I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14894

A Golden Opportunity: The Gold Rush, Entrepreneurship and Culture

Michael Stuetzer
Abel Brodeur
Martin Obschonka
David B. Audretsch
Peter J. Rentfrow
Jeff Potter
Samuel D. Gosling

NOVEMBER 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14894

A Golden Opportunity: The Gold Rush, Entrepreneurship and Culture

Michael Stuetzer
Baden Wuerttemberg Cooperative State University and Ilmenau University of Technology

Abel Brodeur
University of Ottawa and IZA

Martin Obschonka
Queensland University of Technology

David B. Audretsch
Indiana University

Peter J. Rentfrow
University of Cambridge

Jeff Potter
Atof Inc., Cambridge

Samuel D. Gosling
University of Texas at Austin and University of Melbourne

NOVEMBER 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

A Golden Opportunity: The Gold Rush, Entrepreneurship and Culture*

We study the origins of entrepreneurship (culture) in the United States. For the analysis we make use of a quasi-natural experiment – the gold rush in the second part of the 19th century. We argue that the presence of gold attracted individuals with entrepreneurial personality traits. Due to a genetic founder effect and the formation of an entrepreneurship culture, we expect gold rush counties to have higher entrepreneurship rates. The analysis shows that gold rush counties indeed have higher entrepreneurship rates from 1910, when records began, until the present as well as a higher prevalence of entrepreneurial traits in the populace.

JEL Classification: L26, R12, N5, N9

Keywords: gold rush, entrepreneurship, culture

Corresponding author:
Michael Stuetzer
Professor of Economics and Quantitative Methods
Baden-Wuerttemberg Cooperative State University
Coblitzallee 1-9
D-68193 Mannheim
Germany
E-mail: michael.stuetzer@dhbw-mannheim.de

* We thank Michael Fritsch, Michael Wyrwich, Per Davidsson, Jörn Block, and Sanaz Talifar for helpful comments. Special thanks to Samuel Baszi, Martin Fiszbein and Mesay Gebresilasse for data sharing. Joey Cheng and Sebastian Stuetzer provided excellent research assistantance.

A Golden Opportunity: The Gold Rush, Entrepreneurship and Culture

1. Introduction

Entrepreneurship culture is a key determinant of entrepreneurial activity and subsequent economic growth (e.g., Chinitz 1961; Davidsson 1995; Saxenian 1996; Audretsch and Keilbach 2004; Fritsch and Wyrwich 2017; Glaeser, Kerr and Kerr 2015). While it manifestly differs across regions and can persist over decades and centuries (Andersson and Koster 2011; Fritsch and Wyrwich 2014; Fotopoulos and Storey 2017; Cosci, Meliciani and Pini 2019), its origins remain elusive. Thus, with only a few exceptions, the origins of clusters and deserts of entrepreneurship culture have not been analyzed. In this paper, we trace the spatial variation of entrepreneurship and entrepreneurship culture in the United States (U.S.) back to the gold rushes in the western part of the country, such as in California, Colorado, and Oregon. The gold rushes initiated one of the biggest mass migrations in history, with several hundred thousand people moving to an area largely unsettled by non-natives.[1] The people rushing to the gold fields in response to a previously unavailable economic opportunity were often referred to as Argonauts. The economic incentive for the Argonauts to strike gold made the gold rush migration different from the early colonial settlement of the Thirteen Colonies, which was partly driven by deportation, religious persecution, and slavery.

Here, we argue that the Argonauts were entrepreneurs and more likely to possess a unique constellation of personality traits that differed from the general population. The journey to an undeveloped and demanding environment with the hope of striking gold arguably attracted mostly individuals who possessed personality traits associated with openness (e.g., resourceful, innovative, and curious), conscientiousness (e.g., hardworking, persistent, and cautious), and emotional stability (e.g., even-tempered, steady, and confident) (e.g., Rohrbough 1997; Brands 2002). The high risk-high reward structure of the earnings in the gold fields was also arguably more appealing to people with low risk aversion, low fear of failure, and high self-efficacy. Collectively, this constellation of personality traits is consistently associated with entrepreneurial activity (e.g., Kerr, Kerr and Wu 2017).

The influx and concentration of entrepreneurially-minded people may have led to persistently high entrepreneurship rates over time via three inter-related mechanisms: genetic founder effects, formal institutions, and local culture. The founder effect refers to the pattern whereby new populations are often based on non-representative "founders" who originate

[1] This does not imply that the western part of the U.S. was unsettled at the time of the gold rush. Approximately 150,000 Native Americans lived in California alone at the dawn of the gold rush (Brands 2002, p. 306). However, the Native population shrank rapidly due to diseases, war and displacement.

from a parental population (Diamond and Rotter 1987). The population of founders carries only a fraction of the genetic variation of the parental population. During the gold rush, entrepreneurship-prone personality traits were arguably overrepresented among the migrants. As personality traits are partly heritable (Jang et al. 1996; Lo et. 2017; Abdellaoui et al. 2019; Bleidorn et al. 2021), the entrepreneurial traits common among migrants might be over-represented in future generations. A second potential mechanism is formal institutions (North 1991). The Argonauts accounted for a large share of the population, so they had influence in drafting laws, policies, and constitutions in the western U.S. states. Institutions that support entrepreneurship can persist over time, but their inertia is probably weaker than that of culture – the third potential mechanism. The spatial concentration and repeated interaction of entrepreneurially-minded people can establish shared norms conducive to entrepreneurship. Once such an entrepreneurial culture is formed, children from the early migrants, as well as subsequent generations of migrants socialized into this culture. This process can reinforce the culture and enable it to persist (Bisin and Verdier 2001; Nunn 2012; Zelinsky, 1973).

There is a sizable literature linking historical determinants to subsequent effects on culture and formal institutions (e.g., Nunn 2012; Glaeser, Kerr and Kerr 2015; Couttenier, Grosjean and Sangnier 2017) with the known problem of establishing causality. In our case, gold counties differ from non-gold counties on several characteristics (Brodeur and Haddad 2021). To account for this problem, we split counties with gold discoveries into two groups. The first group – termed gold rush counties – had gold discoveries in the western, sparsely settled part of the U.S. between 1848 and 1899 (Figure 1). The second group, which we term gold mining counties, had gold discoveries at a time when the county was already settled by non-natives – after 1899 in the West or at any point in time in the eastern part of the U.S. (Figure 2). Our identification assumption is that gold mining counties (control group) would enjoy the same high entrepreneurship rates as gold rush counties (treated group) if the former would have been unsettled by non-natives at the time of the discoveries. However, because the control group of gold mining counties was already settled when gold was discovered there, the population share of miners was lower compared to gold rush counties, reducing the genetic founder as well as the culture effect, and making the setup of pro-entrepreneurial formal institutions less likely. Brodeur and Haddad (2021) show that gold rush and gold mining counties do not differ from each other on most geographic variables – leaving the extent of pre-gold settlement as the major distinction.

[Insert Figures 1 and 2 about here]

The Mineral Resource Data System (U.S. Geological Survey) provides information on the date and location of gold and other mineral discoveries in the contiguous U.S., which we map onto present-day counties. Because of its elusive nature, we lack direct measures for entrepreneurship culture. However, entrepreneurship and entrepreneurship culture are endogenous and co-evolve. Therefore, following others (Audretsch and Keilbach 2004; Fritsch and Wyrwich 2017), we rely on entrepreneurship rates as the manifest materialization of the entrepreneurship culture for most of the analysis. The focal analysis starts with the historic self-employment rates from 1910 to 1940 from the U.S. Census[2], which shows that gold rush counties had higher self-employment rates compared to gold mining counties and counties without gold with the exception of 1930. On average, the gold rush increased the 1910 self-employment rate by 0.751 percentage points and the 1940 self-employment rate by 0.353 percentage points, which is roughly 58 percent and 28 percent of the respective standard deviations.

For the period starting from 1978, we use startup rates from Business Dynamics Statistics (BDS) as the entrepreneurship indicator. Gold rush counties have significantly higher startup rates from the 1970s until the 2010s, in contrast, gold mining counties do not have significantly higher startup rates than counties without any gold discoveries. The effect size of the gold rush is relatively constant at around 40 percent of the standard deviations. Regarding present-day entrepreneurship indicators, we also exploit a large-scale personality data set providing information on entrepreneurial personality traits for 3+ million citizens (Gosling-Potter-Internet-Personality-Project – henceforth GPIPP). From this data set, we compute an entrepreneurial constellation of the Big Five traits for the individuals and average this indicator over the county's population to obtain an entrepreneurial personality profile for the counties. We find that gold rush counties score higher on this entrepreneurial personality profile, while gold mining counties do not.

Another part of the empirical analysis concerns the quality of entrepreneurship. Start-ups differ in their quality, with the famous Silicon Valley cluster of high-quality firms located near the Californian gold fields. Startup quality data from the Startup Cartography Project (Andrews et al. 2019), however, shows that gold rush counties do not have higher quality startups compared to other counties.

We conducted several robustness checks and alternative specifications. A first robustness check deals with the choice of the treated and control groups. Because gold mining

[2] Data on self-employment prior to 1910 are not available in the census. Full-count Census data are not publicly available from 1950 onwards.

counties in the Appalachians might not make a suitable control group due to their underdeveloped nature, we drop gold mining counties in the eastern part of the U.S. and solely rely on gold mining counties in the western part of the U.S. as a control group. We also considered as a control group counties that were incorporated or initially organized in the same time period from 1848 to 1899, as was used to define the gold rush counties. These robustness checks support our findings.

For the treatment group, we look at the effect of minerals other than gold. Silver discoveries might initialize the same pro-entrepreneurship effects as gold because it initiated some rushes in Nevada and Colorado. Indeed, silver discoveries between 1848 and 1899 also positively affected entrepreneurship. As a placebo test, we would not expect and indeed did not find the same pro-entrepreneurship effect from iron discoveries in this time period. Another robustness check concerns non-representative nature and limited coverage of the GPIPP personality data used in the analysis. Our results remain stable across most robustness checks.

Additional analyses deal with exiting firms. If an entrepreneurial culture emerged in gold rush counties where failure was tolerated as a rather natural byproduct of taking risks, then entrepreneurs are more likely to cancel entrepreneurial projects falling short of owners' expectations. We use exit data from the BDS to test this hypothesis with the result that the exit rate of firms is significantly higher in the gold rush counties. We further investigate the employment effect of entrepreneurship. The empirical results show that new businesses create more jobs in gold rush counties. Lastly, we investigate channels how the impact of the gold rush on entrepreneurship persisted over time. Our empirical analysis yields some support that the culture effect dominates the founder effect.

Our paper makes two key contributions to the literature. First, we identify a causal origin of historic and present-day differences in entrepreneurship and entrepreneurship culture in the U.S. context. Second, we theorize and present empirical evidence on the mechanisms for regional differences in entrepreneurship. Entrepreneurship culture has been defined as the result of either the concentration of people with entrepreneurial personality traits or the collective programming of the mind (Davidsson 1995). The present results suggest that in the long run, the latter dominates the former.

Our research relates to several strands in the existing literature. We add to the entrepreneurship literature on the persistence of regional differences in entrepreneurship and its origins (e.g., Andersson and Koster 2011; Fritsch and Wyrwich 2014; Fotopoulos and Storey 2017; Cosci, Meliciani and Pini 2019; Bennet et al. 2020).

Our research adds to the literature on the effects of natural resources on institutions and socio-economic outcomes (e.g. Papyrakis and Gerlagh 2004; Mehlum, Moehne and Torvik 2006; Dell 2010; Ross 2015; Couttenier, Grosjean and Sangnier 2017). Research has largely found negative effects for oil and mineral wealth and has emphasized the role of weak or unfavorable formal institutions enacted to exploit the resource wealth as a key mechanism. The exploitation of the silver deposits at Potosí in the Andes (Dell 2010) provides a suitable comparison to our study. The Spanish Empire enacted forced labor (mita) which conscripted a substantial share of the local population to mining. This institution was highly extractive in the sense that it set no incentives to work hard, innovate or act entrepreneurial and did not attract migrants. The long-run effects of these extractive institutions materialize today in lower educational attainment and less provision of public goods. In contrast, the gold rush had long-run positive effects which can be partly attributed to the formation of inclusive institutions setting incentives that kept the entrepreneurial engine running.

We proceed in Section 2 by providing a short overview of the gold rushes and theorizing about their impact on entrepreneurship. Data and methods are presented in Section 3. In Section 4, we provide empirical results. Section 5 discusses and concludes.

2. The gold rush

2.1 A short history of the gold rushes in the U.S.

The *Age of Gold* began in January 1848 when James Marshall discovered the precious metal at Sutter's Mill in present day El Dorado county (Brands 2002). The rumors of Marshall's discovery spread throughout California but were amplified when Sam Brannan presented material evidence on the streets of San Francisco. Following his presentation, people rushed to the gold field, which left the city deserted – a pattern that would soon be replicated elsewhere. As news of the discovery spread across the U.S., Latin America, and the Pacific Ocean, Argonauts flocked to California with dreams of striking gold. While the non-Native population in California was around 8,000 in 1840, it swelled to approximately 120,000 in 1850, before reaching 380,000 in 1860 (Brands 2002).

The California version is the most famous gold rush in the U.S., but it was not unique. Prospectors working northward from California discovered gold in several places in what today is Oregon. Discoveries at Josephine Creek and Jackson Creek initiated the Oregon gold rush in 1851 (Potter 1976). Regions east of California also saw significant rushes. Gold was discovered on the eastern side of the Sierra Nevada. However, the silver discoveries of the Comstock lode around 1859 proved to be even more abundant and led to a substantial migration wave that resulted in the founding of present-day Nevada (Bancroft, 1890). Further east, in the Rocky Mountains, gold was discovered in the Kansas and Nebraska territories. What became known as the Pike's Peak Gold Rush started in 1858 and was the first major non-native population push into Colorado (Ubbelohde, Benson and Smith 1988).[3]

There were even gold discoveries prior to the Californian rush. Miners from the southern part of the U.S. were involved in the Pike's Peak Gold Rush as well as in the Californian gold rush. Gold was discovered as early as 1828 in North Carolina. Soon after, thousands of gold diggers worked in gold fields stretching from Virginia over the Carolinas to Alabama. The southern gold bonanza from 1828 to 1836 can be regarded as a prelude to the Age of Gold, but it did not have the lasting impact that their counterparts in the West had (Young 1982) – which is a topic we return to later and empirically exploit.

2.2 Gold rush and entrepreneurship

In the present research, we posit that migration to the gold fields was essentially entrepreneurship, and that the people who chose to pursue this journey possessed certain traits that proved to be conducive for entrepreneurship. While it is not possible to posthumously hand out questionnaires to gold diggers gathering data on their traits or motivations, we turn to secondary sources summarizing the writings of gold diggers and their contemporaries and make inferences based on historical accounts of the gold rush.[4]

The gold rush offered ample opportunities for entrepreneurial activity. Prospecting and digging for gold were occupational choices that required early investments with uncertain profits in an innovative industry. As such, searching for gold constitutes entrepreneurship. People gave up their previous jobs to take their chance on an opportunity. Regarding investments, the arduous journey to the gold fields was expensive. The cheapest and most frequently used overland route cost between $200 and $300, which was equivalent to half a year's salary for an unskilled worker (Rohrbough 1997; Clay and Jones 2008). Equipment was expensive in California too. In the beginning, shovel, picks, and pans were sufficient to dig for gold. However, mining techniques evolved rapidly requiring larger capital investments which led to the formation of stock-based companies. Additional costs arose for those

[3] We do not include Alaska in our empirical analysis despite its famous gold rush. Alaska as well as Hawaii are special cases for which many historic variables are not available.

[4] We contacted the Bancroft, Beinecke, and Huntington libraries for original sources of letters, diaries, and journals of the early migrants to the West. Unfortunately, most of the material is neither digitized nor transcribed. Marni Sandweiss' *Transcribing the West Project* at Princeton is not yet complete either. Thus, we conclude, that there is no set of transcribed material available covering this period.

Argonauts needing to financially support relatives staying home. As a result, many had to borrow money from friends and family to finance the trip to the West (Rohrbough 1997).

The returns for the gold diggers were modest at best. In California, nominal wages (after board) were six times higher than the U.S. average, but real wages were lower due to the inflated California prices for all goods and services (Margo 1997; Clay and Jones 2008). One also has to take into account the high variance of the earnings distribution. Success in the gold fields depended largely on luck rather than skills. The difference between success and failure was sometimes literally a matter of meters, where in the same riverbed one claim proved to be rich and another claim poor. Failure was a common occurrence in the gold fields as well as in the industries serving the miners. Mining camps sprung up like mushrooms becoming small cities that could just as quickly vanish if prospects turned poor. Higher and surprisingly stable earnings were prevalent in many professional occupations (lawyers, doctors, shop owners, and occupations in the entertainment industries) that provided services to the gold diggers. Sam Brannan, mentioned above, is a case in point. Brannan owned a shop near Sutter's fort and learned of the gold discoveries early on. He considered whether to keep his knowledge private and dig for gold on his own but decided that more money could be made from trading picks, shovels, and pans with the miners than from mining (Brands 2002). Lore has it that Brannan purchased all available pans for 20 cents each, only to subsequently resell them for $15 apiece after he made the discovery public (Rivera 2004).

Beyond acting entrepreneurially, the Argonauts arguably possessed a constellation of traits that was conducive for entrepreneurship. The journey West was long and dangerous, and the prospects of success were far from certain. Only people who had the confidence, resourcefulness, persistence, and ingenuity to withstand and overcome such challenges were likely to head to the gold fields. These are characteristics associated with the Big Five traits of openness, conscientiousness, and emotional stability. Rohrbough (1997, p. 89) notes that the Argonauts "wrote home in wonder and amazement at the openness and honesty of the gold camps." Despite the lack of formal institutions and weak law enforcement, the miners developed institutions for regulating claim sizes, property rights, and water access to provide some order and security (Clay and Wright 2005).

As noted above, mining success often depended on luck and failure was a common companion. Charles Crocker – one of the constructors of the first continental railroad – reflected on his business career in California: "One man works hard all his life and ends up a pauper. Another man, no smarter, makes twenty million dollars. Luck has a hell of a lot do with it" (cited in Brands 2002, p. 441). Claims could suddenly not be worked because a mining company up the river diverted water. Lodging houses on the route to failing camps were given up. Cargo from ships rotted in the harbor of San Francisco if an earlier shipment had satisfied the demand. Banks often failed often on humbug construction projects. The famous Pony Express, which delivered mail in 10 days from Missouri to California closed after 1.5 years in business just two days after the first transcontinental telegraph was established (Brands 2002). Business cycles were indeed shorter in the gold fields.

We argue that in this dynamic environment, entrepreneurial traits such as self-efficacy, risk tolerance, and low fear of failure (Kerr, Kerr and Wu 2017; Obschonka and Stuetzer 2017) were essential and thus arguably more common among the Argonauts than among those who did not make the journey west. Miners were regularly jumping from one claim to another, returning to their previous occupations or trying new ones. Rohrbough (1997, p. 70) notes that "farmers from Illinois did day labor on the streets and in construction; lawyers from Philadelphia unloaded ships […]; doctors from St. Louis waited tables in restaurants or dealt cards in gambling houses […] It was an unformed world in which men took new work identities." Arguably, the Argonauts with the highest risk tolerance and fear of failure prolonged their stay for additional mining seasons or stayed permanently, while others quit and returned home. The ambitions and characteristics of the Argonauts were summarized by John Hittel: "The people who come to California are bold adventurers naturally […] We came here to enjoy an exciting life and make money rapidly. It is no uncommon thing to see men who have been wealthy on three of four different occasions and then poor again […] When men fail they do not despair, they hope to be rich again." (cited in Brands 2002, p. 441).

2.3 Persistence and transmission of entrepreneurship (culture)

We have argued that the people who migrated to the gold fields were acting entrepreneurially and that entrepreneurial traits were more common among them than among those who did not make the journey to the gold fields. Three mechanisms can help explain the potential path-dependency of high entrepreneurship rates in the gold rush regions: genetic founder effects, formal institutions, and entrepreneurship culture.

The concept of genetic founder effects comes from evolutionary biology and refers to the establishment of new populations from non-representative 'founders', who originated from a parent population. Because founders are non-representative, their genetic variation represents a small fraction of the full genetic variation in the parental population (Diamond and Rotter 1987). Founder effects can inform our understanding of the impact that human migration has had on geographical differences, as various parts of the world were settled by small migrating groups of people who shared certain traits, which were reproduced for generations and eventually resulted in uneven distributions of those traits across space. Results from twin studies and genome-wide association studies indicate that individual differences in the Big Five traits have a genetic basis (Jang et al. 1996; Terracciano et al. 2010; van den Berg et al. 2016; Bleidorn et al. 2021). Furthermore, evidence suggests that historical migration patterns are responsible for present-day geographical differences in a host of traits, including the Big Five (Abdellaoui et al. 2019; Bazzi, Fiszbein and Gebresilasse 2020). Thus, within the context of the gold rush, the Argonauts likely shared a constellation of personality traits that made the dangerous journey to the gold fields seem worth it. It is thus likely that among those who chose to stay in the West, those traits would have been passed on to their offspring and subsequent generations. Over time, this founder effect would have resulted in a large proportion of residents with traits similar to those of the early gold miners. Because traits such as openness, conscientiousness, self-efficacy and risk tolerance are related to individual entrepreneurial activity (Kerr, Kerr and Wu 2017; Obschonka and Stuetzer 2017), the gold rush regions can enjoy high present-day entrepreneurship rates.

Formal institutions comprise written laws, regulations, and bureaucracy, which define the boundaries of interactions in several domains. They set the benefits and costs of economic behavior and thereby define the individual incentive structures (North 1991). A growing strand in the economic literature has found that formal institutions have long-term economic consequences (Acemoglu and Robinson 2012). The gold rushes in America's West present a special case because states and the accompanying formal institutions were initially formed by the non-native settlers (Couttenier, Grosjean and Sangnier 2017). Historical accounts reveal that early settlers (including miners) played a key role in the drafting of the western states' constitutions and influenced politics via elections (Ubbelohde, Benson and Smith 1988; Brands 2002). Thus, it is reasonable to assume that their pro-entrepreneurial attitudes had a profound impact on the formation of institutions that supported entrepreneurship. However, the impact that formal institutions have in explaining differences in and the persistence of entrepreneurship is of lesser interest in the empirical section of the paper because we analyze county-level variation in entrepreneurship using state-fixed effects. Thus, the regression analysis in section 4 reveals only the lower boundary of the true impact of the gold rush at the county level, ignoring effects on the formal institutions at the state level.

Nunn (2012) defines culture as decision-making heuristics, which typically manifest themselves as values, beliefs, or social norms. Culture is often referred to as an informal institution or set of unwritten rules because it also defines payoffs to human actions (Fritsch and Wyrwich 2014). These decision-making heuristics or mental models provide guidance when making decisions in complex situations by encapsulating other people's evaluation and approval of possible actions. Culture is learned. Quite literally, the programming of individual mental models often happens early in life by family and friends (Hofstede 2001). As a result of this vertical transmission to the next generation, culture can persist over time (Bisin and Verdier 2001; Nunn 2012). Zelinsky (1973) makes a similar prediction with his Doctrine of First Effect Settlement, where he posits that the first dominant settlers – regardless its size – in an area have a major influence on the development of the norms and culture.

The mental programming of the mind approach can explain how an entrepreneurial culture can persist over time and lead to persistently high entrepreneurship rates in some regions. However, it does not explain how an entrepreneurial culture initially formed in the gold rush regions because the regions were largely settled by groups of non-natives coming from different regions, having different occupational backgrounds, and different life experiences. How could a single coherent culture consolidate from a collection of such divergent influences? Denzau and North (1994) explain how different individual mental models can become more similar by repeated interactions and information exchange in the same external environment between individuals. As a result of this horizontal transmission, a shared mental model consolidates from the converging individual mental models. In the gold rush regions, the repeated interactions of the entrepreneurially acting gold rush population can establish shared values and norms that are conducive to entrepreneurship. Indeed, if entrepreneurship is a common way to earn a living, risk taking becomes normalized, and trying again after failure is common, then these often observed behaviors will first become accepted and tolerated and subsequently expected as the norm. The formation of an entrepreneurial culture was arguably accelerated in the early years in the gold rush regions by the absence of other norm-inducing elements, such as religious authorities, and the friends and families of the Argonauts (Brodeur and Haddad 2021).

Taken together, founder effects, entrepreneurship-friendly formal institutions, and the formation and persistence of an entrepreneurial culture can explain present-day entrepreneurial activity in the gold-rush regions. It is important to emphasize that these mechanisms are not independent of each other, but operate and co-evolve together. It is uncontroversial to claim that formal institutions and culture influence each other. However entrepreneurial traits, which are the basis of the founder effects, can also be influenced by formal pro-entrepreneurial institutions and an entrepreneurial culture via selective migration and acculturation (Rentfrow, Gosling and Potter 2008).

With respect to migration, one might reasonably expect that continued migration into a region increases variation in traits. However, even continued migration can be selective. Selective migration can continue if the conditions persist that select for entrepreneurial traits. Beyond persisting conditions, research on the psychology of selective migration has revealed that personality traits play a role in decisions to migrate. High levels of openness and extraversion, and low levels of agreeableness have been associated with migration (Jokela, 2009). One explanation for the link between personality and migration is that individuals selectively migrate to places that satisfy their social and psychological needs. Indeed, people report higher well-being if their individual personality is congruent to the personalities of the people in their region (Bleidorn et al. 2016). Fit between personality traits and characteristics of the environment can thus influence migration decisions. In related research Bazzi, Fiszbein and Gebresilasse (2020) show that 1) people migrating to the Western frontier had more often infrequent names signaling individualism compared to those who stayed in the East and 2) people migrating East from the Western frontier were less individualistic. Gold rush regions with a high prevalence of entrepreneurial traits arguably attracted like-minded people stabilizing and reinforcing the existing trait structure.

Regarding acculturation, even personality traits such as the Big Five have a malleable component that is conducive to change (Jang et al. 1996; Bleidorn et al. 2021). Theoretical models in psychology predict that an individual's trait structure can be influenced by formal institutions as well as social norms and behavioral tendencies (Rentfrow, Gosling and Potter 2008). As a regional aggregate, traits thus behave like modeling clay by having a tendency to persist but are also subject to change.

3. Data and Methods

We use a manifold of variables in the analysis; as such, below we describe only briefly their source and refer readers to the Online Appendix Table A1 for more details. The descriptive statistics and correlations of the variables are presented in the Online Appendix Table A2.

3.1. Dependent variables

Entrepreneurship and entrepreneurship culture co-evolve, so we rely on entrepreneurship rates as the manifest materialization of an entrepreneurial culture. The most suitable indicator for entrepreneurship is the startup rate, which is defined as the number of new businesses per 1,000 employed in each year (Fritsch and Wyrwich 2014). We average the startup rate over decades to avoid spurious fluctuations. Data on new businesses come from the public use

BDS (Census Bureau 1978-2019). Startup data for historic entrepreneurial activity are available from the BDS until 1978, which allows us to compute the average startup rates for the 1970s to the 2010s. For previous time periods, we use self-employment rates which are available for the years 1910, 1920, 1930 and 1940 from the U.S. Census (Ruggles et al. 2020).

We also use data on the quality of startups (Startup Cartography Project 2010-2016). This data set captures the quality of startups based on observable characteristics such as registering as a corporation or registering in Delaware to facilitate equity financing, the name of the firm (which can signal its high-tech nature), and creating innovations measurable as patents (Andrews et al. 2019). From this data set, we extract the quality-adjusted number of startups (RECPI) and compute the high-quality startup rate as the quality-adjusted number of startups per one million employed in the period from 2010 to 2016.

Beyond manifest entrepreneurial activity, we try to measure the more latent entrepreneurial personality structure in counties from the GPIPP (2003-2015) by means of a regional prevalence of an entrepreneurial personality profile. Here we focus on the Big Five personality traits which are central to the entrepreneurial personality system of individuals. They directly relate to individual entrepreneurial activity. Meta-analyses and reviews conclude that individual entrepreneurial activity is more likely for people scoring high in Openness (O), Conscientiousness (C), Extraversion (E), and low in Agreeableness (A) and Neuroticism (N) (Zhao & Seibert 2006; Brandstätter 2011; Kerr, Kerr and Wu 2017). The Big Five traits also affect entrepreneurial activity indirectly via the channel of more task-specific traits such as risk-tolerance, self-efficacy, and locus of control (Obschonka and Stuetzer 2017). A growing stream of research suggests that an entrepreneurial constellation of the Big Five traits (high in O, C, E and low in A and N) is a better predictor of entrepreneurship than the separate effects of the traits (for an overview, see Obschonka and Stuetzer 2017).

Regional analyses have shown that the individual Big Five traits are not equally distributed but clustered in space (Rentfrow 2010; Abdellaoui et al. 2019). Research regarding the entrepreneurial personality profile shows that the regional profile correlates with regional entrepreneurship rates in different countries, including the U.S., United Kingdom, and Germany (Obschonka et al. 2015; Stuetzer et al. 2016; Fritsch et al. 2018). The spatial variation of the entrepreneurial personality profile can be traced back to major events, such as the Industrial Revolution in the case of the United Kingdom (Stuetzer et al. 2016), and city foundations during Roman times in the case of Germany (Fritsch et al. 2020).

Detailed information on the construction of the GPIPP data base is presented in the Appendix Table A1. In a nutshell, the data are collected via a non-commercial webpage operating since 1998. We restrict this data set to respondents providing valid ZIP codes attributable to counties. To ensure measurement precision, we drop counties with fewer than 100 observations in the personality dataset. After these deletions, the dataset contains information from N=3,159,083 respondents in 2,084 counties (mean sample size per county = 1,516). The Big Five personality data were collected using the standard 44-item BFI inventory (John and Srivastava 1999). We first compute the entrepreneurial personality profile at the individual level and then take its average over the counties as the dependent variable.

Figure 1 maps the spatial distribution of the 1910 self-employment rate, the startup rate in the 2010s, and the entrepreneurial personality profile.[5] A visual inspection of the maps and the correlations depicted in Table A2 (Online Appendix) confirm the persistence of the regional distribution of entrepreneurship over time.

3.2. Independent variables

The MRDS provides information on mineral discoveries (U.S. Geological Survey). We restrict the data base to only those discoveries for which the discovery year is available. To discriminate between minor and economically viable discoveries, we further restrict the data set to gold discoveries classified as 'primary commodities' by the MRDS. We also exclude a few discoveries in the Pacific and Mountain Census regions prior to 1848 from the analysis, because the documented history of these discoveries is vague and they did not induce a gold rush.[6] With these restrictions, our dataset includes gold discoveries in 300 counties in the contiguous U.S. with discovery years stretching from the 18th into the 21st century.

Based on the year and location of the gold discoveries, we categorize gold discoveries on the basis of whether they are part of the gold rush or not. We define counties as gold rush counties if they had a gold discovery between 1848 and 1899 in the Pacific or Mountain Census divisions. We also include the few counties with discoveries during that time period in some states in the Midwest Census region, which were sparsely populated in 1850 (Michigan, Minnesota, South Dakota, and Wisconsin). We refer to this spatial definition henceforth as the *West*. We define as a comparison group counties that had a gold discovery in the West after 1899 or in the East at any point in time. Based on these definitions, there are 197 gold rush counties (treated group), 103 gold mining counties (comparison group), and 2,809 counties without gold discoveries in our dataset. Figure 2 maps the regional distribution of the gold

[5] Figure A1 in the Online Appendix maps the distribution of all entrepreneurship rates over time.

[6] An additional reason for this exclusion is the absence of Census data on the inhabitants of the West prior to 1848.

rush and gold mining counties. These maps confirm that there are many gold rush counties in California, Nevada, and Colorado. Gold mining counties are scattered in the West and clustered around the Appalachians in the East.

3.3. Geographic control variables

We use a number of control variables to account for other influences on the spatial distribution of entrepreneurship (culture). The first set of controls includes exogenous features. Beside gold, other resources, such as coal and iron (Tarr and McCurry 1910), as well as petroleum (Lujala, Ketil and Thieme 2007) were highly relevant for the economic development in the U.S. Therefore, we use dummy variables indicating whether a county had at least one iron field, coal field, and petroleum field. We expect these variables to be negatively related to entrepreneurship and culture because these resources were connected to large-scale industries (Glaeser, Kerr and Kerr 2015; Stuetzer et al. 2016).

Market access plays an important role for regional economic development, especially in the 19^{th} century U.S. (Atack et al. 2010). To capture market access, we include dummies indicating coastal counties (NOAA 2017), access to portage sites (Bleakley and Lin 2012), and access to railroads in 1910 (Bazzi, Fiszbein and Gebresilasse 2020). We further control for a county's distance to major lakes, rivers, and its state capital (Manson et al. 2019).

Prior to industrialization, regional wealth depended on agrarian productivity (Combes et al. 2010). We use potential agricultural yields based on climate and water supply to capture agrarian productivity (IIASA and FAO 2012). We additionally use county's elevation mean and standard deviation to capture topological restrictions on agriculture.

What today constitutes the contiguous U.S. was controlled by different colonial powers from the 16^{th} century until the mid-19^{th} century. Besides the British, the Spanish Empire temporarily controlled Florida and the southwest from California to Texas. New France stretched from the northeast over the Great Lakes along the Mississippi to Louisiana. Although these regions were thinly populated by non-natives compared to the 13 British colonies, it is possible that informal institutions spawned by these colonial powers have endured (Zelinsky 1973). We capture the Spanish influence with the share of the population having Hispanic ties for the period between 1850 and 1870. A similar control using historic census data is not available for French influence. We instead use the share of the regional population speaking French (incl. Patois and Cajun) and French Creole in 2010 (American Community Survey 2010).[7]

We lack direct controls for the historical presence of Native Americans, who either died of imported diseases, were decimated rapidly after the arrival of the colonists or which were forced to settle in reservations because the settlement of the U.S. had substantial elements of conquest and displacement. Prior to colonist's arrival, the Native American population density was arguably related to the same geographic controls such as potential agricultural yield, proximity to lakes and rivers, as well as elevation, which we already use. However, we capture the present-day distribution of Native Americans by using its population share in the counties as described below in Section 3.4. Finally, we include counties' latitude and longitude as completely exogenous controls for regional differences.

3.4. Socio-economic control variables

While the above described geographic controls are mainly exogenous in nature, we also employ endogenous socio-economic control variables measuring the industrial structure, unemployment and the racial composition of the population. Our aim is to have a similar set of control variables throughout the decades. The industrial structure is measured with the employment share in construction and manufacturing, services, and government with the employment share in agriculture serving as the reference category in the regressions (Ruggles et al. 2020; Bureau of Economic Analysis 1970-2020). The industrial structure is included as a control because of differing costs and opportunities of entrepreneurship in different sectors (Glaeser and Kerr 2009).

Unemployment is measured with the number of unemployed as the percentage of the laborforce from 1970 onwards (Bureau of Labor Statistics 1976-2020). County level data on unemployment are not uniformly available for the 1910 to 1940 time period. As a substitute, we turn to the laborforce participation rate of the population aged 16 years or older for this period (Ruggles et al. 2020). Unemployment might be positively linked to entrepreneurship because it lowers the opportunity costs of starting up (Koellinger and Thurik 2012).

The racial composition of the population is measured with the population share of Caucasians, Black Americans, Native Americans, and Asian Americans and Pacific Islanders (hereafter referred to as Asian Americans) (Census of Population 1970, 2000-2010, Haines 2018; Ruggles et al. 2020). We include this control because minority groups might be

[7] We do not consider the minor colonial powers such as the Netherlands, Sweden, and Russia, which also had some presence in the contiguous U.S., because their impact was rather limited or their colonies short-lived.

discriminated against in the labor market and pushed into entrepreneurship to earn a living (Waldinger, Aldrich and Ward 1990).

Lastly, we control for population density as a catch-all variable for different regional characteristics such as land prizes and knowledge generation (Manson et al. 2019).

4. Regression Results

4.1 Regression model

We test the hypothesis that gold rush counties have higher entrepreneurship rates and a stronger entrepreneurship culture by comparing counties that experience a gold rush with gold mining counties as well as all other counties. Our baseline equation is:

$$Y_c = \alpha + \beta \cdot gold\ rush + \gamma \cdot gold\ mining + \delta \cdot X_c + \theta_S + \varepsilon_c$$

where Y_c are several entrepreneurship indicators such as start-up rates and the entrepreneurial personality profile. The dummy variable *gold rush* is our variable of interest and indicates whether a county experienced a gold rush between 1848 and 1899 in the West. The dummy variable *gold mining* takes the value of one if the county reported a) a gold discovery in the West in 1900 or after, or b) a gold discovery in the East at any point in time. Our empirical strategy relies on comparing gold rush to gold mining counties because we expect the gold rush dummy to be a significant predictor of entrepreneurship while the gold mining dummy should not be. X_c represents a vector of control variables at the county level and θ_S are state-fixed effects. We use robust standard errors and cluster the error term ε_c at the state level.

4.2 Results regarding entrepreneurship rates

Our basic set of regressions are displayed in Table 1 with the historical self-employment rates from 1910-1940, the startup rates for the decades starting from the 1970s to 2010s, and the regional entrepreneurial personality profile as the dependent variables. In Panel A we estimate a model with only state-fixed effects as control variables while in Panel B we add geographic controls. Panel C additionally takes socio-economic controls into account. The full regression results of Panel C are depicted in the Online Appendix Table A3. We find that counties with a gold rush enjoyed higher entrepreneurship rates, whereas counties with gold mining did not. The only exception is in Model 3 of Panels B and C with the 1930 self-employment rate, where the gold rush does not consistently predict self-employment. However, we note that 1930 was in the midst of the Great Depression, which can be regarded as a temporary structural break in U.S. economic history, affecting risk taking and availability of credit (Calomiris 1993, Cogley and Sargent 2008), which are two major determinants of entrepreneurial activity. Panel B additionally shows the results of a Wald test directly comparing the size of the coefficient of the gold rush dummy and the gold mining dummy. This one-sided test confirms our expectation that gold rush counties have higher entrepreneurship rates than gold mining counties. In order to compute effect sizes, we utilize the Panel B results. In this setting the gold rush increased the 1910 self-employment rate by 0.751 percentage points and the 1940 self-employment rate by 0.353 percentage points, which translates into 58% and 28% of the respective standard deviations. When looking at the more recent startup rates and the regional entrepreneurial personality profile, we find relatively stable effect sizes of the gold rush dummy at around 40% of the standard deviations.

[Insert Table 1 about here]

4.3 Results regarding high quality entrepreneurship

In this section, we look at contemporary high-quality entrepreneurship rather than the quantity of entrepreneurship (Guzman and Stern 2015). Figure 2 shows that gold rush counties are located near clusters of high-quality startups such as Silicon Valley. We use the quality-adjusted number of startups per one million employed as the dependent variable in the analysis presented in Table 2.[8] The results clearly show that the gold rush does not directly predict contemporary high-quality entrepreneurship. Interestingly, very few control variables are related to high quality entrepreneurship, which suggests that the rise of clusters[9] like Silicon Valley in the U.S. cannot be explained by a set of variables for everyday entrepreneurship but are rather exceptional developments. The emergence of clusters in the U.S. was often aided by military spending on new technologies with research universities playing an important role (Saxenian 1996). Such research institutes may offer indirect links between the gold rush and entrepreneurship (culture). For example, Stanford University was founded by Leland Stanford – an Argonaut, successful businessman and politician. These indirect links are not captured by our econometric approach.

[Insert Table 2 about here]

4.4 Robustness checks

Our findings are robust to alternative specifications of the control group, which consisted of counties with gold discoveries in the West after 1899 and a gold discovery in the eastern part of the U.S. at any point in time. Many counties in this control group are located in the

[8] Models displaying the controls are available in the Online Appendix Table A4.

[9] The Gini coefficient of the quality-adjusted number of startups over the period from 2010 to 2016 is 0.94.

Appalachians where gold was discovered in the first half of the 19th century and initiated a gold rush to these regions. However, as these regions were already settled, we do not expect that the influx of gold miners led to a genetic founder effect, the formation of an entrepreneurial culture, nor the creation of entrepreneurship-friendly formal institutions. One might argue that the Appalachians do not make a suitable control group to the gold rush counties due to their underdeveloped nature. Thus, as a robustness check, we restrict our control group to non-Appalachian gold mining counties. Table 3 (Panel A) repeats the main analysis for the entrepreneurship indicators used before. Our results remain unchanged – the gold rush dummy still predicts the entrepreneurship indicators while the gold mining dummy does not. A Wald test also confirms that gold rush counties have higher entrepreneurship rates than gold mining counties. Additionally, we construct another control group based on the county incorporation dates from the Atlas of Historical County Boundaries (2021). The main idea behind this robustness check is to construct a comparison group of counties that were incorporated during the same time period from 1848 to 1899 as was used to define the gold rush counties. Another version of this test uses the initial land organization date of the counties to construct the comparison group of counties that were organized during the gold rush period from 1848 to 1899. Our results in Panels B and C in Table 3 largely shows the gold rush counties still enjoy higher entrepreneurship rates compared to the control groups.[10]

[Insert Table 3 about here]

The second set of robustness checks deals with alternative specifications of the treatment group by exploring the effects of precious metals other than gold. Especially in Arizona, Colorado, and Nevada a silver boom occurred at roughly the same time as the gold rush. Unsuccessful gold diggers often left their gold claims to dig for silver. As the silver boom occurred at roughly the same time, involved partly the same people and set in motion the same processes, we expect a silver rush dummy to positively predict entrepreneurship rates – albeit to a somewhat smaller degree because the silver rush did not have quite the same magnitude as the gold rush. To test this hypotheses, we compute a silver rush dummy taking the value of one in case of a silver discovery in the period between 1848 and 1899 in the West. Accordingly, we also reconstruct a silver mining dummy for silver discoveries later in the West or at any point in time in the East. The results depicted in Table 4 Panel A and show that the silver rush often predicts the entrepreneurship indicators. As expected, the point

[10] Models displaying the controls are available in the Online Appendix Table A5-A7.

estimates are a bit lower compared to the gold rush dummy and in some models the silver rush counties do not enjoy higher entrepreneurship rates compared to silver mining counties.

We also look at another metal besides gold and silver. Iron played a vital role in the industrialization of the economy, but historical accounts do not mention an iron rush. We exploit this by examining the effect of iron discoveries over time. Again, we compute an iron rush dummy for discoveries in the West between 1848 and 1899 and an iron mining dummy for discoveries later in the West or at any point in time in the East.[11] We do not expect in this placebo test that the iron rush dummy predicts entrepreneurship rates, because iron mines are rather large-scale operations, that do not attract an entrepreneurially-minded workforce. The results of this placebo test are shown in Table 4 Panel B and reveal indeed that the iron rush does not positively predict entrepreneurship.[12]

[Insert Table 4 about here]

The third set robustness checks centers on the GPIPP personality data set. Recall that we initially dropped counties with fewer than 100 observations in the GPIPP data set to ensure measurement precision, which overwhelmingly affects the Midwest and Pacific Census region. Another issue of the GPIPP data set is that it is not representative but skewed toward younger and female respondents. We re-compute the entrepreneurship prone personality profile while weighting the individual observations to match the regional age x gender distribution and run a robustness check including all counties regardless of the number of respondents. The results are presented in Table 5 and show that the gold rush still predicts the entrepreneurial personality profile as well as that gold rush counties have higher entrepreneurship rates compared to gold mining counties.[13]

[Insert Table 5 about here]

The fourth set of robustness checks revolves around the spatial nature of the relationship between the gold rush and entrepreneurship. Computing Moran's I revealed the presence of spatial autocorrelation of the dependent variables with a weighing matrix containing the inverse distance of the counties from each other. We, thus, run a series of additional analysis described in more detail in the Online Appendix Section 2 (Tables A11-

[11] Note that we used the dummy iron field in previous settings as control variable. The data on iron fields stem from a digitized map of iron deposits in 1909 in Tarr and McCurry (1910). For the placebo test, however, we use the MRDS data set which has additional information on discovery dates allowing us to discriminate between iron discoveries over time in order to compute iron rush and iron mining dummies.
[12] Models displaying the controls are available in the Online Appendix Tables A8-A9.
[13] Models displaying the controls are available in the Online Appendix Table A10.

A13). In a nutshell, we reran the basic regressions with the spatial weighting matrix with largely similar results. Further analysis also show that gold rushes in nearby counties do not matter for entrepreneurship in the specific county under study underlining the local nature of the processes that made the gold rush counties entrepreneurial.

4.6. Additional results

To further explore the emergence and the effects of an entrepreneurial culture, we undertake additional analyses regarding firm exits, job creation, and transmission channels. First, the exit of firms is an integral part of a vital market economy. However, there is often a negative connotation of exit, with failure resulting in stigmatization of the "failing" entrepreneur (Ucbasaran et al 2013). Even in the U.S. there are regional differences in the sensemaking of exits. For example, the press coverage of exiting firms more often describes the exits as mistakes of the entrepreneur in Chicago, New York, or Washington, but predominantly describe them as misfortunes in Atlanta, Austin, and San Francisco (Cardon, Stevens, and Potter 2011). If the gold rush created an entrepreneurial culture where risk taking is rewarded, failure is more accepted, and restarting is encouraged, then entrepreneurs would presumably be more inclined to terminate businesses falling short of their expectations rather than sticking with them in fear of stigmatization. Thus, we would expect a higher exit rate in the gold rush counties. We use data on exits from the BDS from the 1970s until today and compute the exit rate as the percentage of active firms to be terminated. The results for each decade are depicted in Table 6 and show that the exit rate is higher in gold rush counties. Additionally, the direct comparison of gold rush counties with gold mining counties confirm that the former have higher exit rates than the latter. [14]

[Insert Table 6 about here]

A second additional analysis looks at the job creation effect of entrepreneurship. We compute the job creation rate as the number of jobs created by new births (based on BDS data from the 1970s until the 2010s) over total jobs. Our expectation is that the job creation rate is higher in gold rush counties. The results are displayed in Table 7 and confirm this expectation. The Wald test again shows significant higher job creation rates in gold rush counties compared to gold mining counties.[15]

[Insert Table 7 about here]

[14] Models displaying the controls are available in the Online Appendix Table A14.
[15] Models displaying the controls are available in the Online Appendix Table A15.

The third additional analysis investigates more deeply the mechanisms underlying the higher entrepreneurship rates in gold rush counties. Our theorizing highlighted the importance of rushes to counties unsettled by non-natives prior to the gold discoveries. The idea that first-settler groups are critically important for cultural geography was first popularized by Zelinsky (1973). In his Doctrine of First Effect Settlement, he argues that the first dominant settlers – regardless its size – in an area have a major influence on the development of the norms and culture. The development of the national U.S. culture is a case in point in which the first ten thousand largely British settlers had a much stronger imprint on the national culture than did the millions of migrants arriving later. Our setting of regional differences in entrepreneurship is ideal for testing Zelinsky's doctrine. More specifically, we test whether the population density at the year of the first gold discovery has an effect on subsequent entrepreneurship rates. In counties with few or no settlers when gold is discovered, the arriving entrepreneurial Argonauts should have a big impact on subsequent local culture, eventually resulting in high entrepreneurship rates measured many years later; but in counties that are already settled, the impact of the Argonauts on culture should be weaker. Thus, in line with Zelinsky (1973), we expect a negative non-linear relationship between the population density and the entrepreneurship rates – high entrepreneurship rates for countries with very low population densities at the time of the first gold discovery and a steep decline of the entrepreneurship rates with rising population densities.

For the empirical analysis, we compute the population density in each county at each year based on the Census population data from 1790 to 2010 (Manson et al. 2019). We thereby follow closely the methodology of Bazzi, Fiszbein and Gebresilasse (2020). For intercensal years, we interpolate county-level population by assuming a constant annual population growth rate, which matches the decadal growth rate. We thereby replace initial zeros with one to avoid infinite growth rates. This procedure provides us with a population number for each county each year. We then compute the population density for each county at the time of the first gold discovery. The analysis is therefore restricted to the 300 counties with gold discoveries.

Figure 3 plots the relationship between the population density at the year of the first gold discovery on entrepreneurship rates for selected time periods.[16] The figure shows that entrepreneurship rates can be very high in counties unsettled by non-natives, but drop sharply even for small increases in population densities, which is consistent with Zelinsky's (1973).

[16] Figure S2 contains the plotted relationships for all periods.

[Insert Figure 3 about here]

This pattern is in line with the culture mechanism described in Section 2.3. A very low population density at the year of the first gold discovery is equivalent with the Argonauts being the first dominant settler group in a county. Our theorizing argued that an entrepreneurial culture forms from their shared viewpoints on taking advantage of opportunities and risk taking. Later migrants into these regions then socialize into this existing entrepreneurial culture. In contrast, a population density of larger than zero at the year of the first gold discovery means that farmers had already settled the county prior to the gold discoveries. In this case, a farmer culture emerges from the first settlers into which later migrants socialize. Such a farmer culture is arguably less entrepreneurial, because agriculture does not require the same level of risk taking, innovativeness, and openness to new ideas compared to the rapidly evolving mining business. Thus, the culture in the counties is distinctively different when, for example, 1,000 Argonauts arrive first in a county and with 1,000 farmers arriving later versus when 1,000 farmers arrive first with 1,000 Argonauts arriving later. Figure 3 also shows that there is little difference in entrepreneurship activity regarding the level of prior farmer settlement. The entrepreneurship rates are only slightly lower for pre gold discovery population densities of 50 persons per square mile compared to 30 or 20 persons per square mile. This pattern again is in line Zelinsky's argument that the size of the first settler group matters little for its contribution to the culture.[17]

In addition to the visual inspection of Figure 3, we regress entrepreneurship on several transformations of this population density variable to establish further evidence for the non-linear relationship between the pre-gold population density and entrepreneurship. Fractional polynomial regressions suggest that the square root of the population density contributes to explaining entrepreneurship rates, but not in the case of the entrepreneurial personality profile for which the natural logarithm is a better predictor. Due to these differences, in Table 8 we present the regression results using the linear term of the population density in combination with its square root (Panel A) or the natural logarithm (Panel B).

[Insert Table 8 about here]

[17] Our theorizing also discussed the role of a genetic founder effect, which could potentially also explain the negative relationship between the population density at the year of the first gold discovery and subsequent entrepreneurship. The larger the farmer population prior to gold discoveries, the larger will be the share of the heritable traits of the farmers and the smaller will be the share of the heritable traits of the Argonauts among their offspring, resulting in a more linear relationship between the pre-gold population density and entrepreneurship in Figure 3. We, thus, conclude that the genetic founder effect is likely of lesser importance in the gold rush case.

Note that these regressions do not include any control variables because the population density is strongly correlated with variables, such as the potential agricultural yield, the coastal location, as well as latitude and longitude. We also refrain from including state-fixed effects in this regression. The main reason for this omission is an intricate relationship among the gold discoveries, the early population density, land organization, and statehood in the western U.S. In particular, gold discoveries initiated rushes, which led to higher population figures; then higher populations set in motion the process of forming territories and subsequently states, which organized the territory into counties. Only in these organized counties, was the Census conducted, providing the population figures for our analyses. As a result of this sequence, in many western states the initial land organization, but also to a certain degree the final incorporation of the county, is similar within a state. State dummies are thus correlated with the early population figures and gold discoveries with causality partly running from population numbers and gold discoveries to the states. However, when we control for state-fixed effects, neither the linear term nor the non-linear term of the population density at the year of the first gold discovery are significant predictors of the entrepreneurship rates, which would suggest that neither the genetic founder effect nor the culture effect can explain the higher entrepreneurship rates of the gold rush counties. When, however, the state-fixed effects are excluded, the results of the regression in Table 9 show that the linear term of the population density is small and often statistically insignificant but the squared or logarithmic term predict entrepreneurship rates.

5. Discussion and conclusion

This paper investigated the origins of entrepreneurship and its culture in the U.S. We find a positive effect of gold rushes between 1848 and 1899 on entrepreneurship in the western U.S., which persists both over time with stable effect sizes (Table 1 Panel B). The control group of gold mining counties do not exhibit higher entrepreneurship rates. We argue that the persistence of entrepreneurship in gold-rush counties can be attributed to two effects unique to those counties: genetic founder effects and the formation of an entrepreneurial culture, while a third potential mechanism – the setup of entrepreneurship-friendly formal institutions at the state level – is controlled for via state-fixed effects. We posited that the Argonauts journeying to the gold fields were entrepreneurs pursuing an opportunity to get rich by either striking gold or providing services to the miners. We argue that, compared to those who did not set out to strike gold, the Argonauts would have been likely to score higher on several personality traits such as openness, risk tolerance, and low fear of failure, which are important drivers of entrepreneurial behavior. As the Argonauts constituted a large population share in the gold rush counties, these traits can be overrepresented among their descendants. The repeated interaction of entrepreneurially minded people may have led to the formation of an entrepreneurial culture in which subsequent migrants and their offspring were socialized. The empirical results from our analyses indicate that entrepreneurial culture played a key role in maintaining entrepreneurial activity in gold-rush regions.

One key question left unanswered in the empirical part of the paper is the role of formal institutions. These were controlled away by state-fixed effects to account for omitted factors that could bias regression results. However, in reality, state-level institutions were key to jump start the entrepreneurial engine and keep it running in the states most heavily affected by the rushes such as California, Nevada, and Colorado. As the rush westwards led to a massive increase in population numbers, territories and subsequently states formed. Given the highly democratic nature of the U.S. political system, the institutions and subsequent laws responded to the needs of the Argonauts by being inclusive and setting incentives for entrepreneurship. For example, the ban of non-compete clauses in California stems from the the early 1850s and has been beneficial for access to entrepreneurial opportunities (Gilson 1999). Together with the entrepreneurial culture and the entrepreneurial traits of the population, the early pro-entrepreneurial formal institutions were the fuel initially boosting the entrepreneurial engine. Later on, we do find evidence for a stronger effects of the entrepreneurial culture and a weaker effect of the entrepreneurial traits.

The gold rush case in the western U.S. is in stark contrast to the developments in South America, where rich silver deposits were exploited in the Andes. Firstly, the Andes silver deposits were located in regions that were already densely populated and are, therefore, more comparable to the gold mining U.S. counties instead of the gold rush counties. Second and more importantly, the economic institutions set in place mirrored the political institutions by being highly extractive in a way that they did not encourage entrepreneurship. The mita system of forced labor, which characterized the Andes silver mines, was such an extractive institutions designed to extract rents for members of colonial elite while leaving the indigenous population at a subsistence level. Even today, the mita regions have fewer roads and a lower educational attainment than other regions do (Dell, 2010). These factors hinder entrepreneurship and do not attract an entrepreneurially-minded populace – quite the contrary, as Dell (2010) reports outmigration from those regions. Also at the national level, the extractive institutions in South America persisted over time well into the 20th century

(Acemoglu and Robinson, 2012). Taken together, the extractive institutions did not induce a rush of people with entrepreneurial traits into the region and no entrepreneurial culture could form under this set of institutions.

Comparing the U.S. case with the South American case, points to the important role of formal institutions as an enabling and supportive factor for the genetic founder effect and the culture effect. The vastly different trajectories of the gold-rush regions of the U.S. and the silver regions of South America leaves one to wonder how the development of the Northern America would have turned out if the Spanish Conquistadors exploring the south-western part of the United States in the 16th century had discovered the nearby gold deposits in the Appalachians.

A related topic not analysed in this paper is change. While we find persistently high entrepreneurship rates in the gold rush counties, some gold rush counties have started to lag behind in entrepreneurship, while some non-gold rush counties have been catching up. An in-depth analysis of this change and its determinants is beyond the scope of this paper, but we can briefly return to the California case. The Forbes magazine lists States regarding their business conditions. In recent lists California ranks top in entrepreneurship culture but at the bottom in terms of regulation of business activities (Forbes 2019). This disconnect is reflected in the departure of several innovative firms from the Golden State (Massie, 2021).

What are the potential effects of this disconnect? We have argued in the theory section, that culture and formal institutions co-evolve. As the formal institutions become entrepreneurship prohibitive, the other elements of the fuel (selective migration and entrepreneurial culture) powering the entrepreneurial engine can lose some of their force too. For example, an important channel of how entrepreneurship culture can be transmitted over time is that entrepreneurs serve as role models for the next generation. If entrepreneurship is inhibited, fewer entrepreneurs can serve as role models, ultimately weakening the culture. Additionally, fewer entrepreneurially-minded people will make migrate to this region and some of the entrepreneurially minded populace might even leave it to realize their entrepreneurial aspirations elsewhere. At the end of this evolutionary process, an entrepreneurial culture might vanish. Thus, future research should look more closely into the interplay between formal institutions and culture over time.

References

Acemoglu, D. & Robinson, J. (2012). *Why Nations Fail? The Origins of Power, Prosperity, and Poverty*. New York, Crown Publishing.

Abdellaoui, A., D. Hugh-Jones, L. Yengo, K. et al. (2019). Genetic correlates of social stratification in Great Britain. *Nature Human Behaviour*, 3(12), 1332-1342.

American Community Survey 2010 (5yr average). U.S. Census Bureau. Table B16001: *Language Spoken at Home by Ability to Speak English* (accessed December 1, 2020).

Andersson, M. & Koster, S. (2011). Sources of persistence in regional start-up rates—evidence from Sweden. *Journal of Economic Geography*, 11(1), 179-201.

Andrews, R., Fazio, C., Guzman, J., Liu, Y. & Stern, S. (2019). *The Startup Cartography Project: Measuring and Mapping Entrepreneurial Ecosystems*. https://www.startupcartography.com/research (last accessed May 30, 2019)

Atlas of Historical County Boundaries (2021). *Atlas of Historical County Boundaries* (last accessed May 2, 2021).

Atack, J., Bateman, F., Haines, M. & Margo, R. (2010). Did railroads induce or follow economic growth?: Urbanization and population growth in the American Midwest, 1850–1860. *Social Science History*, 34(2), 171-197.

Audretsch, D. & Keilbach, M. (2004). Does entrepreneurship capital matter? *Entrepreneurship Theory and Practice*, 28(5), 419-429

Bancroft, H. (1890). *History of Nevada, Colorado, and Wyoming*. San Francisco, The History Company.

Bazzi, S., Fiszbein, M. & Gebresilasse, M. (2020). Frontier culture: The roots and persistence of “Rugged Individualism” in the United States. *Econometrica*, 88(6), 2329-2368.

Bennett, R., Smith, H., van Lieshout, C., Montebruno, P. & Newton, G. (2020). *The Age of Entrepreneurship: Business Propriertors, Self-employment and Corporations since 1851*. Oxon and New York, Routledge.

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

Bleidorn, W., Hopwood, C., Back, M., Jaap, J., Denissen, A., Hennecke, M. & Hill, P. (2021). Personality trait stability and change. *Personality Science*, 2(1), 1-20.

Bleidorn, W., Schönbrodt, F., Gebauer, J., Rentfrow, P., Potter, J. & Gosling, S. (2016). To live among like-minded others: Exploring the links between person-city personality fit and self-esteem. *Psychological Science*, 27(3), 419-427.

Bisin, A. & Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97(2), 298-319.

Brands, H. (2002). *The Age of Gold: The California Gold Rush and the New American Dream*. New York, Doubleday.

Brandstätter, H. (2011). Personality aspects of entrepreneurship: A look at five meta-analyses. *Personality and Individual Differences*, 51(3), 222-230.

Brodeur, A. & Haddad, J. (2021). Institutions, attitudes and LBGT: Evidence from the gold rush. *Journal of Economic Behavior & Organization*, 16(5), 1540-1576.

Bureau of Economic Analysis (1970-2019). *Employment by County, Metro and Other Areas*. CAEMP25S file and CAEMP25N file https://www.bea.gov/data/employment/employment-county-metro-and-other-areas (last accessed September 10, 2021)

Bureau of Labor Statitics (1976-2020). *Local Area Unemployment Statistics*. ttps://www.bls.gov/lau/#tables (accessed July 24, 2021).

Calomiris, C. (1993). Financial factors in the Great Depression. *Journal of Economic Perspectives*, 7(2), 61-85.

Cardon, M. Stevens, C. & Potter D. (2011). Misfortunes or mistakes? Cultural sensemaking of entrepreneurial failure. *Journal of Business Venturing*, 26(1), 79-92.

Census Bureau, Center for Economic Studies (1978-2019). *Business Dynamics Statistics*. States Department of Commerce, https://www.census.gov/data/tables/time-series/econ/bds/bds-tables.html (last accessed October 29, 2021)

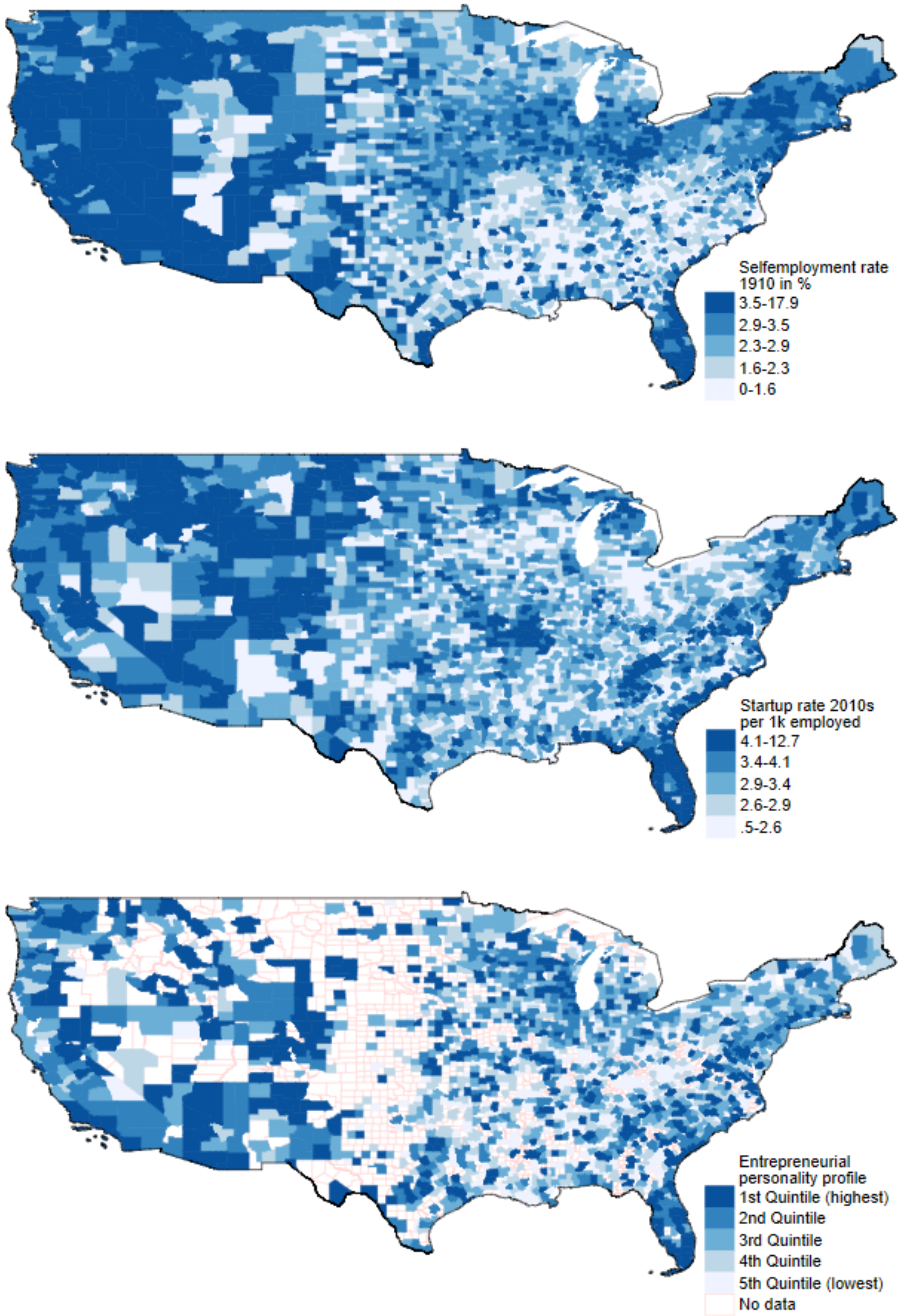
Census of the Population (1970). Supplementary report, *Race of the population by county*. Report number PC(S1)-104.

Census of the Population (2000-2010). U.S. Census Bureau; *Census of Population and Housing*, Census Profiles of General Demographic Characteristics. Census of the population by age, sex, and race data files (last accessed April 8, 2021)
Chinitz, B. (1961). Contrasts in agglomeration: New York and Pittsburgh. *American Economic Review*, 51(2), 279-289.
Clay, K. & Jones, R. (2008). Migrating to riches? Evidence from the California gold rush. *The Journal of Economic History*, 68(4), 997-1027.
Clay, K. & Wright, G. (2005). Order without law? Property rights during the California gold rush. *Explorations in Economic History*, 42(2), 155-183.
Cogley, T. & Sargent, T. (2008). The market price of risk and the equity premium: A legacy of the Great Depression? *Journal of Monetary Economics*, 55(3), 454-476.
Combes, P.-P., Duranton, G., Gobillon, L. & Roux, S. (2010). Estimating agglomeration economies with history, geology and worker efforts. In: Glaeser, E. (Eds.) *The Economic of Agglomeration* (pp. 15-66) Chicago and London, University of Chicago Press.
Cosci, S., Meliciani, V. & Pini, M. (2019). *Persistence in entrepreneurship: An analysis of the determinants of the start-up rate across Italian provinces*. 60th Annual Conference of the Italian Economic Association (SIE). University of Palermo.
Couttenier, M., Grosjean, P. & Sangnier, M. (2017). The Wild West is wild: The homicide resource curse. *Journal of the European Economic Association*, 15(3), 558-585.
Davidsson, P. (1995). Culture, structure and regional levels of entrepreneurship. *Entrepreneurship & Regional Development*, 7(1), 41-62.
Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica*, 78(6), 1863-1903.
Denzau, A. & North, D. (1994). Shared mental models: Ideologies and institutions. *Kyklos*, 47(1), 3-31.
Diamond, J. & Rotter, J. (1987). Observing the founder effect in human evolution. *Nature*, 329(6135), 105-106.
Forbes (2019). Best States for Business. https://www.forbes.com/best-states-for-business/list/#tab:overall (Accessed 26th March 2021).
Fotopoulos, G. & Storey, D. (2017). Persistence and change in interregional differences in entrepreneurship: England and Wales, 1921–2011. *Environment and Planning A*, 49(3), 670-702.
Freytag, A. & Thurik, R. (2007). Entrepreneurship and its determinants in a cross-country setting. *Journal of Evolutionary Economics*, 17(2), 117-131.
Fritsch, M., Obschonka, M., Wahl, F. & Wyrwich, M. (2020). The deep imprint of Roman sandals: Evidence of long-lasting effects of Roman rule on personality, economic performance, and well-being in Germany. *Jena Economic Research Papers*, 2020-005, https://zs.thulb.uni-jena.de/receive/jportal_jparticle_00787103.
Fritsch, M., Obschonka, M., Wyrwich, M. Gosling, S., Rentfrow, P. & Potter, J. (2018). Regionale Unterschiede der Verteilung von Personen mit unternehmerischem Persönlichkeitsprofil in Deutschland – ein Überblick. *Raumforschung und Raumordnung | Spatial Research and Planning*, 76(1), 65-81.
Fritsch, M. & Wyrwich, M. (2014). The long persistence of regional levels of entrepreneurship: Germany, 1925–2005. *Regional Studies,* 48(6), 955-973.
Fritsch, M. & Wyrwich, M. (2017). The effect of entrepreneurship on economic development—An empirical analysis using regional entrepreneurship culture. *Journal of Economic Geography*, 17(1), 157-189.
Gilson, R. (1999). The legal infrastructure of high technology industrial districts: Silicon Valley, Route 128, and covenants not to compete. *New York University Law Review*, 74(3), 575-629.
Glaeser, E., Kerr, S. & Kerr, W. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics & Statistics*, 97(2), 498-520.
Glaeser, E. & Kerr, W. (2009). Local industrial conditions and entrepreneurship: How much of the spatial distribution can we explain? *Journal of Economics & Management Strategy*, 18(3), 623-663.
GPIPP (2003-2015). Personality data set. Data set available from Sam Gosling, private communication.

Greene, F., Mole, K. & Storey, D. (2004). Does more mean worse? Three decades of enterprise policy in the Tees Valley. *Urban Studies*, 41(7), 1207-1228.
Guzman, J. & Stern, S. (2015). Where is Silicon Valley? *Science*, 347(6222), 606-609.
Haden, E. (1973). French dialect geography in North America, in W. Bright, D. Hymes & J. Lotz (Eds.) *Linguistics in North America* (pp. 422-471). Paris, The Hague.
Haines, M. (2018). *U.S. county populations by single years of age, sex, and race, 1970, 1980, 1990* (ICPSR 37115). Inter-university consortium for political and social research [distributor], 2018-07-30. https://doi.org/10.3886/ICPSR37115.v1 (accessed September 15, 2021).
Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations*. Thousand Oaks, Sage Publications, Inc.
IIASA & FAO (2012). *Global Agro-ecological Zones* (GAEZ v3.0). IIASA, Laxenburg, Austria and FAO, Rome, Italy.
Jang, K., Livesley, W. & Vernom, P. (1996). Heritability of the Big Five personality dimensions and their facets: A twin study. *Journal of Personality*, 64(3), 577-592.
John, O. & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In Pervin, L. & John, O. (Eds.), *Handbook of Personality: Theory and Research* (2nd Ed., pp. 102-138). New York: Guilford.
Jokela, M. (2009). Personality predicts migration within and between US states. *Journal of Research in Personality*, 43(1), 79-83.
Kerr, S., Kerr, W. & Xu, T. (2017). Personality Traits of Entrepreneurs: A Review of Recent Literature. *NBER Working Paper*, 24097.
Koellinger, P. & Thurik, A. (2012). Entrepreneurship and the business cycle. *Review of Economics & Statistics*, 94(4), 1143-1156.
Lo, M., Hinds, D., Tung, J. et al. (2017). Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nature Genetics*, 49(1), 152–156.
Lujala, P., Ketil, R. & Thieme, N. (2007). Fighting over oil: Introducing a new dataset. *Conflict Management and Peace Science*, 24(3), 239-256.
Manson, S., Schroeder, J., van Riper, D. & Ruggles, S. (2019). IPUMS National Historical Geographic Information System: Version 14.0 [Database]. Minneapolis, MN: IPUMS. 2019.
Margo, R. (1997). Wages in California Gold rush. *NBER Working Paper*, H0101.
Massie, G. (2021) *Texit: Why high-tech giants are fleeing Silicon Valley for Texas*. https://www.independent.co.uk/news/world/americas/exit-musk-silicon-valley-texas-b1774644.html (accessed November 1, 2021).
Mehlum, H., Moehne, K. & Torvik, R. (2006). Institutions and the resource curse. *The Economic Journal*, 116(508), 1-20.
NOAA (2017). Coastal County Definitions. Downloaded from https://coast.noaa.gov/data/digitalcoast/pdf/qrt-coastal-county-definitions.pdf (accessed April 16, 2020)
North, D. (1991). Institutions. *Journal of Economic Perspectives*, 5(1), 97-112.
Nunn, N. (2012). Culture and the historical process. *Economic History of Developing Regions*, 27(S1), 108-126.
Obschonka, M., Stuetzer, M. Gosling, S., Rentfrow, P., Lamb, M., Potter, J. & Audretsch, D. (2015). Entrepreneurial regions: Do macro-psychological cultural characteristics of regions help solve the "knowledge-paradox" of economics? *PloS One*, 10(6), e0129332.
Obschonka, M. & Stuetzer, M. (2017). Integrating psychological approaches to entrepreneurship: The Entrepreneurial Personality System (EPS). *Small Business Economics*, 49(1), 203-231.
Papyrakis, E. & Gerlagh, R. (2004). The resource curse hypothesis and its transmission channels. *Journal of Comparative Economics*, 32(1), 181-193.
Peri, G. (2005). Determinants of knowledge flows and their effect on innovation. *The Review of Economics and Statistics*, 87(2), 308-322.
Potter, M. (1976). *Oregon's Golden Years: Bonanza of the West*. Caldwell, The Caxton Printers.
Rentfrow, P., Gosling, S. D. & Potter, J. (2008). A theory of the emergence, persistence, and expression of geographic variation in psychological characteristics. *Perspectives on Psychological Science*, 3(5), 339–369.

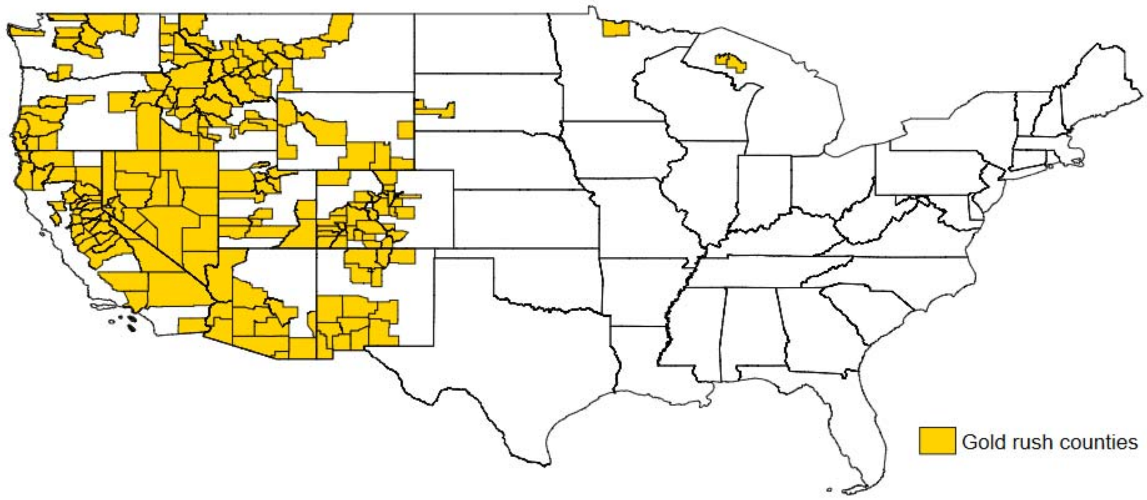
Rentfrow, P. (2010). Statewide differences in personality: Toward a psychological geography of the United States. *American Psychologist*, 65(6), 548-558.
Rivera, S. (2004). *The California Gold Rush*, ABDO Publishing Company.
Rohrbough, M. (1997). *Days of Gold: The California Gold Rush and the American Nation*. Berkeley and Los Angeles (CA), University of California Press.
Ross, M. (2015). What have we learned about the resource curse? *Annual Review of Political Science*, 18(1), 239-259.
Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J. & Sobek, M. (2020). *IPUMS USA*: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. ttps://doi.org/10.18128/D010.V10.0.
Saxenian, A. (1996). *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge (MA), Harvard University Press.
The Startup Cartography Project (2010-2016). *Entrepreneurship_by_County_academic.tab*, https://doi.org/10.7910/DVN/BMRPVH/7A6GVC, Harvard Dataverse, V7, UNF:6:izghknCh01dcMzhftGkmig== [fileUNF] (accessed May 30, 2019)
Stuetzer, M., Obschonka, M., Audretsch, D., Wyrwich, M., Rentfrow, P., Coombes, M., Shaw-Taylor, L. and Satchel, M. (2016). Industry structure, entrepreneurship, and culture: An empirical analysis using historical coalfields. *European Economic Review*, 86, 52-72.
Tarr R. & Mccurry, F. (1910). *New Geographies*. New York: Macmillan Company (Map downloaded from http://etc.usf.edu/maps/pages/2000/2085/2085.htm).
Terracciano, A., Sanna, S., Uda, M. et al. (2010). Genome-wide association scan for five major dimensions of personality. *Molecular Psychiatry*, 15(6), 647–656.
Waldinger, R., Aldrich, H. & Ward, J. (1990). *Ethnic Entrepreneurs: Immigrant Businesses in Industrial Societies*. Newbury Park, CA, Sage.
Ubbelohde, C., Benson, M. & Smith, D. (1988). *A Colorado History*. Boulder (CO), Pruett Publishing Company.
Ucbasaran, D., Shepherd, D. A., Lockett, A. & Lyon, S. J. (2013). Life after business failure: the process and consequences of business failure for entrepreneurs. *Journal of Management*, 39(1), 163-202.
U.S. Geological Survey. Mineral Resource Data System. https://mrdata.usgs.gov/mrds/ (accessed October 14, 2017)
van den Berg, S., de Moor, M, Verweij, K. et al. (2016). Meta-analysis of genome-wide association studies for extraversion: Findings from the Genetics of Personality Consortium. *Behavior Genetics,* 46(2), 170–182.
Young, O. E. (1982). The southern gold rush, 1828-1836. *The Journal of Southern History*, 48(3), 373-392.
Zelinksy, W. (1973). *The Cultural Geography of the United States*. Prentice Hall: Englewood Cliffs
Zhao, H. & Seibert, S. (2006). The Big Five personality dimensions and entrepreneurial status: A meta-analytical review. *Journal of Applied Psychology*, 91(2), 259–271.

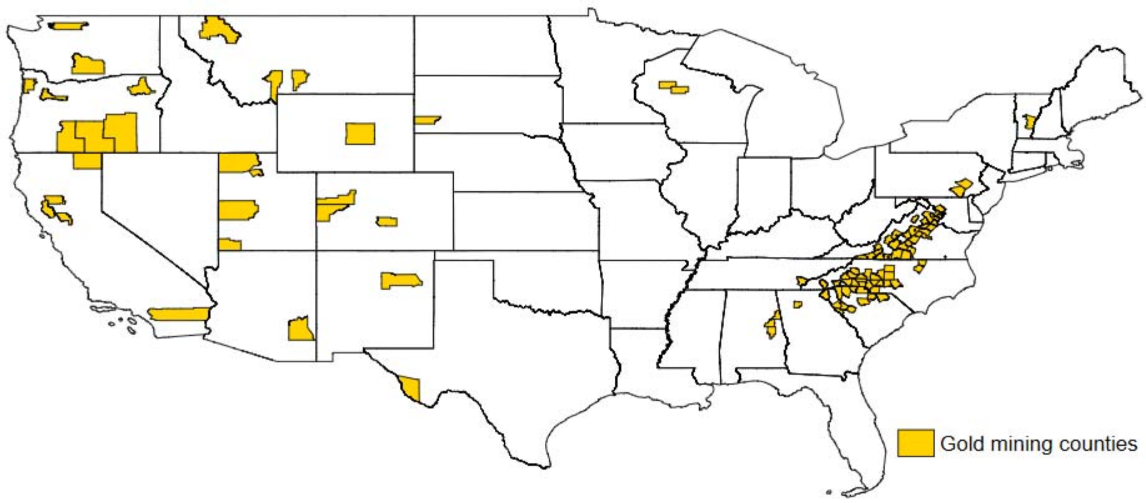
Figure 1: Entrepreneurship rates



Note: Top: 1910 self-employment rate over the age 16+ population; Middle: Startup-rate 2010s per 1,000 employed; Bottom: Entrepreneurial personality profile

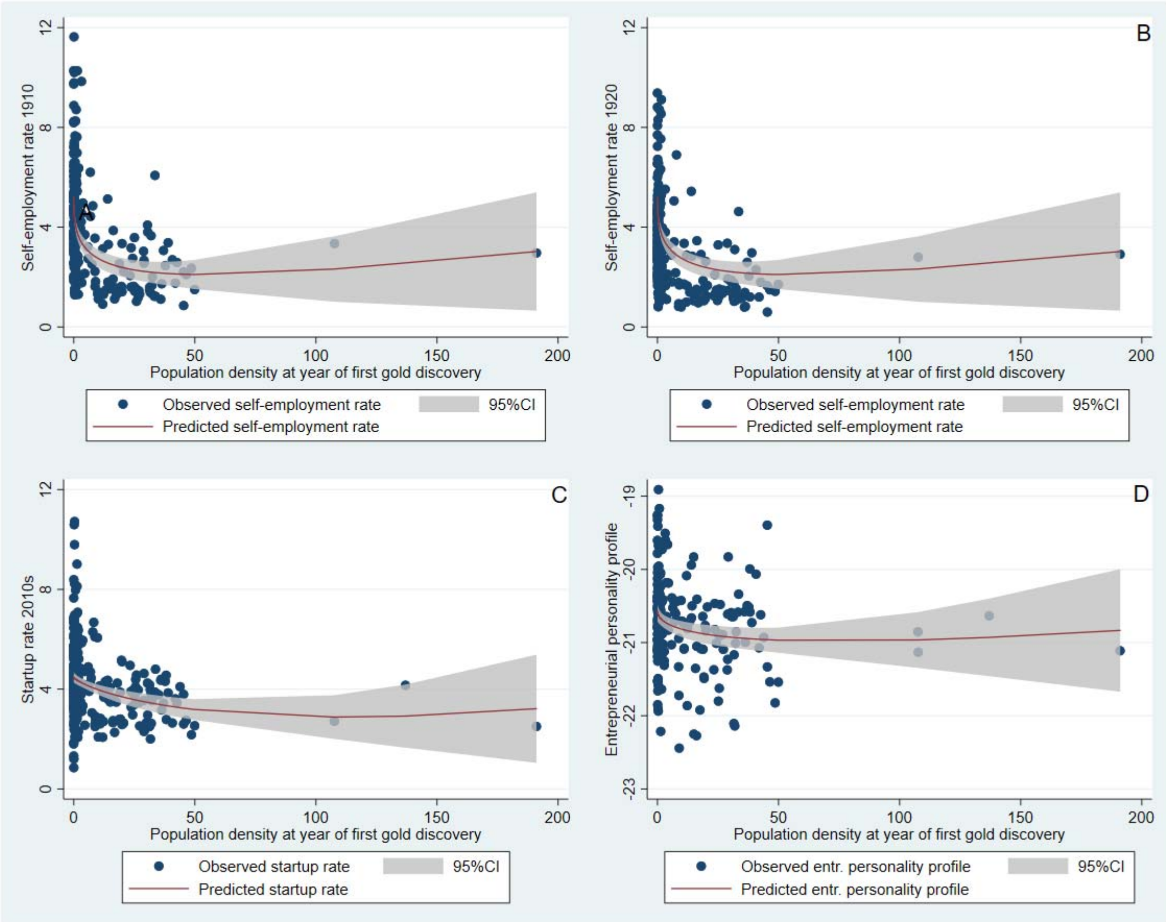
Figure 2: Gold rush and gold mining counties





Notes: Based on the MRDS. Top: Gold indicates gold rush counties as defined as counties in the West having a gold discovery between 1848 and 1899. Bottom: gold indicates gold mining counties as defined as counties in the West having a gold discovery after 1899 or in the East with gold discoveries at any point in time.

Figure 3: Importance of first dominant settler group



Notes: Figure 3A displays the theoretical predicted relationship between the population density at the year of the first gold discovery and entrepreneurship rates. Figure 3B-D show scatterplots and predictions of the entrepreneurship rates with 95% confidence intervals from best fitting fractional polynomial regressions.

Table 1: Basic results regarding entrepreneurship indicators.

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. Profile
Panel A: State-fixed-effects										
Gold rush	0.833***	0.679***	0.271*	0.548***	1.737***	1.399***	1.058***	1.004***	0.737***	0.164***
	(0.199)	(0.219)	(0.156)	(0.195)	(0.416)	(0.294)	(0.304)	(0.253)	(0.226)	(0.060)
Gold mining	-0.188	-0.137	-0.071	-0.064	0.078	-0.340	-0.235	0.023	0.072	0.032
	(0.192)	(0.130)	(0.123)	(0.148)	(0.452)	(0.374)	(0.217)	(0.189)	(0.114)	(0.050)
Geographic controls	No	No	No	No	No	No	No	No	No	No
Socio-econ. Controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.337	0.338	0.383	0.361	0.175	0.209	0.195	0.252	0.259	0.175
Panel B: Geographic controls										
Gold rush	0.751***	0.523***	0.142	0.353**	1.170***	1.068***	0.699***	0.580***	0.462***	0.131**
	(0.167)	(0.132)	(0.096)	(0.149)	(0.288)	(0.201)	(0.191)	(0.167)	(0.143)	(0.053)
Gold mining	-0.195	-0.185	-0.108	-0.123	0.011	-0.330	-0.268**	-0.051	0.033	0.021
	(0.204)	(0.130)	(0.122)	(0.153)	(0.313)	(0.257)	(0.132)	(0.120)	(0.095)	(0.054)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-econ. Controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.369	0.369	0.421	0.386	0.210	0.238	0.225	0.292	0.292	0.194
Effect size	0.580	0.466	0.158	0.280	0.459	0.555	0.446	0.418	0.431	0.205
P-value gold rush > gold mining	0.000	0.000	0.042	0.003	0.006	0.000	0.000	0.001	0.002	0.050
Panel C: Socio-economic controls										
Gold rush	0.527***	0.411***	0.074	0.401***	0.857**	0.686**	0.434**	0.437***	0.341***	0.146**
	(0.151)	(0.103)	(0.101)	(0.132)	(0.333)	(0.270)	(0.202)	(0.157)	(0.121)	(0.056)
Gold mining	-0.221	-0.239***	-0.163	-0.122	-0.041	-0.431**	-0.354***	-0.183*	-0.110	0.059
	(0.147)	(0.080)	(0.102)	(0.111)	(0.219)	(0.184)	(0.095)	(0.096)	(0.079)	(0.051)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-econ. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.420	0.471	0.462	0.445	0.305	0.420	0.386	0.399	0.391	0.327
P-value gold rush > gold mining	0.000	0.000	0.041	0.000	0.018	0.001	0.000	0.001	0.000	0.113

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are the startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. Panel A: State-fixed effects. Panel B: State-fixed effects + geographic controls: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation (mean and sd), potential agricultural yield, Hispanic influence, French influence, latitude, longitude. Panel C: Additional socio-economic controls of the respective decades: Empl. share in manufacturing and construction, empl. share in services, empl. share in government, unempl. rate, Black American pop. share, Native American pop. share, Asian American pop. share, population density. Effect size is computed as the size of the regression coefficient divided by the standard deviation of the DV and estimates the effect size of the gold rush dummy as percent of a standard deviation of the DV. Panel B contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the gold mining coefficient. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 2: Results regarding high quality entrepreneurship.

VARIABLES	(1) High quality startup rate 2010s	(2) High quality startup rate 2010s	(3) High quality startup rate 2010s
Gold rush	-0.539	-0.384	-0.175
	(1.302)	(1.036)	(0.604)
Gold mining	-0.382	-0.156	-0.363
	(0.463)	(0.364)	(0.324)
Geographic controls	No	Yes	Yes
Socio-econ. controls	No	No	Yes
State-fixed effects	Yes	Yes	Yes
Observations	3,094	3,094	3,094
Adjusted R-squared	0.307	0.343	0.566
P-value gold rush > gold mining	0.430	0.376	0.282

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. Dependent variable is the rate of high-quality startups per 1 Mill. employed. Model 1: State-fixed effects. Model 2: State-fixed effects + geographic controls: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude. Model 3: Additional socio-economic controls of the 2010 decade: Employment share in manufacturing and construction, employment share in services, employment share in government, unemployment rate, Black American pop. share, Native American pop. share, Asian American pop. share, population density. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the gold mining coefficient. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 3: Results regarding alternative control groups.

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Panel A: Without Appalachian gold mining counties										
Gold rush	0.775***	0.551***	0.168	0.347**	1.330***	1.172***	0.736***	0.597***	0.463***	0.135**
	(0.180)	(0.137)	(0.102)	(0.164)	(0.302)	(0.218)	(0.201)	(0.181)	(0.156)	(0.062)
Gold mining	-0.008	0.040	0.105	-0.158	1.226**	0.532	0.098	0.139	0.080	0.027
	(0.230)	(0.167)	(0.129)	(0.222)	(0.604)	(0.319)	(0.251)	(0.344)	(0.253)	(0.126)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,036	3,036	3,036	3,036	3,036	3,036	3,036	3,036	3,036	2,022
Adjusted R-squared	0.368	0.363	0.412	0.380	0.212	0.243	0.229	0.297	0.295	0.196
P-value gold rush > gold mining	0.000	0.002	0.292	0.000	0.428	0.006	0.002	0.061	0.030	0.155
Panel B: Counties incorporated between 1848 and 1899 as control group										
Gold rush	0.712***	0.513***	0.135	0.340**	1.162***	1.133***	0.762***	0.613***	0.467***	0.104**
	(0.156)	(0.130)	(0.096)	(0.144)	(0.301)	(0.206)	(0.186)	(0.162)	(0.139)	(0.052)
Incorporated 1848-1899	-0.221***	-0.122***	-0.073**	-0.101*	-0.019	0.040	0.066	0.076	0.030	-0.068**
	(0.043)	(0.039)	(0.032)	(0.053)	(0.117)	(0.083)	(0.077)	(0.057)	(0.080)	(0.031)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.373	0.371	0.421	0.387	0.210	0.237	0.224	0.293	0.292	0.196
P-value gold rush > incorporated	0.000	0.000	0.018	0.002	0.000	0.000	0.000	0.001	0.003	0.003
Panel C: Counties initially organized between 1848 and 1899 as control group										
Gold rush	0.682***	0.485***	0.123	0.335*	1.276***	1.239***	0.803***	0.707***	0.573***	0.110
	(0.195)	(0.158)	(0.120)	(0.177)	(0.357)	(0.235)	(0.230)	(0.178)	(0.139)	(0.072)
Initially organized 1848-1899	-0.197	-0.132	-0.069	-0.072	0.211	0.234	0.121	0.232	0.227	-0.033
	(0.149)	(0.121)	(0.104)	(0.155)	(0.260)	(0.204)	(0.185)	(0.147)	(0.151)	(0.081)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.369	0.369	0.420	0.386	0.210	0.238	0.224	0.293	0.293	0.194
P-value gold rush > organized	0.000	0.000	0.023	0.003	0.000	0.000	0.000	0.002	0.013	0.009

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are startups rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. Panel A drops Appalachian gold mining counties from the control group. Panel B uses as control group all counties without gold rush but an incorporation date between 1848 and 1899. Panel C uses as control group all counties without a gold rush but an initial land organization date between 1848 and 1899. All models use state-fixed effects + geographic controls: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than that of the control group. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 4: Results regarding alternative treatments.

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Panel A: Silver rush										
Silver rush	0.443**	0.437***	0.086	0.121	1.451***	1.202***	0.950***	0.783***	0.600***	0.222***
	(0.173)	(0.142)	(0.107)	(0.218)	(0.327)	(0.205)	(0.217)	(0.213)	(0.154)	(0.053)
Silver mining	0.235	0.149	0.091	0.129	0.262	0.191	0.174	0.142	0.129	-0.061
	(0.215)	(0.213)	(0.135)	(0.142)	(0.360)	(0.242)	(0.198)	(0.263)	(0.197)	(0.075)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.361	0.366	0.420	0.384	0.212	0.237	0.227	0.295	0.295	0.196
P-value silver rush > silver mining	0.230	0.150	0.489	0.487	0.019	0.001	0.012	0.048	0.049	0.001
Panel B: Iron rush										
Iron rush	0.006	-0.215	0.015	-0.538***	0.810	0.697	0.535	0.123	0.114	0.099
	(0.309)	(0.170)	(0.118)	(0.136)	(0.595)	(0.440)	(0.335)	(0.322)	(0.229)	(0.071)
Iron mining	-0.132	0.001	-0.093	0.039	-0.201	-0.362	-0.271	-0.005	0.061	-0.092**
	(0.215)	(0.183)	(0.096)	(0.113)	(0.441)	(0.376)	(0.282)	(0.291)	(0.187)	(0.042)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.357	0.362	0.420	0.385	0.205	0.229	0.219	0.288	0.288	0.191
P-value iron rush > iron mining	0.390	0.267	0.295	0.008	0.153	0.091	0.086	0.413	0.445	0.024

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are the startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. Panel A: The variable silver rush takes the value of one if a county had a silver discovery in the West between 1848 and 1899. The variable silver mining takes the value of one if a county had a silver discovery in the east at any point in time or in the West after 1899. Panel B: The variable iron rush takes the value of one if a county had an iron discovery in the West between 1848 and 1899. The variable iron mining takes the value of one if a county had an iron discovery in the east at any point in time or in the West after 1899. All models use state-fixed effects + geographic controls: Coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude. Last row contains a one-sided Wald test (p-value reported) whether the silver/iron rush coefficient is larger than the silver/iron mining coefficient. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 5: Robustness check regarding the GPIPP data set.

VARIABLES	(1) N<100 + Weighted	(2) N<100 + Weighted	(3) N<100 + Weighted
Gold rush	0.413***	0.308***	0.360***
	(0.140)	(0.101)	(0.093)
Gold mining	0.056	0.039	0.049
	(0.055)	(0.068)	(0.068)
Geographic controls	No	Yes	Yes
Socio-econ. controls	No	No	Yes
State-fixed effects	Yes	Yes	Yes
Observations	3,106	3,106	3,106
Adj. R-squared	0.048	0.057	0.086
P-value gold rush > gold mining	0.002	0.001	0.000

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The models use all counties even with less than 100 respondents per county. Additionally, the regional entrepreneurial personality profile is weighted by age and gender. Model 1 use only state-fixed effects as controls. Model 2 include geographic controls: Ironfield, coalfield, Petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, and longitude. Model 3 additionally includes socio-economic controls of the 2010 decade: Empl. share in manufacturing and construction, empl. share in services, empl. share in government, unempl. rate, Black American pop. share, Native American pop. share, Asian American pop. share, population density. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the gold mining coefficient. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 6: Results regarding exits.

Variables	(1) Exit rate 1970s	(2) Exit rate 1970s	(3) Exit rate 1980s	(4) Exit rate 1980s	(5) Exit rate 1990s	(6) Exit rate 1990s	(7) Exit rate 2000s	(8) Exit rate 2000s	(9) Exit rate 2010s	(10) Exit rate 2010s
Gold rush	0.669***	0.441*	0.824***	0.640***	0.470**	0.446**	0.795***	0.639**	0.671**	0.456**
	(0.245)	(0.259)	(0.177)	(0.174)	(0.180)	(0.168)	(0.223)	(0.282)	(0.250)	(0.196)
Gold mining	-0.510**	-0.540**	-0.286	-0.296	-0.329***	-0.270***	-0.051	-0.048	0.151	0.141
	(0.211)	(0.250)	(0.203)	(0.180)	(0.120)	(0.078)	(0.090)	(0.115)	(0.170)	(0.201)
Geographic controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109
Adj. R-squared	0.119	0.144	0.232	0.273	0.138	0.168	0.028	0.038	0.073	0.089
P-value gold rush > gold mining	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.002	0.012	0.054

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variable are exit rates (number of exits) per 100 firms. Models 1, 3, 5, 7, 9 use only state-fixed effects as controls. Models 2, 4, 6, 8, 10 include state-fixed effects + geographic controls: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the gold mining coefficient. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 7: Results regarding job creation through firm births.

Variables	(1) Job creation rate 1970s	(2) Job creation rate 1970s	(3) Job creation rate 1980s	(4) Job creation rate 1980s	(5) Job creation rate 1990s	(6) Job creation rate 1990s	(7) Job creation rate 2000s	(8) Job creation rate 2000s	(9) Job creation rate 2010s	(10) Job creation rate 2010s
Gold rush	12.197***	8.753***	4.859***	4.446***	5.714**	5.407***	3.384***	3.336***	2.743***	2.546***
	(2.805)	(2.377)	(1.765)	(1.300)	(2.579)	(1.797)	(1.025)	(1.075)	(0.885)	(0.782)
Gold mining	-1.155	-1.219	-0.289	0.448	-1.730	-1.011	0.045	0.554	-0.000	0.323
	(4.702)	(4.501)	(4.313)	(4.156)	(2.878)	(2.833)	(2.194)	(2.015)	(1.373)	(1.364)
Geographic controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109
Adj. R-squared	0.042	0.058	0.109	0.132	0.175	0.214	0.172	0.215	0.128	0.152
P-value gold rush > gold mining	0.003	0.013	0.108	0.176	0.015	0.024	0.044	0.103	0.019	0.048

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables are job creation rates through firm births (in non-agricultural industries) in percent of all jobs. Models 1, 3, 5, 7, 9 use only state-fixed effects as controls. Models 2, 4, 6, 8, 10 include state-fixed effects + geographic controls: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the gold mining coefficient. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

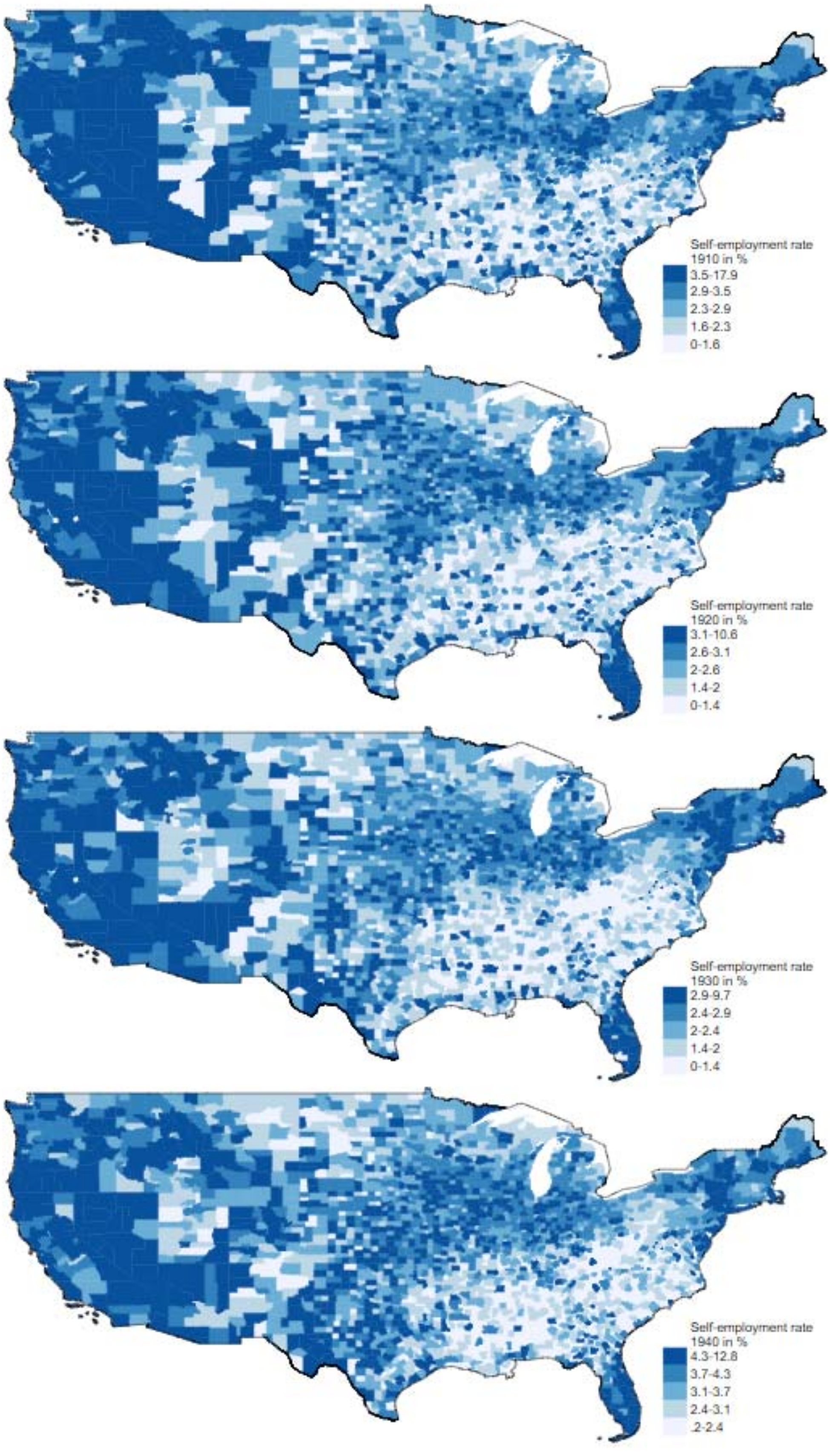
Table 8: Results regarding the importance of early settler groups.

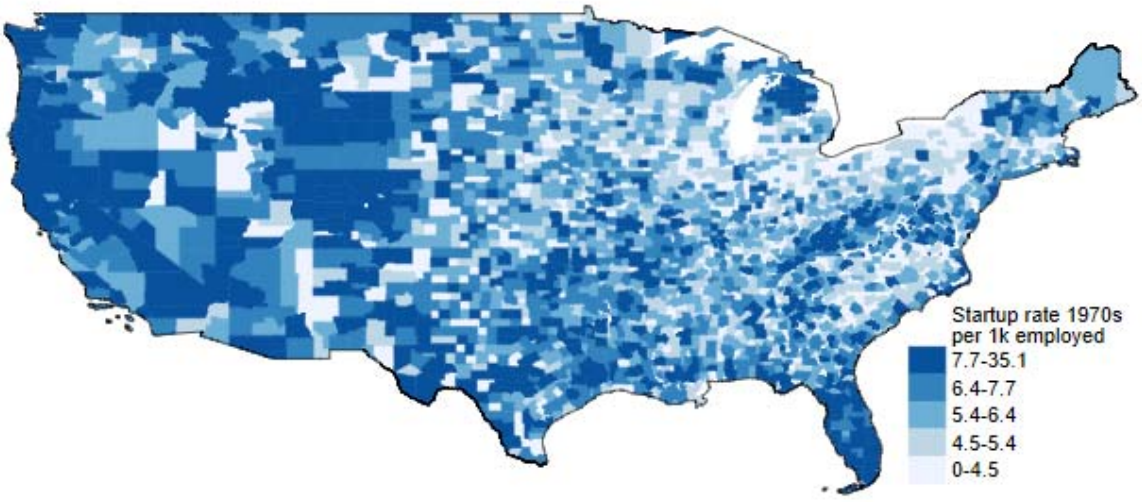
Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Panel A: Square root pop. density										
Population density	0.054***	0.046***	0.029***	0.040***	-0.006	0.004	0.013	0.010	0.007	0.005*
	(0.016)	(0.012)	(0.009)	(0.013)	(0.034)	(0.016)	(0.011)	(0.009)	(0.008)	(0.003)
Square root pop. density	-0.798***	-0.677***	-0.457***	-0.630***	-0.475	-0.395**	-0.371**	-0.311**	-0.225**	-0.095***
	(0.170)	(0.129)	(0.098)	(0.151)	(0.296)	(0.180)	(0.138)	(0.118)	(0.104)	(0.032)
Geographic controls	No	No	No	No	No	No	No	No	No	No
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	No	No	No	No	No	No	No	No	No	No
Observations	273	280	283	288	293	296	300	300	300	213
Adjusted R-squared	0.229	0.234	0.224	0.189	0.079	0.093	0.080	0.074	0.068	0.040
Panel B: Ln pop density										
Population density	0.019*	0.015*	0.008	0.011	-0.031	-0.014	-0.003	-0.004	-0.003	0.001
	(0.009)	(0.007)	(0.006)	(0.009)	(0.026)	(0.010)	(0.007)	(0.005)	(0.004)	(0.002)
Ln pop. density	-0.908***	-0.763***	-0.506***	-0.694***	-0.488	-0.440*	-0.425**	-0.359**	-0.255**	-0.109***
	(0.205)	(0.157)	(0.121)	(0.187)	(0.372)	(0.217)	(0.161)	(0.139)	(0.120)	(0.037)
Geographic controls	No	No	No	No	No	No	No	No	No	No
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	No	No	No	No	No	No	No	No	No	No
Observations	273	280	283	288	293	296	300	300	300	213
Adjusted R-squared	0.226	0.228	0.216	0.180	0.077	0.092	0.080	0.075	0.068	0.040

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. The number of observations differ across the models, because we excluded gold discoveries after the respective decades from the analysis. Panel A uses the square root of the population density as IV. Panel B uses the natural logarithm of the population density as IV. No controls are used in these models. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

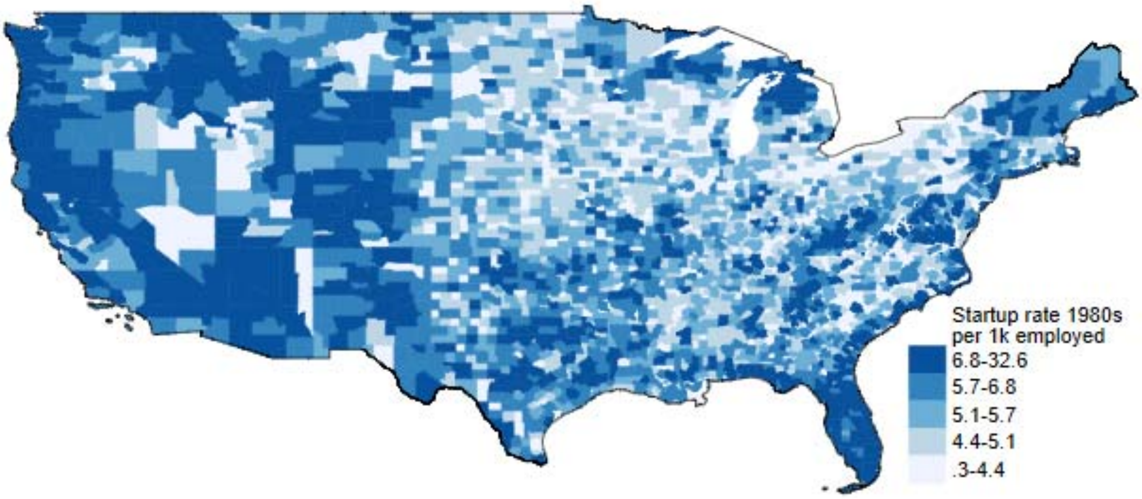
Online Appendix – For Online Publication

Figure A1: Entrepreneurship rates 1910-2010s

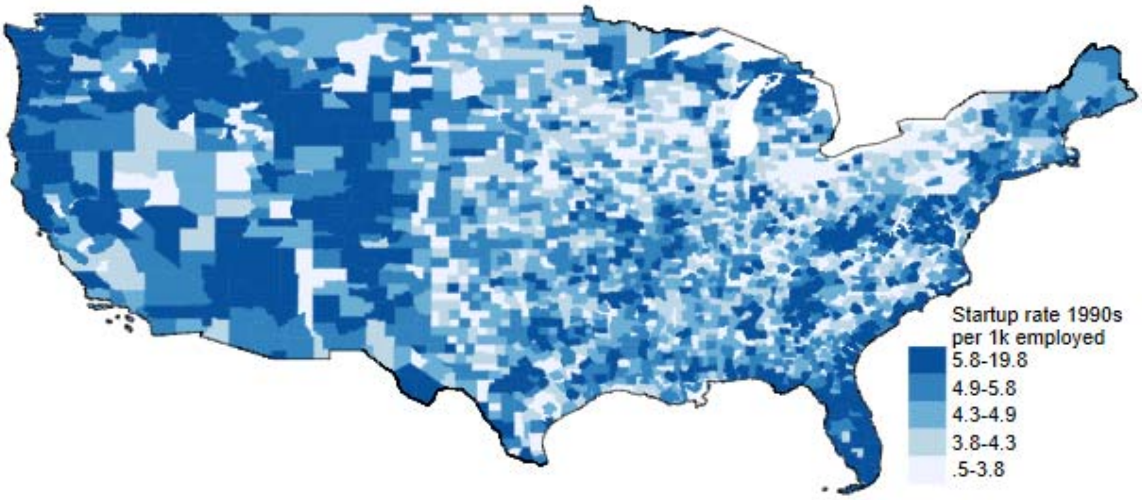




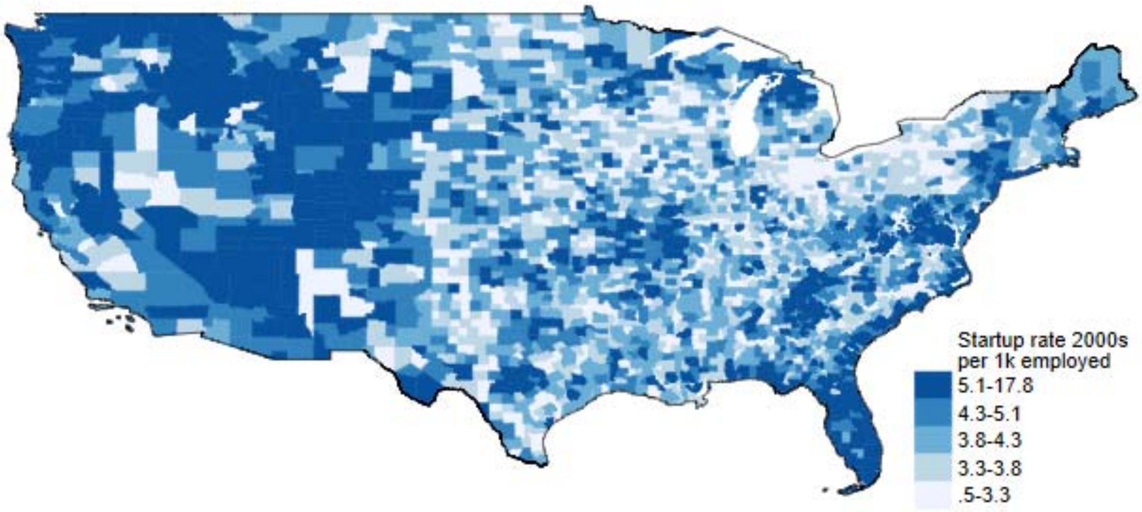
Startup rate 1970s
per 1k employed
7.7-35.1
6.4-7.7
5.4-6.4
4.5-5.4
0-4.5



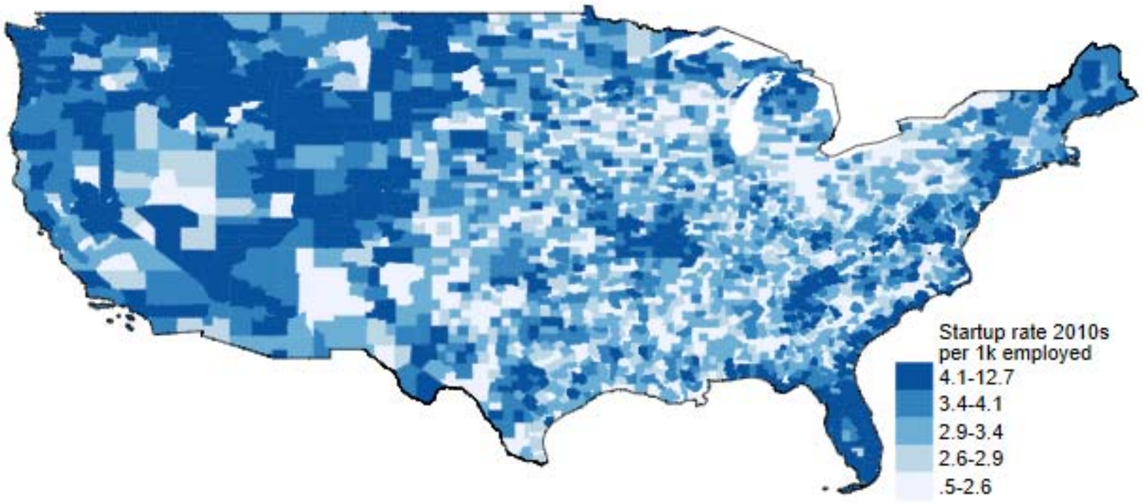
Startup rate 1980s
per 1k employed
6.8-32.6
5.7-6.8
5.1-5.7
4.4-5.1
.3-4.4



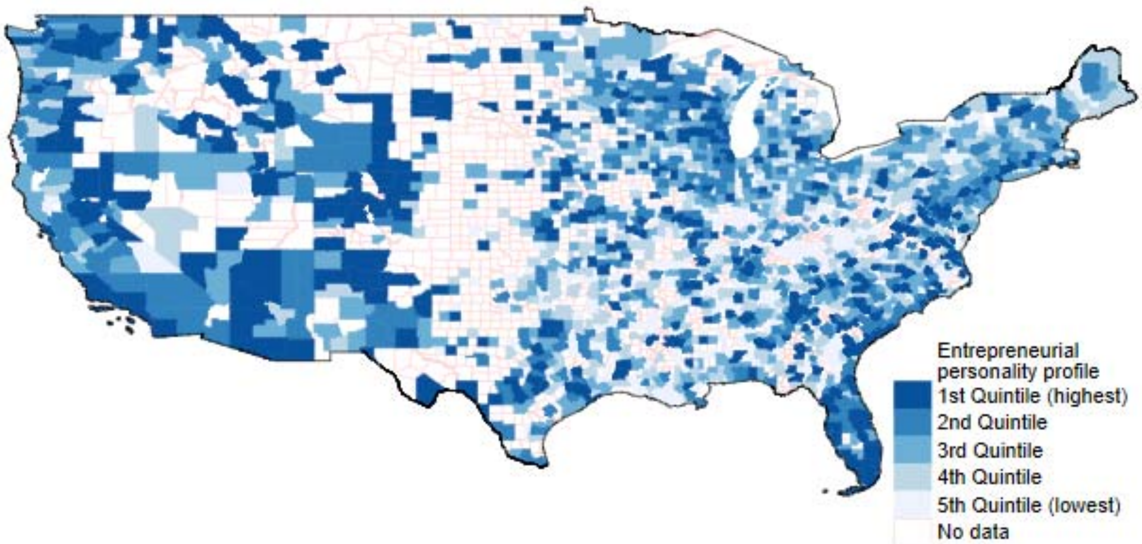
Startup rate 1990s
per 1k employed
5.8-19.8
4.9-5.8
4.3-4.9
3.8-4.3
.5-3.8



Startup rate 2000s
per 1k employed
5.1-17.8
4.3-5.1
3.8-4.3
3.3-3.8
.5-3.3

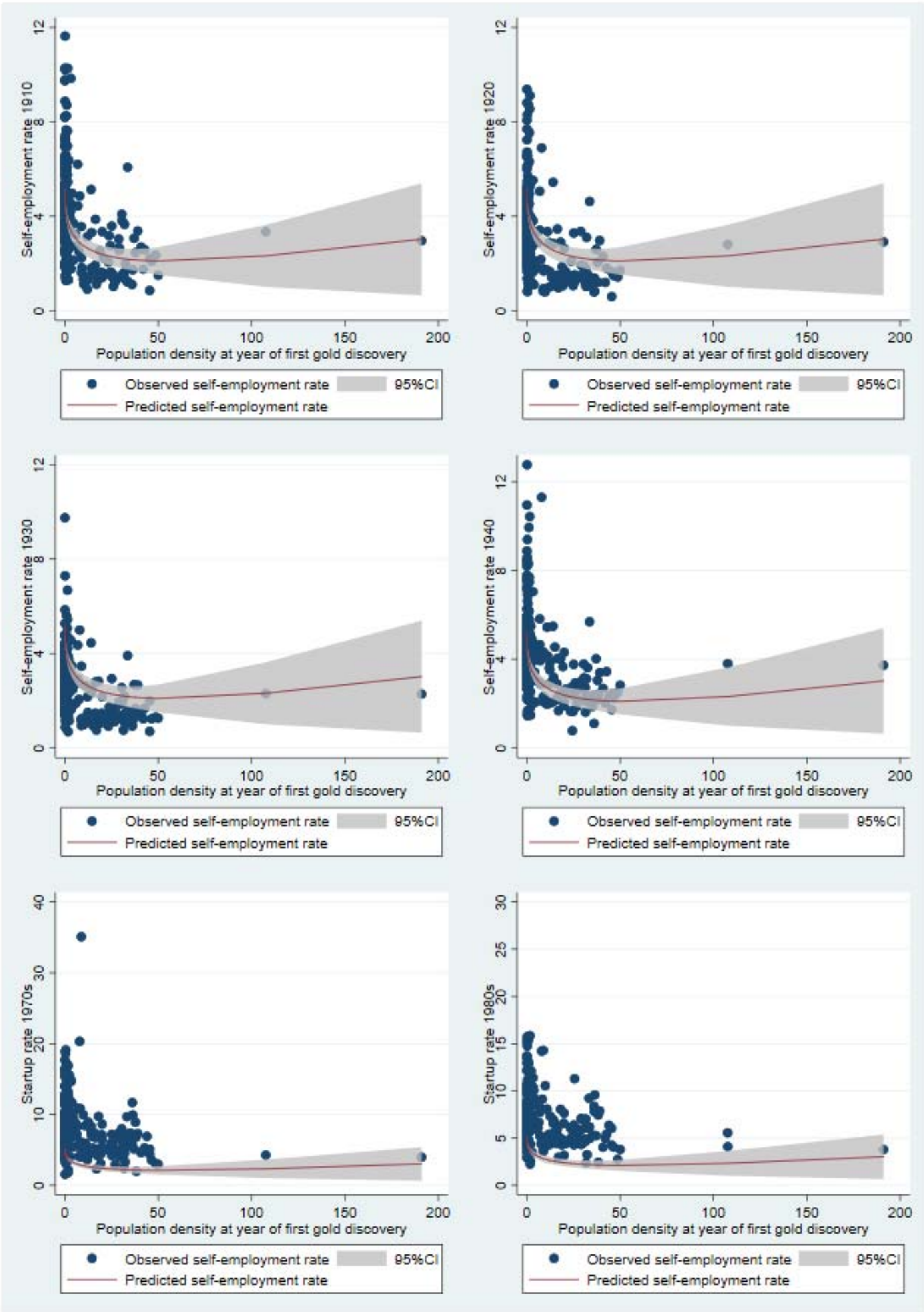


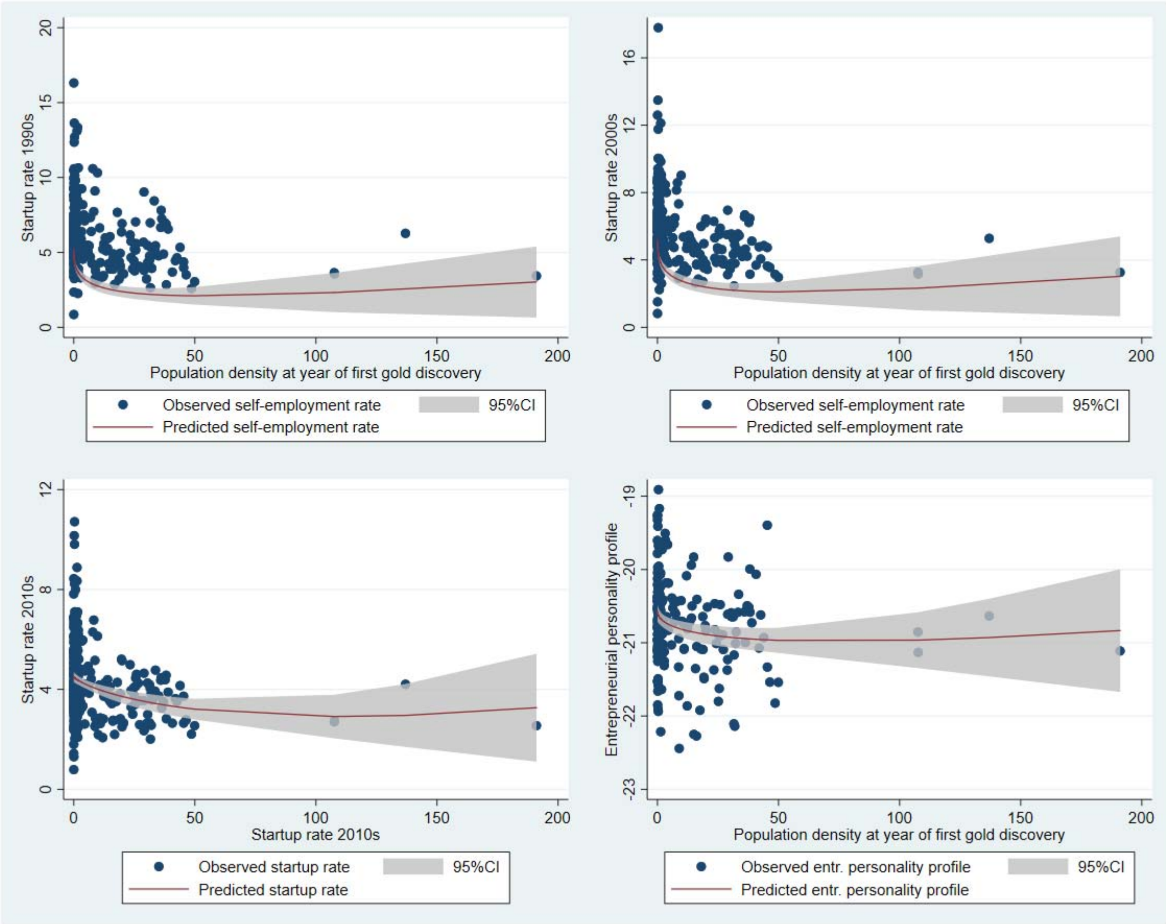
Startup rate 2010s
per 1k employed
4.1-12.7
3.4-4.1
2.9-3.4
2.6-2.9
.5-2.6



Entrepreneurial
personality profile
1st Quintile (highest)
2nd Quintile
3rd Quintile
4th Quintile
5th Quintile (lowest)
No data

Figure A2: Entrepreneurship rates and population density





Startup rate 1990s
Population density at year of first gold discovery
Observed self-employment rate
95%CI
Predicted self-employment rate
Startup rate 2000s
Population density at year of first gold discovery
Observed self-employment rate
95%CI
Predicted self-employment rate
Startup rate 2010s
Startup rate 2010s
Observed startup rate
95%CI
Predicted startup rate
Entrepreneurial personality profile
Population density at year of first gold discovery
Observed entr. personality profile
95%CI
Predicted entr. personality profile

Table A1: Variables used and their sources

Variable	Description and Source
Startup rate	Number of non-agricultural startups per 1,000 employed. We average the yearly startup rates for each decade starting from the 1970s (1979, 1980-1989, 1990-1999, 2000-2009, 2010-2019). Data prior to 1978 are not available. Source: Census Bureau 1978-2019. Startup data are taken from the public use Business Dynamics Statistics (BDS). The BDS suppresses some startup data because of confidentiality reasons in case of very few startups in a region. These few missing data were imputed with values of 1 in case a random number was smaller than 0.5 and with 2 in case the random number was larger than 0.5 (Stata set seed 1001). The startup data refer to all non-agricultural industries. Employment data refer to the number of total jobs from the CAEMP25S file (Line Code 10) of the Bureau of Economic Analysis (BEA). The BEA data reports combined data of the independent Virginia cities with the surrounding counties. The combined data were split up using the population share on the base of the population share of the respective regions. Following Haltiwanger, Jarmin and Miranda (2013) the startup rate is computed based on the average employment of the current year and the previous year. $\text{startup rate}_t = 1.000 \cdot \frac{\text{entries}_t}{0{,}5 \cdot (\text{employment}_t + \text{employment}_{t-1})}$
Self-employment rate	Number of individuals working on own account or employers in non-agricultural sectors in percent of the population aged 16 years or older for the years 1910, 1920, 1930 and 1940. Source: Data on own account workers, employers and the age 16+ population are taken from the U.S. Census (Ruggles et al. 2020). As the spatial structure of the U.S. has changed since the early 20th century, we use a look-up table from Eckert et al. (2020) who use GIS methods to attribute historic counties to contemporary ones based on the overlaying area. Based on this look-up table we get the number of self-employed and the population at the contemporary county structure and compute the respective self-employment rate. Note, that a few counties have either missing observations in some years or are not covered properly by the look-up table. In these cases, we impute the data based on the trend from earlier or later census data.
High-quality startup rate	Quality-adjusted number of startups (RECPI) per 1 Mill. employed. We average the yearly startup rates for the period between 2010 and 2016. Source: Startup Cartography Project (2010-2016). This data captures the quality of startups based on observable characteristics such as registering as a corporation or in Delaware for equity financing, the name of the firm which can signal the high-tech nature of a firm, creating innovations measurable as patents (Andrews et al., 2019). The data cover in general 49 states and DC. However, data from a few counties are missing. Data are not yet available for years after 2016. Data from Delaware across the whole time span is missing as well as data from Illinois, South Carolina and Illinois for the period of 2014-2016 due to increased administrative burden and costs. Quality-adjusted quantity of startups can be interpreted as the expected number of start-ups with growth potential based on observable characteristics at the time of business registration (Guzman and Stern, 2019).

	Following Haltiwanger, Jarmin and Miranda (2013) the startup rate is computed based on the average employment of the current year and the previous year. $$\text{high quality startup rate}_t = 1\text{ Mill.}\cdot\frac{\text{high quality entries}_t}{0{,}5\cdot(\text{employment}_t+\text{employment}_{t-1})}$$
Entrepreneurial personality profile	Regional average of an individual entrepreneurial personality profile high in O, C, E and low in A and N. Source: GPIPP 2003-2015: The Gosling-Potter-Internet-Project collects the data via a non-commercial internet webpage. The most recent version of the dataset spans the years 1998 and 2015. Since 2003 respondents are asked to provide ZIP codes of their residence allowing for regional analysis. We conducted several quality checks: First, responses that provided contradictory information on their residence were excluded. For instance, some respondents state they lived in a U.S. state but at the same time in a different county. Second, responses were also excluded when respondents did not consent for their answers to be used for research purposes. Third, responses were excluded where respondents stated that they had completed the study before. Fourth, responses were excluded when respondents stated they did not answer their questions truthfully. Fifth, responses were excluded when respondents failed to respond to all items that measured personality traits. Sixth, following Meade and Craig's (2012) criteria, responses were dropped when respondents gave careless responses. Seventh, we excluded counties with less than 100 observations to ensure measurement precision. As most internet surveys the dataset (N=3,159,083) is skewed towards younger (average age=26.6 years) and female respondents (64.4%). Thus in a robustness check we use an entrepreneurial personality profile weighting the individual observations by age and gender. We do not include more variables in the weighting procedure, nor do we employ the weighted entrepreneurial personality profile in the main regressions because the weighting procedure has its drawbacks. Many counties especially in the Mountain and West North Central Census divisions are sparsely populated. This leads to almost empty strata and even sometimes truly empty strata (e.g., number of male respondents aged 50-65). Producing weights in almost empty strata results in very variation weights. An undesired outcome of high variation in weights is an increased variance, which can offset any advantages of the weighting procedure and ultimately increases the mean square error and negatively affect regression results (Elliot 2008). Thus, we trim our weights in the robustness check ($-3 \leq w \leq 3$) and only use it as robustness check and not in the main analysis. The Big Five personality data were collected with the 44-item BFI inventory (John & Srivastava, 1999). An individual entrepreneurial personality profile is computed in several steps. First, the sum of the squared difference between the statistical reference point (5 for O, C, E and 1 for A, N) and the individual empirical Big Five values is computed for each Big Five trait. Second, the sum of these squared differences is computed. Third, the sign is reversed in order that higher values reflect a more entrepreneurial personality profile. The entrepreneurial personality profile is computed as the regional average of the individual profile.
Exit rate	Number of non-agricultural exits per 100 firms. We average the yearly exit rates for each decade starting from the 1970s (1978-1979, 1980-

	1989, 1990-1999, 2000-2009, 2010-2018). Exit data prior to 1978 or after 2018 are not available. Source: Census Bureau 1978-2019. Exit data are taken from the Business Dynamics Statistics (BDS). The BDS suppresses some exit data because of confidentiality reasons in case of very few exits in a region. These few missing data were imputed with values of 1 in case a random number was smaller than 0.5 and with 2 in case the random number was larger than 0.5 (Stata set seed 1001). The exit data refer to all non-agricultural industries. The denominator is the average number of establishments in the year t and the prior year $t-1$. Following guidance from the BDS, the longitudinally consistent number of establishments in year $t-1$ is computed by the number of establishments in year t plus establishment exits in year t minus establishment entries in year t. The exit rate in year t is finally computed as follows $\text{exit rate}_t = 100 \cdot \frac{\text{exits}_t}{0{,}5 \cdot (\text{establishments}_t + \text{establishments}_{t-1})}$
Job creation rate	Number of new jobs created through firm births per 100 jobs. We average the yearly job creation rates for each decade starting from the 1970s (1979, 1980-1989, 1990-1999, 2000-2009, 2010-2019). Data prior to 1978 are not available. Source: Census Bureau 1978-2019. Job creation data are taken from the Business Dynamics Statistics (BDS). The BDS suppresses some job data because of confidentiality reasons in case of very few jobs in a region. These few missing data were imputed with values of 1 in case a random number was smaller than 0.5 and with 2 in case the random number was larger than 0.5 (Stata set seed 1001). The job creation data refer to all non-agricultural industries. Overall employment data refer to the number of total jobs from the CAEMP25S file (Line Code 10) of the Bureau of Economic Analysis (BEA). The BEA data reports combined data of the independent Virginia cities with the surrounding counties. The combined data were split up using the population share on the base of the population share of the respective regions. In the spirit of Haltiwanger, Jarmin and Miranda (2013) the job creation rate is computed based on the average employment of the current year and the previous year. $\text{job creation rate}_t = 1.000 \cdot \frac{\text{Job creation through firm births}_t}{0{,}5 \cdot (\text{employment}_t + \text{employment}_{t-1})}$
Gold rush	Dummy variable: 1=County with a gold discovery between 1848 and 1899 in the western part of the United States. Source: U.S. Geological Survey. Data on gold discoveries and their location are taken from the Mineral Resource Data System (U.S. Geological Survey). Gold discoveries are deposits that include gold as primary commodity either exclusively or together with other minerals. The gold rush variable is set to 1 if there was at least one gold discovery in the county between 1848 and 1899 in contiguous West Census Region. We also define counties as gold rush counties in the states Michigan, Minnesota, South Dakota and Wisconsin in the Midwest Census region, which were sparsely populated in 1850.
Gold mining	Dummy variable: 1=County with a) a gold discovery later than 1900 in the Western Census Region as well as Michigan, Minnesota, South

	Dakota and Wisconsin or b) a gold discovery at any point in time in the Eastern part of the U.S.
Iron field	Dummy variable: 1=County with iron deposits. Source: Tarr and McCurry (1910). We digitized the map of the iron deposits in 1909.
Coal field	Dummy variable: 1=County with coal deposits. Source: Source: Tarr and McCurry (1910). We digitized the map of the coalfields in 1909.
Petroleum field	Dummy variable: 1=County with Petroleum deposits. Source: Lujala, Ketil and Thieme. (2007).
Elevation (Mean)	County-level average terrain elevation in km. Source: IIASA and FAO 2012: See Fischer et al. 2012.
Elevation (SD)	Standard deviation of counties terrain elevation in km. Source: IIASA and FAO 2012: See Fischer et al. 2012.
Coastal county	Dummy variable: 1=Coastal watershed county. Source: NOAA (2017).
Distance to lake	Distance of a county (centroid) to the nearest lake in 100 km. Source: Manson et al. (2019).
Distance to river	Distance of a county (centroid) to the nearest river in 100 km. Source: Manson et al. (2019).
Portage	Dummy variable: Within 15km to portage site. Source: Bleakley and Lin (2012).
Railroad access by 1910	Dummy variable: 1=County with railroad crossing by 1910. Source: Bazzi, Fiszbein and Gebresilasse (2020) based on the original dataset of Atack et al. (2010).
Distance to state capital	Distance of a county (centroid) to the State capital. Source: Brodeur and Haddad (2021).
Potential agricultural yield	Average Potential Agricultural Yield: 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. Source: Source: IIASA and FAO 2012: See Fischer et al. 2012. 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 0.083X0.083 degrees. We use FAO's measures of agro-climatic yields (based solely on climate, not on soil conditions) for intermediate levels of inputs/technology and rain-fed conditions.
Hispanic influence	Percentage of the regional population with Hispanic ties for the period between 1850 and 1870. Source: U.S. Census (Ruggles et al. 2020) HISPRULE is a pre-computed variable in the Census showing why a person in the Census was coded as Spanish/Hispanic/Latino. HISPRULE has a code ranging from e.g., 0=Person is not Hispanic, 1=Person was born in a Hispanic area, 5=Person is a relative of a householder who is Hispanic because of rules 1-4, 6=The person has a Spanish surname. For

	a complete list of the coding, see https://usa.ipums.org/usa-action/variables/HISPRULE#description_section. We compute the population share in a region with a HISPRULE value larger than 0 in the States that were under Spanish/Mexican rule for an extended period of time (AL, AZ, CA, CO, FL, LO, MI, NM, NV, TX, UT). To avoid spurious correlations, the value of the variable Hispanic Influence is coded as 0 in the remaining U.S. States.
French influence	Percentage of the regional population speaking French (incl. Patois and Cajun) and French Creole in 2010. Source: American Community Survey 2010 (5yr average). U.S. Census Bureau. Table B16001: Language Spoken at Home by Ability to Speak English. We compute the population share of French speakers in a region in the States that were under French rule for an extended period of time (AL, AR, IL, IO, KS, KY, LA, ME, MI, MN, MS, MO, NH, NY, ND, OH, PA, VT, WI). To avoid spurious correlations, the value of the variable French Influence is coded as 0 in the remaining U.S. States. One might argue that historic data on French speakers are to be preferred over contemporary data. However, the 19th century Census data do not include information on language spoken at home. A visual inspection of the 2010 data we use reveals that the hotspots of French speakers in Louisiana and Vermont in 2010 are the same as in the mid-20th century (Haden, 1973).
Latitude	Geographical latitude of a counties centroid.
Longitude	Geographical longitude of a counties centroid.
Employment share for Agriculture Manu. and construction Services Government	Employment share in agriculture; manufacturing, mining and construction; private services; government. We average the yearly employment shares for each decade starting from the 1970s (1970-1979, 1980-1989, 1990-1999, 2000-2009, 2010-2019). Sources: 1910-1940: U.S. Census (Ruggles et al. 2020). 1970-2019: Bureau of Economic Analysis 1970-2020. Employment by County, Metro and Other Areas: CAEMP25S and CAEMP25N file. We apply the same GIS methods to attribute historic counties to contemporary ones as described above. Note, that a few counties have either missing observations in the years 1910 to 1940. In these cases, we imputed the 1910-1940 data based on the trend from earlier or later census data. Regarding the data from 1970 onwards, we imputed the data based on population shares. Despite these efforts, especially data on employment in agriculture, but also for many service industries, from 2010 onwards suffer from missings as the BEA aims to avoid disclosure of identifiable data points. For this reason, the agriculture employment share is not included in the models but serves as reference category.
Unemployment rate	Number of unemployed as the percentage of the laborforce. We average the yearly unemployment rates for each decade starting from the 1970s (1976-1979, 1980-1989, 1990-1999, 2000-2009, 2010-2019). Source: Bureau of Labor Statistics 1976-2020: Local Area Unemployment Statistics. Note that official BLS time series for unemployment begin 1990. However, from the BLS unofficial unemployment data are available from 1976 onwards. These data are marked as unofficial as they are on methodological grounds incomparable to the series from 1990. As we use the unemployment data not as time series but separately for the decades, this issue does not apply to us.
Laborforce participation rate	Number of people in the laborforce in percent of the population aged 16 years or older for the years 1910, 1920, 1930 and 1940.

	Source: Data on laborforce and the age 16+ population are taken from the U.S. Census (Ruggles et al. 2020). We apply the same GIS methods to attribute historic counties to contemporary ones as described above. Note, that a few counties either have missing observations in some years or are not covered properly by the look-up table. In these cases, we impute the data based on the trend from earlier or later census data.
Population shares for Caucasian pop. Black American pop. Native American pop. Asian American pop.	Percentage of the Caucasian Population; Black American Population; Native American Population; Asian American + Pacific Islanders Population. Sources: 1910-1940: U.S. Census (Ruggles et al. 2020) 1970: Census of the Population (1970), Supplementary report, Race of the population by county. Report number PC(S1)-104. 1980-1990: Haines (2018). U.S. County populations by single years of age, sex, and race, 1970, 1980, 1990 (ICPSR 37115). 2000-2010: Census of the Population. 2000-2010. U.S. Census Bureau; Census of Population and Housing, Census Profiles of General Demographic Characteristics. Census of the population by age, sex, and race data files. Note, that a few counties either have missing observations in some years or are not covered properly by the look-up tables. In these cases, we imputed the 1910-1940 data based on the trend from earlier or later census data. Regarding the data from 1970 onwards, we imputed the data based on population shares. In the regressions, the Caucasian pop. share is not included in the models but serves as reference category.
Population density	Population densities are measured across the decades as 1,000 people per square mile. Sources: Manson et al. (2019). We use the crosswalk files from Eckert, F. et al (2020) to map historic counties to present-day counties. We conducted several robustness checks to ensure accuracy of the population numbers over time. 1) Quite often a county reports zero population even if has a non-zero population number in the census before or after. This is often the case due to county splitups and changing allocation procentages over the decades. We replace the zero population number with a value based on the growth rates between the previous and the following census 2) Sometimes, the reported population number switches from high values to low values back to high values, which is clearly an error in the data. Thus, we checked for a population drop of a factor of 1/10 and an increase by a factor of 10 between two Census data points. In such a case the errorenous number will be imputed from the census prior and after the respective observation. When computing the population densities, we follow closely the procedure by Bazzi, Fiszbein and Gebresilasse (2020). We interpolate county-level population for inter-censal years by assuming a constant annual population growth rate, which matches the decadal growth rate. We thereby replace initial zeros with one to avoid infinite growth rates. This procedure provides us with a population number for each county each year.

Table A2: Correlations regarding main variables

	Variables	Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1	Gold rush county	0.06	0.24	1.00																											
2	Gold mining county	0.03	0.18	-0.05	1.00																										
3	Self-emp. rate 1910	2.71	1.30	0.35	-0.05	1.00																									
4	Self-emp. rate 1920	2.37	1.12	0.31	-0.07	0.75	1.00																								
5	Self-emp. rate 1930	2.20	0.90	0.24	-0.08	0.71	0.79	1.00																							
6	Self-emp. rate 1940	3.46	1.26	0.25	-0.08	0.65	0.72	0.78	1.00																						
7	Startup rate 1970s	6.34	2.55	0.27	0.00	0.12	0.12	0.15	0.18	1.00																					
8	Startup rate 1980s	5.79	1.92	0.27	0.00	0.16	0.15	0.18	0.19	0.82	1.00																				
9	Startup rate 1990s	4.92	1.57	0.26	0.01	0.10	0.11	0.13	0.13	0.72	0.87	1.00																			
10	Startup rate 2000s	4.28	1.39	0.29	0.05	0.11	0.12	0.13	0.14	0.61	0.75	0.87	1.00																		
11	Startup rate 2010s	3.35	1.07	0.27	0.04	0.13	0.15	0.16	0.17	0.55	0.66	0.76	0.87	1.00																	
12	Entr. pers. profile	-21.11	0.64	0.20	0.05	0.31	0.30	0.30	0.27	0.16	0.19	0.18	0.20	0.24	1.00																
13	Iron field	0.11	0.32	0.06	0.00	-0.06	-0.05	-0.08	-0.09	0.00	0.01	0.02	0.00	-0.01	-0.11	1.00															
14	Coal field	0.29	0.45	-0.06	-0.09	-0.09	-0.08	-0.10	-0.08	0.03	-0.04	-0.06	-0.10	-0.08	-0.11	0.09	1.00														
15	Petroleum field	0.30	0.46	-0.10	-0.11	-0.05	-0.07	-0.04	-0.03	-0.02	-0.06	-0.11	-0.15	-0.14	-0.16	0.01	0.31	1.00													
16	Coastal county	0.21	0.41	-0.06	-0.01	0.11	0.09	0.12	0.05	0.05	0.15	0.14	0.13	0.12	0.10	-0.09	-0.24	-0.05	1.00												
17	Distance lakes	1.65	1.11	-0.10	-0.09	0.03	0.07	0.08	0.12	-0.04	-0.08	-0.07	-0.08	-0.07	-0.06	-0.06	0.22	0.07	-0.18	1.00											
18	Distance rivers	0.44	0.39	0.05	-0.03	0.07	0.09	0.13	0.13	0.13	0.15	0.11	0.09	0.13	0.10	-0.09	-0.05	-0.01	0.22	-0.03	1.00										
19	Portage	0.01	0.11	-0.03	0.01	-0.04	-0.04	-0.05	-0.05	0.03	0.05	0.04	0.02	-0.00	-0.01	-0.01	-0.01	-0.05	-0.00	-0.04	0.02	1.00									
20	Railroad access	0.11	0.31	0.04	-0.02	-0.13	-0.13	-0.12	-0.05	0.10	0.07	0.08	0.09	0.10	-0.02	-0.02	-0.01	0.08	-0.04	-0.03	0.10	-0.03	1.00								
21	Distance to State capital	2.26	1.55	0.04	-0.02	0.03	0.05	0.06	0.08	0.05	0.04	0.03	0.04	0.07	0.01	-0.00	-0.04	0.12	-0.09	-0.17	0.13	-0.07	0.18	1.00							
22	Elevation (Mean)	0.40	0.49	0.59	0.06	0.27	0.27	0.24	0.25	0.25	0.22	0.25	0.32	0.32	0.23	0.08	0.01	-0.04	-0.28	-0.10	0.05	-0.06	0.16	0.28	1.00						
23	Elevation (SD)	0.02	0.03	0.12	0.12	0.06	0.06	0.02	0.02	0.04	0.08	0.09	0.12	0.10	0.06	0.17	-0.04	-0.10	-0.06	-0.04	-0.10	-0.02	0.01	-0.01	0.21	1.00					
24	Potential agri. yield	0.46	0.19	-0.48	0.02	-0.31	-0.32	-0.34	-0.34	-0.21	-0.17	-0.15	-0.19	-0.24	-0.25	0.04	0.05	-0.04	0.18	0.12	-0.20	0.08	-0.24	-0.42	-0.78	-0.02	1.00				
25	Hispanic influece	3.77	14.79	0.23	-0.01	0.12	0.10	0.12	0.10	0.07	0.11	0.06	0.04	-0.00	0.13	-0.02	-0.03	0.07	-0.03	-0.02	0.10	-0.02	0.14	0.24	0.34	-0.02	-0.48	1.00			
26	French influence	0.36	1.98	-0.05	-0.03	-0.03	-0.03	-0.01	-0.04	-0.01	-0.02	-0.03	-0.03	-0.04	-0.08	-0.02	-0.05	0.06	0.15	-0.09	-0.00	-0.01	-0.02	-0.08	-0.09	-0.03	0.07	-0.04	1.00		
27	Latitude	38.28	4.84	0.13	-0.02	0.21	0.23	0.24	0.23	-0.07	-0.10	-0.04	0.00	0.09	0.07	0.03	-0.02	-0.17	-0.07	-0.03	-0.03	-0.09	-0.06	-0.00	0.25	0.08	-0.34	-0.25	-0.04	1.00	
28	Longitude	-91.66	11.49	-0.49	0.03	-0.24	-0.25	-0.25	-0.29	-0.24	-0.14	-0.13	-0.16	-0.20	-0.18	0.12	-0.04	-0.06	0.26	-0.05	-0.12	0.06	-0.17	-0.30	-0.60	0.05	0.77	-0.30	0.10	-0.19	1.00

Note: Correlations above |0.05| are significant at the 1% level.

1. Results regarding historic and contemporary entrepreneurship

In Table S2 we present the full regression results including the socio-economic controls. We expected that other natural resources such as coal, iron, and petroleum have a negative relationship with entrepreneurship because of the large-scale nature of its processing industries. In Tables S2, the signs of the respective regression coefficients are always negative but only rarely significant different from zero. These findings somewhat support other work on industrialization on entrepreneurship (Glaeser, Kerr and Kerr 2015; Stuetzer et al. 2016).

We expected higher levels of entrepreneurship in regions with better connections to the large U.S. domestic market and find some evidence for this. Self-employment, startup rates and the regional entrepreneurial personality profile are consistently greater in coastal counties. Railroad access affect entrepreneurship rates from the 1970s onwards. However, many of the infrastructure variables do not predict contemporary entrepreneurship and only weakly predict historic self-employment, which is in line with findings in Germany (Audretsch, Heger and Veith, 2015).

We find that regions with higher potential agricultural yields had somewhat lower entrepreneurship rates during the 1910 to 1940 period. This result is somewhat surprising as other studies argue that soil quality and agricultural yields to be a predictor for early high population density and productivity prior to industrialization (Combes et al. 2010; Falck et al. 2012). The negative correlation in the U.S. context might be explained by the nature of the agricultural use. A visual inspection of geographic distribution reveals that counties with the highest potential agricultural yield are clustered in the south east where large slave-operated plantations cultivated cash crops. In these regions, local institutions denied economic freedom to many blacks prior to the civil war which only gradually improved over time arguably resulting in overall lower entrepreneurship rates (Walker 1986).

Another set of historic controls were included to capture the institutional and cultural imprint from other colonial nations in the U.S. Our results in Table B3 (Model 1-4) show a stronger French or Spanish imprint does not seem to influence entrepreneurship between 1910 to 1940. In later periods, counties with a stronger Hispanic and French influence have lower entrepreneurship rates. However, as there is no continuous effect of both variables over time, the more recent significant effects are probably due to more recent developments taking place in these counties rather than an effect of the Hispanic and French cultural roots.

A final set of controls was included to capture the impact of the more recent socio-economic structure on entrepreneurship. Among the consistent findings in Table S2 is that counties with higher minority population shares have lower entrepreneurship rates.

Table A3: Basic results regarding entrepreneurship indicators.

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Gold rush	0.527***	0.411***	0.074	0.401***	0.857**	0.686**	0.434**	0.437***	0.341***	0.146**
	(0.151)	(0.103)	(0.101)	(0.132)	(0.333)	(0.270)	(0.202)	(0.157)	(0.121)	(0.056)
Gold mining	-0.221	-0.239***	-0.163	-0.122	-0.041	-0.431**	-0.354***	-0.183*	-0.110	0.059
	(0.147)	(0.080)	(0.102)	(0.111)	(0.219)	(0.184)	(0.095)	(0.096)	(0.079)	(0.051)
Iron field	-0.268***	-0.072	-0.042	-0.067	-0.030	-0.063	-0.064	-0.064	-0.070	-0.046
	(0.084)	(0.079)	(0.062)	(0.074)	(0.171)	(0.143)	(0.133)	(0.122)	(0.090)	(0.046)
Coal field	-0.153	-0.112*	-0.116**	-0.084	0.303**	0.024	-0.021	-0.080	-0.025	-0.033
	(0.093)	(0.057)	(0.053)	(0.061)	(0.120)	(0.092)	(0.083)	(0.090)	(0.090)	(0.033)
Petroleum field	-0.036	-0.089	-0.052	-0.025	-0.076	-0.088	-0.157**	-0.123*	-0.057	-0.018
	(0.082)	(0.058)	(0.047)	(0.047)	(0.120)	(0.093)	(0.069)	(0.073)	(0.064)	(0.038)
Coastal county	0.193**	0.112	0.198***	0.246**	0.471**	0.534**	0.508***	0.495***	0.308***	0.029
	(0.087)	(0.077)	(0.073)	(0.102)	(0.203)	(0.202)	(0.163)	(0.147)	(0.109)	(0.039)
Distance to lake	0.067	0.089*	0.063*	0.084	-0.023	0.017	-0.028	-0.030	-0.053	-0.052***
	(0.040)	(0.046)	(0.035)	(0.050)	(0.093)	(0.073)	(0.067)	(0.065)	(0.060)	(0.018)
Distance to river	-0.004	0.020	0.051	0.053	0.292	0.280*	0.200	0.104	0.156*	0.059*
	(0.091)	(0.075)	(0.069)	(0.073)	(0.201)	(0.147)	(0.132)	(0.111)	(0.089)	(0.035)
Portage	-0.009	0.108	0.036	0.092	0.659	0.804	0.513	0.349*	0.195**	-0.136
	(0.231)	(0.093)	(0.095)	(0.105)	(0.713)	(0.673)	(0.426)	(0.193)	(0.088)	(0.103)
Railroad access	-0.517***	-0.465***	-0.428***	-0.370***	0.676***	0.451***	0.366***	0.318**	0.210**	-0.090
	(0.087)	(0.078)	(0.061)	(0.100)	(0.194)	(0.112)	(0.129)	(0.129)	(0.089)	(0.078)
Distance to State capital	-0.024	-0.038	-0.036	-0.063*	-0.068*	-0.033	-0.025	-0.036	-0.032	-0.026**
	(0.029)	(0.029)	(0.026)	(0.036)	(0.036)	(0.025)	(0.023)	(0.025)	(0.019)	(0.012)
Elevation (Mean)	0.130	0.210	0.103	0.343*	1.096	0.537	0.635	1.026**	0.669**	0.148*
	(0.224)	(0.186)	(0.134)	(0.191)	(0.712)	(0.483)	(0.440)	(0.472)	(0.331)	(0.083)
Elevation (SD)	1.280	0.300	0.210	1.133*	-1.280	-0.650	-0.342	0.537	-0.106	0.488
	(0.832)	(0.513)	(0.670)	(0.627)	(2.263)	(1.963)	(1.804)	(1.242)	(0.773)	(0.318)
Potential agricultural yield	-1.336**	-0.515	-1.208**	-0.751	0.416	0.731	1.436*	1.057	-0.146	-0.526
	(0.631)	(0.623)	(0.540)	(0.630)	(0.960)	(0.925)	(0.745)	(0.793)	(0.657)	(0.323)
Hispanic influence	-0.003	0.002	-0.001	-0.006	-0.042***	-0.021***	-0.013***	-0.008**	-0.009**	0.003
	(0.008)	(0.007)	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)
French influence	0.003	0.005	-0.002	-0.012	-0.014	-0.039***	-0.030***	-0.024***	-0.021***	-0.015***
	(0.007)	(0.004)	(0.006)	(0.009)	(0.011)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)
Latitude	0.001	-0.056	-0.021	-0.021	-0.030	0.008	-0.004	-0.010	-0.007	-0.002
	(0.040)	(0.037)	(0.039)	(0.045)	(0.055)	(0.052)	(0.046)	(0.043)	(0.038)	(0.019)
Longitude	0.027	0.006	-0.002	0.002	-0.053	-0.030	-0.047*	-0.013	0.001	-0.006
	(0.025)	(0.025)	(0.020)	(0.031)	(0.042)	(0.030)	(0.026)	(0.025)	(0.021)	(0.010)
Empl. share manu. & cons.	0.024***	0.009***	0.005**	-0.008	-0.012*	-0.012*	-0.016***	-0.059***	-0.054***	-0.001
	(0.005)	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.005)	(0.013)	(0.014)	(0.023)
Empl. share services	0.045***	0.054***	0.023***	0.049***	0.074***	0.073***	0.048***	-0.026*	-0.035**	0.016
	(0.010)	(0.007)	(0.004)	(0.014)	(0.007)	(0.005)	(0.004)	(0.015)	(0.017)	(0.022)

Empl. share government	0.018**	0.005	0.003	-0.030***	0.003	-0.008	-0.015**	-0.070***	-0.063***	0.009
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.006)	(0.007)	(0.016)	(0.017)	(0.023)
Laborforce participation rate	0.017*	0.057***	0.044***	0.087***	----	----	----	----	----	----
	(0.009)	(0.009)	(0.006)	(0.008)						
Unemp. rate	----	----	----	----	0.085**	0.049**	0.046***	0.026	0.044	-0.094***
					(0.034)	(0.020)	(0.017)	(0.032)	(0.031)	(0.019)
Black American pop. share	0.004	-0.005	-0.003	-0.013***	-0.036***	-0.037***	-0.031***	-0.021***	-0.018***	0.011***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.002)
Native American pop. share	-0.003	0.014	-0.008	0.001	-0.022	-0.016***	-0.016***	-0.011***	-0.012***	0.000
	(0.019)	(0.013)	(0.009)	(0.016)	(0.019)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
Asian American pop. share	0.083	-0.061***	-0.093***	-0.199***	-0.588*	-0.424***	-0.190***	-0.082***	-0.001	0.019*
	(0.070)	(0.022)	(0.024)	(0.047)	(0.333)	(0.105)	(0.055)	(0.027)	(0.015)	(0.009)
Pop. density	0.000	0.000	0.000	-0.000	-0.000	-0.000	0.000**	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	3.155	0.289	-0.346	-0.738	-0.241	0.199	-0.933	7.285**	7.900**	-22.185***
	(3.148)	(2.952)	(2.770)	(3.667)	(4.910)	(3.698)	(3.051)	(3.340)	(3.632)	(2.383)
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.420	0.471	0.462	0.445	0.305	0.420	0.386	0.399	0.391	0.327

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, employment share in manufacturing, mining and construction, employment share in private services, employment share in government, laborforce participation rate (1910-1940), unemployment rate (1970-present day), Black American population share, Native American population share, Asian American population share, and population density. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table A4: Results regarding high-quality entrepreneurship.

VARIABLES	(1) High quality startup rate 2010s	(2) High quality startup rate 2010s	(3) High quality startup rate 2010s
Gold rush	-0.539	-0.384	-0.175
	(1.302)	(1.036)	(0.604)
Gold mining	-0.382	-0.156	-0.363
	(0.463)	(0.364)	(0.324)
Iron field		-0.021	0.126
		(0.149)	(0.126)
Coal field		-0.023	0.079
		(0.114)	(0.108)
Petroleum field		0.039	0.072
		(0.126)	(0.152)
Coastal county		0.974**	0.425***
		(0.374)	(0.115)
Distance to lake		0.133	0.048
		(0.112)	(0.049)
Distance to river		0.333	0.410*
		(0.289)	(0.214)
Portage		0.455	0.209
		(0.410)	(0.238)
Railroad access		-0.188	0.152
		(0.185)	(0.121)
Distance to State capital		-0.052	0.003
		(0.048)	(0.037)
Elevation (Mean)		-0.602	0.098
		(0.640)	(0.304)
Elevation (SD)		10.293	3.611*
		(6.555)	(1.953)
Potential agricultural yield		-0.417	-0.347
		(1.022)	(0.762)
Hispanic influence		-0.006	-0.005
		(0.011)	(0.008)
French influence		-0.016	-0.031***
		(0.015)	(0.009)
Latitude		-0.001	-0.012
		(0.063)	(0.042)
Longitude		-0.036	-0.018
		(0.068)	(0.035)

Empl. share manu. & cons.			-0.027
			(0.016)
Empl. share services.			-0.017
			(0.011)
Empl. share government.			-0.038***
			(0.014)
Unemp. rate			-0.078**
			(0.037)
Black American pop. share			-0.003
			(0.004)
Native American pop. share			0.005
			(0.004)
Asian American pop. share			0.554***
			(0.158)
Pop. density			0.000
			(0.000)
Constant	0.669***	-2.492	1.899
	(0.021)	(5.968)	(3.980)
State-fixed effects	Yes	Yes	Yes
Observations	3,094	3,094	3,094
Adjusted R-squared	0.307	0.343	0.566

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. Dependent variable is the rate of high-quality startups per 1 Mill. employed. Note that data from Delaware across the whole time span is not available as well as data from Illinois, South Carolina and Illinois for the period of 2014-2016. Model 1: State-fixed effects. Model 2: State-fixed effects + geographic controls: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude. Model 3: Additional socio-economic controls of the 2010 decade: Employment share in manufacturing and construction, employment share in services, employment share in government, unemployment rate, Black American pop. share, Native American pop. share, Asian American pop. share, population density. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table A5: Results regarding alternative control groups: Without Appalachian gold mining counties

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Gold rush	0.544***	0.439***	0.109	0.415***	0.971***	0.764**	0.447**	0.444***	0.336**	0.159**
	(0.165)	(0.108)	(0.105)	(0.147)	(0.350)	(0.291)	(0.211)	(0.164)	(0.128)	(0.059)
Gold mining	-0.111	-0.012	0.110	0.004	0.744	0.119	-0.207	-0.023	-0.080	0.109
	(0.230)	(0.154)	(0.122)	(0.224)	(0.553)	(0.297)	(0.201)	(0.295)	(0.212)	(0.120)
Iron field	-0.261***	-0.071	-0.044	-0.062	-0.043	-0.061	-0.062	-0.066	-0.073	-0.037
	(0.087)	(0.080)	(0.065)	(0.076)	(0.181)	(0.145)	(0.135)	(0.124)	(0.093)	(0.043)
Coal field	-0.155	-0.114*	-0.118**	-0.086	0.302**	0.023	-0.022	-0.082	-0.025	-0.036
	(0.093)	(0.058)	(0.054)	(0.062)	(0.119)	(0.093)	(0.083)	(0.089)	(0.091)	(0.034)
Petroleum field	-0.034	-0.087	-0.051	-0.027	-0.079	-0.095	-0.166**	-0.130*	-0.062	-0.020
	(0.082)	(0.058)	(0.047)	(0.047)	(0.121)	(0.092)	(0.068)	(0.073)	(0.065)	(0.038)
Coastal county	0.206**	0.113	0.203**	0.242**	0.440**	0.508**	0.497***	0.489***	0.307***	0.026
	(0.091)	(0.079)	(0.077)	(0.107)	(0.197)	(0.200)	(0.163)	(0.148)	(0.110)	(0.039)
Distance to lake	0.066	0.089*	0.063*	0.083	-0.022	0.023	-0.022	-0.025	-0.049	-0.054***
	(0.041)	(0.046)	(0.036)	(0.051)	(0.092)	(0.073)	(0.067)	(0.064)	(0.061)	(0.019)
Distance to river	-0.009	0.014	0.049	0.054	0.299	0.278*	0.200	0.098	0.152*	0.058
	(0.091)	(0.074)	(0.068)	(0.073)	(0.201)	(0.149)	(0.133)	(0.111)	(0.090)	(0.036)
Portage	0.038	0.132	0.051	0.126	0.610	0.775	0.472	0.355*	0.198**	-0.158
	(0.227)	(0.100)	(0.095)	(0.099)	(0.704)	(0.669)	(0.404)	(0.202)	(0.095)	(0.121)
Railroad access	-0.522***	-0.468***	-0.431***	-0.370***	0.670***	0.445***	0.364***	0.315**	0.205**	-0.086
	(0.086)	(0.077)	(0.061)	(0.101)	(0.195)	(0.110)	(0.129)	(0.130)	(0.088)	(0.078)
Distance to State capital	-0.024	-0.037	-0.035	-0.063*	-0.062	-0.023	-0.016	-0.028	-0.027	-0.027**
	(0.029)	(0.029)	(0.026)	(0.036)	(0.037)	(0.025)	(0.022)	(0.024)	(0.019)	(0.012)
Elevation (Mean)	0.127	0.199	0.093	0.337*	1.070	0.542	0.659	1.052**	0.687**	0.125
	(0.227)	(0.185)	(0.133)	(0.190)	(0.725)	(0.481)	(0.438)	(0.472)	(0.333)	(0.085)
Elevation (SD)	1.213	0.331	0.163	0.905	-1.774	-1.011	-0.409	0.593	-0.047	0.238
	(0.921)	(0.544)	(0.735)	(0.709)	(2.127)	(1.809)	(1.724)	(1.141)	(0.740)	(0.298)
Potential agricultural yield	-1.342**	-0.481	-1.186**	-0.700	0.533	0.906	1.560**	1.171	-0.082	-0.554*
	(0.635)	(0.614)	(0.535)	(0.621)	(0.962)	(0.921)	(0.743)	(0.796)	(0.659)	(0.321)
Hispanic influence	-0.003	0.002	-0.001	-0.006	-0.041***	-0.021***	-0.013***	-0.008**	-0.009**	0.003
	(0.008)	(0.007)	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.002)
French influence	0.003	0.005	-0.002	-0.012	-0.014	-0.039***	-0.029***	-0.024***	-0.021***	-0.015**
	(0.007)	(0.004)	(0.006)	(0.009)	(0.012)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)
Latitude	0.000	-0.057	-0.021	-0.021	-0.032	0.003	-0.009	-0.014	-0.009	-0.002
	(0.041)	(0.038)	(0.039)	(0.045)	(0.054)	(0.051)	(0.045)	(0.042)	(0.038)	(0.019)
Longitude	0.027	0.005	-0.002	0.002	-0.053	-0.033	-0.050*	-0.016	-0.001	-0.005

	(0.025)	(0.025)	(0.020)	(0.031)	(0.042)	(0.031)	(0.026)	(0.025)	(0.021)	(0.009)
Empl. share manu. & cons.	0.024***	0.009***	0.006**	-0.008	-0.011	-0.011*	-0.015***	-0.060***	-0.054***	-0.000
	(0.005)	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.005)	(0.014)	(0.014)	(0.023)
Empl. share services	0.044***	0.054***	0.024***	0.049***	0.074***	0.073***	0.047***	-0.027*	-0.036**	0.017
	(0.010)	(0.007)	(0.005)	(0.014)	(0.008)	(0.005)	(0.004)	(0.015)	(0.017)	(0.022)
Empl. share government	0.017**	0.005	0.003	-0.031***	0.003	-0.009	-0.016**	-0.071***	-0.064***	0.010
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.006)	(0.007)	(0.017)	(0.017)	(0.023)
Laborforce participation rate	0.017*	0.058***	0.044***	0.089***	----	----	----	----	----	----
	(0.009)	(0.009)	(0.006)	(0.008)						
Unemp. rate	----	----	----	----	0.083**	0.048**	0.047***	0.036	0.048	-0.095***
					(0.034)	(0.020)	(0.017)	(0.031)	(0.031)	(0.019)
Black American pop. share	0.004	-0.005*	-0.003	-0.013***	-0.039***	-0.039***	-0.033***	-0.023***	-0.019***	0.011***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.006)	(0.005)	(0.004)	(0.004)	(0.003)	(0.002)
Native American pop. share	-0.002	0.015	-0.007	0.001	-0.021	-0.015***	-0.016***	-0.011***	-0.012***	0.001
	(0.019)	(0.013)	(0.009)	(0.015)	(0.019)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
Asian American pop. share	0.083	-0.060***	-0.092***	-0.200***	-0.555*	-0.398***	-0.180***	-0.075**	0.002	0.018*
	(0.070)	(0.022)	(0.024)	(0.047)	(0.324)	(0.099)	(0.057)	(0.029)	(0.017)	(0.010)
Pop. density	0.000	0.000	0.000	-0.000	-0.000	-0.000	0.000*	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	3.218	0.240	-0.371	-0.900	-0.179	0.092	-1.128	7.133**	7.828**	-22.123***
	(3.183)	(2.950)	(2.775)	(3.681)	(4.935)	(3.742)	(3.079)	(3.397)	(3.685)	(2.379)
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,036	3,036	3,036	3,036	3,036	3,036	3,036	3,036	3,036	2,022
Adjusted R-squared	0.418	0.466	0.454	0.439	0.308	0.427	0.389	0.404	0.395	0.332
P-value gold rush>gold mining	0	0.003	0.499	0.008	0.337	0.005	0.001	0.042	0.016	0.325

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. Alternative control group: Dropping Appalachian gold mining counties from the control group. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, employment share in manufacturing, mining and construction, employment share in private services, employment share in government, laborforce participation rate (1910-1940), unemployment rate (1970-present day), Black American population share, Native American population share, Asian American population share, and population density. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the gold mining coefficient. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A6: Results regarding alternative control groups: With counties incorporated between 1848 and 1899 as control group

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. Profile
Gold rush	0.511***	0.426***	0.082	0.400***	0.855**	0.759***	0.502**	0.479***	0.359***	0.122**
	(0.141)	(0.103)	(0.099)	(0.128)	(0.343)	(0.267)	(0.192)	(0.151)	(0.114)	(0.057)
Incorporated 1848-1899	-0.167***	-0.073*	-0.058*	-0.064	-0.026	0.014	0.032	0.036	-0.001	-0.034
	(0.036)	(0.042)	(0.033)	(0.045)	(0.092)	(0.073)	(0.066)	(0.045)	(0.071)	(0.024)
Iron field	-0.260***	-0.067	-0.038	-0.064	-0.028	-0.059	-0.061	-0.063	-0.069	-0.047
	(0.082)	(0.078)	(0.061)	(0.073)	(0.172)	(0.144)	(0.134)	(0.122)	(0.091)	(0.046)
Coal field	-0.152	-0.108*	-0.114**	-0.084	0.303**	0.031	-0.015	-0.076	-0.023	-0.034
	(0.091)	(0.055)	(0.051)	(0.060)	(0.121)	(0.093)	(0.083)	(0.088)	(0.089)	(0.034)
Petroleum field	-0.028	-0.082	-0.046	-0.021	-0.075	-0.075	-0.147**	-0.119	-0.054	-0.018
	(0.080)	(0.057)	(0.046)	(0.046)	(0.120)	(0.094)	(0.068)	(0.073)	(0.064)	(0.037)
Coastal county	0.182**	0.114	0.198***	0.243**	0.469**	0.554***	0.527***	0.508***	0.313***	0.022
	(0.082)	(0.076)	(0.072)	(0.104)	(0.205)	(0.206)	(0.164)	(0.147)	(0.107)	(0.038)
Distance to lake	0.071*	0.092**	0.065*	0.085*	-0.023	0.019	-0.026	-0.030	-0.052	-0.052***
	(0.038)	(0.045)	(0.034)	(0.049)	(0.093)	(0.075)	(0.068)	(0.065)	(0.061)	(0.018)
Distance to river	0.003	0.023	0.054	0.056	0.293	0.278*	0.197	0.102	0.156*	0.062*
	(0.089)	(0.075)	(0.068)	(0.072)	(0.201)	(0.147)	(0.132)	(0.110)	(0.091)	(0.035)
Portage	-0.007	0.115	0.040	0.094	0.659	0.822	0.529	0.357*	0.200**	-0.139
	(0.234)	(0.098)	(0.097)	(0.106)	(0.718)	(0.679)	(0.430)	(0.194)	(0.087)	(0.104)
Railroad access	-0.517***	-0.462***	-0.425***	-0.369***	0.676***	0.458***	0.371***	0.321**	0.212**	-0.091
	(0.087)	(0.077)	(0.062)	(0.101)	(0.195)	(0.115)	(0.131)	(0.130)	(0.089)	(0.077)
Distance to State capital	-0.021	-0.036	-0.035	-0.062*	-0.068*	-0.032	-0.024	-0.035	-0.031	-0.026**
	(0.028)	(0.029)	(0.026)	(0.035)	(0.036)	(0.025)	(0.023)	(0.025)	(0.019)	(0.012)
Elevation (Mean)	0.107	0.190	0.090	0.331*	1.092	0.509	0.612	1.014**	0.661*	0.150*
	(0.222)	(0.185)	(0.133)	(0.190)	(0.713)	(0.487)	(0.440)	(0.472)	(0.329)	(0.082)
Elevation (SD)	1.396*	0.396	0.281	1.192*	-1.258	-0.523	-0.236	0.593	-0.066	0.477
	(0.806)	(0.524)	(0.664)	(0.629)	(2.298)	(2.029)	(1.835)	(1.262)	(0.782)	(0.311)
Potential agricultural yield	-1.398**	-0.569	-1.245**	-0.785	0.407	0.670	1.390*	1.035	-0.163	-0.518
	(0.616)	(0.620)	(0.543)	(0.632)	(0.968)	(0.941)	(0.751)	(0.795)	(0.654)	(0.328)
Hispanic influence	-0.003	0.002	-0.001	-0.006	-0.042***	-0.021***	-0.013***	-0.008**	-0.009**	0.003
	(0.008)	(0.007)	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)
French influence	0.004	0.005	-0.002	-0.012	-0.014	-0.039***	-0.030***	-0.025***	-0.021***	-0.015***
	(0.007)	(0.004)	(0.006)	(0.009)	(0.011)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)
Latitude	0.005	-0.053	-0.019	-0.019	-0.029	0.010	-0.003	-0.009	-0.006	-0.002
	(0.040)	(0.038)	(0.039)	(0.045)	(0.055)	(0.052)	(0.045)	(0.043)	(0.039)	(0.019)
Longitude	0.029	0.008	-0.001	0.003	-0.053	-0.026	-0.044*	-0.011	0.002	-0.007

	(0.025)	(0.026)	(0.021)	(0.031)	(0.042)	(0.031)	(0.026)	(0.025)	(0.021)	(0.009)
Empl. share manu. & cons.	0.024***	0.009***	0.005**	-0.008	-0.012*	-0.012*	-0.017***	-0.060***	-0.054***	-0.001
	(0.005)	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.006)	(0.014)	(0.014)	(0.023)
Empl. share services	0.044***	0.053***	0.023***	0.049***	0.073***	0.073***	0.047***	-0.027*	-0.035**	0.016
	(0.010)	(0.007)	(0.004)	(0.014)	(0.007)	(0.005)	(0.004)	(0.015)	(0.017)	(0.022)
Empl. share government	0.018**	0.005	0.003	-0.030***	0.003	-0.008	-0.015**	-0.070***	-0.063***	0.010
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.006)	(0.007)	(0.016)	(0.017)	(0.023)
Laborforce participation rate	0.016*	0.057***	0.043***	0.087***	----	----	----	----	----	----
	(0.009)	(0.009)	(0.006)	(0.008)						
Unemp. rate	----	----	----	----	0.085**	0.049**	0.045**	0.024	0.043	-0.093***
					(0.034)	(0.019)	(0.017)	(0.032)	(0.031)	(0.019)
Black American pop. share	0.005	-0.004	-0.003	-0.013***	-0.036***	-0.036***	-0.031***	-0.021***	-0.018***	0.010***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.002)
Native American pop. share	-0.004	0.014	-0.008	0.000	-0.022	-0.016***	-0.015***	-0.010***	-0.011***	0.000
	(0.019)	(0.013)	(0.009)	(0.016)	(0.019)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
Asian American pop. share	0.088	-0.057**	-0.091***	-0.196***	-0.587*	-0.423***	-0.191***	-0.083***	-0.001	0.019*
	(0.069)	(0.023)	(0.024)	(0.048)	(0.334)	(0.107)	(0.056)	(0.027)	(0.015)	(0.009)
Pop. density	0.000	0.000	0.000	-0.000	-0.000	-0.000	0.000**	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	3.344	0.450	-0.230	-0.641	-0.217	0.463	-0.716	7.422**	7.987**	-22.281***
	(3.105)	(2.950)	(2.764)	(3.654)	(4.919)	(3.729)	(3.059)	(3.352)	(3.632)	(2.376)
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.422	0.470	0.462	0.446	0.305	0.419	0.384	0.398	0.390	0.328
P-value gold	0.000	0.000	0.099	0.001	0.006	0.005	0.013	0.004	0.005	0.005

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. Alternative control group: We use as control group all counties without gold rush but an incorporation date between 1848 and 1899. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, employment share in manufacturing, mining and construction, employment share in private services, employment share in government, laborforce participation rate (1910-1940), unemployment rate (1970-present day), Black American population share, Native American population share, Asian American population share, and population density. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the incorporated coefficient. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A7: Results regarding alternative control groups: With counties initially organized between 1848 and 1899 as control group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	Self-empl. rate 1910	Self-empl. rate 1920	Self-empl. rate 1930	Self-empl. rate 1940	Startup rate 1970s	Startup rate 1980s	Startup rate 1990s	Startup rate 2000s	Startup rate 2010s	Entr. pers. profile
Gold rush	0.516***	0.457***	0.085	0.379**	0.932**	0.852***	0.517**	0.539***	0.433***	0.078
	(0.164)	(0.127)	(0.105)	(0.143)	(0.403)	(0.271)	(0.220)	(0.167)	(0.122)	(0.066)
Initially organized 1848-1899	-0.093	0.016	-0.032	-0.082	0.133	0.191	0.051	0.140	0.145	-0.110**
	(0.110)	(0.101)	(0.095)	(0.117)	(0.267)	(0.187)	(0.170)	(0.134)	(0.115)	(0.052)
Iron field	-0.266***	-0.069	-0.040	-0.065	-0.030	-0.059	-0.061	-0.063	-0.070	-0.046
	(0.084)	(0.079)	(0.062)	(0.073)	(0.171)	(0.143)	(0.133)	(0.122)	(0.090)	(0.046)
Coal field	-0.149	-0.107*	-0.113**	-0.082	0.304**	0.031	-0.016	-0.077	-0.022	-0.034
	(0.093)	(0.057)	(0.052)	(0.060)	(0.120)	(0.094)	(0.084)	(0.090)	(0.090)	(0.033)
Petroleum field	-0.030	-0.083	-0.047	-0.021	-0.075	-0.075	-0.147**	-0.118	-0.054	-0.020
	(0.082)	(0.058)	(0.047)	(0.046)	(0.120)	(0.094)	(0.068)	(0.073)	(0.064)	(0.037)
Coastal county	0.205**	0.123	0.207***	0.254**	0.469**	0.547**	0.522***	0.500***	0.309***	0.029
	(0.087)	(0.078)	(0.073)	(0.103)	(0.208)	(0.209)	(0.166)	(0.148)	(0.109)	(0.038)
Distance to lake	0.069*	0.091*	0.064*	0.085*	-0.024	0.018	-0.026	-0.030	-0.053	-0.053***
	(0.040)	(0.046)	(0.035)	(0.050)	(0.093)	(0.074)	(0.068)	(0.065)	(0.060)	(0.018)
Distance to river	-0.005	0.020	0.051	0.052	0.292	0.279*	0.199	0.104	0.157*	0.060*
	(0.091)	(0.074)	(0.069)	(0.073)	(0.201)	(0.146)	(0.132)	(0.111)	(0.090)	(0.035)
Portage	-0.004	0.120	0.041	0.093	0.667	0.831	0.530	0.363*	0.207**	-0.143
	(0.230)	(0.094)	(0.096)	(0.107)	(0.717)	(0.677)	(0.428)	(0.190)	(0.084)	(0.103)
Railroad access	-0.514***	-0.460***	-0.425***	-0.368***	0.678***	0.460***	0.371***	0.322**	0.213**	-0.091
	(0.086)	(0.076)	(0.062)	(0.100)	(0.195)	(0.114)	(0.131)	(0.130)	(0.089)	(0.078)
Distance to State capital	-0.023	-0.037	-0.036	-0.063*	-0.068*	-0.033	-0.024	-0.036	-0.032	-0.026**
	(0.029)	(0.029)	(0.026)	(0.035)	(0.036)	(0.025)	(0.023)	(0.025)	(0.019)	(0.012)
Elevation (Mean)	0.105	0.195	0.090	0.327*	1.107	0.528	0.615	1.027**	0.677**	0.142*
	(0.223)	(0.185)	(0.132)	(0.190)	(0.709)	(0.493)	(0.441)	(0.477)	(0.332)	(0.081)
Elevation (SD)	1.348	0.370	0.266	1.183*	-1.277	-0.531	-0.230	0.594	-0.079	0.474
	(0.829)	(0.528)	(0.670)	(0.634)	(2.310)	(2.038)	(1.840)	(1.269)	(0.785)	(0.312)
Potential agricultural yield	-1.388**	-0.559	-1.239**	-0.776	0.420	0.685	1.390*	1.044	-0.148	-0.531
	(0.635)	(0.619)	(0.544)	(0.631)	(0.958)	(0.950)	(0.753)	(0.805)	(0.659)	(0.332)
Hispanic influence	-0.003	0.002	-0.001	-0.006	-0.042***	-0.021***	-0.013***	-0.008*	-0.009**	0.003
	(0.008)	(0.007)	(0.005)	(0.005)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)
French influence	0.003	0.005	-0.002	-0.012	-0.014	-0.039***	-0.030***	-0.024***	-0.021***	-0.015***
	(0.007)	(0.004)	(0.006)	(0.009)	(0.011)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)
Latitude	0.003	-0.055	-0.020	-0.019	-0.031	0.008	-0.003	-0.010	-0.008	-0.002
	(0.041)	(0.038)	(0.039)	(0.045)	(0.055)	(0.053)	(0.046)	(0.044)	(0.038)	(0.020)
Longitude	0.028	0.008	-0.001	0.003	-0.052	-0.026	-0.044	-0.011	0.002	-0.008

	(0.025)	(0.026)	(0.021)	(0.031)	(0.042)	(0.031)	(0.027)	(0.026)	(0.021)	(0.009)
Empl. share manu. & cons.	0.024***	0.009***	0.005**	-0.008*	-0.012	-0.012*	-0.017***	-0.060***	-0.054***	-0.002
	(0.005)	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.006)	(0.014)	(0.014)	(0.023)
Empl. share services	0.045***	0.054***	0.023***	0.049***	0.074***	0.073***	0.047***	-0.027*	-0.035**	0.016
	(0.010)	(0.007)	(0.004)	(0.014)	(0.007)	(0.005)	(0.004)	(0.015)	(0.017)	(0.022)
Empl. share government	0.018**	0.005	0.003	-0.030***	0.004	-0.008	-0.015**	-0.070***	-0.063***	0.009
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.006)	(0.007)	(0.016)	(0.017)	(0.023)
Laborforce participation rate	0.016*	0.057***	0.044***	0.087***	----	----	----	----	----	----
	(0.009)	(0.009)	(0.006)	(0.008)						
Unemp. rate	----	----	----	----	0.085**	0.049**	0.045**	0.025	0.043	-0.094***
					(0.034)	(0.019)	(0.017)	(0.032)	(0.031)	(0.019)
Black American pop. share	0.005	-0.004	-0.003	-0.013***	-0.036***	-0.036***	-0.031***	-0.021***	-0.018***	0.011***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.002)
Native American pop. share	-0.003	0.014	-0.008	0.000	-0.022	-0.015***	-0.015***	-0.010***	-0.011***	-0.000
	(0.019)	(0.013)	(0.009)	(0.015)	(0.019)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
Asian American pop. share	0.084	-0.059**	-0.092***	-0.199***	-0.591*	-0.426***	-0.191***	-0.084***	-0.001	0.019**
	(0.069)	(0.022)	(0.024)	(0.048)	(0.333)	(0.106)	(0.056)	(0.027)	(0.015)	(0.010)
Pop. density	0.000	0.000	0.000	-0.000	-0.000	-0.000	0.000**	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	3.265	0.438	-0.257	-0.695	-0.172	0.519	-0.703	7.469**	8.008**	-22.229***
	(3.157)	(2.965)	(2.772)	(3.654)	(4.965)	(3.761)	(3.079)	(3.379)	(3.635)	(2.393)
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.420	0.469	0.461	0.445	0.305	0.419	0.384	0.399	0.391	0.328
P-value gold rush>organized	0.000	0.000	0.159	0.000	0.006	0.011	0.012	0.005	0.012	0.000

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. Alternative control group: We use as control group all counties without gold rush but with initial land organization between 1848 and 1899. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, employment share in manufacturing, mining and construction, employment share in private services, employment share in government, laborforce participation rate (1910-1940), unemployment rate (1970-present day), Black American population share, Native American population share, Asian American population share, and population density. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the organized coefficient. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table A8: Results regarding alternative treatment: Silver rush

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Silver rush	0.240	0.326***	0.046	0.218	1.219***	0.935***	0.758***	0.694***	0.506***	0.160***
	(0.168)	(0.118)	(0.102)	(0.202)	(0.345)	(0.221)	(0.225)	(0.210)	(0.146)	(0.059)
Silver mining	0.093	-0.013	0.098	0.263*	0.237	0.160	0.104	0.105	0.030	0.005
	(0.208)	(0.239)	(0.141)	(0.142)	(0.363)	(0.213)	(0.186)	(0.259)	(0.189)	(0.072)
Iron field	-0.253***	-0.063	-0.036	-0.052	-0.034	-0.061	-0.067	-0.067	-0.072	-0.049
	(0.087)	(0.081)	(0.062)	(0.074)	(0.170)	(0.142)	(0.131)	(0.122)	(0.089)	(0.046)
Coal field	-0.153	-0.107*	-0.113**	-0.082	0.318**	0.039	-0.007	-0.069	-0.018	-0.034
	(0.094)	(0.057)	(0.053)	(0.062)	(0.122)	(0.094)	(0.085)	(0.091)	(0.090)	(0.033)
Petroleum field	-0.040	-0.089	-0.047	-0.025	-0.066	-0.071	-0.140**	-0.112	-0.051	-0.018
	(0.082)	(0.058)	(0.047)	(0.046)	(0.121)	(0.094)	(0.067)	(0.071)	(0.064)	(0.037)
Coastal county	0.211**	0.125	0.209***	0.263**	0.466**	0.550**	0.517***	0.499***	0.309***	0.025
	(0.087)	(0.079)	(0.074)	(0.105)	(0.207)	(0.210)	(0.164)	(0.146)	(0.108)	(0.038)
Distance to lake	0.064	0.087*	0.064*	0.082	-0.031	0.014	-0.030	-0.033	-0.055	-0.054***
	(0.041)	(0.047)	(0.035)	(0.051)	(0.093)	(0.073)	(0.067)	(0.064)	(0.060)	(0.018)
Distance to river	-0.004	0.018	0.051	0.054	0.281	0.271*	0.191	0.097	0.151*	0.058
	(0.090)	(0.075)	(0.068)	(0.073)	(0.203)	(0.145)	(0.131)	(0.110)	(0.089)	(0.035)
Portage	0.002	0.117	0.045	0.103	0.661	0.819	0.526	0.355*	0.198**	-0.137
	(0.227)	(0.093)	(0.095)	(0.103)	(0.722)	(0.682)	(0.431)	(0.195)	(0.088)	(0.103)
Railroad access	-0.508***	-0.456***	-0.424***	-0.366***	0.698***	0.474***	0.383***	0.333**	0.220**	-0.089
	(0.087)	(0.077)	(0.062)	(0.100)	(0.197)	(0.115)	(0.131)	(0.130)	(0.089)	(0.078)
Distance to State capital	-0.029	-0.040	-0.037	-0.067*	-0.068*	-0.034	-0.023	-0.035	-0.032	-0.027**
	(0.029)	(0.030)	(0.026)	(0.036)	(0.036)	(0.025)	(0.023)	(0.025)	(0.019)	(0.012)
Elevation (Mean)	0.185	0.219	0.103	0.372*	0.968	0.429	0.521	0.932**	0.608**	0.148*
	(0.230)	(0.192)	(0.142)	(0.218)	(0.656)	(0.455)	(0.394)	(0.422)	(0.298)	(0.082)
Elevation (SD)	1.274	0.327	0.270	1.188*	-1.135	-0.435	-0.133	0.697	-0.017	0.476
	(0.952)	(0.551)	(0.689)	(0.650)	(2.209)	(1.953)	(1.731)	(1.170)	(0.708)	(0.332)
Potential agricultural yield	-1.329**	-0.563	-1.227**	-0.739	0.216	0.526	1.253*	0.909	-0.254	-0.528
	(0.641)	(0.625)	(0.532)	(0.624)	(0.933)	(0.903)	(0.707)	(0.745)	(0.622)	(0.330)
Hispanic influence	-0.003	0.001	-0.001	-0.006	-0.043***	-0.022***	-0.014***	-0.009**	-0.010***	0.002
	(0.008)	(0.007)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)
French influence	0.003	0.005	-0.002	-0.012	-0.014	-0.039***	-0.030***	-0.024***	-0.021***	-0.015***
	(0.007)	(0.004)	(0.006)	(0.009)	(0.011)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)
Latitude	-0.001	-0.056	-0.020	-0.021	-0.028	0.011	-0.001	-0.008	-0.006	-0.003
	(0.041)	(0.038)	(0.039)	(0.045)	(0.053)	(0.051)	(0.044)	(0.042)	(0.037)	(0.019)
Longitude	0.025	0.006	-0.002	0.001	-0.051	-0.027	-0.043	-0.011	0.002	-0.007

	(0.026)	(0.026)	(0.020)	(0.031)	(0.042)	(0.031)	(0.026)	(0.025)	(0.022)	(0.009)
Empl. share manu. & cons.	0.025***	0.009***	0.005**	-0.008	-0.013*	-0.013*	-0.018***	-0.060***	-0.054***	-0.001
	(0.005)	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.006)	(0.014)	(0.015)	(0.023)
Empl. share services	0.045***	0.053***	0.023***	0.049***	0.073***	0.073***	0.047***	-0.027*	-0.035**	0.016
	(0.010)	(0.007)	(0.004)	(0.014)	(0.007)	(0.005)	(0.004)	(0.015)	(0.017)	(0.022)
Empl. share government	0.019**	0.006	0.003	-0.030***	0.003	-0.007	-0.015**	-0.071***	-0.063***	0.010
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.006)	(0.007)	(0.016)	(0.017)	(0.023)
Laborforce participation rate	0.017*	0.057***	0.044***	0.087***	----	----	----	----	----	----
	(0.009)	(0.009)	(0.006)	(0.008)						
Unemp. rate	----	----	----	----	0.087**	0.051***	0.046***	0.027	0.045	-0.092***
					(0.033)	(0.019)	(0.016)	(0.031)	(0.030)	(0.019)
Black American pop. share	0.005	-0.004	-0.003	-0.013***	-0.037***	-0.037***	-0.031***	-0.021***	-0.018***	0.010***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.002)
Native American pop. share	-0.003	0.014	-0.008	0.000	-0.021	-0.016***	-0.015***	-0.010***	-0.011***	0.000
	(0.019)	(0.013)	(0.009)	(0.015)	(0.019)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
Asian American pop. share	0.086	-0.052**	-0.096***	-0.216***	-0.590*	-0.432***	-0.191***	-0.086***	-0.003	0.019*
	(0.069)	(0.023)	(0.025)	(0.048)	(0.323)	(0.105)	(0.056)	(0.027)	(0.015)	(0.010)
Pop. density	0.000*	0.000	0.000	-0.000	-0.000	-0.000	0.000**	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	2.965	0.292	-0.303	-0.877	0.002	0.477	-0.537	7.589**	8.019**	-22.196***
	(3.221)	(2.966)	(2.763)	(3.667)	(4.897)	(3.735)	(3.044)	(3.422)	(3.664)	(2.375)
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.415	0.467	0.461	0.443	0.307	0.420	0.387	0.401	0.393	0.327
P-value silver rush>silver	0.280	0.125	0.389	0.430	0.038	0.002	0.032	0.067	0.049	0.048

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. The variable silver rush takes the value of one if a county had a silver discovery in the West between 1848 and 1899. The variable silver mining takes the value of one if a county had a silver discovery in the east at any point in time or in the West after 1899. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, employment share in manufacturing, mining and construction, employment share in private services, employment share in government, laborforce participation rate (1910-1940), unemployment rate (1970-present day), Black American population share, Native American population share, Asian American population share, and population density. Last row contains a one-sided Wald test (p-value reported) whether the silver rush coefficient is larger than the silver mining coefficient. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table A9: Results regarding alternative treatment: Placebo test: Iron rush

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Iron rush	-0.156	-0.235	0.040	-0.509***	0.615	0.625	0.552*	0.113	0.068	0.002
	(0.292)	(0.165)	(0.110)	(0.156)	(0.596)	(0.393)	(0.319)	(0.326)	(0.229)	(0.060)
Iron mining	-0.151	0.007	-0.095	0.080	-0.172	-0.374	-0.273	-0.044	0.028	-0.041
	(0.191)	(0.162)	(0.091)	(0.115)	(0.421)	(0.352)	(0.273)	(0.297)	(0.193)	(0.050)
Coal field	-0.169*	-0.115*	-0.117**	-0.085	0.283**	0.009	-0.031	-0.087	-0.031	-0.035
	(0.094)	(0.060)	(0.052)	(0.062)	(0.120)	(0.092)	(0.082)	(0.085)	(0.085)	(0.034)
Petroleum field	-0.049	-0.099*	-0.051	-0.037	-0.102	-0.104	-0.164**	-0.131*	-0.062	-0.025
	(0.083)	(0.059)	(0.047)	(0.047)	(0.117)	(0.091)	(0.068)	(0.074)	(0.066)	(0.037)
Coastal county	0.225**	0.142*	0.208***	0.278***	0.483**	0.561***	0.526***	0.520***	0.326***	0.031
	(0.088)	(0.078)	(0.072)	(0.103)	(0.205)	(0.206)	(0.162)	(0.148)	(0.109)	(0.039)
Distance to lake	0.068	0.087*	0.064*	0.081	-0.025	0.019	-0.026	-0.032	-0.054	-0.054***
	(0.043)	(0.048)	(0.036)	(0.052)	(0.095)	(0.076)	(0.068)	(0.064)	(0.059)	(0.018)
Distance to river	0.003	0.024	0.053	0.056	0.296	0.288*	0.205	0.106	0.157*	0.062*
	(0.090)	(0.077)	(0.069)	(0.074)	(0.200)	(0.146)	(0.133)	(0.113)	(0.092)	(0.035)
Portage	0.015	0.122	0.046	0.099	0.660	0.828	0.535	0.361*	0.202**	-0.134
	(0.226)	(0.097)	(0.096)	(0.105)	(0.733)	(0.690)	(0.437)	(0.202)	(0.093)	(0.101)
Railroad access	-0.522***	-0.468***	-0.431***	-0.380***	0.675***	0.448***	0.366***	0.318**	0.212**	-0.092
	(0.085)	(0.077)	(0.060)	(0.099)	(0.189)	(0.107)	(0.128)	(0.127)	(0.091)	(0.079)
Distance to State capital	-0.032	-0.044	-0.037	-0.070*	-0.078**	-0.040	-0.029	-0.042	-0.037*	-0.027**
	(0.030)	(0.030)	(0.026)	(0.037)	(0.037)	(0.024)	(0.023)	(0.025)	(0.019)	(0.013)
Elevation (Mean)	0.234	0.307	0.119	0.440**	1.292*	0.692	0.728	1.125**	0.738**	0.184**
	(0.238)	(0.209)	(0.141)	(0.211)	(0.719)	(0.511)	(0.459)	(0.486)	(0.348)	(0.081)
Elevation (SD)	1.019	0.217	0.228	1.040	-1.502	-0.680	-0.350	0.430	-0.238	0.413
	(0.911)	(0.544)	(0.663)	(0.641)	(2.217)	(1.946)	(1.767)	(1.216)	(0.735)	(0.325)
Potential agricultural yield	-1.296**	-0.486	-1.215**	-0.725	0.607	0.855	1.525*	1.138	-0.093	-0.516
	(0.642)	(0.616)	(0.534)	(0.626)	(0.971)	(0.976)	(0.772)	(0.812)	(0.668)	(0.332)
Hispanic influence	-0.003	0.001	-0.001	-0.006	-0.043***	-0.022***	-0.013***	-0.009**	-0.010**	0.002
	(0.008)	(0.007)	(0.005)	(0.005)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)
French influence	0.003	0.005	-0.002	-0.012	-0.014	-0.039***	-0.030***	-0.025***	-0.021***	-0.015***
	(0.007)	(0.004)	(0.006)	(0.009)	(0.011)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)
Latitude	-0.003	-0.058	-0.021	-0.023	-0.038	0.003	-0.007	-0.013	-0.009	-0.004
	(0.042)	(0.038)	(0.039)	(0.046)	(0.056)	(0.053)	(0.045)	(0.043)	(0.038)	(0.019)
Longitude	0.022	0.003	-0.002	-0.002	-0.062	-0.034	-0.049*	-0.017	-0.003	-0.008
	(0.027)	(0.026)	(0.020)	(0.031)	(0.045)	(0.033)	(0.028)	(0.026)	(0.022)	(0.010)
Empl. share manu. & cons.	0.025***	0.010***	0.006**	-0.008	-0.011	-0.011*	-0.016***	-0.059***	-0.054***	-0.001

	(0.005)	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.005)	(0.013)	(0.015)	(0.022)
Empl. share services	0.045***	0.053***	0.023***	0.047***	0.075***	0.075***	0.048***	-0.026*	-0.035**	0.017
	(0.010)	(0.007)	(0.004)	(0.014)	(0.008)	(0.005)	(0.004)	(0.014)	(0.017)	(0.022)
Empl. share government	0.017**	0.006	0.003	-0.030***	0.005	-0.006	-0.014**	-0.070***	-0.063***	0.010
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.006)	(0.007)	(0.016)	(0.017)	(0.022)
Laborforce participation rate	0.017*	0.057***	0.044***	0.087***	----	----	----	----	----	----
	(0.010)	(0.009)	(0.006)	(0.008)						
Unemp. rate	----	----	----	----	0.094***	0.055***	0.048***	0.033	0.049	-0.093***
					(0.033)	(0.020)	(0.017)	(0.032)	(0.031)	(0.019)
Black American pop. share	0.005	-0.004	-0.003	-0.012***	-0.036***	-0.036***	-0.031***	-0.021***	-0.018***	0.011***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.002)
Native American pop. share	-0.003	0.013	-0.008	-0.000	-0.025	-0.018***	-0.017***	-0.012***	-0.012***	-0.001
	(0.019)	(0.013)	(0.009)	(0.015)	(0.019)	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)
Asian American pop. share	0.078	-0.064**	-0.097***	-0.221***	-0.616*	-0.443***	-0.196***	-0.085***	-0.002	0.019*
	(0.071)	(0.024)	(0.024)	(0.048)	(0.329)	(0.104)	(0.055)	(0.026)	(0.016)	(0.010)
Pop. density	0.000*	0.000	0.000	0.000	-0.000	-0.000	0.000**	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	2.693	-0.010	-0.317	-1.107	-1.082	-0.283	-1.174	6.865**	7.548**	-22.313***
	(3.276)	(3.004)	(2.749)	(3.692)	(5.210)	(4.000)	(3.232)	(3.414)	(3.612)	(2.328)
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	3,109	2,084
Adjusted R-squared	0.413	0.466	0.461	0.444	0.303	0.416	0.383	0.395	0.387	0.326
P-value iron rush>iron mining	0.496	0.224	0.239	0.0130	0.204	0.0850	0.0770	0.396	0.460	0.330

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables for Models 1-4 are self-employment (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are the number of startups (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. The variable iron rush takes the value of one if a county had an iron discovery in the West between 1848 and 1899. The variable iron mining takes the value of one if a county had an iron discovery in the east at any point in time or in the West after 1899. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, employment share in manufacturing, mining and construction, employment share in private services, employment share in government, laborforce participation rate (1910-1940), unemployment rate (1970-present day), Black American population share, Native American population share, Asian American population share, and population density. Last row contains a one-sided Wald test (p-value reported) whether the iron rush coefficient is larger than the iron mining coefficient. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table A10: Robustness check regarding the GPIPP data set.

VARIABLES	(1) N<100 + Weighted	(2) N<100 + Weighted	(3) N<100 + Weighted
Gold rush	0.413***	0.308***	0.360***
	(0.140)	(0.101)	(0.093)
Gold mining	0.056	0.039	0.049
	(0.055)	(0.068)	(0.068)
Iron field		-0.067	-0.011
		(0.068)	(0.062)
Coal field		-0.100*	-0.073
		(0.060)	(0.058)
Petroleum field		0.031	0.051
		(0.065)	(0.067)
Coastal county		0.156**	0.075
		(0.064)	(0.049)
Distance to lake		-0.006	-0.014
		(0.026)	(0.024)
Distance to river		0.101	0.120**
		(0.061)	(0.057)
Portage		-0.135	-0.161
		(0.176)	(0.131)
Railroad access		-0.140	-0.045
		(0.095)	(0.095)
Distance to State capital		-0.073***	-0.059***
		(0.021)	(0.022)
Elevation (Mean)		0.248*	0.319**
		(0.138)	(0.124)
Elevation (SD)		1.685***	1.120**
		(0.503)	(0.431)
Potential agricultural yield		0.272	0.351
		(0.476)	(0.466)
Hispanic influence		0.000	0.002
		(0.005)	(0.005)
French influence		-0.003	-0.006
		(0.008)	(0.006)
Latitude		0.002	0.004
		(0.030)	(0.033)
Longitude		0.007	0.009
		(0.018)	(0.015)
Empl. share manu. & cons.			-0.028***
			(0.009)
Empl. share services.			-0.019*
			(0.010)
Empl. share government.			-0.031***
			(0.010)
Unemp. rate			-0.063**
			(0.025)
Black American pop. share			0.006***
			(0.002)
Native American pop. share			0.003
			(0.004)
Asian American pop. share			0.036**
			(0.016)
Pop. density			0.000
			(0.000)
Constant	-20.964***	-20.479***	-17.922***
	(0.002)	(2.035)	(1.690)
State-fixed effects	Yes	Yes	Yes
Observations	3,106	3,106	3,106
Adjusted R-squared	0.048	0.057	0.086

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The models use all counties even with less than 100 respondents per county. Additionally, the regional entrepreneurial personality profile is weighted by age and gender. Model 1 use only state-fixed effects as controls. Model 2 include geographic controls: Ironfield, coalfield, Petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, and longitude. Model 3 additionally includes socio-economic controls of the 2010 decade: Empl. share in manufacturing and construction, empl. share in services, empl. share in government, unempl. rate, Black American pop. share, Native American pop. share, Asian American pop. share, population density. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

2. Spatial auto-correlation

This robustness check revolves around the spatial nature of the relationship between the gold rush and entrepreneurship. Computing Moran's revealed the presence of spatial auto-correlation of the dependent variables with a weighing matrix containing the inverse distance of the counties from each other. We reran the basic regressions with the spatial weighting matrix. The results in Table S11 show that the gold rush predicts entrepreneurship rates but that spatially lagged gold rushes also have an impact in more recent decades.

However, this weighting matrix gives weight to distant regions. Therefore, we directly checked whether gold rushes in nearby counties matter for entrepreneurship in the specific county under study. If one truly believes in the local nature – especially of the culture effect – the concentration of entrepreneurially-minded people in the specific county has much more influence on the formation of an entrepreneurship culture than does the concentration of entrepreneurially-minded people in other counties nearby. To test this idea, we compute a dummy variable taking the value of one if there is a gold rush in an adjacent county. The results in Table S12 show that gold rushes in neighboring counties are rarely related with the entrepreneurship indicators.

Additionally, we checked whether an amplifier effect exists in the sense that a gold rush in the specific county and a gold rush in an adjacent county together have a positive effect on entrepreneurship. The results for these tests are shown in Table S13 and reveal that the respective interaction term is positive and significant in only a few models. We thus conclude that spatial autocorrelation does not impact our results and that the gold rush has a predominantly local effect on entrepreneurship.

Table A11: Results regarding spatial auto-correlation

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Gold rush (direct effect)	0.591***	0.374***	0.095	0.256**	0.737***	0.748***	0.411***	0.351***	0.308***	0.105
	(0.113)	(0.097)	(0.074)	(0.108)	(0.245)	(0.183)	(0.151)	(0.128)	(0.100)	(0.077)
Gold rush (indirect effect)	14.236***	15.505***	-0.874	12.249**	35.890***	25.696***	21.944***	15.938**	11.470**	-3.268
	(5.531)	(4.747)	(3.622)	(5.301	(12.213)	(9.018)	(7.415)	(6.285)	(4.929)	(3.953)
Gold mining (direct effect)	-0.184	-0.162*	-0.063	-0.107	-0.136	-0.420**	-0.365**	-0.132	12.506***	0.047
	(0.114)	(0.098)	(0.075)	(0.110)	(0.249)	(0.186)	(0.153)	(0.128)	(0.102)	(0.071)
Gold mining (indirect effect)	-1.436	-1.730	-1.029	0.091	31.046**	24.012***	20.023***	15.392***	12.505***	0.417
	(2.371)	(2.035)	(1.553)	(2.275)	(5.236)	(3.866)	(3.179)	(2.695)	(2.113)	(1.150)
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-econ. controls	No	No	No	No	No	No	No	No	No	No
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,108	3,108	3,108	3,108	3,108	3,108	3,108	3,108	3,108	2,083
Adjusted R-squared	0.413	0.422	0.472	0.429	0.258	0.292	0.278	0.339	0.334	0.256

Note: Spatial regression with inverse distance weighting matrix. The dependent variables for Models 1-4 are self-employment (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are the number of startups (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. All models use state-fixed effects + geographic controls: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude. These effects are not displayed here in detail, because Stata is unable to export the indirect effects. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table A12: Results regarding gold rushes in neighboring counties

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Gold rush	0.570***	0.449***	0.093	0.442***	0.796**	0.630**	0.375*	0.386**	0.335***	0.155***
	(0.149)	(0.105)	(0.093)	(0.121)	(0.316)	(0.247)	(0.195)	(0.147)	(0.115)	(0.057)
Gold rush in neighboring county	-0.189	-0.130	0.006	-0.104	0.451*	0.424***	0.354***	0.326***	0.045	-0.020
	(0.144)	(0.107)	(0.113)	(0.168)	(0.262)	(0.149)	(0.118)	(0.110)	(0.095)	(0.068)
Gold mining	-0.194	-0.181**	-0.115	-0.055	0.005	-0.375*	-0.337***	-0.161	-0.091	0.078
	(0.143)	(0.074)	(0.090)	(0.097)	(0.240)	(0.196)	(0.105)	(0.100)	(0.084)	(0.050)
Gold mining in neighboring county	-0.047	-0.150*	-0.159**	-0.196**	-0.207	-0.242*	-0.091	-0.124	-0.043	-0.055
	(0.091)	(0.083)	(0.071)	(0.088)	(0.277)	(0.131)	(0.097)	(0.078)	(0.069)	(0.040)
Iron field	-0.267***	-0.074	-0.049	-0.072	-0.049	-0.078	-0.074	-0.074	-0.071	-0.047
	(0.085)	(0.081)	(0.065)	(0.075)	(0.173)	(0.144)	(0.134)	(0.122)	(0.091)	(0.046)
Coal field	-0.151	-0.111*	-0.117**	-0.086	0.300**	0.022	-0.024	-0.082	-0.024	-0.033
	(0.093)	(0.058)	(0.054)	(0.062)	(0.118)	(0.091)	(0.082)	(0.089)	(0.090)	(0.033)
Petroleum field	-0.040	-0.097*	-0.058	-0.033	-0.080	-0.099	-0.156**	-0.125*	-0.060	-0.021
	(0.083)	(0.057)	(0.047)	(0.048)	(0.124)	(0.095)	(0.068)	(0.072)	(0.064)	(0.037)
Coastal county	0.194**	0.105	0.183**	0.232**	0.442**	0.504**	0.492***	0.476***	0.303***	0.026
	(0.089)	(0.081)	(0.074)	(0.103)	(0.205)	(0.200)	(0.164)	(0.146)	(0.109)	(0.039)
Distance to lake	0.064	0.084*	0.059	0.079	-0.021	0.019	-0.025	-0.028	-0.054	-0.052***
	(0.040)	(0.046)	(0.036)	(0.051)	(0.095)	(0.076)	(0.069)	(0.066)	(0.061)	(0.018)
Distance to river	-0.006	0.017	0.045	0.051	0.291	0.275*	0.199	0.105	0.156*	0.057
	(0.091)	(0.075)	(0.069)	(0.072)	(0.200)	(0.146)	(0.131)	(0.110)	(0.090)	(0.035)
Portage	-0.000	0.126	0.041	0.107	0.673	0.814	0.521	0.355*	0.192**	-0.128
	(0.232)	(0.096)	(0.096)	(0.102)	(0.708)	(0.675)	(0.426)	(0.192)	(0.086)	(0.097)
Railroad access	-0.521***	-0.473***	-0.434***	-0.377***	0.680***	0.452***	0.368***	0.322**	0.210**	-0.091
	(0.084)	(0.078)	(0.059)	(0.099)	(0.195)	(0.113)	(0.129)	(0.129)	(0.090)	(0.079)
Distance to State capital	-0.025	-0.039	-0.038	-0.065*	-0.068*	-0.034	-0.024	-0.035	-0.032	-0.026**
	(0.029)	(0.030)	(0.026)	(0.036)	(0.036)	(0.025)	(0.023)	(0.025)	(0.019)	(0.012)
Elevation (Mean)	0.165	0.235	0.106	0.371*	1.029	0.479	0.579	0.978**	0.665**	0.153*
	(0.240)	(0.200)	(0.144)	(0.200)	(0.700)	(0.482)	(0.429)	(0.461)	(0.329)	(0.081)
Elevation (SD)	1.267	0.377	0.303	1.190*	-1.144	-0.527	-0.298	0.657	-0.096	0.509
	(0.836)	(0.486)	(0.594)	(0.592)	(2.295)	(1.919)	(1.800)	(1.232)	(0.799)	(0.342)
Potential agricultural yield	-1.296**	-0.476	-1.210**	-0.708	0.400	0.677	1.382*	1.037	-0.148	-0.537
	(0.643)	(0.635)	(0.546)	(0.644)	(0.943)	(0.926)	(0.743)	(0.792)	(0.657)	(0.322)
Hispanic influence	-0.003	0.002	-0.001	-0.006	-0.042***	-0.021***	-0.013***	-0.008**	-0.009**	0.003
	(0.008)	(0.007)	(0.006)	(0.005)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)
French influence	0.003	0.004	-0.003	-0.013	-0.014	-0.039***	-0.030***	-0.024***	-0.021***	-0.015***

	(0.007)	(0.004)	(0.007)	(0.009)	(0.011)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)
Latitude	0.001	-0.055	-0.020	-0.020	-0.028	0.009	-0.003	-0.008	-0.006	-0.002
	(0.040)	(0.037)	(0.039)	(0.044)	(0.053)	(0.051)	(0.045)	(0.043)	(0.038)	(0.020)
Longitude	0.024	0.003	-0.003	-0.000	-0.049	-0.027	-0.043	-0.010	0.001	-0.007
	(0.026)	(0.026)	(0.021)	(0.032)	(0.040)	(0.031)	(0.026)	(0.025)	(0.022)	(0.009)
Empl. share manu. & cons.	0.024***	0.009***	0.005**	-0.008	-0.012	-0.011*	-0.017***	-0.059***	-0.054***	-0.000
	(0.005)	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.006)	(0.014)	(0.014)	(0.023)
Empl. share services	0.045***	0.054***	0.023***	0.048***	0.073***	0.073***	0.047***	-0.027*	-0.035**	0.017
	(0.010)	(0.007)	(0.004)	(0.013)	(0.007)	(0.005)	(0.004)	(0.015)	(0.017)	(0.022)
Empl. share government	0.019***	0.005	0.003	-0.029***	0.003	-0.009	-0.015**	-0.070***	-0.063***	0.010
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.006)	(0.007)	(0.016)	(0.017)	(0.023)
Laborforce participation rate	0.017*	0.057***	0.044***	0.088***	----	----	----	----	----	----
	(0.009)	(0.009)	(0.006)	(0.008)						
Unemp. rate	----	----	----	----	0.083**	0.048**	0.045**	0.024	0.043	-0.094***
					(0.033)	(0.020)	(0.017)	(0.032)	(0.031)	(0.019)
Black American pop. share	0.004	-0.005	-0.003	-0.014***	-0.037***	-0.038***	-0.032***	-0.022***	-0.018***	0.010***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.002)
Native American pop. share	-0.002	0.015	-0.007	0.001	-0.021	-0.016***	-0.015***	-0.011***	-0.012***	0.000
	(0.019)	(0.013)	(0.009)	(0.016)	(0.019)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
Asian American pop. share	0.089	-0.057**	-0.090***	-0.194***	-0.585*	-0.417***	-0.190***	-0.081***	-0.000	0.019*
	(0.069)	(0.022)	(0.023)	(0.047)	(0.333)	(0.105)	(0.056)	(0.027)	(0.015)	(0.010)
Pop. density	0.000	0.000	0.000	-0.000	-0.000	-0.000	0.000**	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	2.859	0.030	-0.442	-0.975	0.127	0.542	-0.553	7.565**	7.907**	-22.316***
	(3.189)	(2.988)	(2.816)	(3.744)	(4.779)	(3.708)	(3.053)	(3.325)	(3.664)	(2.364)
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,108	3,108	3,108	3,108	3,108	3,108	3,108	3,108	3,108	2,083
Adjusted R-squared	0.421	0.472	0.463	0.446	0.305	0.421	0.386	0.399	0.390	0.327

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. Gold rush in neighboring county equals one if a neighboring county had a gold rush. Gold mining in neighboring county equals one if a neighboring county is a gold mining county. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, employment share in manufacturing, mining and construction, employment share in private services, employment share in government, laborforce participation rate (1910-1940), unemployment rate (1970-present day), Black American population share, Native American population share, Asian American population share, and population density. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A13: Results regarding gold rush interactions

Variables	(1) Self-empl. rate 1910	(2) Self-empl. rate 1920	(3) Self-empl. rate 1930	(4) Self-empl. rate 1940	(5) Startup rate 1970s	(6) Startup rate 1980s	(7) Startup rate 1990s	(8) Startup rate 2000s	(9) Startup rate 2010s	(10) Entr. pers. profile
Gold rush	0.378	0.305	0.151	-0.184	0.770	0.401	-0.071	0.271	0.369	-0.102
	(0.271)	(0.251)	(0.225)	(0.226)	(0.821)	(0.376)	(0.195)	(0.268)	(0.221)	(0.124)
Gold rush in neighboring county	-0.184	-0.130	0.008	-0.136	0.384	0.363***	0.299***	0.286***	0.030	-0.044
	(0.155)	(0.114)	(0.112)	(0.172)	(0.264)	(0.132)	(0.110)	(0.103)	(0.095)	(0.070)
Interaction gold rush	0.185	0.144	-0.059	0.655***	0.108	0.302	0.509*	0.161	-0.015	0.281**
	(0.344)	(0.269)	(0.194)	(0.211)	(0.933)	(0.372)	(0.274)	(0.311)	(0.226)	(0.122)
Gold mining	-0.432*	-0.290*	-0.089	-0.038	1.173**	0.505	0.257	0.442	0.205	0.224
	(0.257)	(0.145)	(0.108)	(0.191)	(0.488)	(0.345)	(0.298)	(0.379)	(0.287)	(0.149)
Gold mining in neighboring county	-0.071	-0.161*	-0.156**	-0.191*	-0.086	-0.150	-0.028	-0.060	-0.013	-0.038
	(0.096)	(0.090)	(0.073)	(0.098)	(0.285)	(0.152)	(0.109)	(0.083)	(0.071)	(0.040)
Interactin gold mining	0.320	0.149	-0.036	-0.018	-1.567***	-1.179**	-0.793**	-0.808*	-0.397	-0.193
	(0.262)	(0.173)	(0.151)	(0.199)	(0.529)	(0.487)	(0.357)	(0.418)	(0.296)	(0.200)
Iron field	-0.263***	-0.072	-0.049	-0.070	-0.063	-0.088	-0.079	-0.082	-0.075	-0.047
	(0.085)	(0.081)	(0.066)	(0.075)	(0.169)	(0.144)	(0.134)	(0.122)	(0.091)	(0.047)
Coal field	-0.151	-0.111*	-0.117**	-0.086	0.301**	0.023	-0.023	-0.082	-0.024	-0.033
	(0.094)	(0.058)	(0.054)	(0.062)	(0.120)	(0.093)	(0.082)	(0.090)	(0.090)	(0.033)
Petroleum field	-0.042	-0.098*	-0.057	-0.033	-0.069	-0.090	-0.150**	-0.119	-0.057	-0.020
	(0.083)	(0.057)	(0.047)	(0.048)	(0.126)	(0.096)	(0.068)	(0.073)	(0.065)	(0.037)
Coastal county	0.197**	0.106	0.183**	0.235**	0.434**	0.499**	0.490***	0.471***	0.301***	0.028
	(0.090)	(0.081)	(0.074)	(0.104)	(0.210)	(0.205)	(0.167)	(0.150)	(0.110)	(0.039)
Distance to lake	0.064	0.084*	0.059	0.079	-0.022	0.018	-0.026	-0.029	-0.055	-0.052***
	(0.040)	(0.046)	(0.036)	(0.051)	(0.095)	(0.075)	(0.069)	(0.066)	(0.061)	(0.018)
Distance to river	-0.005	0.018	0.045	0.052	0.290	0.274*	0.198	0.103	0.156*	0.057
	(0.091)	(0.075)	(0.069)	(0.073)	(0.202)	(0.147)	(0.132)	(0.110)	(0.090)	(0.036)
Portage	0.003	0.127	0.041	0.107	0.659	0.804	0.514	0.347*	0.188**	-0.131
	(0.232)	(0.096)	(0.097)	(0.103)	(0.708)	(0.673)	(0.425)	(0.192)	(0.086)	(0.099)
Railroad access	-0.522***	-0.473***	-0.434***	-0.373***	0.691***	0.462***	0.376***	0.328**	0.213**	-0.089
	(0.085)	(0.078)	(0.060)	(0.099)	(0.195)	(0.114)	(0.130)	(0.129)	(0.090)	(0.079)
Distance to State capital	-0.026	-0.040	-0.038	-0.065*	-0.064*	-0.032	-0.023	-0.033	-0.031	-0.026**
	(0.030)	(0.030)	(0.026)	(0.036)	(0.036)	(0.026)	(0.022)	(0.025)	(0.019)	(0.012)
Elevation (Mean)	0.168	0.236	0.106	0.364*	1.002	0.457	0.560	0.962**	0.658*	0.146*
	(0.241)	(0.199)	(0.143)	(0.199)	(0.706)	(0.490)	(0.434)	(0.468)	(0.333)	(0.081)
Elevation (SD)	1.287	0.387	0.300	1.195**	-1.226	-0.586	-0.336	0.611	-0.118	0.498
	(0.828)	(0.489)	(0.592)	(0.593)	(2.243)	(1.879)	(1.774)	(1.203)	(0.782)	(0.333)
Potential agricultural yield	-1.308**	-0.483	-1.208**	-0.709	0.446	0.714	1.407*	1.065	-0.133	-0.524
	(0.647)	(0.636)	(0.546)	(0.645)	(0.963)	(0.930)	(0.745)	(0.797)	(0.659)	(0.323)
Hispanic influence	-0.003	0.002	-0.001	-0.006	-0.042***	-0.021***	-0.013***	-0.008**	-0.009**	0.003
	(0.008)	(0.007)	(0.006)	(0.005)	(0.004)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)
French influence	0.002	0.004	-0.003	-0.013	-0.013	-0.038***	-0.029***	-0.024***	-0.021***	-0.015**

	(0.007)	(0.004)	(0.007)	(0.009)	(0.011)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)
Latitude	0.001	-0.055	-0.020	-0.019	-0.026	0.011	-0.001	-0.007	-0.005	-0.001
	(0.040)	(0.037)	(0.039)	(0.044)	(0.053)	(0.051)	(0.045)	(0.043)	(0.038)	(0.020)
Longitude	0.024	0.003	-0.003	-0.001	-0.049	-0.027	-0.044	-0.010	0.001	-0.008
	(0.026)	(0.026)	(0.021)	(0.032)	(0.040)	(0.031)	(0.026)	(0.025)	(0.022)	(0.009)
Empl. share manu. & cons.	0.024***	0.009***	0.005**	-0.008	-0.012	-0.011*	-0.017***	-0.060***	-0.054***	-0.000
	(0.005)	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)	(0.006)	(0.014)	(0.014)	(0.023)
Empl. share services	0.045***	0.054***	0.023***	0.048***	0.073***	0.072***	0.047***	-0.027*	-0.035**	0.017
	(0.010)	(0.007)	(0.004)	(0.013)	(0.007)	(0.005)	(0.004)	(0.015)	(0.017)	(0.022)
Empl. share government	0.018**	0.005	0.003	-0.030***	0.003	-0.009	-0.016**	-0.071***	-0.063***	0.010
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.016)	(0.017)	(0.023)
Laborforce participation rate	0.017*	0.057***	0.044***	0.088***	----	----	----	----	----	----
	(0.009)	(0.009)	(0.006)	(0.008)						
Unemp. rate	----	----	----	----	0.081**	0.046**	0.043**	0.024	0.043	-0.094***
					(0.034)	(0.020)	(0.017)	(0.032)	(0.031)	(0.019)
Black American pop. share	0.005	-0.005	-0.004	-0.014***	-0.037***	-0.038***	-0.032***	-0.022***	-0.018***	0.010***
	(0.004)	(0.003)	(0.002)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.002)
Native American pop. share	-0.003	0.015	-0.007	0.002	-0.021	-0.015***	-0.015***	-0.010***	-0.011***	0.001
	(0.019)	(0.013)	(0.009)	(0.016)	(0.019)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)
Asian American pop. share	0.087	-0.058**	-0.089***	-0.196***	-0.572*	-0.414***	-0.189***	-0.080***	0.000	0.019*
	(0.070)	(0.022)	(0.023)	(0.047)	(0.327)	(0.104)	(0.055)	(0.027)	(0.015)	(0.010)
Pop. density	0.000	0.000	0.000	-0.000	-0.000	-0.000	0.000**	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	2.869	0.031	-0.441	-1.045	0.053	0.476	-0.619	7.557**	7.899**	-22.432***
	(3.193)	(2.989)	(2.820)	(3.754)	(4.715)	(3.714)	(3.078)	(3.372)	(3.690)	(2.390)
State-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,108	3,108	3,108	3,108	3,108	3,108	3,108	3,108	3,108	2,083
Adjusted R-squared	0.421	0.471	0.463	0.446	0.307	0.422	0.387	0.400	0.390	0.327

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. Interaction gold rush equals one if there is a gold rush in the respective county + a gold rush in a neighboring county. Interaction gold mining equals one if the respective county is a gold mining county + a neighboring county is a gold mining county. The dependent variables for Models 1-4 are self-employment rates (in non-agricultural industries) in percent of the population aged 16 years or older. The dependent variables for Models 5-9 are startup rates (in non-agricultural industries) per 1,000 employed. The dependent variable for Model 10 is the regional average of an entrepreneurial constellation of the Big Five traits. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, employment share in manufacturing, mining and construction, employment share in private services, employment share in government, laborforce participation rate (1910-1940), unemployment rate (1970-present day), Black American population share, Native American population share, Asian American population share, and population density. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table A14: Results regarding exit rates.

VARIABLES	(1) Exit rate 1970s	(2) Exit rate 1980s	(3) Exit rate 1990s	(4) Exit rate 2000s	(5) Exit rate 2010s
Gold rush	0.329	0.685***	0.600***	0.710*	0.473*
	(0.287)	(0.237)	(0.208)	(0.353)	(0.263)
Gold mining	-0.487**	-0.173	-0.120	0.086	0.204
	(0.239)	(0.128)	(0.093)	(0.140)	(0.217)
Iron field	0.145	-0.010	-0.043	-0.139	-0.120
	(0.182)	(0.107)	(0.097)	(0.109)	(0.079)
Coal field	0.117	0.184	0.116	-0.224	-0.104
	(0.234)	(0.147)	(0.113)	(0.242)	(0.137)
Petroleum field	0.095	0.325*	0.214	0.205	0.100
	(0.188)	(0.175)	(0.141)	(0.276)	(0.193)
Coastal county	0.308*	0.276**	0.238**	0.276*	0.196
	(0.173)	(0.123)	(0.116)	(0.157)	(0.174)
Distance to lake	0.131	0.083	-0.032	0.019	-0.093
	(0.090)	(0.064)	(0.092)	(0.052)	(0.079)
Distance to river	-0.273	-0.005	-0.125	-0.235	-0.200
	(0.265)	(0.173)	(0.254)	(0.360)	(0.239)
Portage	0.342	0.240	0.436*	0.642***	0.324**
	(0.344)	(0.219)	(0.228)	(0.202)	(0.154)
Railroad access	0.653***	0.765***	0.666***	0.617**	0.634***
	(0.213)	(0.192)	(0.143)	(0.245)	(0.144)
Distance to State capital	0.085	0.099*	0.045	-0.066	-0.066
	(0.059)	(0.059)	(0.043)	(0.043)	(0.051)
Elevation (Mean)	0.622*	0.569	0.320	0.778*	0.800*
	(0.353)	(0.440)	(0.285)	(0.400)	(0.412)
Elevation (SD)	0.370	0.249	0.841	2.398	1.937
	(1.239)	(1.471)	(1.473)	(1.568)	(1.555)
Potential agricultural yield	-0.862	-0.238	-0.886	0.770	-0.369
	(1.399)	(0.931)	(1.195)	(1.068)	(0.966)
Hispanic influence	0.006	-0.019***	-0.003	0.002	-0.003
	(0.007)	(0.006)	(0.006)	(0.008)	(0.007)
French influence	-0.056***	-0.012**	-0.031***	-0.038***	-0.012*
	(0.008)	(0.005)	(0.011)	(0.012)	(0.007)
Latitude	-0.040	-0.129***	-0.092***	0.056	0.003
	(0.061)	(0.038)	(0.032)	(0.067)	(0.047)
Longitude	-0.077*	-0.091**	-0.048	-0.019	0.027
	(0.046)	(0.041)	(0.034)	(0.029)	(0.028)

Empl. share manu. & cons.	-0.018**	-0.036***	-0.046***	-0.252**	-0.256
	(0.008)	(0.011)	(0.009)	(0.121)	(0.174)
Empl. share services.	-0.064**	-0.054***	-0.044**	-0.237*	-0.246
	(0.025)	(0.016)	(0.017)	(0.125)	(0.173)
Empl. share government.	-0.011	-0.017	-0.020*	-0.193*	-0.229
	(0.014)	(0.013)	(0.012)	(0.103)	(0.169)
Unemp. rate	0.208***	0.105***	0.068**	0.137***	0.057
	(0.037)	(0.028)	(0.027)	(0.047)	(0.050)
Black American pop. share	-0.008*	-0.006	-0.003	-0.023**	-0.016**
	(0.005)	(0.004)	(0.003)	(0.011)	(0.006)
Native American pop. share	0.006	-0.005	0.010*	-0.017	-0.005
	(0.012)	(0.007)	(0.006)	(0.018)	(0.011)
Asian American pop. share	-0.290	-0.086	0.116*	0.082	0.069**
	(0.192)	(0.055)	(0.067)	(0.055)	(0.028)
Pop. density	0.000***	0.000***	0.000**	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	9.102**	11.079***	12.704***	28.895**	34.890*
	(3.946)	(3.795)	(3.254)	(11.677)	(18.923)
State-fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109
Adjusted R-squared	0.183	0.324	0.203	0.075	0.128
P-value gold rush>gold mining	0.007	0.000	0.001	0.007	0.120

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variable are exit rates (number of exits) per 100 firms. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, unemployment rate, Black American population share, Native American population share, Asian American population share, and population density. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the gold mining coefficient. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A15: Results regarding job creation rates.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Job creation rate 1970s	Job creation rate 1980s	Job creation rate 1990s	Job creation rate 2000s	Job creation rate 2010s
Gold rush	6.884***	1.041	3.108**	2.662**	2.230***
	(1.849)	(1.172)	(1.463)	(1.085)	(0.657)
Gold mining	-3.388	-1.745	-2.350	-0.588	-0.658
	(3.148)	(3.166)	(2.329)	(1.795)	(1.219)
Iron field	-1.747	-0.518	0.204	-0.769*	-0.136
	(2.184)	(0.786)	(0.490)	(0.439)	(0.455)
Coal field	1.272	0.704	0.164	0.054	0.931**
	(1.845)	(0.837)	(0.482)	(0.531)	(0.406)
Petroleum field	-0.770	0.221	0.319	0.230	0.734
	(1.183)	(0.777)	(0.643)	(0.507)	(0.456)
Coastal county	-0.308	0.970	1.673***	1.073**	0.655
	(2.154)	(0.848)	(0.578)	(0.469)	(0.520)
Distance to lake	0.911	-0.267	-0.195	-0.032	-0.235
	(0.942)	(0.496)	(0.402)	(0.349)	(0.274)
Distance to river	2.711	1.237	-0.584	-1.323***	-0.420
	(2.128)	(1.143)	(0.587)	(0.457)	(0.381)
Portage	6.582	7.783	5.274	5.200*	1.398
	(4.230)	(4.829)	(3.256)	(2.630)	(1.222)
Railroad access	9.009***	1.053	0.367	0.337	-0.212
	(2.796)	(0.812)	(0.646)	(0.435)	(0.635)
Distance to State capital	-1.074*	0.034	0.042	0.118	-0.059
	(0.594)	(0.281)	(0.244)	(0.179)	(0.134)
Elevation (Mean)	7.485*	0.862	0.098	1.440	1.075
	(4.420)	(1.519)	(1.429)	(1.255)	(1.372)
Elevation (SD)	-8.542	6.810	-1.731	6.645	1.447
	(19.952)	(13.171)	(11.586)	(8.460)	(6.377)
Potential agricultural yield	-11.195	-2.574	-1.130	-0.424	-9.628**
	(11.738)	(4.879)	(4.381)	(2.973)	(4.364)
Hispanic influence	-0.314***	-0.069**	-0.096**	-0.029	-0.021
	(0.060)	(0.028)	(0.039)	(0.019)	(0.013)
French influence	0.360	-0.133	-0.008	-0.081	-0.052
	(0.411)	(0.107)	(0.087)	(0.083)	(0.053)
Latitude	-0.384	-0.093	-0.277	-0.471***	-0.528**
	(0.528)	(0.302)	(0.196)	(0.152)	(0.237)
Longitude	0.398	0.161	0.062	0.146	0.064
	(0.442)	(0.178)	(0.144)	(0.159)	(0.126)

Empl. share manu. & cons.	0.411***	0.514***	0.267***	0.403***	0.281***
	(0.048)	(0.033)	(0.033)	(0.108)	(0.054)
Empl. share services.	0.808***	0.879***	0.689***	0.690***	0.409***
	(0.053)	(0.040)	(0.037)	(0.095)	(0.043)
Empl. share government.	0.245***	0.382***	0.085*	0.223**	0.142***
	(0.072)	(0.073)	(0.047)	(0.103)	(0.046)
Unemp. rate	-0.315	-0.115	-0.048	-0.196	0.004
	(0.240)	(0.086)	(0.103)	(0.209)	(0.181)
Black American pop. share	-0.207***	-0.160***	-0.113***	-0.016	-0.015
	(0.075)	(0.046)	(0.027)	(0.023)	(0.015)
Native American pop. share	0.298*	-0.010	0.079**	0.114***	0.079**
	(0.150)	(0.045)	(0.030)	(0.030)	(0.033)
Asian American pop. share	1.822	0.269	0.623	0.894**	0.665***
	(2.034)	(0.606)	(0.393)	(0.387)	(0.209)
Pop. density	-0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	53.208	4.799	11.312	8.399	14.871
	(44.687)	(20.227)	(17.322)	(19.198)	(17.039)
State-fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	3,109	3,109	3,109	3,109	3,109
Adjusted R-squared	0.115	0.319	0.489	0.481	0.338
P-value gold rush>gold mining	0.001	0.220	0.020	0.045	0.007

Note: OLS regression with robust standard errors. Standard errors clustered by state are reported parentheses. The dependent variables are job creation rates through firm births (in non-agricultural industries) in percent of all jobs. All models use state-fixed effects, geographic controls and the socio-economic controls of the decade: Iron field, coal field, petroleum field, coastal county, distance to lake, distance to river, portage, railroad access by 1910, distance to state capital, elevation mean and standard deviation, potential agricultural yield, Hispanic influence, French influence, latitude, longitude, unemployment rate, Black American population share, Native American population share, Asian American population share, and population density. Last row contains a one-sided Wald test (p-value reported) whether the gold rush coefficient is larger than the gold mining coefficient. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

References for Online Appendix

American Community Survey 2010 (5yr average). U.S. Census Bureau. Table B16001: *Language Spoken at Home by Ability to Speak English* (accessed December 1, 2020).

Andrews, R., Fazio, C., Guzman, J., Liu, Y. & Stern, S. (2019). *The Startup Cartography Project: Measuring and Mapping Entrepreneurial Ecosystems*. https://www.startupcartography.com/research (last accessed May 30, 2019)

Atack, J., Bateman, F., Haines, M. & Margo, R. (2010). Did railroads induce or follow economic growth?: Urbanization and population growth in the American Midwest, 1850–1860. *Social Science History*, 34(2), 171-197.

Audretsch, D., Heger, D. & Veith, T. (2015). Infrastructure and entrepreneurship. *Small Business Economics*, 44(2), 219-230.

Bazzi, S., Fiszbein, M. & Gebresilasse, M. (2020). Frontier culture: The roots and persistence of "Rugged Individualism" in the United States. *Econometrica*, 88(6), 2329-2368.

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

Brodeur, A. & Haddad, J. (2021). Institutions, attitudes and LBGT: Evidence from the gold rush. *Journal of Economic Behavior & Organization*, 16(5), 1540-1576.

Bureau of Economic Analysis (1970-2019). *Employment by County, Metro and Other Areas*. CAEMP25S file and CAEMP25N file https://www.bea.gov/data/employment/employment-county-metro-and-other-areas (last accessed September 10, 2021)

Bureau of Labor Statitics (1976-2020). *Local Area Unemployment Statistics*. ttps://www.bls.gov/lau/#tables (accessed July 24, 2021).

Census Bureau, Center for Economic Studies. 1978-2019. *Business Dynamics Statistics*. States Department of Commerce, https://www.census.gov/data/tables/time-series/econ/bds/bds-tables.html (last accessed October 29, 2021)

Census of the Population (1970). Supplementary report, *Race of the population by county*. Report number PC(S1)-104.

Census of the Population (2000-2010). U.S. Census Bureau; *Census of Population and Housing*, Census Profiles of General Demographic Characteristics. Census of the population by age, sex, and race data files (last accessed April 8, 2021)

Combes, P.-P., Duranton, G., Gobillon, L. & Roux, S. (2010). Estimating agglomeration economies with history, geology and worker efforts. In: Glaeser, E. (Eds.) *The Economic of Agglomeration* (pp. 15-66) Chicago and London, University of Chicago Press.

Eckert, F., Gvirtz, A., Liang, J. & Peters, M. (2020). A method to construct geographical crosswalks with an application to U.S. counties since 1790. *NBER Working Paper* 26770.

Elliot, M. (2008). Model averaging methods for weight trimming. *Journal of Official Statistics*, 24, 517–540.

Falck, O., Heblich, S., Lameli, A. & Südekum, J. (2012). Dialects, cultural identity, and economic exchange. *Journal of Urban Economics*, 72(2–3), 225-239.

Fischer, G., Nachtergaele, F., Prieler, S., Teixeira, E., Tóth, G. van Velthuizen, H., Verelst, L. & Wiberg, D. (2012). *Global Agro-Ecological Zones (GAEZ v3.0): Model documentation*. International Institute for Applied Systems Analysis. Laxenburg, Austria and Rome, IIASA and FAO.

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

Guzman, J. & Stern, S. (2015). Where is Silicon Valley? *Science*, 347(6222), 606-609.

GPIPP (2003-2015). Personality data set. Data set available from Sam Gosling, private communication.

Haden, E. (1973). French dialect geography in North America, in Bright, W., Hymes, W. & Lotz, J. (Eds.) *Linguistics in North America* (pp. 422-471). Paris, The Hague.
Haines, M. (2018). *U.S. county populations by single years of age, sex, and race, 1970, 1980, 1990* (ICPSR 37115). Inter-university consortium for political and social research [distributor], 2018-07-30. https://doi.org/10.3886/ICPSR37115.v1 (accessed September 15, 2021).
Haltiwanger, J., Jarmin, R. & Miranda, J. (2013). Who creates jobs? Small versus large versus young. *The Review of Economics and Statistics*, 95(2), 347-361.
IIASA & FAO (2012). *Global Agro-ecological Zones* (GAEZ v3.0). IIASA, Laxenburg, Austria and FAO, Rome, Italy.
John, O. & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In Pervin, L & John, O. (Eds.), *Handbook of Personality: Theory and Research* (2nd Ed., pp. 102-138). New York: Guilford.
Lujala, P., Ketil, R. & Thieme, N. (2007). Fighting over oil: Introducing a new dataset. *Conflict Management and Peace Science*, 24(3), 239-256.
Manson, S., Schroeder, J., van Riper, D. & Ruggles, S. (2019). *IPUMS National Historical Geographic Information System*: Version 14.0 [Database]. Minneapolis, MN: IPUMS. 2019.
Meade, A. & Craig, S. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437-455.
NOAA (2017). *Coastal County Definitions*. Downloaded from https://coast.noaa.gov/data/digitalcoast/pdf/qrt-coastal-county-definitions.pdf (accessed April 16, 2020)
Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J. & Sobek, M. (2020). *IPUMS USA*: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. ttps://doi.org/10.18128/D010.V10.0.
Stuetzer, M., Obschonka, M., Audretsch, D., Wyrwich, M., Rentfrow, P., Coombes, M., Shaw-Taylor, L. & Satchel, M. (2016). Industry structure, entrepreneurship, and culture: An empirical analysis using historical coalfields. *European Economic Review*, 86, 52-72.
Tarr R. & Mccurry, F. (1910). *New Geographies*. New York: Macmillan Company (Map downloaded from http://etc.usf.edu/maps/pages/2000/2085/2085.htm).
The Startup Cartography Project (2010-2016). *Entrepreneurship_by_County_academic.tab*, https://doi.org/10.7910/DVN/BMRPVH/7A6GVC, Harvard Dataverse, V7, UNF:6:izghknCh01dcMzhftGkmig== [fileUNF] (accessed May 30, 2019)
U.S. Geological Survey. *Mineral Resource Data System*. https://mrdata.usgs.gov/mrds/ (accessed October 14, 2017)
Walker, J. (1986). Racism, slavery, and free enterprise: Black entrepreneurship in the United States before the civil war. *The Business History Review*, 60(3), 343-382.