.50	0.73	0.40	0.44
	Panel C: County Property Tax Rate in 2010							
total frontier experience	-0.034***	-0.032***	-0.018***	-0.020***	-0.010*	-0.022*	-0.014***	-0.028***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.005)	(0.012)	(0.005)	(0.008)
total low density experience								-0.008
								(0.005)
Oster δ for $\beta = 0$	1.67	1.55	0.82	1.01	0.58	2.17	0.85	1.00
Number of Counties	2,029	2,029	2,014	2,029	2,014	223	1,806	2,029
Mean of Dependent Variable	1.02	1.02	1.02	1.02	1.02	1.34	0.98	1.02
R^2	0.82	0.82	0.84	0.85	0.87	0.90	0.86	0.82
	Panel D: Republican Vote Share, Average 2000–16							
total frontier experience	2.055***	1.905***	1.333***	1.517***	1.545***	1.886*	1.490***	1.255***
	(0.349)	(0.347)	(0.353)	(0.347)	(0.351)	(0.987)	(0.353)	(0.404)
total low density experience								1.256***
								(0.290)
Oster δ for $\beta = 0$	13.01	8.89	2.61	4.40	3.46	3.71	11.16	2.06
Number of Counties	2,036	2,036	2,021	2,036	2,021	223	1,813	2,036
Mean of Dependent Variable	60.04	60.04	60.11	60.04	60.11	49.73	61.30	60.04
R^2	0.33	0.35	0.39	0.40	0.37	0.29	0.38	0.35
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table disentangles the effects of TFE from the effects of historical and contemporary population density. Column 1 reproduces the baseline estimates for the four outcomes. Column 2 and 3 control for historical population density and within state population density deciles in 1890, the year the frontier closed according to the Census, respectively. Column 4 and 5 control for contemporary population density and within state population density deciles in 1940 or 2010 depending on the outcome, respectively. Columns 6 and 7 split the sample into counties above and below the 90th percentile of urban population shares in 2010. Column 8 controls for the total number of years that the country had population density less than 6 people/mi^2 from 1790–1890. This is one of the components of total frontier experience, the other being the total number of years that the county was within 100 km of the frontier line during that period. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. The Oster (2016) tests are with reference to a baseline specification with no controls.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 8: Robustness (III): Additional Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Baseline controlling for...														
ruggedness		✓												✓
rainfall risk			✓											✓
distance to nearest portage site				✓										✓
distance to nearest mine					✓									✓
distance to nearest Indian battle						✓								✓
slave population share, 1860							✓							✓
sex ratio, 1890								✓						✓
immigrant share, 1890									✓					✓
Scottish and Irish immigrant share, 1890										✓				✓
birthplace diversity, 1890											✓			✓
years connected to railroad by 1890												✓		✓
manufacturing employment share, 1890													✓	✓
	Panel A: Infrequent Boy Name Share in 1940 (normalized)													
total frontier experience	0.112***	0.113***	0.109***	0.113***	0.112***	0.114***	0.080***	0.112***	0.105***	0.079***	0.099***	0.072***	0.120***	0.065***
	(0.022)	(0.021)	(0.022)	(0.021)	(0.021)	(0.022)	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.023)	(0.021)	(0.021)
Oster δ for $\beta = 0$	1.70	1.75	1.65	1.71	1.71	1.70	1.01	1.71	1.54	1.07	1.42	0.92	1.91	0.70
Number of Counties	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036
Mean of Dependent Variable	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
R^2	0.61	0.63	0.61	0.61	0.61	0.61	0.63	0.61	0.61	0.67	0.62	0.65	0.66	0.74
	Panel B: Infrequent Girl Name Share in 1940 (normalized)													
total frontier experience	0.161***	0.164***	0.159***	0.167***	0.162***	0.166***	0.125***	0.162***	0.162***	0.133***	0.154***	0.112***	0.170***	0.123***
	(0.024)	(0.023)	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)	(0.024)	(0.023)	(0.023)	(0.023)	(0.025)	(0.024)	(0.022)
Oster δ for $\beta = 0$	3.42	3.56	3.28	3.56	3.44	3.45	2.05	3.44	3.32	2.44	3.06	1.84	3.87	1.78
Number of Counties	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036
Mean of Dependent Variable	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R^2	0.42	0.45	0.42	0.43	0.43	0.42	0.44	0.43	0.42	0.47	0.43	0.48	0.47	0.58
	Panel C: County Property Tax Rate in 2010													
total frontier experience	-0.034***	-0.034***	-0.034***	-0.033***	-0.034***	-0.033***	-0.033***	-0.034***	-0.028***	-0.026***	-0.027***	-0.023***	-0.036***	-0.023***
	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Oster δ for $\beta = 0$	1.67	1.72	1.71	1.58	1.69	1.60	1.52	1.67	1.36	1.32	1.33	1.10	1.93	1.07
Number of Counties	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029
Mean of Dependent Variable	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
R^2	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.83	0.83	0.83	0.83	0.83	0.85
	Panel D: Republican Vote Share, Average 2000–16													
total frontier experience	2.055***	2.050***	2.115***	2.095***	2.055***	2.172***	1.399***	2.060***	1.715***	1.717***	1.689***	1.640***	2.137***	0.931***
	(0.349)	(0.349)	(0.338)	(0.344)	(0.350)	(0.351)	(0.361)	(0.347)	(0.328)	(0.340)	(0.327)	(0.361)	(0.350)	(0.316)
Oster δ for $\beta = 0$	13.01	12.84	15.11	13.78	13.02	16.49	3.45	13.25	6.17	6.19	5.92	5.14	17.16	1.57
Number of Counties	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036
Mean of Dependent Variable	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04
R^2	0.33	0.33	0.34	0.33	0.33	0.34	0.38	0.34	0.38	0.38	0.39	0.36	0.36	0.49
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table augments the baseline specification, reproduced in column 1, with a set of controls aimed at clarifying potentially confounding channels by which TFE affects four key outcomes. The variables are defined in Section 5.5 and at the end of Appendix C, but we note here that the measure in column 5 is based on the known mining sites pre-1890. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. The Oster (2016) tests are with reference to a baseline specification with no controls.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 9: Racial Differences in the Long-Run Effects of Frontier Experience

Dependent Variable:	Prefers Cut Public Spending on Welfare (1)	Prefers Reduce Debt by Spending Cuts (2)	Opposes Affordable Care Act (3)	Opposes Increasing Minimum Wage (4)	Opposes Banning Assault Rifles (5)	Opposes Regulation of CO_2 Emissions (6)
total frontier experience × white	0.009***	0.016***	0.027***	0.025***	0.018***	0.018***
	(0.003)	(0.002)	(0.004)	(0.008)	(0.004)	(0.004)
total frontier experience × black	-0.008	-0.002	-0.000	0.005	-0.002	-0.008
	(0.006)	(0.005)	(0.007)	(0.014)	(0.008)	(0.007)
total frontier experience × other	0.010	0.015**	0.014	0.009	0.019**	0.029***
	(0.007)	(0.007)	(0.009)	(0.019)	(0.008)	(0.008)
white	0.044***	0.065***	0.047***	-0.064**	0.006	0.046***
	(0.012)	(0.011)	(0.015)	(0.032)	(0.012)	(0.015)
black	-0.177***	-0.065***	-0.215***	-0.285***	-0.148***	-0.067***
	(0.014)	(0.011)	(0.021)	(0.038)	(0.021)	(0.025)
Number of Individuals	53,472	111,853	29,446	5,134	29,404	29,215
Number of Counties	1,863	1,963	1,728	1,066	1,723	1,718
TFE(black)=TFE(white), p-value	0.010	0.000	0.000	0.125	0.022	0.000
Mean of Dependent Variable, Whites	0.43	0.44	0.58	0.32	0.39	0.35
Share White Respondents	0.79	0.77	0.76	0.86	0.76	0.76
Share Black Respondents	0.11	0.13	0.11	0.09	0.11	0.11
Share Other Respondents	0.10	0.11	0.13	0.05	0.13	0.13
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table allows the effects of TFE to vary by (self-identified) race of respondents for the six CCES outcomes used in Tables 3 and 5. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 10: Instrumental Variables (IV) Estimates for Summary Outcomes

Dependent Variable:	Infrequent Name Share normalized		County Property	Republican Vote Share
	Boys	Girls	Tax Rate	Avg.
	1940	1940	2010	2000–16
	(1)	(2)	(3)	(4)
	Panel A: IV = Log Average Actual National Migration Inflows Over 30 Years			
total frontier experience	0.207***	0.207***	-0.045***	3.407***
	(0.042)	(0.041)	(0.014)	(0.585)
Number of Counties	2,036	2,036	2,029	2,036
Mean of Dependent Variable	-0.00	0.00	1.02	60.04
First Stage F Statistic	193.64	193.64	194.13	193.64
	Panel B: IV = Log Average Predicted National Migration Inflows Over 30 Years			
total frontier experience	0.239***	0.235***	-0.049***	3.177***
	(0.047)	(0.048)	(0.014)	(0.624)
Number of Counties	2,036	2,036	2,029	2,036
Mean of Dependent Variable	-0.00	0.00	1.02	60.04
First Stage F Statistic	195.84	195.84	196.31	195.84
State Fixed Effects	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes

Notes: This table reports instrumental variables estimates of equation (5) based on the instruments described in Section 5.6. We again report results for the four summary outcomes examined in prior tables, and total frontier experience is measured in decades. Panel A reports the IV estimates for the baseline sample and specification using the log of the average national annual actual migration inflows over the 30 years since the frontier is within 110km from the county centroid. Panel B reports the estimates using the IV constructed based on annual migration inflows to the US predicted by weather shocks in Europe. The details on the construction of both instrumental variables are presented in the Appendix Section B.9. The first-stage F statistics are cluster-robust, and standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 11: Frontier Individualism and Selective Migration

	Dependent Variable: Child Has Infrequent Name in 1880			
	(1)	(2)	(3)	(4)
omitted reference group:	non-frontier resident, 1870–80	non-frontier resident, 1880	non-frontier resident, 1880	frontier immigrant, 1870–80
frontier county resident in 1880, arrived between 1870 and 1880	0.042***		0.055***	
	(0.012)		(0.011)	
frontier county resident in 1880		0.075***		
		(0.018)		
frontier county resident in 1880, arrived before 1870			0.186***	0.118***
			(0.035)	(0.026)
mean infrequent name, omitted group:	0.707	0.708	0.708	0.767
Number of Individuals	1,223,600	1,239,513	1,239,513	12,630
R^2	0.02	0.02	0.02	0.05
Gender×Age Fixed Effects	Yes	Yes	Yes	Yes
Birth Order Fixed Effects	Yes	Yes	Yes	Yes

Notes: This table reports estimates of equation (6) based on the linked historical Census data from 1870 to 1880 for households with white, native-born fathers age 30–50 and children aged 0–10 in 1880. This linked sample is detailed in footnote 45 in the paper. The dependent variable is an indicator equal to one if the child is given a name that falls outside the top 10 most popular names nationally in the 1870s. The top of each column reports the omitted reference group and the mean infrequent name share among them. We define immigrant status here based on whether the father switched counties between 1870 and 1880. Frontier counties are as defined in 1870 and 1880 based on the main east-to-west frontier line. Column 1 reports the selective migration differential between migrants from non-frontier to frontier counties and those that remained in non-frontier counties in both 1870 and 1880. Column 2 reports the overall differential in infrequent names between frontier and non-frontier counties in 1880, i.e., inclusive of stayers in frontier counties. Column 3 breaks down the overall differential into the component due to migrants between 1870 and 1880 and those that resided in the frontier county prior to 1870 (either by birth or earlier migration). Column 4 then restricts to frontier county residents, identifying the differential between recent immigrants and longer-term residents. In addition to gender×age and birth order fixed effects, all regressions control for indicators for whether the mother and father have infrequent names. Standard errors are clustered by county in 1870.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table 12: Individualism, Socioeconomic Success, and Endurance on the Frontier

	(1)	(2)	(3)	(4)
	Panel A: Dep. Var.: Father's Economic Status in 1880 (normalized)			
	sei			*occscore*
at least one child 0-10 has infrequent name	0.077***		0.080***	0.065***
	(0.007)		(0.007)	(0.007)
at least one child 0-10 has infrequent name × frontier county	0.077*		0.073*	0.066
	(0.042)		(0.042)	(0.042)
father has infrequent name		-0.047***	-0.049***	-0.044***
		(0.005)	(0.005)	(0.005)
father has infrequent name × frontier county		0.072**	0.069**	0.075**
		(0.033)	(0.033)	(0.033)
Number of Individuals	264,038	264,038	264,038	264,038
R^2	0.09	0.09	0.09	0.10
County Fixed Effects	Yes	Yes	Yes	Yes
	Panel B: Dep. Var.: Migrated from Frontier County in 1870 to Non-Frontier County in 1880			
father has infrequent name	-0.042***			-0.041***
	(0.009)			(0.009)
mother has infrequent name		-0.010		-0.009
		(0.007)		(0.007)
at least one child 0-10 has infrequent name			-0.026**	-0.023**
			(0.011)	(0.011)
Number of Individuals	27,066	27,066	27,066	27,066
Mean of Dependent Variable	0.410	0.410	0.410	0.410
R^2	0.08	0.07	0.07	0.08
Origin County Fixed Effects	Yes	Yes	Yes	Yes

Notes: This table reports estimates of equation (7) in Panel A and equation (8) in Panel B based on the same linked sample of households from the 1870 and 1880 Census described in the notes to the previous table and at length in Section 6.1. Infrequent name measures are as defined elsewhere and based on the top 10 nationally for all family members. The frontier dummy in both panels is as defined earlier. In Panel A, the dependent variable in columns 1–3 is the normalized Duncan (1961) socioeconomic index (*sei*) and in column 4 is the normalized occupational score (*occscore*), both as observed in 1880 and as provided by the North Atlantic Population Project (NAPP). The sample in Panel A includes all white native-born male household heads (fathers) aged 30–50 that did not migrate across counties between 1870 and 1880. Standard errors are clustered at the county level. The sample in Panel B is restricted to all white native-born households aged 30–50 residing in frontier counties in 1870, and the dependent variable equals one if the household moved to a non-frontier county by 1880. Standard errors are clustered at the origin county. All regressions include dummies for the number of children born in the 1870s.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Appendix

Contents

A Retracing the Frontier

This section provides a step-by-step description of how we construct the frontier lines for each year between 1790–1890.

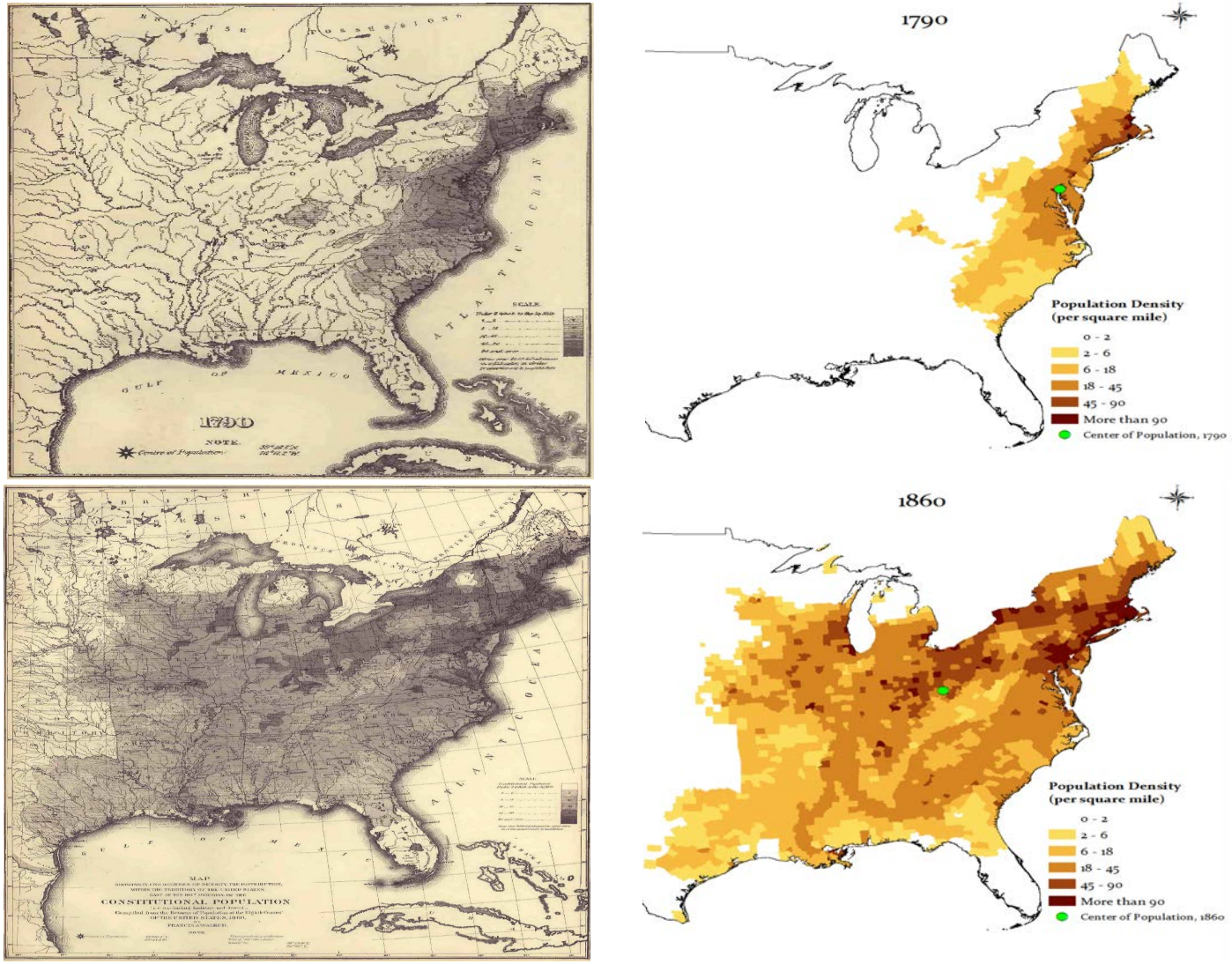
1. Calculate county level population density per square mile for each year in 1790–1890 using the 2010 county boundaries. First, we harmonize the county-level population data from each year to the 2010 county boundaries using the procedure discussed in Section 3. For intercensal years, we interpolate county-level population by assuming a constant annual population growth rate that matches the decadal growth rate (replacing initial zeros with 0.01 to avoid infinite growth rates). Then, using the 2010 county boundaries shapefile, we calculate the county-level population density as the ratio of population over county area in square miles.

2. Draw a contour line at population density equal to 2 people per square mile for each year. We use ArcGIS and the 2010 county boundaries. First, for each year, we convert the polygon containing the county level population density data into a raster file using ***PolygonToRaster*** tool and set population density for the given year as the ***"value field"*** for the conversion. Then, using the ***ContourList*** tool, select the raster file created in the preceding step as an input and set the ***"contour value"*** to "2" to create contour lines at population density equal to 2. The resulting lines delineate the counties that have a population density below 2 people per square mile from those counties that have a population density above 2.

3. Clean the contour lines to retain only the significant frontier lines. With the purpose of capturing historical notions of the frontier as "margins of civilization," we discard all contour line segments less than 500 km, as well as discard isolated pockets of relatively sparse populations within the main area of settled territory. These isolated pockets are the "inner islands" formed by counties with population density below 2 people per square mile surrounded by counties with population density above 2 people per square mile. A second set of frontier lines emerge in the West Coast in mid-19th century. This process of settlement was marked by the Gold Rush and different historical forces than the main east-to-west expansion, so for our baseline analysis we focus on the territory spanned by east-to-west expansion. We do this by keeping only those frontier lines that are east of the westernmost east-to-west frontier line in 1890. In the robustness analysis, we add the West Coast to our baseline sample.

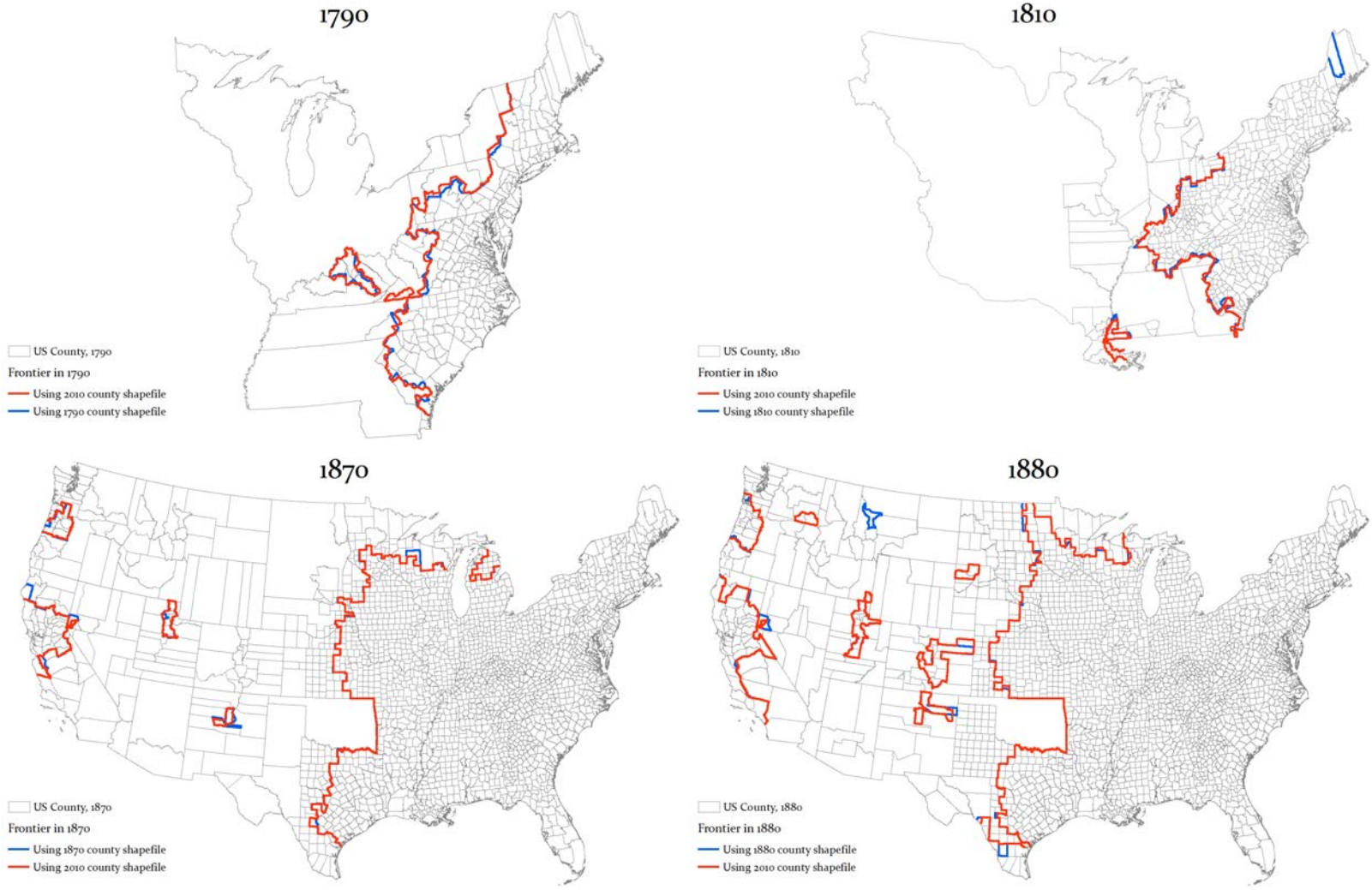
We select line segments based on length and location (e.g., X centroid of the line midpoint) in ArcGIS using the ***SelectLayerByAttribute*** tool, and apply ***CopyFeatures*** to keep only the selected lines. In the detailed robustness checks in Section 5.5, we also consider various alternatives to the frontier definition such as changing the line cutoffs, restricting to single westernmost frontier line, including the "inner island" lines, and considering the frontier lines that emerge from the West Coast.

Figure A.1: Population Density Maps from the 1890 Census Report and Our Maps (1790 and 1860)



Notes: This figure compares our estimates on the right of population density in 1790 and 1860 at the (2010) county level using the harmonization procedure described in Section 3 to the historical estimates in on the left based on the the noteworthy *Progress of the Nation* Census report on which Turner based his thesis.

Figure A.2: Frontier Lines Using Contemporaneous vs 2010 County Boundaries for selected years



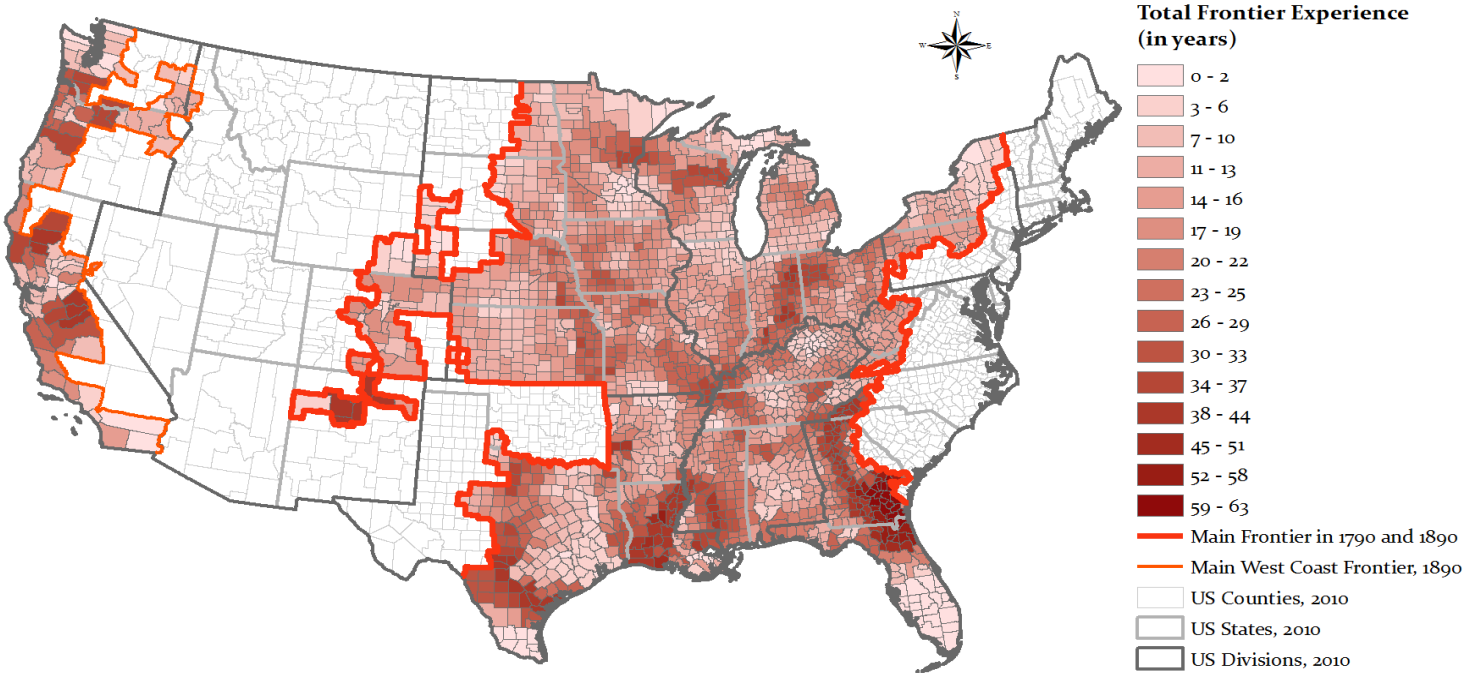
Notes: Based on county level Population Census data from 1790-1880 and NHGIS county shapefiles. The figures provide the county boundaries for selected years and the frontier lines for the corresponding years drawn using the contemporaneous county boundaries as well as the 2010 county boundary. The frontier lines delineate the counties that had population density of two persons or higher. The frontier lines in blue are drawn using the contemporaneous county boundaries where as the frontier lines in red are drawn using the 2010 county boundaries (after the data harmonization discussed in Section 3.1).

B Additional Results and Robustness Checks

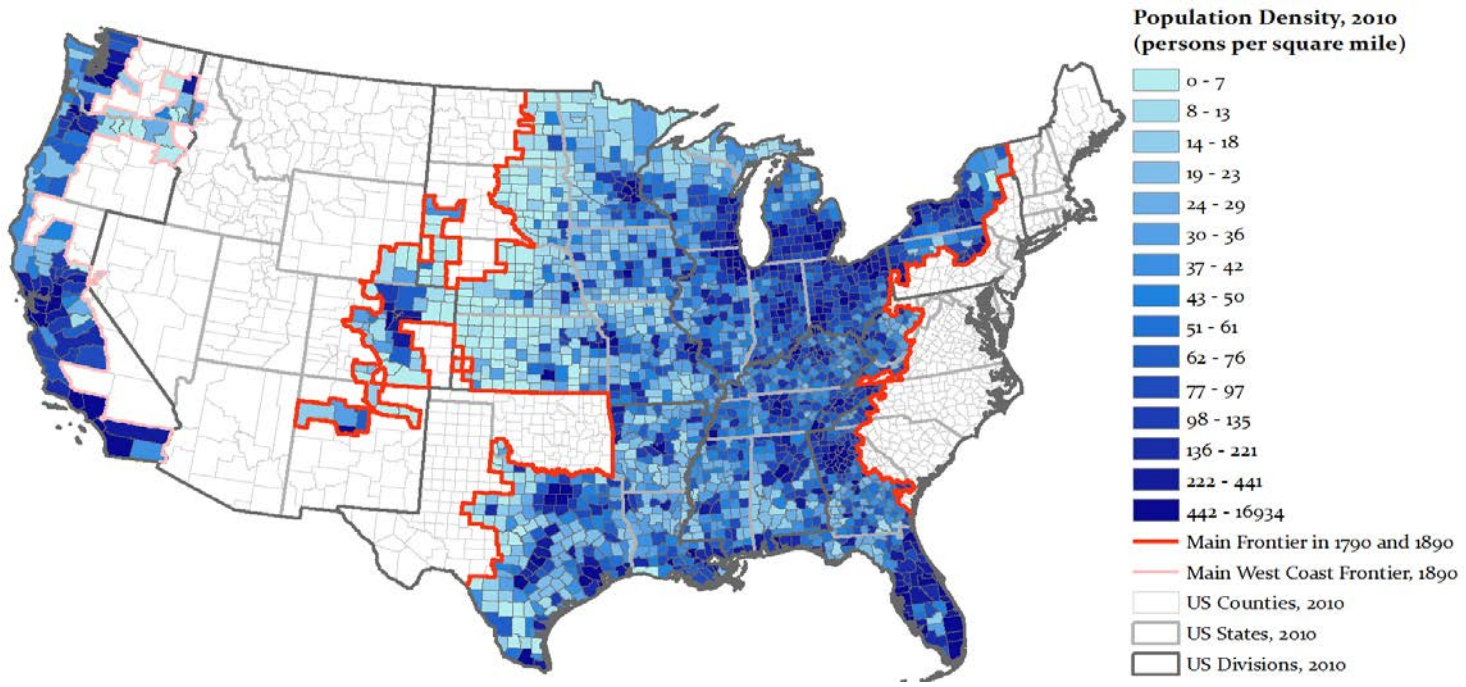
B.1 Comparing Total Frontier Experience and Current Population Density

Figure B.1: TFE is Distinct from Current Population Density

(a) Total Frontier Experience, 1790 to 1890



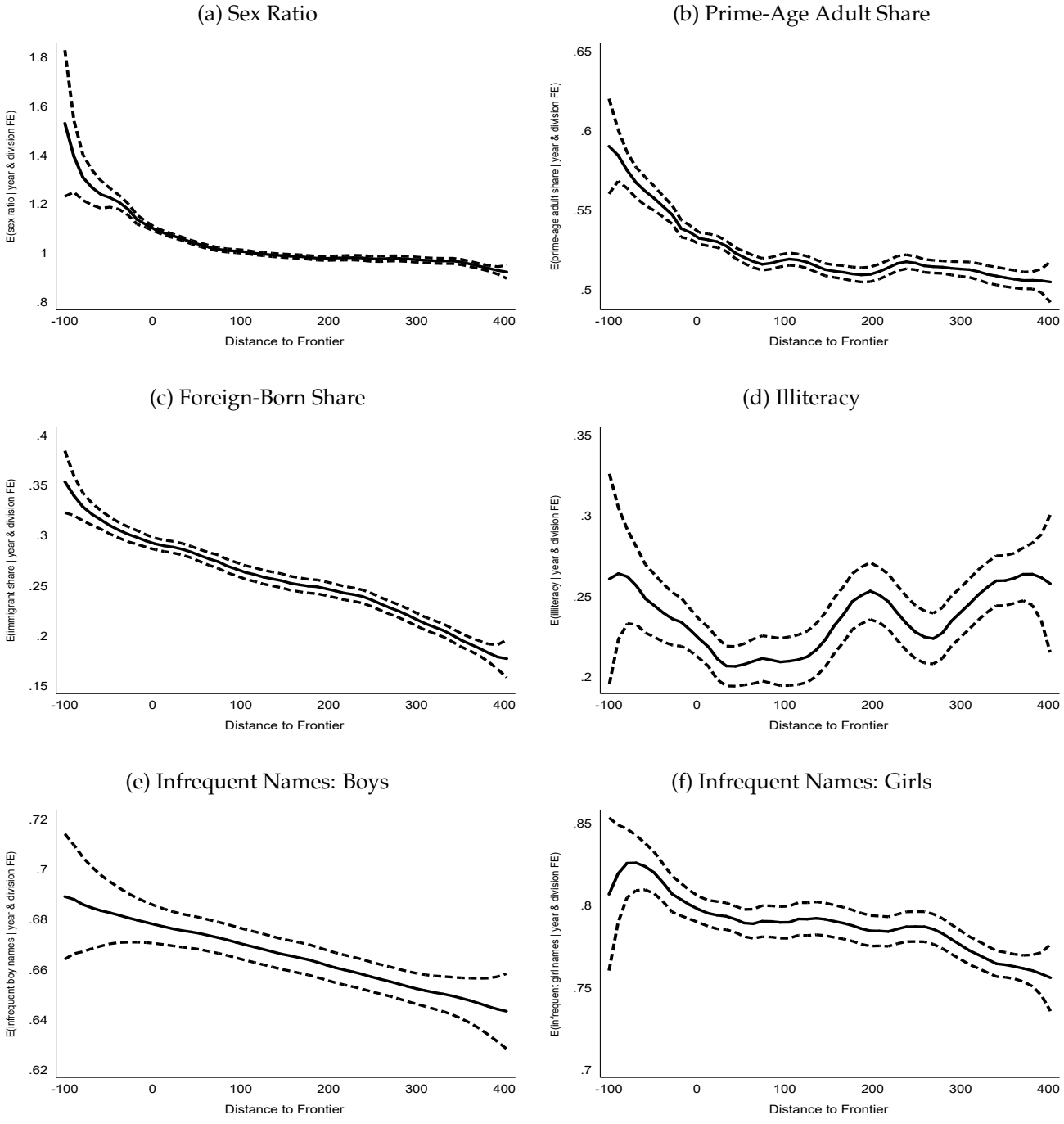
(b) Population Density, 2010



Notes: Panel (a) reproduces Figure 3, and (b) presents a similarly scaled map of population density in 2010 for the same counties.

B.2 Further Background Characterizing Frontier Life

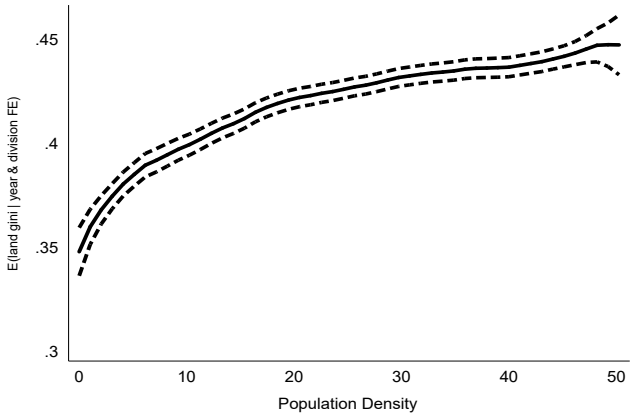
Figure B.2: Distribution of Demographics and Individualism by Distance to the Frontier



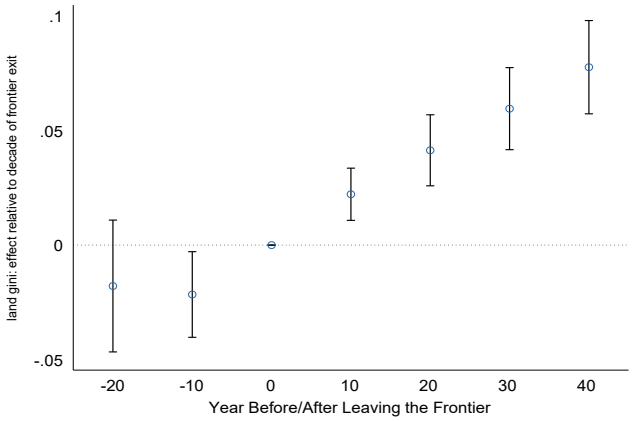
Notes: Distance to the frontier, measured in kilometers, is the distance from the county's centroid to the nearest frontier line. The distance is negative if the county centroid is to the west of the nearest main frontier line. Figures (a)-(f) provide the semiparametric estimates of the corresponding dependent variables, with 95 percent confidence intervals, as a function of distance to the frontier estimated using county-level pooled data and applying a nonlinear function recovered using the partially linear Robinson (1988) estimator. The specification includes Census division and year fixed effects and are based on an Epanechnikov kernel and rule-of-thumb bandwidth.

Figure B.2: Inequality is Lower on the Frontier

(a) Semiparametric



(b) Event Study



Notes: Based on county level data from National Historical Geographic Information System: Version 11.0 Database from 1790-1890. Land inequality is measured using the county level gini coefficient based on the number of farms in seven bins of farm size. The semiparametric specification in (a) is the same as in Figure 4, and the event study specification in (b) is the same as in Figure 5. See the notes therein for details.

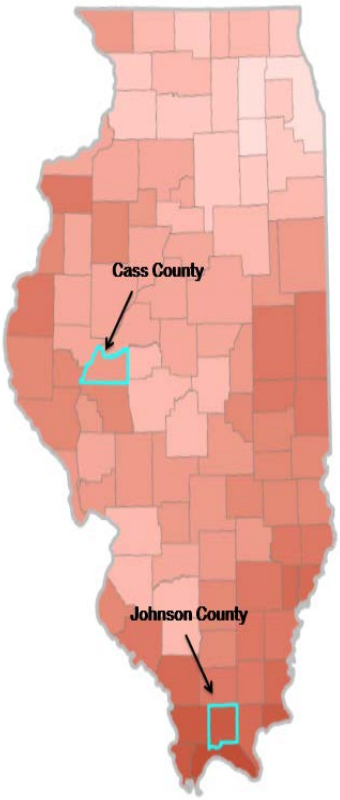
Table B.1: Occupational Composition in Frontier and Non-Frontier Counties

	Employment Share
Frontier Counties	
Farmers (owners and tenants)	.606
Laborers (n.e.c.)	.125
Farm laborers, wage workers	.047
Managers, officials, and proprietors (n.e.c.)	.040
Carpenters	.023
Truck and tractor drivers	.014
Blacksmiths	.013
Operative and kindred workers (n.e.c.)	.012
Other non-occupational response	.012
Lawyers and judges	.009
Non-Frontier Counties	
Farmers (owners and tenants)	.534
Managers, officials, and proprietors (n.e.c.)	.063
Laborers (n.e.c.)	.061
Operative and kindred workers (n.e.c.)	.0465
Carpenters	.037
Farm laborers, wage workers	.028
Salesmen and sales clerks (n.e.c.)	.015
Blacksmiths	.014
Other non-occupational response	.012
Physicians and surgeons	.010

Notes: This table reports the top 10 occupational shares in frontier and non-frontier counties in 1870 using the 1870–1880 linked sample that we use in our main analysis in Section 6.

B.3 Case Study Illustrating Long-Run Effects

To fix ideas, consider the two counties of Cass and Johnson mentioned in Section 3.2 and seen in the TFE map on the right, which is a snapshot of Illinois from Figure 3. Both are roughly equidistant from the Mississippi River and the important historical city of St. Louis. Today, the two rural counties look very similar. Cass has 36.3 people/mi^2, median income is around US$ 41,544, and 86 percent of the population is white. Johnson also has 36.6 people/mi^2, median income around US$ 41,619, and 89 percent white population. These two counties had very similar population density in 1890 as well. However, they differ significantly in their total frontier experience historically. Cass was on the frontier for 10 years, and Johnson for 32 years. This difference may be explained by any number of factors shaping the westward movement of the frontier through this area of the midwest in the early 1800s as seen in Figure 2. One potentially important contributor lies in our instrumental variable. Johnson entered the frontier in 1803 whereas Cass entered in 1818. While only 15 years apart, this implied a considerable difference in exposure to subsequent immigration-induced pressure on the westward expansion of the frontier over the next few decades as evidenced in Figures B.3, B.4, and especially B.5 below.



These historical differences in TFE translate into substantial long-run differences in the prevalence of rugged individualism in local culture. In Cass, 75 (64) percent of girls (boys) have infrequent names in 1940, Republican presidential candidates captured 55 percent of the vote in the average election since 2000, and local property tax rates are around 1.9 percent in 2010. Meanwhile, in Johnson, 78 (71) percent of girls (boys) have infrequent names in 1940, 68 percent average Republican vote shares since 2000, and 1.3 percent local property tax rates in 2010. This is striking insomuch as the two counties have such similar contemporary population density, median income, and racial composition.

B.4 Further Robustness Checks on Individualism

Tables B.2 and B.3 show the robustness of the OLS and IV results, respectively, for infrequent names to alternative measures of infrequency and restrictions on (grand)parental ancestry. Table B.4 shows that the baseline OLS results for infrequent names look similar in each decade before 1940 but after the official closing of the frontier. Table B.5 validates the long-run relationship of TFE with individualism using an alternative survey-based proxy from 1990 ANES data.

Power Law Property of Names. One important dimension of robustness that we corroborate in Tables B.2 and B.3 is that the results are not sensitive to the cutoff for defining infrequent names (10, 25, 100, ...). This is likely due to the fact that naming frequencies in the United States follow a power law (see Hahn and Bentley, 2003; Gureckis and Goldstone, 2009), and hence the share of people with each of the top 10 names can be characterized by the same shape parameter as the share with the top 25 names, and so on. Adopting this parametrization, the literature on names has documented an increase in the national power law exponent over time, which suggests a growing trend towards less concentration among popular names. Our results here identify mid-20th century differences in name (in)frequency across counties.

Auxiliary Measures of Individualism. Beyond infrequent names, we draw upon a well-suited measure from the ANES data to provide further evidence of the link between TFE and high levels of individualism. Specifically, we use the 1990 ANES round in which respondents were asked whether (1) "it is more important to be a cooperative person who works well with others", or (2) "it is more important to be a self-reliant person able to take care of oneself." While this question was designed explicitly for studies of American individualism (see Markus, 2001), unfortunately, it was only asked in a single round.

Table B.5 provides evidence that self-reliant preferences are stronger today in counties with longer exposure to the frontier historically. Around 55 percent of individuals respond in support of the cooperative answer. However, across different specifications, each decade of additional TFE is associated with around 2–6 percentage points lower support for cooperation over self-reliance.[1] While the results with the full set of controls are noisy, we nevertheless view these findings as at least suggestive of longstanding claims about the rugged individualism pervasive on the frontier. In linking to results elsewhere in the paper, it is worth noting that individuals that identify as Republican in the ANES data are around 15–20 percent more likely to believe that it is better to be a self-reliant than a cooperative person.

[1] In a related result using the CCES, we find that residents in counties with greater TFE are significantly less likely to have ever belonged to a union. While this result may be explained in part by differences in sectoral composition, it is also consistent with weaker collectivist tendencies.

Table B.2: Robustness to Other Measures of Infrequent Names (OLS Estimates)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Sample: White Children Aged 0-10, 1940 Census									
Further Sample Restriction:	None	Native Father	Native Parents	Native Grand-parents	None	None	None	None	None	None
Infrequent Measure:	Top 10 Division				Top 10 National	Top 10 State	Top 10 County	Top 25 Division	Top 100 Division	Top 10 National Non-Biblical
	Panel A: Share of Boys with Infrequent Names									
total frontier experience	0.109***	0.113***	0.112***	0.089***	0.123***	0.121***	0.098***	0.098***	0.115***	0.100***
	(0.022)	(0.022)	(0.022)	(0.024)	(0.020)	(0.022)	(0.034)	(0.023)	(0.022)	(0.017)
Number of Counties	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036
R^2	0.63	0.62	0.61	0.25	0.63	0.61	0.20	0.64	0.67	0.73
	Panel B: Share of Girls with Infrequent Names									
total frontier experience	0.167***	0.162***	0.161***	0.085***	0.166***	0.175***	0.047	0.148***	0.157***	0.155***
	(0.025)	(0.024)	(0.024)	(0.023)	(0.023)	(0.023)	(0.034)	(0.020)	(0.020)	(0.021)
Number of Counties	2,036	2,036	2,036	2,035	2,036	2,036	2,036	2,036	2,036	2,036
R^2	0.43	0.43	0.42	0.11	0.48	0.41	0.22	0.58	0.64	0.53

Notes: This table reports analogous estimates of column 4 in Table 2 for alternative measures of the prevalence of individualistic naming patterns. The dependent variable is normalized across all columns. Column 1 show the results with no further sample restriction. Column 2 restricts the measure to children with native-born fathers, column 3 restricts to those with native-born mothers and fathers (the baseline in Table 2), and column 4 restricts to those with native-born grandparents. Column 5 changes the definition of infrequency of names to be based on the top 10 nationally, column 6 changes to the top 10 at the state level, and column 7 to the top 10 at the given county level. Column 8 increases the uncommon threshold to the top 25, and column 9 increases that to the top 100. Column 10 restricts to top names that do not have biblical roots. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table B.3: Robustness to Other Measures of Infrequent Names (IV Estimates)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Sample: White Children Aged 0-10, 1940 Census									
Further Sample Restriction:	None	Native Father	Native Parents	Native Grand-parents	None	None	None	None	None	None
Infrequent Measure:	Top 10 Division				Top 10 National	Top 10 State	Top 10 County	Top 25 Division	Top 100 Division	Top 10 National Non-Biblical
	Panel A: Share of Boys with Infrequent Names									
total frontier experience	0.208***	0.209***	0.207***	0.171***	0.201***	0.210***	0.227***	0.188***	0.207***	0.157***
	(0.041)	(0.041)	(0.042)	(0.046)	(0.041)	(0.041)	(0.078)	(0.041)	(0.037)	(0.035)
Number of Counties	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036
First Stage F Statistic	193.6	193.6	193.6	193.6	193.6	193.6	193.6	193.6	193.6	193.6
	Panel B: Share of Girls with Infrequent Names									
total frontier experience	0.214***	0.209***	0.207***	0.084**	0.165***	0.205***	0.095	0.217***	0.223***	0.162***
	(0.041)	(0.041)	(0.041)	(0.037)	(0.039)	(0.043)	(0.066)	(0.036)	(0.037)	(0.039)
Number of Counties	2,036	2,036	2,036	2,035	2,036	2,036	2,036	2,036	2,036	2,036
First Stage F Statistic	193.6	193.6	193.6	193.6	193.6	193.6	193.6	193.6	193.6	193.6

Notes: This table reports analogous instrumental variables estimates of the OLS specifications in Table B.2. See the notes to Table 10 for details on the IV. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table B.4: Persistence of Individualism

	Sample: White Children Aged 0-10 with Native-Born Parents							
	1910		1920		1930		1940	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Share of Boys with Infrequent Names							
total frontier experience	0.155***	0.157***	0.150***	0.239***	0.119***	0.174***	0.112***	0.207***
	(0.029)	(0.059)	(0.026)	(0.047)	(0.025)	(0.040)	(0.022)	(0.042)
Number of Counties	2,029	2,029	2,036	2,036	2,036	2,036	2,036	2,036
R^2	0.28	0.03	0.39	0.02	0.49	0.02	0.61	0.01
First Stage F Statistic		197.3		193.6		193.6		193.6
	Panel B: Share of Girls with Infrequent Names							
total frontier experience	0.114***	0.128***	0.152***	0.178***	0.146***	0.169***	0.161***	0.207***
	(0.022)	(0.043)	(0.021)	(0.047)	(0.024)	(0.043)	(0.024)	(0.041)
Number of Counties	2,029	2,029	2,036	2,036	2,036	2,036	2,036	2,036
R^2	0.49	0.02	0.55	0.04	0.51	0.04	0.42	0.04
First Stage F Statistic		197.3		193.6		193.6		193.6

Notes: This table reports analogous OLS and IV estimates of Table 2 but for each year since 1910. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table B.5: Total Frontier Experience and Contemporary Cooperation vs. Self-Reliance

	OLS	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)
total frontier experience	-0.019*	-0.025**	-0.041***	-0.026**	-0.025
	(0.009)	(0.009)	(0.014)	(0.012)	(0.029)
Oster δ for $\beta = 0$	-2.77	-2.61	-15.37	-249.36	
Number of Individuals	567	567	567	567	567
Number of Counties	48	48	48	48	48
Mean of Dependent Variable	0.549	0.549	0.549	0.549	0.549
First Stage F Statistic					9.9
R^2	0.01	0.02	0.02	0.03	0.01
Individual Demographic Controls	Yes	Yes	Yes	Yes	Yes
Division Fixed Effects	No	Yes	Yes	Yes	Yes
State Fixed Effects	No	No	Yes	No	No
Geographic/Agroclimatic Controls	No	No	No	Yes	Yes

Notes: This table reports estimates for a dependent variable based on a proxy for individualism in the 1990 round of ANES, covering 567 individuals in 48 counties across 17 states in our sample. The measure asks individuals whether (1) "it is more important to be a cooperative person who works well with others", or (2) "it is more important to be a self-reliant person able to take care of oneself." The dependent variable equals one if they answer (1). We report the same set of specifications in columns 1–4 as in Table 2 to demonstrate the statistically and economically significant effect sizes despite the coverage limitations. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. The Oster (2016) tests are with reference to a baseline specification with no controls. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

B.5 Addressing Potential Individual-Level Confounders of Policy Preferences

Many of the policies in Tables 3 and 5 elicit strong partisanship within the U.S. as Republicans and Democrats hew closely to the party line. However, as seen in Tables B.6, greater TFE is associated with stronger opposition to government intervention even after controlling for the strength of Republican party support reported in the CCES. Moreover, these results survive further controls for individual education and family income. Again, although these covariates are "bad controls," their inclusion helps rule out the concern that all of the observed effects are driven by prolonged frontier experience simply leading to tribal party- and class-based identity unrelated to the deep roots of frontier culture.

Table B.6: Robustness to Controls for Income, Education, and Partisan Identification

Dependent Variable:	Prefers Cutting Public Spending on Welfare		Prefers Balancing Budget By Cutting Spending		Prefers Repealing Affordable Care Act		Opposes Increasing Minimum Wage		Opposes Banning Assault Rifles		Opposes EPA Regulation of CO_2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
total frontier experience	0.007**	0.006*	0.014***	0.009***	0.022***	0.019***	0.023***	0.015**	0.015***	0.015***	0.016***	0.016***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)	(0.004)	(0.008)	(0.007)	(0.004)	(0.004)	(0.004)	(0.003)
identifies as very strong Republican		0.299***		0.379***		0.415***		0.457***		0.284***		0.338***
		(0.008)		(0.006)		(0.009)		(0.017)		(0.008)		(0.009)
family income > USD 50,000		0.099***		0.048***		-0.019***				-0.004		0.019***
		(0.006)		(0.005)		(0.007)				(0.005)		(0.006)
education > high school		0.007		-0.007		-0.080***		0.076***		0.015**		0.018***
		(0.006)		(0.005)		(0.006)		(0.014)		(0.007)		(0.006)
Oster δ for $\beta = 0$	6.86	4.63	2.40	1.76	2.96	2.26	3.05	1.85	2.46	2.19	2.22	2.16
Number of Counties	53,472	47,851	169,630	80,155	29,446	26,131	5,134	4,618	29,404	26,093	29,215	25,938
Mean of Dependent Variable	0.40	0.39	0.43	0.38	0.53	0.53	0.31	0.30	0.37	0.36	0.32	0.31
R^2	0.04	0.10	0.04	0.11	0.06	0.14	0.06	0.21	0.09	0.13	0.08	0.14

Notes: This table subjects the results in Table 5 to additional, non-predetermined controls for education, family income, and Republican Party identification as described in Appendix C. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. The Oster (2016) tests are with reference to a baseline specification with no controls.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

B.6 Full Elaboration of Additional Controls in Table 8

Table B.7: Robustness (III): Additional Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Panel A: Infrequent Boy Name Share in 1940 (normalized)													
total frontier experience	0.112***	0.113***	0.109***	0.113***	0.112***	0.114***	0.080***	0.112***	0.105***	0.079***	0.099***	0.072***	0.120***	0.065***
	(0.022)	(0.021)	(0.022)	(0.021)	(0.021)	(0.022)	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.023)	(0.021)	(0.021)
ruggedness		4.944***												3.619***
		(0.821)												(0.635)
rainfall risk			2.772											0.039
			(1.854)											(1.509)
distance to nearest portage site				0.036										0.036
				(0.047)										(0.039)
distance to nearest mineral discovery site					0.088***									0.095***
					(0.026)									(0.023)
distance to nearest Indian battle						-0.026								-0.040
						(0.037)								(0.027)
slave population share, 1860							-0.011***							-0.008***
							(0.002)							(0.001)
sex ratio, 1890								-0.355***						-0.213**
								(0.137)						(0.087)
immigrant share, 1890									-0.007*					0.025***
									(0.004)					(0.009)
scottish/irish immigrant share, 1890										-0.260***				-0.197***
										(0.019)				(0.019)
birthplace diversity, 1890											-1.102***			-1.181*
											(0.257)			(0.659)
years connected to railroad by 1890												-0.018***		-0.009***
												(0.002)		(0.002)
manufacturing employment share, 1890													-0.073***	-0.052***
													(0.006)	(0.005)
Oster δ for $\beta = 0$	1.70	1.75	1.65	1.71	1.71	1.70	1.01	1.71	1.54	1.07	1.42	0.92	1.91	0.70
Number of Counties	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036
Mean of Dependent Variable	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
R^2	0.61	0.63	0.61	0.61	0.61	0.61	0.63	0.61	0.61	0.67	0.62	0.65	0.66	0.74
	Panel B: Infrequent Girl Name Share in 1940 (normalized)													
total frontier experience	0.161***	0.164***	0.159***	0.167***	0.162***	0.166***	0.125***	0.162***	0.162***	0.133***	0.154***	0.112***	0.170***	0.123***
	(0.024)	(0.023)	(0.024)	(0.024)	(0.024)	(0.025)	(0.025)	(0.024)	(0.023)	(0.023)	(0.023)	(0.025)	(0.024)	(0.022)
ruggedness		6.192***												4.778***
		(1.002)												(0.778)
rainfall risk			3.462											-0.063
			(2.202)											(1.881)
distance to nearest portage site				0.119**										0.089**
				(0.055)										(0.041)
distance to nearest mineral discovery site					0.132***									0.139***
					(0.031)									(0.027)
distance to nearest Indian battle						-0.050								-0.083**
						(0.041)								(0.032)
slave population share, 1860							-0.013***							-0.007***
							(0.002)							(0.002)
sex ratio, 1890								-0.341**						-0.265**
								(0.133)						(0.104)
immigrant share, 1890									0.000					0.041***
									(0.006)					(0.012)
scottish/irish immigrant share, 1890										-0.229***				-0.157***
										(0.023)				(0.022)
birthplace diversity, 1890											-0.639**			-1.905**
											(0.322)			(0.789)
years connected to railroad by 1890												-0.022***		-0.014***
												(0.002)		(0.002)
manufacturing employment share, 1890													-0.074***	-0.055***
													(0.006)	(0.006)
Oster δ for $\beta = 0$	3.42	3.56	3.28	3.56	3.44	3.45	2.05	3.44	3.32	2.44	3.06	1.84	3.87	1.78
Number of Counties	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036
Mean of Dependent Variable	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R^2	0.42	0.45	0.42	0.43	0.43	0.42	0.44	0.43	0.42	0.47	0.43	0.48	0.47	0.58
	Panel C: County Property Tax Rate in 2010													
total frontier experience	-0.034***	-0.034***	-0.034***	-0.033***	-0.034***	-0.033***	-0.033***	-0.034***	-0.028***	-0.026***	-0.027***	-0.023***	-0.036***	-0.023***
	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
ruggedness		-0.959***												-0.724***
		(0.219)												(0.208)
rainfall risk			0.959*											0.736
			(0.581)											(0.516)
distance to nearest portage site				0.026										0.006
				(0.019)										(0.016)
distance to nearest mineral discovery site					0.026***									0.024***
					(0.009)									(0.008)
distance to nearest Indian battle						-0.004								0.002
						(0.011)								(0.009)
slave population share, 1860							0.000							-0.001
							(0.000)							(0.001)
sex ratio, 1890								-0.010						-0.024
								(0.027)						(0.024)
immigrant share, 1890									0.007***					-0.003
									(0.001)					(0.003)
scottish/irish immigrant share, 1890										0.059***				0.029***
										(0.008)				(0.006)
birthplace diversity, 1890											0.636***			0.541**
											(0.084)			(0.244)
years connected to railroad by 1890												0.005***		0.003***
												(0.000)		(0.000)
manufacturing employment share, 1890													0.016***	0.008***
													(0.002)	(0.002)
Oster δ for $\beta = 0$	1.67	1.72	1.71	1.58	1.69	1.60	1.52	1.67	1.36	1.32	1.33	1.10	1.93	1.07
Number of Counties	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029	2,029
Mean of Dependent Variable	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
R^2	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.83	0.83	0.83	0.83	0.83	0.85
	Panel D: Republican Vote Share, Average 2000–16													
total frontier experience	2.055***	2.050***	2.115***	2.095***	2.055***	2.172***	1.399***	2.060***	1.715***	1.717***	1.689***	1.640***	2.137***	0.931***
	(0.349)	(0.349)	(0.338)	(0.344)	(0.350)	(0.351)	(0.361)	(0.347)	(0.328)	(0.340)	(0.327)	(0.361)	(0.350)	(0.316)
ruggedness		-11.710												-34.668***
		(11.247)												(9.508)
rainfall risk			-77.882*											-108.600***
			(42.010)											(36.779)
distance to nearest portage site				0.930										0.966
				(0.777)										(0.623)
distance to nearest mineral discovery site					0.191									0.137
					(0.406)									(0.344)
distance to nearest Indian battle						-1.393**								-0.794*
						(0.603)								(0.471)
slave population share, 1860							-0.237***							-0.270***
							(0.035)							(0.034)
sex ratio, 1890								-5.000***						-1.948**
								(1.856)						(0.940)
immigrant share, 1890									-0.417***					-0.217
									(0.051)					(0.198)
scottish/irish immigrant share, 1890										-2.738***				-1.427***
										(0.263)				(0.243)
birthplace diversity, 1890											-32.582***			-4.970
											(3.563)			(13.286)
years connected to railroad by 1890												-0.186***		-0.089***
												(0.026)		(0.025)
manufacturing employment share, 1890													-0.689***	-0.422***
													(0.090)	(0.075)
Oster δ for $\beta = 0$	13.01	12.84	15.11	13.78	13.02	16.49	3.45	13.25	6.17	6.19	5.92	5.14	17.16	1.57
Number of Counties	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036	2,036
Mean of Dependent Variable	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04	60.04
R^2	0.33	0.33	0.34	0.33	0.33	0.34	0.38	0.34	0.38	0.38	0.39	0.36	0.36	0.49
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reproduces the estimates from Table 8, showing the coefficient estimates for the additional variables listed at the top of that table. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

B.7 Alternative Measures of Total Frontier Experience

Our baseline measure of TFE closely followed definitions in the historical literature as discussed in Section 3. In Table B.8, we demonstrate the robustness of our results to three relevant margins of adjustment to our measure of TFE. In each case, we redefinie what it means for county c to be on the frontier at time t. First, we reduce the catchment area from 100 km to 50 km in proximity to the frontier line. Second, we adjust the density restriction to include counties with > 2 people/mi^2 but still less than 6, counties with ≤ 18 people/mi^2, and then remove the population density restriction altogether. Finally, we consider defining the frontier line as including only the main, westernmost extent of all contour lines identified by the GIS algorithm. The overall message is that our particular choice of the frontier definition based on the historical record is not driving the main findings.

Table B.8: Robustness to Alternative Measures of TFE for Summary Outcomes

Dependent Variable:	Infrequent Name Share normalized		County Property Tax Rate	Republican Vote Share
	Boys 1940	Girls 1940	2010	Avg. 2000–16
	(1)	(2)	(3)	(4)
TFE: 100 km, ≤ 6/mi^2, all contour lines (baseline)	0.112***	0.161***	-0.034***	2.055***
	(0.022)	(0.024)	(0.007)	(0.349)
TFE: 50 km, ≤ 6/mi^2, all contour lines	0.117***	0.173***	-0.035***	2.051***
	(0.022)	(0.026)	(0.007)	(0.358)
TFE: 100 km, ≤ 18/mi^2, no inner island lines	0.096***	0.105***	-0.027***	1.575***
	(0.022)	(0.025)	(0.007)	(0.339)
TFE: 50 km, ≤ 18/mi^2, no inner island lines	0.085***	0.105***	-0.025***	1.458***
	(0.021)	(0.023)	(0.006)	(0.351)
TFE: 100 km, 2-6/mi^2, no inner island lines	0.081**	0.110**	-0.014*	1.877***
	(0.033)	(0.044)	(0.008)	(0.485)
TFE: 50 km, 2-6/mi^2, no inner island lines	0.063*	0.105**	-0.012	1.771***
	(0.038)	(0.049)	(0.009)	(0.530)
TFE: 100 km, no density restriction, no inner island lines	0.033	0.030	-0.011	1.001***
	(0.021)	(0.025)	(0.007)	(0.335)
TFE: 50 km, no density restriction, no inner island lines	0.054***	0.068***	-0.018***	1.078***
	(0.020)	(0.023)	(0.006)	(0.339)
TFE: 100 km, ≤ 6/mi^2, including inner island lines	0.132***	0.188***	-0.032***	2.048***
	(0.020)	(0.023)	(0.006)	(0.320)
TFE: 50 km, ≤ 6/mi^2, including inner island lines	0.143***	0.205***	-0.035***	2.098***
	(0.022)	(0.025)	(0.007)	(0.335)
TFE: 100 km, ≤ 6/mi^2, main single contour line	0.087***	0.117***	-0.037***	1.872***
	(0.026)	(0.032)	(0.008)	(0.436)
TFE: 50 km, ≤ 6/mi^2, main single contour line	0.082***	0.113***	-0.043***	1.787***
	(0.028)	(0.034)	(0.008)	(0.460)
TFE: 50 km, ≤ 6/mi^2, no inner or outer island lines	0.111***	0.149***	-0.034***	2.133***
	(0.022)	(0.025)	(0.007)	(0.357)
TFE: 50 km, ≤ 6/mi^2, no inner or outer island lines	0.111***	0.159***	-0.035***	2.116***
	(0.023)	(0.026)	(0.007)	(0.373)
State Fixed Effects	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes

Notes: This table reports estimates of equation (5) for three measures of infrequent names for white children, age 0–10 in the 1940 Census. Each cell is a different regression based on the given dependent variable in the column and the measure of total frontier experience in the given row. The frontier lines considered in the baseline are countour lines longer than 500km after removing all "inner island lines" that are east of the main frontier line. The alternative measures of frontier experience considered above vary (i) the catchment area from 100 to 50 km from the contour lines, (ii) the density restriction from ≤ 6 people/mi^2 to $2 \geq$people/mi$^2 \leq 6$ to no restriction, (iii) including inner island lines, and (iv) including only the longest single contour line. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

B.8 Robustness Checks for Additional Survey Outcomes

The following tables report several robustness checks for the additional survey-based outcomes not reported in the main robustness checks tables.

Table B.9: Robustness (I): Adding West Coast and Extended Time Frame for Survey Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Data Source:	ANES	CCES	GSS	CCES	GSS	CCES			
Dependent Variable:	Prefers Cut on Poor	Public Spending on Welfare	Government Should Redistribute	Prefer Cut Debt by Spending Cuts	Index of Preferences for Cut Spending	Opposes Affordable Care Act	Opposes Increasing Min. Wage	Opposes Banning Assault Rifle	Opposes EPA Regulation of CO_2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Adding West Coast to the Baseline Sample								
total frontier experience	0.008*	0.009***	-0.007	0.015***	0.033***	0.023***	0.023***	0.021***	0.017***
	(0.005)	(0.003)	(0.014)	(0.003)	(0.011)	(0.004)	(0.007)	(0.005)	(0.004)
Number of Individuals	2,810	66,254	11,271	139,618	7,109	36,768	6,553	36,711	36,479
Number of Counties	108	1,963	290	2,064	288	1,828	1,157	1,823	1,818
Mean of Dependent Variable	0.10	0.39	-0.00	0.40	0.00	0.51	0.30	0.36	0.31
	Panel B: Extending Historical Frontier Period to 1950								
total frontier experience	0.007	0.011***	0.006	0.013***	0.035***	0.021***	0.016***	0.020***	0.016***
	(0.005)	(0.003)	(0.011)	(0.002)	(0.009)	(0.003)	(0.005)	(0.003)	(0.003)
Number of Individuals	3,035	74,260	12,566	157,131	7,916	41,211	7,252	41,151	40,895
Number of Counties	113	2,241	319	2,389	317	2,076	1,294	2,072	2,067
Mean of Dependent Variable	0.10	0.39	0.00	0.40	0.00	0.52	0.31	0.37	0.32
Survey Wave Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State/Division Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table estimates the specifications in columns 1 (Panel A) and 5 (Panel B) of Table 6 for the other, survey-based outcomes examined in Section 5.4. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1.

Table B.10: Robustness (II) and (III): Population Density and Other Controls

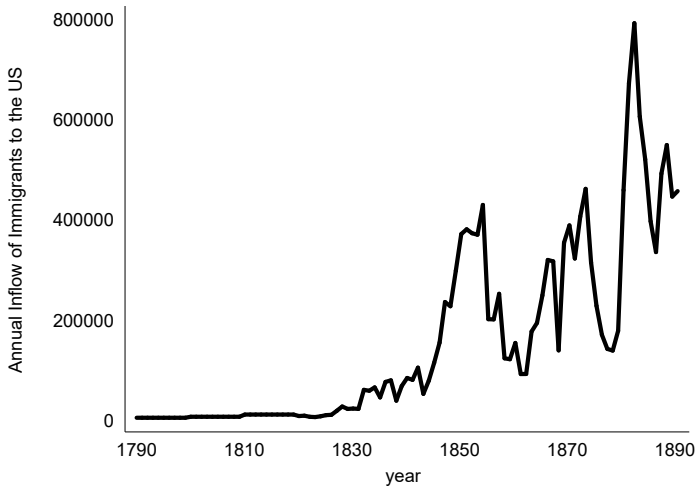
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Data Source:	ANES	CCES	GSS	CCES	GSS	CCES			
Dependent Variable:	Prefers Cut on Poor	Public Spending on Welfare	Government Should Redistribute	Prefer Cut Debt by Spending Cuts	Index of Preferences for Cut Spending	Opposes Affordable Care Act	Opposes Increasing Min. Wage	Opposes Banning Assault Rifle	Opposes EPA Regulation of CO_2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Adding Population Density in 1890								
total frontier experience	0.012***	0.005	-0.025*	0.011***	0.028**	0.020***	0.022***	0.012***	0.014***
	(0.004)	(0.003)	(0.012)	(0.002)	(0.011)	(0.004)	(0.008)	(0.004)	(0.004)
Number of Individuals	2,322	53,472	9,085	111,853	5,739	29,446	5,134	29,404	29,215
Number of Counties	95	1,863	255	1,963	253	1,728	1,066	1,723	1,718
Mean of Dependent Variable	0.09	0.40	0.00	0.41	-0.00	0.53	0.31	0.37	0.32
	Panel B: Adding Table 8, Column 14 Controls								
total frontier experience	0.002	0.005	-0.017	0.010***	0.010	0.016***	0.019**	0.010**	0.011***
	(0.007)	(0.003)	(0.014)	(0.003)	(0.013)	(0.004)	(0.008)	(0.004)	(0.004)
Number of Individuals	2,188	51,171	8,466	106,964	5,382	28,165	4,905	28,128	27,941
Number of Counties	87	1,711	242	1,792	240	1,591	1,004	1,589	1,582
Mean of Dependent Variable	0.09	0.40	0.00	0.41	-0.00	0.53	0.31	0.37	0.31
Survey Wave Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State/Division Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic/Agroclimatic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table estimates the specifications in column 2 of Table 7 (Panel A) and column 10 (Panel B) of Table 8 for the other, survey-based outcomes examined in Section 5.4. Standard errors are clustered based on the grid-cell approach of Bester, Conley and Hansen (2011) as detailed in Section 5.1.

B.9 Instrumental Variables Strategy: Further Background and Additional Results

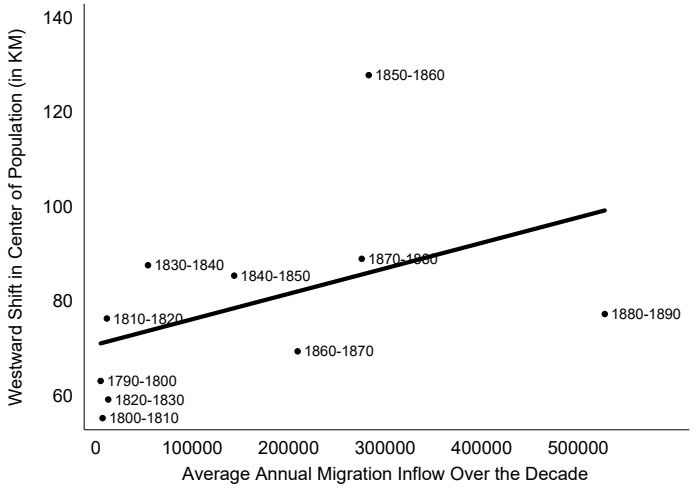
We augment the widely-used Migration Policy Institute (2016) data on annual migration inflows (collected by the Office of Immigration Statistics) with data from Tucker (1843) for the pre-1820 period (see Appendix C). Figure B.3 shows the ups and downs of immigration to the U.S. over the study period. Figure B.4 then shows the strong positive correlation between these immigrant inflows by decade and the speed of westward expansion, proxied by the east-to-west distance traveled by the country's population centroid (the green dot in Figure A.1(b) for 1860). This simple scatterplot helps visualize the process by which immigrants arriving in the U.S. (largely on the Eastern seaboard) pushed the edges of settlement farther westward, which in turn hastened the forward march of the frontier line. In periods with low immigrant inflows, this push slowed down, leading some counties to remain part of the frontier for longer than those that just happen to be getting closer to the frontier line at a time of rapid inflows into the U.S. Table B.11 demonstrates the strong first stage in our main IV regressions from Table 10.

Figure B.3: Annual Migration Inflows



Notes: This figure plots the total number of migrants entering the United States, 1790-1890. The data for 1820–1890 is available from the Migration Policy Institute (2016), while the data for 1790-1819 is imputed from Tucker (1843).

Figure B.4: Immigration and Westward Expansion



Notes: This figure plots the length of the decadal westward shift of the center of population (in km) against the average annual immigrant inflow during the decade. The center of population is the point at which weights of equal magnitude corresponding to the location of each person in an imaginary flat surface representing the U.S. would balance out. This measure was reported historically by the U.S. Census Bureau (see footnote 10 in the paper).

Table B.11: First Stage Results for the Instrumental Variables Estimates in Table 10

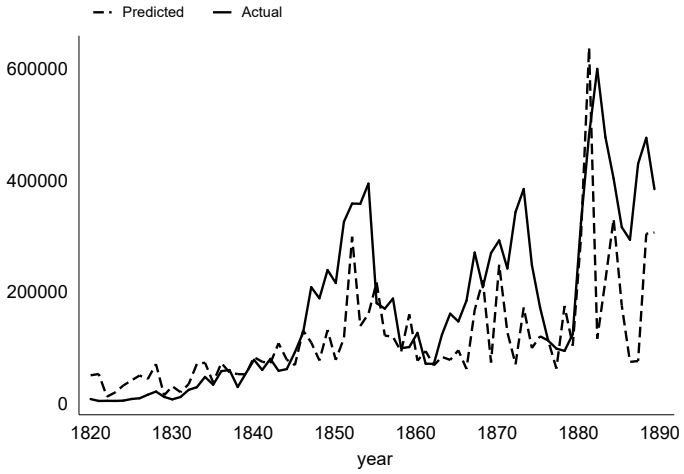
Dependent Variable (in first stage):	total frontier experience	
	(1)	(2)
Log Average Actual National Migration Inflows	-1.016***	
	(0.073)	
Log Average Predicted National Migration Inflows		-2.010***
		(0.144)
Log county area	0.234***	0.232***
	(0.072)	(0.073)
Latitude	-0.091	-0.076
	(0.079)	(0.075)
Longitude	-0.153***	-0.176***
	(0.029)	(0.030)
Mean Annual Temperature	-0.144**	-0.102
	(0.071)	(0.068)
Mean Annual Rainfall	-0.002***	-0.002***
	(0.001)	(0.001)
Mean of Median Altitude	-0.001**	-0.001**
	(0.000)	(0.000)
Distance to Coast	-0.001**	-0.001**
	(0.000)	(0.000)
Distance to Rivers	0.003***	0.004***
	(0.001)	(0.001)
Distance to Lakes	-0.000	-0.000
	(0.000)	(0.000)
Average Agricultural Suitability	2.686***	0.876
	(0.788)	(0.752)
Number of Counties	2036	2036
First Stage F Statistic	193.6	195.8
State Fixed Effects	Yes	Yes

Notes: This table reports the first stage results corresponding to the baseline IV regressions presented in Table 10. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Section 5.6 shows that the main results are robust to an instrumental variable estimation exploiting time series variation in national migration inflows. To address concerns regarding the excludability of the baseline instrument due to pull factors associated with immigrant inflows, we show in Panel B of Table 10 that the IV results are qualitatively unchanged when using an instrument based on push factors unrelated to frontier conditions. For this version of the IV, we draw on the approach in Nunn, Qian and Sequeira (2017), using country-year level data on migrant inflows from 16 European countries to the US from 1820–1890 and constructing predicted migration outflows induced by weather shocks. First, using country-specific regressions, we predict the annual migrant outflows from each country to the US as a function of country-specific shocks to temperature and rainfall in the prior year (see Nunn, Qian and Sequeira, 2017, for details on these measures). Second, we aggregate across countries to obtain the total predicted migrant inflows to the US for each year. Analogous to our baseline instrument, we then construct the IV for each county in our sample by calculating the average annual predicted migrant

inflow to the US over the 30 years starting from the first year in which the given county is just west of the frontier. Figure B.5 shows how the predicted inflows, which isolate push factors, compare to the actual inflows, which naturally include both push and pull. While the data on migrant inflows from Europe to the US is available only starting in 1820, we retain the full sample of counties in the IV regressions by imputing the inflows for 1790-1819 using linear extrapolation of the post-1819 predicted inflows.[2]

Figure B.5: Actual vs. Predicted Immigration Inflows from Europe to the United States



Notes: This figure compares the actual migration inflows from Europe from 1820–1890 to the predicted flows based on the total country-specific predicted outflows using the climatic shocks approach in Nunn, Qian and Sequeira (2017) as described above.

[2]Restricting the sample to counties just west of the frontier after 1820—for which the IV is solely based on predicted flows without extrapolation—delivers similar results, though the estimates are noisier due to the smaller sample size.

B.10 The Parasite-Stress Theory of Values

The parasite-stress theory of values due to Thornhill and Fincher (2014) argues that the prevalence of infectious diseases leads to higher levels of in-group assortative sociality, which they associate with collectivism, as an adaptive response that minimizes contagion. In the context of our study, this theory might suggest that frontier individualism resulted from the low prevalence of infectious diseases on the frontier. However, this potential mechanism does not arise in historical narratives. Nor do we find evidence of differential disease prevalence or morbidity on the frontier. As seen in Table B.12 below, the prevalence of pathogens—associated with tuberculosis, malaria, and typhoid, among other diseases considered in Gorodnichenko and Roland (2016)—does not exhibit any differential intensity on the frontier. We can measure the incidence of these specific infectious diseases as well as a broad array of other illnesses for the first time in the 1880 Population Census. Adopting specifications similar to Table 1, we find little evidence that individuals living on the frontier had differential (infectious) disease or illness. If the parasite-stress mechanism were salient, we would find that frontier locations exhibit significantly less prevalence of infectious diseases. While the relatively precise zeros in the table may be specific to 1880, this provides suggestive evidence that the parasite-stress channel is not a first-order factor in explaining the differential individualism on the frontier.

Table B.12: No Differential Infectious Diseases or Sickness on the Frontier

Dependent Variable:	Share of Pop. with Infectious Disease		Share of Pop. with Any Illness	
	(1)	(2)	(3)	(4)
on the frontier	0.0001		0.0009	
	(0.0003)		(0.0013)	
near frontier line		-0.0001		0.0011
		(0.0001)		(0.0007)
low population density		0.0000		0.0001
		(0.0002)		(0.0010)
Mean Dependent Variable	0.001	0.001	0.009	0.009
Number of County-Years	1,780	1,780	1,780	1,780
R^2	0.05	0.05	0.08	0.08

Notes: This table reports estimates of the relationship between frontier definitions and the share of the county with any of the infectious diseases considered in Gorodnichenko and Roland (2016) (columns 1–2) and any illness (column 3–4). The infectious diseases of interest include tuberculosis, malaria, and typhus. The specification is otherwise similar to that in Table 1, with Census division FE and standard errors clustered using the grid-cell approach of Bester, Conley and Hansen (2011).
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

B.11 Further Results on Individualism, Success, and Endurance on the Frontier

Table B.13: Individualism and Changes in Socioeconomic Status

Dependent Variable:	$sei_{1880} - sei_{1870}$				$occscore_{1880} - occscore_{1870}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
has infrequent name	0.374	0.338	0.205	-0.240	0.731	0.739	0.626	0.573
	(1.029)	(0.989)	(0.982)	(0.878)	(0.558)	(0.524)	(0.517)	(0.488)
has infrequent name × frontier county	21.463**	20.116**	20.944***	18.681***	9.012**	8.281**	8.990**	8.537**
	(9.689)	(8.088)	(7.616)	(6.466)	(4.517)	(4.139)	(4.212)	(4.050)
Number of Individuals	4,268	4,268	4,268	4,268	4,268	4,268	4,268	4,268
Mean of Dependent Variable	-0.245	-0.245	-0.245	-0.245	-0.548	-0.548	-0.548	-0.548
R^2	0.23	0.25	0.27	0.40	0.23	0.27	0.29	0.39
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age, Age Squared Control	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Number of Children Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Baseline Farmer Dummy	No	No	Yes	Yes	No	No	Yes	Yes
Baseline *sei* or *occscore*	No	No	No	Yes	No	No	No	Yes

Notes: This table reports estimates analogous to those in Panel A of Table 12 but based on the NAPP Linked Sample based on a 1 percent of the population in the 1870 and 1880 Population Censuses. The sample is restricted relative to the one in Table 12 as we are interested in the change in socioeconomic status (*sei*) and occupational standing (*occscore*) between 1870 and 1880, and this information is only available in both years for this smaller linked sample. The estimates are based on white, male household heads that reside in the same county in 1870 and 1880. The dependent variable is the change in *sei* or *occscore*; results are similar taking logs, but the levels allow us to retain individuals that switch from zero valued status to positive or vice versa. The frontier dummy equals one if the given county lies in the frontier in 1880. The infrequent name measure is based on the top 10 names nationally. The controls listed at the bottom of the table include a dummy for farmer occupations in 1870 in columns 3/4 and 7/8, and the baseline dependent variable in columns 4 and 8. Standard errors are clustered at the county level. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table B.14: Individualism and Socioeconomic Success on the Frontier | Farmer Dummy

	Dep. Var.: Father's Economic Status in 1880 (normalized)			
	sei			*occscore*
	(1)	(2)	(3)	(4)
at least one child 0-10 has unique name	0.056***		0.058***	0.036***
	(0.006)		(0.006)	(0.005)
at least one child 0-10 has unique name × frontier county	0.051		0.047	0.032
	(0.034)		(0.034)	(0.029)
father has unique name		-0.032***	-0.033***	-0.023***
		(0.005)	(0.005)	(0.004)
father has unique name × frontier county		0.063**	0.061**	0.065***
		(0.025)	(0.025)	(0.024)
Number of Individuals	264,038	264,038	264,038	264,038
R^2	0.30	0.30	0.30	0.45
County Fixed Effects	Yes	Yes	Yes	Yes
Farmer Dummy	Yes	Yes	Yes	Yes

Notes: This table re-estimates the regressions in Panel A of Table 12, controlling for an indicator of whether or not the individual is in a farming occupation. Standard errors are clustered at the county level. All regressions include dummies for the number of children born in the 1870s. Significance levels: $*$: 10% $**$: 5% $***$: 1%.

Table B.15: Individualism and Endurance on the Frontier

	(1)	(2)	(3)	(4)
	Dependent Variable:			
	Panel A: Emigrated from Frontier County in 1870			
father has infrequent name	-0.051***			-0.050***
	(0.009)			(0.009)
mother has infrequent name		-0.005		-0.005
		(0.007)		(0.007)
at least one child 0-10 has infrequent name			-0.024**	-0.020*
			(0.010)	(0.010)
Number of Individuals	27,066	27,066	27,066	27,066
Mean of Dependent Variable	0.583	0.583	0.583	0.583
R^2	0.07	0.07	0.07	0.07
	Panel B: Migrated from Frontier County in 1870 to Non-Frontier County in 1880			
father has infrequent name	-0.042***			-0.041***
	(0.009)			(0.009)
mother has infrequent name		-0.010		-0.009
		(0.007)		(0.007)
at least one child 0-10 has infrequent name			-0.026**	-0.023**
			(0.011)	(0.011)
Number of Individuals	27,066	27,066	27,066	27,066
Mean of Dependent Variable	0.410	0.410	0.410	0.410
R^2	0.08	0.07	0.07	0.08
	Panel C: Onward Migrated from Frontier County in 1870 to Frontier County in 1880			
father has infrequent name	-0.006			-0.006
	(0.006)			(0.006)
mother has infrequent name		0.006		0.006
		(0.005)		(0.005)
at least one child 0-10 has infrequent name			0.001	0.001
			(0.008)	(0.008)
Number of Individuals	27,066	27,066	27,066	27,066
Mean of Dependent Variable	0.167	0.167	0.167	0.167
R^2	0.11	0.11	0.11	0.11

Notes: This table reproduces the estimates from Panel B of Table 12 alongside other outcomes in Panels A and C that clarify that the "return migration" effect comprises the full effect on outmigration destinations discussed with respect to that finding. Standard errors are clustered at the origin county level.
Significance levels: $*$: 10% $**$: 5% $***$: 1%.

C Data Sources and Construction

Harmonization to 2010 Boundaries

We harmonize all historical Census data to the 2010 boundaries using an approach suggested in Hornbeck (2010). First, we intersect the county shapefiles from each of the decadal census years with the 2010 county shapefile and calculate the area of each intersection. When the 2010 county falls in one or more counties of the earlier shapefile, each piece of the 2010 county is assigned a value equal to the share of the area of the piece in the earlier county multiplied by the total value of the data for the earlier county. Then, the data for each county in 2010 is the sum of all the pieces falling within its area. This harmonization procedure would be exact if all the data from the various years are evenly distributed across county areas.

Demographic Variables and Individualism

Population density. Population/area. Digitized U.S. Census data on population for every decade in 1790–2010, from Minnesota Population Center (2016). The data on area is calculated using the 2010 county shapefiles from NHGIS (Minnesota Population Center, 2011) using GIS software. The county level population data along with other pre-2010 data are harmonized to the 2010 county boundaries and the data for intercensal years is imputed using the procedure detailed in Section A.

Sex Ratio. Whites males/white females. The data is available for every decade in 1790-1860 and 1890. Data source: (Minnesota Population Center, 2011).

Prime Age Adult Share. Whites aged 15–49/all whites. The data used is consistently available for every decade in 1830-1860. Data source: (Minnesota Population Center, 2011).

Illiteracy. Illiterate whites aged above 20/whites aged over 20. The variable is available consistently for 1830-1860. Data source: (Minnesota Population Center, 2011).

Immigrant Share. Foreign born/population. The variable used is available for every decade in 1820-1890 (excluding 1840). Data source: (Minnesota Population Center, 2011).

Out of State Born Share. Out-of-state born/population. The variable is consistently available for every decade in 1850-1880. Data source: (Minnesota Population Center, 2011).

Land inequality. Gini index using distribution of farm sizes, based on county level data on the number of farms of sizes 0–10, 10–19, 20–49, 50–99, 100–499, 500–1000, and above 1000 acres. Available for every decade in 1860-1890. (Minnesota Population Center, 2011).

Infrequent Children Names. White Children Aged 0–10 with Non-Top 10 First Names in Division/White Children Aged 0-10. We also construct similar variables further restricting to children aged 0–10 with native parents, and native grandparents. In addition, for the same sample, we construct additional variables by calculating the popularity of names at the national level instead of the Census division. We use the following procedure to generate the name shares: start by restricting the sample as desired (e.g. white children aged 0-10 with native parents), then calculate the number of children in the county for each given name, then using that value identify the top 10 given names within the census division (or nationally), and then accordingly count the number of children in that county with the identified top 10 names in their corresponding census division. The variables restricting to white children aged 0-10 is available for every decade in 1850–1940 (excluding 1890), with further native-parent restriction for 1850 and 1880-1940 (excluding 1890), and with grandparent restriction for 1880–1940 (excluding 1890). To give some examples, in 1850 the top 10 boy names nationally in descending order of popularity were John, William, James, George, Charles, Henry, Thomas, Joseph, Samuel and David. Meanwhile, a

random sample of less common names (outside the top 25) includes ones like Alfred, Nathan, Patrick, Reuben, Herbert, Matthew, Thaddeus and Luke. For girls, the top 10 include Mary, Sarah, Elizabeth, Martha, Margaret, Nancy, Ann, Susan, Jane, and Catherine while less common names (outside the top 25) include ones like Rachel, Susannah, Nina, Olive, Charlotte, Lucinda, and Roxanna. By 1880, the rankings shifted only slightly for boys with Samuel falling outside the top 10 and Harry entering. For girls, the changes were a bit more dramatic with the new top 10 list being Mary, Sarah, Emma, Ida, Minnie, Anna, Annie, Martha, Cora, and Alice. Data source: The NAPP full count census data for 1850 and the *Ancestry* data collected by NBER for 1860–1940.

Economic Status. We measure economics status using either the socioeconomic index (*sei*) or the occupational score (*occscore*) measures provided by the North Atlantic Population Project: Complete Count Microdata. Both measures range from 0 to 100, and capture the income returns associated with specific occupations in the 1950 Census while the *sei* measure additionally captures notions of prestige as well as educational attainment. Data source: (Minnesota Population Center, 2011).

Survey-Based Cultural Outcomes

Some of our key measures of contemporary preferences for government policy are based on data from multiple rounds of three widely used, nationally representative surveys: the Cooperative Congressional Election Study (CCES), the General Social Survey (GSS), and the American National Election Study (ANES). These surveys are staples in the social science literature on political preferences and social norms. For instance, Acharya, Blackwell and Sen (2016) uses CCES and ANES in a related methodological setting, and Alesina and Giuliano (2010) conducts a thorough investigation of the determinants of preferences for redistribution using the GSS. The CCES is a web-based survey conducted every two years, the ANES is an in-person survey conducted annually since 1948, and the GSS is an in-person survey conducted annually since 1972. All three are repeated cross-sections.

One advantage of working with three surveys is that we can cross-validate the findings across surveys that ask different questions about similar underlying preferences. For example, the CCES asks respondents if and how respondents would like state-level welfare spending to change whereas the ANES asks respondents if and how federal spending on the poor should change. The CCES also includes a set of questions on policy issues such as gun ownership that are particularly relevant to some of the mechanisms driving the persistence of frontier culture. For all measures, we link county-level identifiers in the underlying data to the 2010 county boundaries.

Despite their rich level of detail, these surveys have one important limitation for our purposes, namely the limited geographic scope. The three surveys are nationally representative, but their coverage differs. While the CCES has broad spatial coverage, the GSS and ANES do not (see Appendix Figures C.1). Despite its broader coverage, the CCES has the potential disadvantage that it captures an internet-savvy sample that may not be reflective of the underlying population in the way that an in-person survey generally would. This is particularly disadvantageous given our focus on county-level variation in TFE across a swathe of the United States outside of major coastal population centers.

Prefers Cutting Public Spending On Poor. The Prefers Cutting Public Spending On Poor is an indicator variable based on the following survey question: "*Should federal spending be increased, decreased, or kept about the same on poor people?*" The variable takes a value of 1 if the respondent answered "*decreased*" and 0 otherwise, and it is available for 1992 and 1996. Data source: The American National Election Studies Cumulative Data (2012). The ANES is a large, nationally-representative survey of the American electorate in the United States taken during the presidential and midterm election years. See Appendix Figure C.1(a) for the map of the maximum survey coverage in the final sample of ANES data merged with the frontier related data.

Prefers State Decrease Welfare Spending. This is an indicator variable based on the following survey question: "*State legislatures must make choices when making spending decisions on important state programs. Would*

you like your legislature to increase or decrease spending on Welfare? 1. Greatly Increase 2. Slightly Increase 3. Maintain 4. Slightly Decrease 5. Greatly Decrease." Prefers Cut Public Spending on Welfare takes a value of 1 if the respondent answered "*Slightly Decrease*" or "*Greatly Decrease*" and 0 otherwise. The data is available in the 2014 and 2016 waves. Data source: Cooperative Congressional Election Study (Ansolabehere and Schaffner, 2017) Common Content surveys. The CCES was formed in 2006, through the cooperation of several academic institutions, to study how congressional elections, representation and voters' behavior and experiences vary with political geography and social context using very large scale national surveys. The 2014 and 2016 CCES surveys were conducted over the Internet by YouGov using a matched random sample methodology. The Common Content portion of the survey, which contains our variables of interest, surveyed 56,200 adults in 2014 and 64,600 adults in 2016. See Appendix Figure C.1(b) for the map of the maximum survey coverage in the final sample of CCES data merged with frontier related data.

Believes Government Should Redistribute. Based on the following survey question: "*Some people think that the government in Washington ought to reduce the income differences between the rich and the poor, perhaps by raising the taxes of wealthy families or by giving income assistance to the poor. Others think that the government should not concern itself with reducing this income difference between the rich and the poor. Here is a card with a scale from 1 to 7.*" We have recoded the variable so that it is increasing in preference for redistribution, where a score of 1 means that the government should not concern itself with reducing income differences and a score of 7 means the government ought to reduce the income differences between rich and poor. The Believes Government Should Redistribute is a normalized version of the above variable, and it is available in our sample for 1993 and all even years between 1994-2016. Data source: The General Social Survey (Smith, Marsden, Hout and Kim, 2015). The GSS is a repeated cross-sectional survey of a nationally representative sample of non-institutionalized adults who speak either English or Spanish. The surveys has been conducted since 1972, almost every year between 1972-1993 and biennial since 1994. While the sample size for the annual surveys was 1500, since 1994 the GSS administers the surveys to two samples in even-numbered years, each with a target sample size of 1500. The surveys provide detailed questionnaires on issues such as national spending priorities, intergroup relations, and confidence in institutions. See Appendix Figure C.1(c) for the map of the maximum survey coverage in the final sample of CCES merged with frontier related data.

Prefers Reducing Debt by Cutting Spending. The variable is based on the CCES survey question: *"The federal budget deficit is approximately [$ year specific amount] this year. If the Congress were to balance the budget it would have to consider cutting defense spending, cutting domestic spending (such as Medicare and Social Security), or raising taxes to cover the deficit. Please rank the options below from what would you most prefer that Congress do to what you would least prefer they do: Cut Defense Spending; Cut Domestic Spending; Raise Taxes."*. While this question varies slightly from year to year, the underlying theme is the same. The Prefers Reducing Debt by Cutting Spending variable takes a value of 1 if the respondent chose "Cut Domestic Spending" as a first priority. The data is available for 2006-2014 (excluding 2013). Data source: (Ansolabehere and Schaffner, 2017).

Index of Preferences for Spending Cuts. The index is the principal component of nine dummy variables that take the value of 1 if the respondents answers "too much" to the following questions: "*We are faced with many problems in this country, none of which can be solved easily or inexpensively. I'm going to name some of these problems, and for each one I'd like you to name some of these problems, and for each one I'd like you to tell me whether you think we're spending too much money on it, too little money, or about the right amount. First (READ ITEM A) . . . are we spending too much, too little, or about the right amount on (ITEM)?*". The items considered are improving and protecting the environment, improving healthcare, solving big city problems, halting increasing crimes, dealing with drug addictions, improving the education system, improving conditions for blacks, military spending, foreign aid, welfare, and roads. The variable is available in our sample for 1993 and all even years between 1994-2016. Data source: (Smith, Marsden, Hout and Kim, 2015).

Prefers Repealing Affordable Care Act. Based on the CCES survey question: *"The Affordable Health Care Act*

was passed into law in 2010. It does the following: Requires all Americans to obtain health insurance, Prevents insurance companies from denying coverage for pre-existing condition, Allows people to keep current health insurance and care provider, and Sets up national health insurance option for those without coverage, but allows states the option to implement their own insurance system. Would you have voted for the Affordable Care Act if you were in Congress in 2010?" The Prefers Repealing Affordable Care Act variable takes a value of 1 if the respondent answers "Yes" and 0 if the answer is "No". The data is available for 2014. Data source: (Ansolabehere and Schaffner, 2017).

Opposes Increasing Minimum Wage. Based on the survey question:*"As you may know, the federal minimum wage is currently $5.15 an hour. Do you favor or oppose raising the minimum wage to $7.25 an hour over the next two years, or not?"*. The variable Opposes Increasing Minimum Wage takes a value of 1 if the respondent choses "oppose" and 0 otherwise. Available in 2007. Data source: (Ansolabehere and Schaffner, 2017).

Opposes Banning Assault Rifles. Based on the CCES survey question: *"On the issue of gun regulation, are you for or against for each of the following proposal? proposal: banning assault rifles"*. Opposes Banning Assault Rifles takes value 1 if the respondent is against banning assault rifles and 0 otherwise. Available for 2014. Data source: (Ansolabehere and Schaffner, 2017).

Opposes EPA Regulations of CO_2 Emissions. Based on the CCES survey question *"Do you support or oppose each of the following proposals? proposal: Environmental Protection Agency regulating Carbon Dioxide emissions."* The Opposes EPA Regulations of CO_2 Emissions takes one if the respondent supports the proposal and 0 the respondent opposes. Available for 2014. Data source: (Ansolabehere and Schaffner, 2017).

Cooperation vs. Self-Reliance. Based on the survey question: *"I am going to ask you to choose which of two statements I read comes closer to your own opinion. You might agree to some extent with both, but we want to know which one is closer to your views: ONE, it is more important to be a cooperative person who works well with others; or TWO, it is more important to be a self-reliant person able to take care of oneself"*. The Cooperation vs. Self-Reliance variable takes a value of 1 if the respondent chooses "cooperative" and 0 otherwise. Available in 1990. Data source: The American National Election Studies.

Identifies As A Strong Republican. An indicator variable that takes 1 if the respondent identifies as a "Strong Republican." Available for 2007, 2012, 2014 and 2016. Data source: (Ansolabehere and Schaffner, 2017).

Other Long-run Outcomes

County Property Tax Rate. The average effective property tax rates per $100 of value, calculated at the county level as the ratio of the average real estate tax over the average house value. Data source: The data is obtained from the National Association of Home Builders, which calculated the average effective property tax rates based on the 2010-2014 American Community Survey (ACS) data from the Census Bureau.

Republican Vote Share in Presidential Elections. Votes for a GOP candidate/total votes, at the county level. For simplicity, we only consider the five presidential elections since 2000. Data source: Dave Leip's Atlas of U.S. Presidential Elections (2017).

Geographic and Agroclimatic Controls

Land productivity measures. Average of attainable yields for alfalfa, barley, buckwheat, cane sugar, carrot, cabbage, cotton, ax, maize, oats, onion, pasture grasses, pasture legumes, potato, pulses, rice, rye, sorghum, sweet potato, tobacco, tomato, and wheat. We normalize each product's values dividing it by the maximum value for that product in the sample. Measures of attainable yields were constructed by the FAO's Global Agro-Ecological Zones project v3.0 (IIASA/FAO, 2012) using climatic data, including

precipitation, temperature, wind speed, sunshine hours and relative humidity (based on which they determine thermal and moisture regimes), together with crop-specific measures of cycle length (i.e. days from sowing to harvest), thermal suitability, water requirements, and growth and development parameters (harvest index, maximum leaf area index, maximum rate of photosynthesis, etc). Combining these data, the GAEZ model determines the maximum attainable yield (measured in tons per hectare per year) for each crop in each grid cell of 0.083×0.083 degrees. We use FAO's measures of agroclimatic yields (based solely on climate, not on soil conditions) for intermediate levels of inputs/technology and rain-fed conditions.

Area. The log of surface area in square miles, calculated using the 2010 county shapefiles from NHGIS (Minnesota Population Center, 2011) using GIS software.

Temperature. County-level mean annual temperature measured in Celsius degrees. Data source: (IIASA/FAO, 2012).

Rainfall. County-level average annual precipitation measured in mm. Data source: (IIASA/FAO, 2012).

Elevation. County-level average terrain elevation in km. Data source: (IIASA/FAO, 2012).

Latitude. Absolute latitudinal distance from the equator in decimal degrees, calculated from the centroid of each county using GIS software and county shapefiles from NHGIS (Minnesota Population Center, 2011).

Longitude. Absolute longitudinal distance from the Greenwich Meridian in decimal degrees, calculated from the centroid of each county using GIS software and county shapefiles from NHGIS (Minnesota Population Center, 2011).

Distance to the coastline, rivers, and lakes. Minimum distance to a point in the coastline, rivers, and lakes in km, calculated from the centroid of each county using GIS software and county shapefiles from NHGIS. Data source: (Minnesota Population Center, 2011)

Additional Variables

Annual Migration Inflow. Total number of migrants entering the United States every year. The data for 1820–1890 is available from the Migration Policy Institute (2016), which tabulates data from the Office of Immigration Statistics, while the data for 1790–1819 is imputed from Tucker (1843). To construct the instrumental variable based on annual migration inflows predicted by weather shocks in Europe, we use the annual migration inflows to the U.S. from Belgium, Denmark, England, France, Germany, Greece, Ireland, Italy, Norway, Poland, Portugal, Russia, Scotland, Spain, Sweden, Wales from 1820–1890. Data source: Willcox (1929).

Years Connected to Railroad by 1890. The number of years since the county is first intersected by railroad to 1890. Data source: Atack, Bateman, Haines and Margo (2010).

Birthplace diversity, 1890. We take $1 - \sum_o (birthplace_{oc}/population_c)^2$, which is simply 1 minus the Herfindahl concentration index for origin o birthplace diversity in county c in 1890. Birthplaces include US or a given country or country grouping abroad. Data source: (Minnesota Population Center, 2011)

Ruggedness. County-level average Terrain Ruggedness Index computed using 30-arc grid data on terrain variability. Data source: (Nunn and Puga, 2012b).

Distance to nearest portage site. Minimum distance from county centroid to the nearest portage site, which is defined as the location where a river basin intersects the fall line. Data source: (Bleakley and Lin, 2012).

Manufacturing Employment Share. County-level percent of employment in manufacturing industries in 1890. Data source: (Minnesota Population Center, 2011).

Distance to Nearest Mine. Minimum distance from county centroid to a site where there was a mineral discovery before 1890. The data is from the Mineral Resources Data System (MRDS) edited by the US Geological Survey. Data source: (McFaul et al., 2000).

Distance to nearest Indian battle sites. Minimum distance from county centroid to major Indian battle sites. The battles sites are digitized using a map from (McFaul et al., 2000).

Immigrant share, 1890. County-level percent of foreign born population in 1890. Data source: (Minnesota Population Center, 2011).

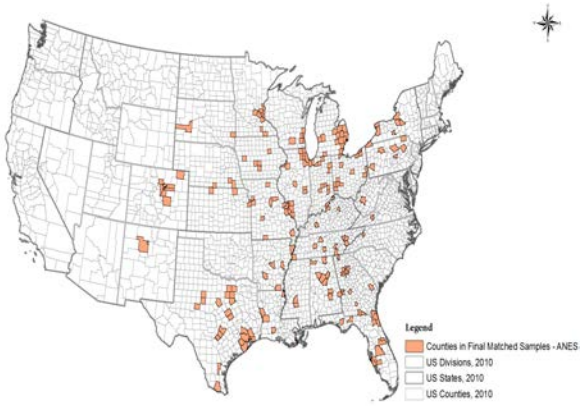
Scottish and Irish immigrant share, 1890. County-level percent of population born in Scotland or Ireland in 1890. Data source: (Minnesota Population Center, 2011)

Slave population share, 1860. County-level percent of slave population in 1860. Data source: (Minnesota Population Center, 2011).

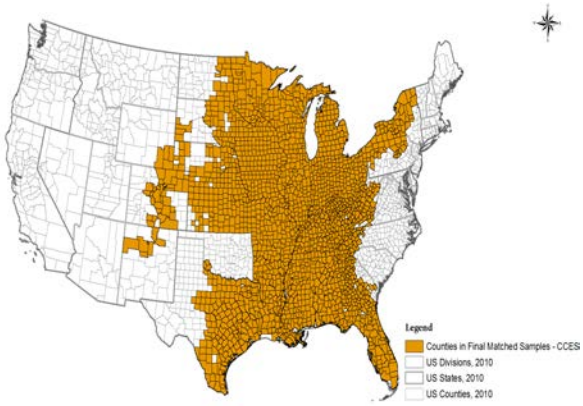
Rainfall Risk. Following Ager and Ciccone (forthcoming), county level rainfall risk is constructed as the variance of the annual average log monthly rainfall from 1895-2000. Data source: (Oregon State University, 2018)

Figure C.1: Data Availability For Main Survey Data Sources

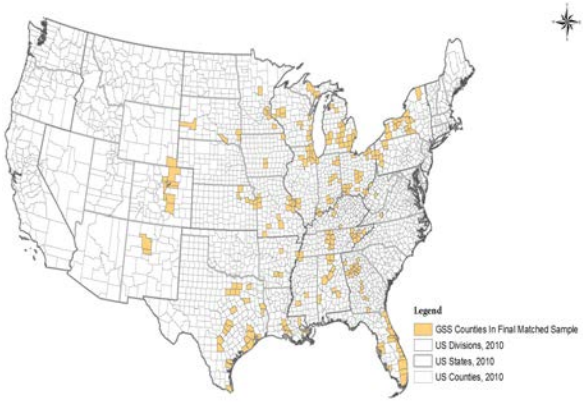
(a) ANES



(b) CCES



(c) GSS



Notes: Figures (a), (b), and (c) provide the geographical distribution of the maximum number of counties available in our baseline sample matched with the ANES, CCES, and GSS data, respectively. Coverage expands to additional counties when incorporating the West Coast sample or extending the historical frontier window to 1950 (see Section 5.5). Note that not all the counties in the above map are included in every baseline regression using the corresponding survey data.