The Review of Financial Studies

Household Inequality, Entrepreneurial Dynamism, and Corporate Financing

Fabio Braggion
Tilburg University

Mintra Dwarkasing
Erasmus University Rotterdam

Steven Ongena
University of Zurich, Swiss Finance Institute, KU Leuven, and CEPR

Economic theories posit conflicting hypotheses on how wealth inequality affects entrepreneurial dynamism. We investigate the impact of wealth inequality on business dynamics by constructing local measures of household wealth inequality based on financial rents, home equity, and 1880 farmland. We then identify the effect of wealth inequality on entrepreneurship by instrumenting it with land distribution under the 1862 Homestead Act. Wealth inequality decreases firm entry and exit, and the proportion of high-tech businesses across metropolitan statistical areas. Wealth inequality also lowers the supply of public goods, such as education. Growth in income per capita consequently lags. (*JEL* D31, G3, H41, L26)

Received July 5, 2018; editorial decision June 23, 2020 by Editor Itay Goldstein.

We thank two anonymous referees; Cédric Argenton, Daniel Berkowitz, Adam Bonica, Fabio Castiglionesi, James Choi, Francesca Cornelli (the editor), Julio Crego, Marco Da Rin, Joost Driessen, Narly Dwarkasing, Rik Frehen, Itay Goldstein (the editor), Thomas Hellmann, William Kerr, Mauricio Larrain, Alberto Manconi, Felix von Meyerinck, Lyndon Moore, Ramana Nanda, Fabiana Peñas, Rodney Ramcharan, Johann Reindl, David Robinson, Jonathan Rork, Moritz Schularick, Karen Smith Conway, Christophe Spaenjers, Nic Schaub, Florian Schuett, Per Strömberg, Joacim Tåg, Vadym Volosovych, and Haikun Zhu; participants at the 2015 ASSA-Meetings (Boston), the 2015 Duke Kauffman Survey Conference, the 2015 European Finance Association Conference, the 2015 FIRS Conference (Reykjavík), the Second European Workshop on Entrepreneurship Economics (Cagliari), the 5th Annual Scientific Conference of Romanian Academic Economists from Abroad (Iaşi), the Sveriges Riksbank and EABCN Conference on Inequality and Macroeconomics (Stockholm), the 15th Workshop on Corporate Governance and Investment (Oslo), and the Bundesbank-CFS-ECB Joint Lunchtime Seminar; and seminar participants at BI Norwegian Business School (Oslo), Bournemouth University, Eastern Mediterranean University (Famagusta), the Einaudi Institute for Economics and Finance (Rome), Erasmus University Rotterdam, the European Investment Bank (Luxembourg), Ghent University, Goethe University Frankfurt, New York University Shanghai, Rotterdam School of Management, Tilburg University, the Tilburg Center for Law and Economics, Özyeõin University (Istanbul), and the universities of Basel and Bonn for their very valuable comments. We thank Adam Bonica for sharing with us a confidential and up-to-date version of his political campaign data. We thank Dave Brooks (the English Language Coaching Service) for excellent editorial assistance. During the writing of this paper, the Netherlands Organisation for Scientific Research (NWO) generously supported Braggion through its VIDI Grant Program and Dwarkasing through its Mosaic Grant Program. Ongena acknowledges financial support from the European Research Council [Grant ADG 2016 - GA 740272]. Braggion is also affiliated with CEPR and ECGI. A previous draft of the paper circulated with the title “Household Inequality, Corporate Capital Structure and Entrepreneurial Dynamism.” Send correspondence to Fabio Braggion, f.braggion@uvt.nl.

The Review of Financial Studies 34 (2021) 2448–2507

doi:10.1093/rfs/hhaa097 Advance Access publication August 26, 2020

Households' wealth inequality is a defining societal characteristic with important implications for financial and real activity.[1] The growth of wealth inequality over recent decades has returned the issue to the top of the agendas of policy makers and social leaders, but despite such renewed attention, economists are still far from reaching a consensus about its implications. In this paper, we study the relationship between household wealth inequality and the main drivers of economic activity: entrepreneurship, corporate financing, and growth in income per capita. Starting with Schumpeter, entrepreneurship has been considered an important component of economic growth (Schumpeter 1934; Aghion and Howitt 1992; Akcigit and Kerr 2018), and it is therefore no surprise that the study of its determinants has gained attention in the academic literature (e.g., Dunne, Roberts, and Samuelson 1988; Glaeser and Kerr 2009; Kerr and Nanda 2010).

A perspective that dates to at least Adam Smith conjectures a negative relationship between inequality and entrepreneurship.[2] Wealth inequality may prevent the development of institutions supportive of entrepreneurship and new businesses, such as schools and the judiciary, thus impairing economic growth. The wealthy may consider the financing of these public goods as a form of redistribution and refuse to fund them or they may distort their functioning to suit their interests.[3] Another view, rooted in the work of Malthus and Kaldor (1961), predicts a positive relationship between wealth inequality and economic development as wealthy individuals may generate both high aggregate demand and high savings, which can ultimately support entrepreneurship.[4]

The related evidence is mixed. A group of papers show a negative relationship between inequality and economic development (Alesina and Rodrik 1994; Persson and Tabellini 1994; Easterly 2007), whereas other works find a positive or a nonlinear relationship (Barro 2000; Forbes 2000; Banerjee and Duflo 2003; Braun, Parro, and Valenzuela 2019). All these studies, however, focus on cross-sections of countries, and they often employ income inequality as a proxy for wealth inequality.

Our study brings these alternative perspectives to the data and focuses on a novel setting based on households' wealth inequality measured at the

[1] Recent academic work consequently seeks to define, measure, analyze, and explain inequality (e.g., Piketty 2014).

[2] In *The Wealth of Nations*, Adam Smith expressed concern that an unequal distribution of land may have negatively affected the development of the New World colonies. In his words: "The engrossing of land, in effect, destroys this plenty and cheapness" (Smith 1776, p. 726).

[3] See Engerman and Sokoloff (1997), Glaeser, Scheinkman, and Shleifer (2003), Sonin (2003), Berkowitz and Clay (2011), Rajan and Ramcharan (2011), Acemoglu and Robinson (2013, pp. 152–8), and Kumhof, Rancière, and Winant (2015).

[4] Malthus (1836, p. 466) writes "there must therefore be a considerable class of persons who have both the will and power to consume more material wealth than they produce, or the mercantile classes could not continue profitably to produce so much more than they consume. In this class the landlords no doubt stand pre-eminent." Stiglitz (1969) and Bourguignon (1981) formalize the notion that wealthy individuals may help economic growth via their higher saving rates.

U.S. metropolitan statistical area (MSA) level.[5] Studying MSAs allows us to compare locations that are similar in terms of quality of the labor force and demography, thus reducing concerns that omitted variables may be driving our results. At the same time, MSAs provide us with a sizable proportion of new business formation, as 80% of new firms are located in MSAs (see Robb and Robinson 2014).

Our contributions are threefold. We first provide different measures of local wealth inequality that identify the fundamental elements of wealth: households' financial rents and housing equity. Despite originating from different sources, these measures are highly correlated and they robustly document the pattern of regional wealth inequality within the United States. We then exploit the U.S. institutional setting to design an identification strategy that allows us to connect wealth inequality to new business formation, corporate financing, and the local degree of economic development, thus minimizing endogeneity concerns. Last, we explore the mechanisms that connect wealth inequality to business formation, focusing in particular on a channel based on political preferences for redistribution and the provision of public goods.

Our first measure of local wealth inequality is based on total dividends earned by U.S. households in 2004 (the first year these data is available). Assuming that every household invests in the market index, financial wealth is directly proportional to the total in dividends they receive. Since the Internal Revenue Service (IRS) provides data on dividends earned in every ZIP code, we are able to back out the distribution of financial wealth in each MSA. Our second measure relies on households' housing equity data, available from the U.S. Census Bureau and the Federal Home Loan Mortgage Corporation (Freddie Mac). The U.S. census provides data, at the MSA level, on the value of housing units that are free of mortgages (i.e., the value of the house corresponds to the household's housing equity), whereas Freddie Mac provides data on all mortgages (and underlying house values) it has acquired between 1999 and 2004.

We use this information to construct the distribution of financial rents and housing equity at the local level and to compute Gini coefficients of financial and housing wealth inequality.[6] In Figure 1, we observe that, despite being originated from different sources, both Ginis produce a very similar map of wealth inequality. Inland California and the South appear to be among the most unequal areas along both inequality measures, whereas the Midwest, and especially Minnesota and Wisconsin, are the least unequal. Figure 2 provides us with a preview of our main result, that is, a negative relationship between

[5] Our investigation at this level of geographical disaggregation is appropriate because U.S. local administrations are often coresponsible (with state-level authorities) for many important elements of public life, such as the organization of schooling, the judiciary and enforcement of the law, and taxation.

[6] Mian, Rao, and Sufi (2013) likewise construct local measures of U.S. households' net worth, and Saez and Zucman (2016) measure U.S.-wide wealth inequality over a longer time period. Additionally, we compute inequality measures based on the wealth holdings of the top 10% wealth owners.

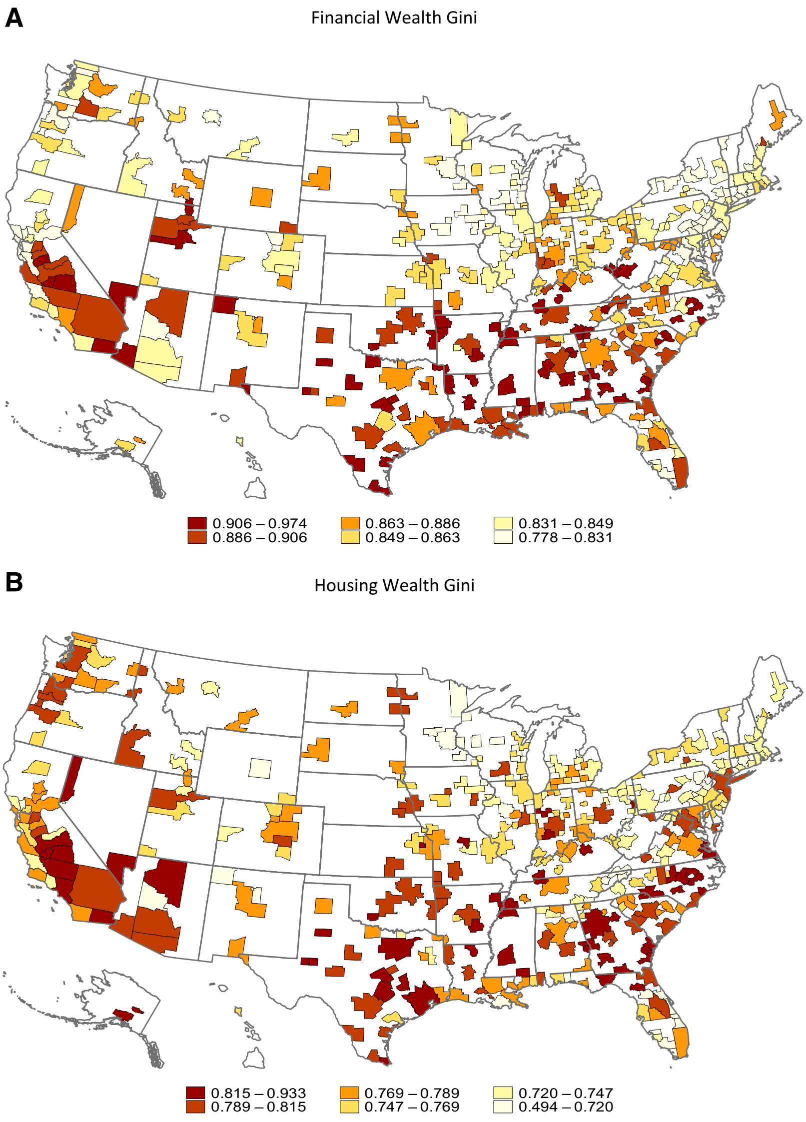

Figure 1
Financial and housing wealth inequality
The figure maps our measures of household wealth inequality across U.S. metropolitan statistical areas. The Financial Wealth Gini is the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the metropolitan statistical area. The Housing Wealth Gini is the Gini coefficient of the distribution of wealth measured as households' home equity. A higher Gini (i.e., darker colored) corresponds to a higher degree of wealth inequality.

wealth inequality (on the x-axis) and business formation from 2004 to 2012 (on the y-axis) across MSAs. Hence, increased wealth disparities are associated with less entrepreneurial activity.[7]

Establishing a causal relationship between wealth inequality and entrepreneurial activities, however, is difficult, as wealth inequality itself may correlate with unobserved factors that are likely to affect our estimates. We alleviate this concern with an empirical strategy that relies on two ingredients. First, as our measures of wealth inequality are local (and because we know the location of new "establishments"—defined here as firms that are created), we can saturate our specifications with state, year, and/or state-year fixed effects to account for any unobserved heterogeneity at those aforementioned levels. This procedure allows us to control for competing explanations of the deeply rooted determinants of institutions, such as individual states' type of colonization and legal traditions (see Acemoglu, Johnson, and Robinson 2001; Berkowitz and Clay 2011, pp. 16–59), as well as changes in their legislation and regulation. Importantly, while our focus is wealth inequality, the detail of our data always allows us to control for the average level of wealth as well as of income in each MSA.

Second, we instrument the contemporary measure of wealth inequality with the proportion of land assigned in each MSA via the 1862 Homestead Act and exploit the Act's distribution of equal plots of land to early settlers in the nineteenth century. Indeed, the Homestead Act entitled any person who was head of a family or 21 years of age to apply for a plot of 160 acres. The only requirement was that the homesteader should either live on the land or cultivate it for the next 5 years.

We exploit the quasi-exogenous variation coming from early settlers' demand for land during the 10 years following first settlement, when the quality of the land was still unknown and the settlers reached the western frontier for the first time, having little or no reliable information on the climate and the geography of the area. This allows us to consider the distribution of land in this period to be quasi-random, an assumption that we then verify in the data. The federal government intended to distribute public land via homesteading quickly, but this made the process chaotic: early settlers, since they had to compete with wealthy speculators to obtain land first, filed applications for plots in areas whose features were still unknown. At the same time, homesteading, by assigning equal plots of land to households, may have biased the distribution of wealth toward equality from the very beginning, in most cases before communities were formed and before local institutions, such as schools and the judiciary, even existed. As a result, the Homestead Act, as the first element to influence inequality, allows us to capture the effects of

[7] In Figure 2, panel B, we obtain a very similar result if we winsorize the employed housing wealth inequality measure at the 1% level.

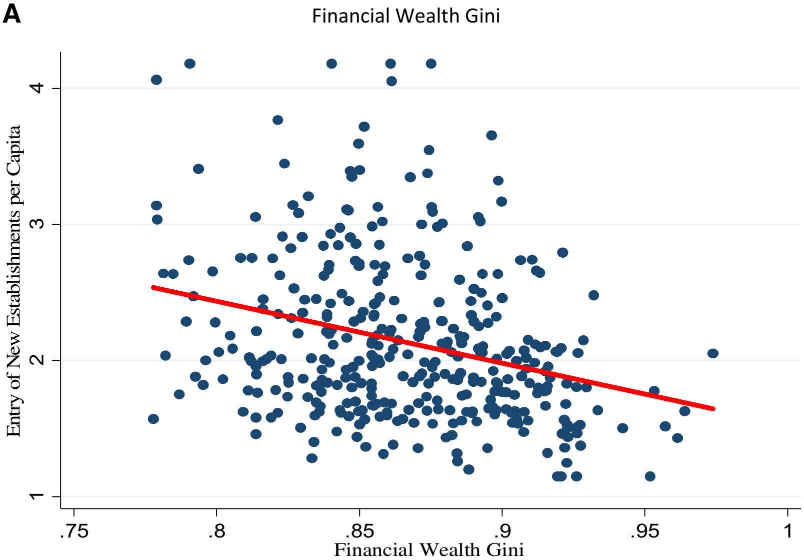

Figure 2
Financial and housing wealth inequality and entry of new establishments
The figure maps the entry of new establishments per capita versus our measures of household wealth inequality across U.S. metropolitan statistical areas. Total Establishments Entry per Capita is the yearly total number of new establishments in the MSA between 2005 and 2014 divided by total MSA population. The Financial Wealth Gini is the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the metropolitan statistical area. The Housing Wealth Gini is the Gini coefficient of the distribution of wealth measured as home equity.

initial inequality as opposed to inequality created later by the local institutions themselves.

A natural concern with our instrument is whether homesteaders possessed different human capital than an average settler or whether the land assigned via the Homestead Act was intrinsically different from an average lot of farmland. In either case, we would have homesteading directly affecting entrepreneurship, hence violating the exclusion restriction. We will address these and other issues with a series of tests intended to corroborate the validity of the identification strategy. The estimated coefficients robustly suggest that MSA-level inequality decreases firm entry into and exit from the MSA, indicating that wealth inequality has a negative effect on local business formation. Our estimates are not only statistically significant but also economically relevant. A one-standard-deviation increase in MSA-level wealth inequality leads to an approximate 12% decline in new establishments' entry and exit per capita. We also find that new establishments created in more unequal areas are more likely to be units generated by already existing firms than new ventures genuinely formed *ex novo*, and that establishments in unequal MSAs are more likely to be engaged in traditional rather than high-tech sectors. More unequal MSAs also have lower growth in income per capita.

Last, we strengthen the causal interpretation of our results by studying what mechanism can explain them. We find that wealth inequality comes with less provision of public goods considered important for business formation and economic growth. In particular, more unequal MSAs are characterized by lower school expenditure per pupil, lower proportion of public school revenue obtained from local sources, and a more inefficient justice system. These findings can explain the negative effect of inequality on entrepreneurship we document: a lower supply of public goods may deter entrepreneurs from setting up a firm in the first place. We also find evidence suggesting that these policy outcomes are the result of stronger political preferences against redistribution. Using detailed data on political campaign contributions from Bonica (2014), we show that in more unequal MSAs larger campaign contributions flow to local and state election candidates with platforms proposing a lower supply of public goods.

In sum, we provide a number of findings that consistently suggest that wealth inequality is an important determinant of entrepreneurship and of the type and amount of financing entrepreneurs receive.[8] Our analysis also adds to a growing literature on finance and inequality. While most of the work in this area studies how finance may affect the degree of income or wealth inequality (for a review, see Demirgüç-Kunt and Levine 2009; Beck, Levine, and Levkov 2010), our paper studies how wealth inequality affects financial outcomes (and,

[8] See Black and Strahan (2002), Berkowitz and White (2004), and Perotti and von Thadden (2006), as well as a review by Carlino and Kerr (2015), among others.

in this sense, is more similar to Rajan 2009; Bagchi and Svejnar 2015; Degryse, Lambert, and Schwienbacher 2018).

We also contribute to the literature that studies the long-run dynamics of business formation in the United States (Decker et al. 2018; Foster et al. 2018), and we relate them to wealth inequality. Using simple time-series data, Figure 3 shows that while U.S. business formation has been in decline since the 1970s, wealth inequality has increased: a pattern broadly consistent with our findings.

1. Inequality and Entrepreneurial Outcomes

1.1 Background

1.1.1 Negative relationship between wealth inequality and business formation. A first group of theories predicts a negative relationship between wealth inequality and economic outcomes. These explanations do not imply a direct effect of inequality on entrepreneurship, but they rely on mediating factors in the form of public goods, income or capital taxation, and credit markets. We will discuss and classify them below.

1.1.1.1. Negative redistribution channels. Negative redistribution channels work through the notion that inequality may affect redistributive policies and the provision of public goods, which in turn may adversely affect entrepreneurial dynamism.[9] For instance, Engerman and Sokoloff (2002), Galor, Moav, and Vollrath (2009), and Acemoglu and Robinson (2013) state that inequality may decrease the amount of redistribution in a society. As a result, unequal societies will be characterized by a lower level of provision of local public goods as their financing is seen by the wealthy as an unwarranted form of redistribution to the poor. For example, when various states introduced some form of redistribution of school financing during the 1980s (Card and Payne 2002), wealthy neighborhoods lobbied to pay a lower contribution to the schooling service because they did not want to redirect resources to poor districts (Stark and Zasloff 2003).[10] The wealthy may also use their influence to distort the development of local public goods in order to promote their own interests. For instance, the fact that state judges are elected may enable wealthy individuals to twist judicial decisions in their favor by contributing to judges' electoral campaigns. Supporting this possibility, a *New York Times* article published in 2006 documents that Ohio Supreme Court judges ruled in favor of those who had contributed to their campaigns 70% of the time (Liptak and Roberts 2006; see also Berkowitz and Clay 2011, p. 133).

[9] For simplicity, we also define education as a public good even if it also can be considered a private good that can be provided by the public sector.

[10] Consistent with the idea that inequality could be detrimental to educational outcomes, Goldin and Katz (1998) and Galor, Moav, and Vollrath (2009) show that both high school graduation rates and expenditures on public education in the twentieth century were lower in states that displayed higher degrees of wealth inequality.

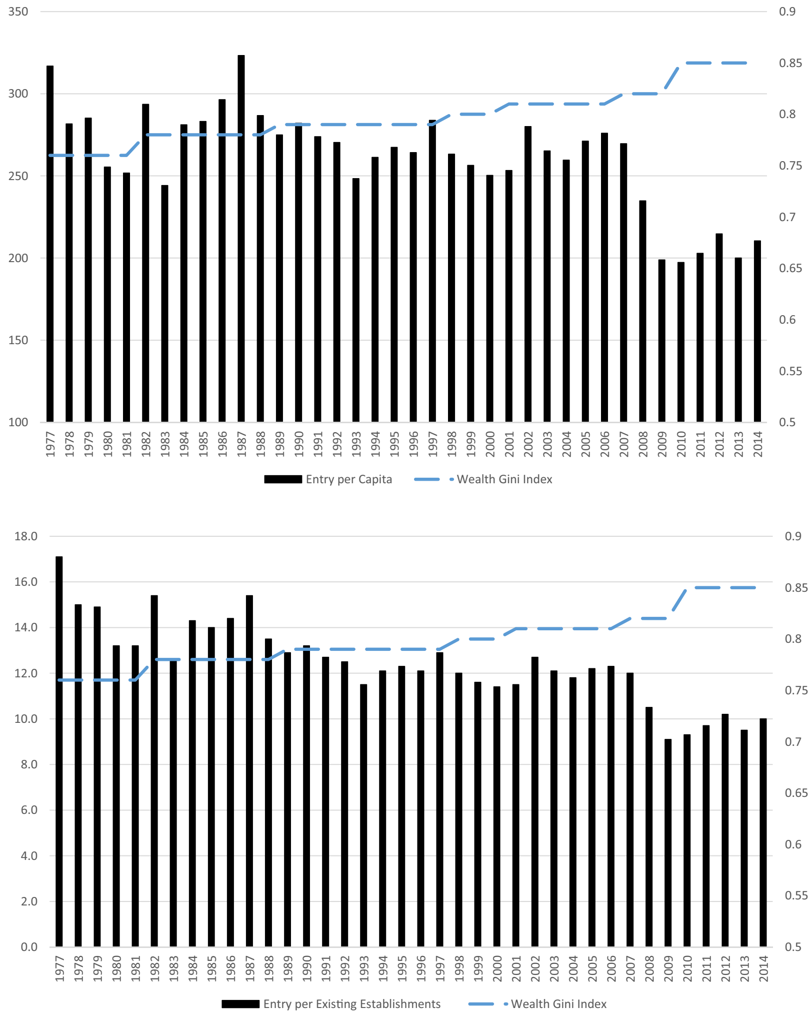

Figure 3
Business formation and wealth inequality from 1977 to 2014
The top panel displays the number of establishments entry per capita and the U.S.-wide Gini index of wealth inequality. The bottom panel displays the number of establishments entry per existing establishments and also displays the U.S.-wide Gini index of wealth inequality. Wealth inequality data ara available for 1977, 1982, 1989, 1992, 1995, 1998, 2001, 2007, and 2013. We interpolated the missing years by using the value of the earlier available year. Sources: U.S. Census Bureau and Kuhn, Schularick, and Steins (2019).

1.1.1.2. Positive redistribution channel. Another set of explanations, put forward by Alesina and Rodrik (1994) and Persson and Tabellini (1994), states that inequality may be detrimental for economic growth because it increases the amount of redistribution and hence future taxation on income and/or capital. If entrepreneurs expect to see a larger share of their gains taken away by the tax authorities, they will be less willing to invest or start a business altogether.

1.1.1.3. Credit market channels. Credit market channels have been studied in more unequal areas, where the quality of the banking and credit markets could be especially poor. Rajan and Ramcharan (2011), for example, provide evidence that in the 1930s, the wealthy may have contributed to keeping banking markets underdeveloped in U.S. counties so as to maintain their grip on power.[11]

1.1.2 A positive relationship between wealth inequality and business formation. An alternative view to those just presented considers wealth inequality as a positive factor for the economy. With their deep pockets, wealthy individuals may help stabilize aggregate demand and reduce the average cost of production of new, high-tech goods, making them more accessible to the middle class (Matsuyama 2002). The existence of inequality in income and wealth may encourage individuals to work harder; rewarding employees with equal wages may discourage them from producing the desired level of effort (Mirrlees 1971). Philanthropy and charity may produce enough resources to extend schooling and economic opportunities to the poor, enhancing human capital formation. In 2001, *Forbes*, for example, espoused the view that, traditionally, charity in the United States "concentrated in education and acculturation" and "stressed the skills and attitudes of self-reliance and personal responsibility" (Magney 2001).

2. Empirical Strategy

2.1 Main specification

In our main analysis, we will estimate the impact of wealth inequality on the number of establishments' entries and exits to and from the market, using data at the MSA level.[12] In particular, for entry and exit, we will estimate the following equation:

$$Y_{j,t} = \alpha + \alpha_s + \alpha_t + \beta Wealth\ Inequality_j + Controls_{j,t-1} + \varepsilon_{j,t}, \tag{1}$$

where $Y_{j,t}$ indicates the natural logarithm for the number of establishments' entries and exits in the MSA j at year t. In line with the focus of Kerr and Nanda (2010), we will focus on gross business entry and exit per

[11] Moreover, even if regions display the same degree of financial development as one another, wealth inequality still could directly negatively affect entrepreneurship as suggested by Banerjee and Newman (1993) and Aghion and Bolton (1997). In an economy with credit constraints, the likelihood of starting a business increases with one's wealth. This implies that, for any level of wealth, more unequal areas will have lower business formation figures as fewer individuals (possibly only the very rich) will have enough wealth to overcome credit market imperfections.

[12] In analyzing the entry and exit of new firms, we use data at the MSA level, as data at the county level are not publicly available.

capita. The variable *Wealth Inequality* indicates one of our measures of local wealth inequality. *Controls* stands for a set of MSA controls, such as total financial and housing wealth, population growth, the Catholic-to-Protestant ratio, and housing prices—all detailed in Tables A.1 and A.2 in the appendix. These variables will allow us to control for time-varying MSA characteristics: for instance, the Catholic-to-Protestant ratio may proxy for differences in the degree of risk aversion in the MSAs (see Hilary and Hui 2009). Finally, the level and growth of house prices allow us to control for the possibility that homeowners may borrow against the value of their homes in order to set up a new business.

2.2 Measuring wealth inequality

Obtaining representative measures of wealth inequality at the local level is challenging. We construct two proxies for local wealth inequality that embody the main avenues through which wealth can be accumulated: financial capital and housing. The first one is based on current levels of financial wealth and is broadly based on a methodology introduced by Mian, Rao, and Sufi (2013) and Saez and Zucman (2016); it intends to construct local-level measures of household net worth. The second measure is based on the available information on housing equity.

The measure of financial wealth inequality looks at the totals in dividends and interest earned by U.S. households in 2004, the first year in our sample period, as reported in Internal Revenue Service (IRS) Statistics of Income (SOI) data. The IRS/SOI data report the total dividend and interest income received by U.S. households in a certain ZIP code. The information is reported as a total amount per ZIP code and is divided into five household income groups, ranging from low income to high income. Under the assumption that a typical household owns the market index for stocks and bonds, the amount of financial rents it receives depends only on the quantity of stocks and bonds it holds. We use this information to construct a Gini index of wealth inequality based on financial rents. The procedure we adopted to construct the index is detailed in Table A.3 and Section A1 in the appendix.[13]

The second measure of inequality relies on housing equity. From the Census Bureau, we obtain data on the values of all houses that are fully owned by a household and do not have an outstanding mortgage. In providing this information, the census divides houses into bins corresponding to their different values in 2004. From Freddie Mac, we obtain information on housing equity for a more recent period of time. Freddie Mac has made available the details of

[13] As an addition to the Gini index, we construct an alternative measure of wealth inequality: the proportion of households in the MSA/county without financial wealth. The results we obtain with this alternative measure are the same as those obtained with the Gini index and, if anything, stronger (i.e., with a lower p-value in terms of statistical significance and larger in terms of economic magnitude).

a large selection of mortgages it purchased between 1999 and 2004.[14] These data include the MSA where the house was purchased, the value of the house at the time of the purchase, and the value of the mortgage originated. Note that between 1999 and 2004 Freddie Mac covered a large proportion of the mortgage market, about 27%. This information allows us to compute housing equity, divide households into the same housing value bins used in the census, and compute a Housing Wealth Gini index per MSA considering both housing equity between 1999 and 2004 and the value of any house that is free of any mortgage (using the latter set we will also construct a No Mortgage Housing Wealth Gini).

Both our inequality measures are imperfect proxies. The Financial Wealth Gini relies on the assumption that every individual owns the same (or very similar classes of) stocks (and, as a result, they obtain the same return).[15] The Housing Wealth Gini lacks information on housing equities for houses purchased before 1999 with a mortgage still attached and relies on the information provided by Fannie Mae. Yet Figure 1 reveals that the maps of wealth inequality originating from our measures are remarkably similar. We observe quite some variation in the level of wealth inequality even within states. MSAs with a higher level of wealth inequality tend to be concentrated in the South and to some extent in inland California. The Midwest appears to have a more equitable distribution of resources.[16]

The average Gini coefficients we obtain are 0.87 for financial wealth inequality and 0.77 in the case of housing wealth inequality. These figures are also in line with measures of household wealth inequality obtained at the aggregated level. For example, Kuhn, Schularick, and Steins (forthcoming) provide U.S.-wide figures of the Gini index that are remarkably similar to ours. In 2007, they find the U.S. total wealth Gini index to be 0.82. A Gini index based only on financial wealth was 0.96; 0.67 when based only on housing wealth. For a comparison, the first decile of our MSA housing wealth distribution is about 0.67 and indicates a level of inequality comparable to France and the Netherlands in the year 2000 (see Davies, Lluberas, and Shorrocks 2017). The top decile of housing wealth inequality is about 0.82, a level similar to Brazil

14 The mortgages we consider in our data are provided by the Fannie Mae Loan Performance data set, which includes over 35 million mortgage loans originated after January 1, 1999. The mortgages considered in the data are fixed rate, fully amortizing mortgage loans, with a maturity less than 5 years but more than 35 years. These mortgages were also supported by full borrower documentation. The data exclude riskier mortgages, such as Alt-A, interest-only mortgages, and mortgages, with a loan-to-value ratio greater than 97%.

15 Piketty (2014, pp. 439–43) describes how wealthy investors tend to obtain higher returns on their investments. Using detailed data on Swedish households, Bach, Calvet, and Sodini (forthcoming) show that the top 10% of wealth owners obtain returns on their investments that are 2.5 percentage points higher than the median household. These higher returns are driven by a larger exposure to various risk factors.

16 Unfortunately, as we do not have household-level data on wealth, we cannot compute a Gini index that simultaneously places the same household in different wealth categories. For instance, a household may have high housing equity but low financial wealth.

or Peru, as recorded in 2014, or to preindustrial societies in Europe (see Credit Suisse, *Global Wealth Book* 2014; Scheidel 2017, pp. 98, 337).[17]

Previous contributions to the literature have also considered land as another possible measure of wealth inequality (Alesina and Rodrik 1994; Galor, Moav, and Vollrath 2009; Ramcharan 2010; Rajan and Ramcharan 2011; Vollrath 2013). To construct such a historical measure, we obtain information on farmland plot sizes at the county level from the 1880 U.S. census. More precisely, for each county, we have information on the total number of farms that—based on their total acreage of farmland—fall within a certain size bin. Farms are assigned to one of seven bins: under 10 acres, from 10 to 19 acres, 20 to 49 acres, 50 to 99 acres, 100 to 499 acres, 500 to 999 acres, and 1,000 or more acres. The average 1880 Gini index is 0.44. It is significantly lower than the contemporary measures as it looks only at the distribution of land and does not consider individuals that are not landowners.[18]

The housing and the financial wealth Ginis have a positive correlation of about 48%. We also find that 1880 land inequality displays a 37% and 26% positive correlation with the Financial Wealth Gini and Housing Gini, respectively.[19]

2.3 Identification

Wealth inequality could be correlated either with omitted factors or with the degree of entrepreneurship itself. The possibility of reverse causality is based on the idea that entrepreneurs are a small fraction of the population but hold a large share of the total wealth (Cagetti and De Nardi 2008). When entrepreneurial activities are successful, most of the rewards are accrued among a limited number of individuals, an outcome that in turn increases wealth inequality. From this perspective, we may expect to find a positive correlation between wealth inequality and entrepreneurship.

Wealth inequality also could be correlated with policies introduced by states, local income, and the racial composition of the geographical area. While we introduce variables that specifically control for these factors, we also address this problem in several additional ways.

---

17 Using the same information, we construct proxies of the total in dividends received by the top 10% of dividend earners and the housing wealth owned by the top 10% of house owners in value terms (see Section A1 in the appendix for details about their construction). Since we do not have household-level data on dividends and housing equity, we are forced to construct proxies of these measures relying on more aggregate data.

18 The Financial Wealth Gini is on average higher than the other two Ginis because a relatively large proportion of households do not report any wealth held in equity, that is, about 78%. The IRS reports change with respect to the number and types of income bins after 2007; as a result, comparing Gini indexes throughout our time period is difficult to do. However, we find that the proportion of individuals not holding any financial wealth in 2012 is 78% and very similar to what we obtained for 2004. This also suggests that, at least in this dimension, the 2008 financial crisis and the Great Recession did affect the distribution of financial wealth in a large way.

19 Table A.4 in the appendix presents the correlations.

2.3.1 State fixed effects and trends. First, as we discussed above, our measure of wealth inequality is constructed at either the MSA or the county level, which allows us to control for state fixed effects and state trends in the analysis. These are a relevant feature of our identification strategy, as other main, deeply rooted determinants of institutions, such as legal and colonial origins, are defined at the state level. Berkowitz and Clay (2011) give a precise overview of which U.S. states have civil law (rather than common law) traditions and link their legal traditions to the countries of origin of early settlers.

Importantly, a study of MSAs also enables us to better control for omitted variable bias as we compare locations that are similar in terms of quality of the labor force and demography.

2.3.2 Instrumental variable analysis: The Homestead Act. Second, we construct an instrumental variable analysis based on the distribution of land to early settlers via the 1862 Homestead Act.[20] To encourage migration into the largely unexplored areas of the West, the federal government allocated free plots of land, all equal in size, to settlers. The Act of 1862 entitled any persons who were head of a family or 21 years of age to apply for a homestead. The only requirement was that the settler should either live on the land or cultivate it for the next 5 years (Gates 1936, pp. 393–4).[21] The idea is that the Homestead Act, by assigning equal-sized plots of land to early settlers, may have biased the local initial conditions toward a more equitable distribution of resources, before communities were even formed and any local institution, such as schools or the judiciary, were established.[22] Most of the land assigned via Homestead was, in fact, beyond the 1860 frontier line as indicated in Figure 4 and was scarcely populated. As a result, the Homestead Act allows us to focus on the initial level of inequality that characterized a certain area, rather than subsequent inequality that could have been generated by local institutions themselves. We conjecture

20 As an alternative set of instruments for wealth inequality, we also employ weather and soil conditions because different climates and geographical environments may favor the production of one type of crop over another. Engerman and Sokoloff (1997, 2002) suggest that crops, such as sugar or tobacco, best suited to large plantations will induce relatively high economic inequality. The production of these crops comes at a high fixed cost; as a result, in equilibrium, the market can support only a few farms. The outcome is thus a society controlled by few wealthy landowners. Conversely, climates supporting crops, such as wheat, will result in a society with low wealth inequality. The production of these crops does not require high fixed costs; hence, the market can "bear" more producers. These societies will be more equal and mainly comprise small landowners.

21 Settlers who decided not to meet the 5-year requirement could obtain the full title to the land by paying a minimum price of $1.25/acre.

22 We consider land assigned via the original Homestead Act and subsequent acts, such as the 1873 Timber Culture Act, the Desert Land Act, the 1877 Desert Land Act, the 1909 Enlarged Homestead Act, and the 1916 Stock-Raising Homestead Act. Since the 1862 Homestead Act corresponds to about 91% of land distributed via homesteading, results do not change.

Source. Walton, G. M., and H. Rockoff. *History of American economy,* 11th edition, p. 134. Independence, KY: South-Western Cengage Learning.

Figure 4
American moving frontier

that areas with a larger proportion of land assigned via the Homestead Act should be more equal today.[23]

Naturally, an instrument ought not to be correlated with other factors that may directly explain the pattern of contemporary entrepreneurship. For instance, the equal distribution of land may directly influence business formation today by affecting subsequent land and farm values. To the extent that land can be used as collateral to receive a loan and start a new business, this would violate our exclusion restriction.[24] To alleviate this problem, we focus on the proportion of land assigned via the Homestead Act during its first 10 years for MSAs already

[23] The idea that by promoting equality the Homestead Act could have produced better local institutions was also recognized by the U.S. Federal government in the nineteenth century. The 1884 report of the Public Land Commission described the Act in the following terms: "the Homestead Act is now the approved and preferred method of acquiring title to the public lands. […] It protects the Government, it fills the States with homes, it builds up communities, and lessens the chances of social and civil disorder by giving ownership of the soil, in small tracts, to the occupants thereof. It was copied from no other nation's system. It was originally and distinctively American, and remains a monument to its originators" (Donaldson 1884, p. 350).

[24] A few recent papers have related real estate value to entrepreneurship: for instance, Adelino, Schoar, and Severino (2015), Corradin and Popov (2015), and Schmalz, Sraer, and Thesmar (2017). Kerr, Kerr, and Nanda (2015), however, provide evidence suggesting that land and home values may be not be a particularly important determinant in one's decision to become an entrepreneur.

settled in 1862, or during the first 10 years following the first settlement of the MSA for MSAs whose first settlement occurred after 1862.[25]

In the early years, land was assigned both via direct sales by the federal government and via homesteading. Both land speculators and early settlers struggled to understand what exactly were the best plots of land in a given area (Gates 1936). Early settlers were confronted with a new and unknown environment, and they lacked knowledge of the quality of the land and the type of weather that characterized the areas they moved into. As late as 1897, the National Weather Bureau admitted the primitive state of knowledge about rainfall and recognized that there were few observed patterns allowing for an accurate prediction of droughts (Hansen and Libecap 2002). The settlers also lacked the knowledge required to interpret the consequences of weather and soil quality for agriculture. All these circumstances suggest that the proportion of land assigned via the Homestead Act may have had an important random component to it determined by settlers' beliefs on the quality of soil and weather of the desired locations, which many times turned out to be faulty.

It also has been observed that homesteading itself may have affected the subsequent value of farming and land (Hansen and Libecap 2004). For instance, in the Great Planes, the small plots assigned via the Homestead Act were not adequate for the type of soil and weather that characterized the area. The large number of farm failures that followed the 1930s dust bowl phenomena has been attributed to the small size of farms in the Great Planes, a legacy of the Homestead movement. Larger plots would have been more appropriate and would have allowed farmers to make necessary investments to prevent drought and strong winds damaging their crops. We also run the analysis excluding areas that were affected by the 1930s Dust Bowl phenomenon and the entire Great Planes area, as—if any—the effect of homesteading on farm values should be stronger in these areas.

Another possible concern relates to the quality of the human capital of early settlers. It could be that homesteaders possessed different human capital than an average migrant (different entrepreneurial spirit, education, or cultural background). To the extent that such differences are persistent over years, homesteading could also affect the pattern of entrepreneurship via individuals' skills.

Our regressions will include a measure of the literacy of the local population as a proxy for human capital, as obtained from the U.S. census of 1860. We also tackle this issue by studying a distribution policy where land was randomly allocated to the local community via a lottery system: the Georgia

[25] We compute the proportion of land assigned via the Homestead Act as the total amount of land assigned via Homesteading divided by the total amount of land assigned in any form to any subject by the federal government. This measure has the advantage of capturing the strength of homesteading as a tool of land distribution vis-à-vis other available means. At the same time, it automatically excludes land that the federal government could not assign because it was uninhabitable or not suited to economic activity. We rerun our analysis using 5, 15, or 20 years for the homesteading assessment window.

land lotteries of 1827–1835. Restricting our analysis to Georgia and focusing on the proportion of land distributed via lotteries allows us to abstract from concerns related to the quality of human capital, as land was given randomly and nearly the entire eligible population participated in the lottery (Bleakley and Ferrie 2016).[26]

3. Business Dynamics

3.1 Data sources

We obtain data on establishments' entry and exit from the Business Dynamics Statistics (BDS), a database set up by the U.S. census that provides annual measures of, among other things, establishment "births" and "deaths" and firm startups and shutdowns.[27] The BDS data are available only at the U.S. MSA level and provide information from 1977 until 2014. They cover a wide range of industrial sectors including agriculture, manufacturing, wholesale trade, retail trade, and services. Following Kerr and Nanda (2009), we define entrepreneurship as the entry of new establishments, but we will also distinguish whether these establishments are related to newly created firms as opposed to already existing companies. Table 1 provides the resultant descriptive statistics for all variables.

We collect the data on our main dependent variables for our financing and employment in high tech analyses from several sources. We access the Thompson 1 database for information on the reliance on venture capital financing (in U.S. dollar amounts per capita) in a given area. From the U.S. Census County Business Patterns, we obtain data on the industry composition of the full set of establishments in the MSAs. This information is useful to construct the proportion of establishments (and the related proportion of the labor force) engaged in high-tech activities. Like Decker et al. (2018) and Foster et al. (2018), we use the level 1 (i.e., narrower) definition of high tech, which includes, among other activities, internet publishing and broadcasting, semiconductor and other electronic component manufacturing, and scientific research and development services (for a full list, see table 4 in Hecker 2005).

Turning to possible channels through which inequality can affect entrepreneurial dynamism, we collect schooling information on the percentage and total amount of school financing coming from local sources, as well as the total amount of school expenditures per pupil from the Annual Survey of School System Finances set up by the Census Bureau. We obtain judges' sentencing data from 2005 from the Bureau of Justice Statistics (BJS) for the 75

[26] As these regressions rely on a small sample, we present them in Table A.9 in the appendix.

[27] The BDS describes an *establishment* as a fixed physical location where economic activity occurs. *Firms* are defined at the enterprise level such that all establishments under the operational control of the enterprise are considered part of the firm (see https://www.census.gov/ces/dataproducts/bds/methodology.html#coverage). The BDS data exclude self-employed and domestic service workers who do not necessarily engage in entrepreneurial activities.

Table 1
Descriptive statistics for the empirical analysis of business formation, job creation, financing, and institutions

Variable name	Number of observations	Units	Mean	Median	Standard deviation	1st percentile	99th percentile
Dependent variables							
Total establishments entry per capita	4,026		2.15	2.00	0.83	1.00	4.59
Total establishments exit per capita	4,026		1.95	1.84	0.67	1.01	3.96
Total establishments exit per capita (Churning)	3,285		0.70	0.63	0.32	0.28	1.77
Total establishments exit per capita (Shumpeterian)	3,285		1.31	1.26	0.42	0.68	2.44
Proportion of new establishments created by existing firms	3,285		0.30	0.30	0.06	0.17	0.46
Proportion of new establishments created by newly founded firms	3,276		0.65	0.65	0.06	0.50	0.78
Venture capital investment divided by MSA population	4,026		42.22	0.00	212.07	0.00	836.49
Proportion of high-tech establishments	3,960		0.03	0.03	0.02	0.01	0.09
Proportion of employment in high-tech	3,960		0.03	0.02	0.03	0.00	0.16
Income per capita growth	3,540	%	0.03	0.04	0.04	-0.07	0.11
Entrepeneur's net worth	525	USD	781,815	127,000	7,523,007	0	5,275,000
Amount of campaign contribution (local)	80,801	USD	5,441	500	66,201	10	77,574
Amount of campaign contribution (state)	737,211	USD	12,504	500	362,178	10.00	139,389
Proportion of public school revenue coming from local sources	3,948		0.41	0.40	0.13	0.13	0.72
MSA school expenditure per pupil	3,948	USD	5707	5145	4934	3170	13796
Inflow of individuals with at least a college degree	2,196	%	54	53	8	36	76
Judicial efficiency	2,417		801	676	510	185	2774
Violent crimes per capita	2,483		51	46	27	11	142
Main independent variable							
Financial wealth Gini	3,276	–	0.87	0.86	0.04	0.78	0.96
Housing wealth Gini	3,276	–	0.77	0.77	0.05	0.59	0.86
Mortgaged housing wealth Gini	3,213	–	0.77	0.77	0.05	0.64	0.86
No mortgage housing wealth Gini	3,240	–	0.75	0.79	0.13	0.34	0.91
Land 1880 Gini	2,979	–	0.44	0.42	0.13	0.13	0.75
Proportion of financial wealth owned by the top 10 percent	3,204	-	0.26	0.22	0.15	0.05	1.00
Proportion of housing wealth owned by the top 10 percent	3,105	-	0.53	0.53	0.07	0.38	0.71
Other measures of inequality							
Earned income Gini	3,938		0.41	0.41	0.05	0.26	0.54
Instrumental variables							
Homesteading in the first 10 years of incorporation of the MSA area	4,026		0.05	0.00	0.12	0.00	0.63

The table provides the number of observations, mean, standard deviation, 1st percentile, the median (50th percentile), and the 99th percentile of all variables used in the empirical analysis. The definition of the variables can be found in Appendix Table A.1. For the sake of brevity, we do not include MSA characteristics separately. Their corresponding descriptives can be found in Table A.2.

Table 2
Main specifications explaining business formation

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			ln(MSA total establishment entries per capita)			*ln(MSA total establishment exits per capita)*	*ln(MSA total establishment exits per capita (Churning))*	*ln(MSA total establishment exits per capita (Schumpeterian))*
Dependent variable			*Benchmark*					
Financial wealth Gini	−2.395***	−2.722***	−2.625***	−2.473***	−2.528***	−2.261***	−2.200***	−2.290***
	(0.401)	(0.446)	(0.437)	(0.462)	(0.538)	(0.454)	(0.534)	(0.469)
MSA percentage of inhabitants living in poverty			−0.002	−0.002	−0.000	−0.003	−0.004	−0.003
			(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Whites-to-total-population ratio			0.002***	0.004***	0.003**	0.002	0.003*	0.002
			(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Earned income Gini					0.734**			
					(0.311)			
State fixed effects	No	Yes	Yes	–	–	–	–	–
Year fixed effects	No	Yes	Yes	–	–	–	–	–
State * Year fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes
MSA controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	4,026	3,773	3,761	3,728	3,728	3,728	3,050	3,050
R-squared	0.075	0.723	0.724	0.741	0.764	0.658	0.714	0.588
Economic relevancy (Gini standard deviation)	−8.7%	−9.8%	−9.5%	−9.0%	−9.2%	−8.2%	−8.0%	−8.3%

The dependent variables are defined as the natural logarithm of the yearly total number of new establishments in the MSA between 2004 and 2014 divided by total MSA population; the natural logarithm of the yearly total number of establishments that became inactive in the MSA between 2004 and 2014 divided by total MSA population; the natural logarithm of the yearly total number of establishments that became inactive in the MSA within 36 months of opening between 2004 and 2012 divided by total MSA population; and the natural logarithm of the yearly total number of establishments that became inactive in the MSA after 36 months of opening between 2004 and 2012 divided by total MSA population. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. All other variables are defined in Appendix Table A.I. All Models are estimated with a linear regression (OLS) model. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic to Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, and MSA Population Growth and a dummy variable (Urban MSA) that takes a value of one if the MSA population density is above the sample median. “Yes” indicates that the set of fixed effects or controls is included. “No” indicates that the set of fixed effects or controls is not included. “–” indicates that the indicated set of fixed effects or controls are comprised in the wider included set of fixed effects. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. *$p<.1$; **$p<.05$; ***$p<.01$.

most-populous U.S. counties. What is particularly useful from this data source is the sentencing length, which we use as a proxy for judicial efficiency, also relating it to our inequality measures. We download the various state and MSA characteristics from the U.S. Census Bureau.

We retrieve data on homesteading intensity, which we use as instrumental variable for our contemporary inequality measures, from the Bureau of Land Management General Land Office Records (BLM GLO records). The records contain information on individual patents (name of the individual, plot size, and location, as well as year granted) granted by the Bureau of Land Management. We restrict our data set to those patents that fall under the authority of the original Homestead entries from 1862.

3.2 Results

3.2.1 Business dynamics. We begin by testing how local wealth inequality affects business dynamics in U.S. MSAs. Table 2 reports the first estimation results. We relate the local measure of inequality based on financial wealth, the Financial Wealth Gini, to the yearly number of new establishments per capita, as well as to the total number of establishments that become inactive in a given year.

In column 1, we start with a univariate ordinary least squares (OLS) estimation that relates establishment entry per capita to financial wealth inequality. In column 2, we repeat the estimation and include state and year fixed effects and control for the MSA population as well as a set of basic MSA characteristics that also includes the level of MSA wealth and MSA real gross domestic product (GDP).[28] The latter specification is our benchmark specification; it serves as the basis from which all-inclusive saturation, instrumentation, and alteration can be pursued. Among other things, it controls for the level of and change in local house prices to make sure that neither the housing boom of the first decade of the twenty-first century nor the subsequent crisis drives the results.[29]

In both regressions, we see a negative correlation between financial wealth inequality and the number of new establishments per capita, which is statistically significant at the 1% level. The economic significance is sizable: a one-standard-deviation increase in wealth inequality reduces the number of new establishments per capita by between 8.7% and 9.8%, corresponding to a reduction of about 19 establishments per 100,000 of MSA inhabitants. In Specification (3), we explore whether our basic result is driven by a "fat left tail," that is, that more unequal MSAs tend to have higher poverty rates or

[28] The basic MSA controls are MSA Financial and Housing Wealth, MSA Catholic-to-Protestant ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, and Urban MSA, a dummy variable that equals one if the population density of the MSA is above the sample average, and zero otherwise.

[29] Recall that the inclusion of year fixed effects controls for annual changes in the stock market index.

because more unequal MSA may be ethnically more diverse (Alesina, Baqir, and Easterly 1999). We control for these factors by including the percentage of the population in poverty and the percentage of white people living in the MSA. While we still find a negative (and statistically significant) coefficient for financial wealth inequality we find little effect of poverty and the percentage of white people living in the MSA on business formation. In Specification 4, we add state-year fixed effects.

In column 5, we control for MSA's income inequality. We add a new independent variable, Earned Income Gini, which is defined as the Gini coefficient of the distribution of income measured as the sum total of wages and total profit and losses coming from entrepreneurial activities as reported in the IRS/SOI data.[30] The result is very interesting. First, the estimated coefficient on the Financial Wealth Gini is, as before, negative and statistically significant. Second, the Earned Income Gini has a positive coefficient that is also economically relevant: a one-standard-deviation increase in the Income Gini increases establishment entry by 3.4%, corresponding to an increase of about 10 establishments per 100,000 MSA inhabitants.

In sum, wealth inequality slows entrepreneurship while income inequality spurs it. The latter, positive effect exerted by income inequality would be consistent with "a supply-side scenario" (Saez 2017), whereby local opportunities to earn more (as reflected in recent local income distribution) incentivize people to work arduously and to create more companies.

Column 6 looks at total establishment exits per capita. These variables' economic relevancy is practically similar: a one-standard-deviation increase in MSA wealth inequality decreases the number of establishments that become inactive by approximately 8%, indicating that business formation (i.e., establishment entries and exits) is less dynamic in more unequal metropolitan areas. Additionally, this result confirms that larger values of wealth inequality are related to lower business dynamics. In the spirit of Kerr and Nanda (2009), we also divide establishment exits between Churning exits (establishment closures within 36 months of formation) and Schumpeterian exits (closures beyond 36 months).[31] We present the results in columns 7 and 8. Remarkably, wealth inequality negatively affects both Churning and Schumpeterian exits. The economic significance is also similar: a one-standard-deviation increase in wealth inequality reduces both the Churning and Schumpeterian exits by approximately 8%.[32]

[30] To measure earned income, we sum the information about salary, stipends, and profit and losses (as reported in Schedules C and F) reported by the IRS/SOI.

[31] The U.S. census provides this information through 2012. This availability explains the lower number of observations in Columns 7 and 8.

[32] We consider the results with other measures of wealth inequality in Table A.5 in the appendix. Table A.6 presents the full set of results for wealth and income inequality.

Table 3
Main specifications explaining business formation: Instrumented

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	*ln(MSA total establishment entries per capita)*						
Sample/Specification					Excluding Dust Bowl states	Including illiteracy	Excluding no homesteading states
Financial wealth Gini	−4.315**	−3.555**	−3.797***	−3.710***	−3.334***	−3.698***	−3.693**
	(1.638)	(1.439)	(1.296)	(1.280)	(1.194)	(1.340)	(1.492)
Railroad line in the MSA in 1860		−0.046	−0.046	−0.042	−0.034	−0.042	−0.061
		(0.030)	(0.031)	(0.031)	(0.031)	(0.031)	(0.054)
Water canal in the MSA in 1860		−0.033	−0.033	−0.036	−0.038	−0.036*	−0.074**
		(0.023)	(0.022)	(0.021)	(0.023)	(0.021)	(0.030)
Native American tribes in the MSA in 1860		0.020	0.019	0.020	0.017	0.020	0.027
		(0.019)	(0.020)	(0.019)	(0.023)	(0.019)	(0.035)
Historical farm values			−0.037	−0.028	−0.028	−0.029	0.022
			(0.041)	(0.040)	(0.040)	(0.039)	(0.054)
Illiteracy						−0.013	−0.346
						(0.235)	(0.264)
State fixed effects	Yes	Yes	Yes	–	–	–	–
Year Fixed Effects	Yes	Yes	Yes	–	–	–	–
State × Year fixed effects	No	No	No	Yes	Yes	Yes	Yes
MSA controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV coefficient first stage	−0.059***	−0.063***	−0.066***	−0.065***	−0.069***	−0.062***	−0.060***
F-statistic first stage	11.611	12.577	14.707	15.140	16.566	14.932	14.582
Number of observations	3,773	3,773	3,740	3,707	3,267	3,707	1,936
R-squared	0.709	0.721	0.719	0.736	0.737	0.736	0.726
Economic relevancy (Gini standard deviation)	−15.1%	−12.6%	−13.4%	−13.1%	−11.6%	−13.1%	−12.1%

The dependent variable is defined as the natural logarithm of the yearly total number of new establishments in the MSA between 2004 and 2014 divided by total MSA population. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. All other variables are defined in Appendix Table A.1. All Models are estimated with a 2SLS estimator. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic-to-Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, and MSA Population Growth and a dummy variable (Urban MSA) that takes the value of one if the MSA population density is above the sample median. "Yes" indicates that the set of fixed effects or controls is included. "No" indicates that the set of fixed effects or controls is not included. "–" indicates that the indicated set of fixed effects or controls are comprised in the wider included set of fixed effects. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. *$p<.1$; **$p<.05$; ***$p<.01$.

3.2.2 Instrumental variables analysis. In this section, we tackle the endogeneity issue by performing an IV analysis. We use the proportion of land distributed via the Homestead Act in the first 10 years following 1862 if the MSA was already settled, and the first 10 years since the first settlement for MSAs that were settled later. Table 3 presents the regression with the Financial Wealth Gini, whereas we report the results on the Housing Wealth Gini in Table A.7 and Land Gini 1880 in Table A.8 in the appendix.

Table 3 presents both the coefficient of the instrumental variables and the F-statistic of the first stage. In every specification, the coefficient of Proportion of Homesteaded Land is negative and statistically significant. This means that a higher proportion of land homesteaded in the early years of the Act (or the early years of the settlement of the area) is negatively related to the contemporary measure of wealth inequality. This is in line with our conjecture that more homesteading may have promoted a more equal distribution of resources. The F-statistics are always above 10, which indicates that the proportion of land assigned via homesteading yields a powerful first stage. This result is also confirmed by the partial R-squared of the excluded instrument in the first stage regression: 5%, a large value when we compare it to the values reported by Jiang (2017).

Table 3, column 1, presents the baseline specification that corresponds to the controls we use in our specification in Table 2, column 2. We observe that the coefficient of the Financial Wealth Gini is negative and statistically significant at the 1% level confirming the OLS results. The economic significance is a bit larger than in the OLS regressions: the number of new establishments per capita declines with about 15% when we increase wealth inequality by one standard deviation. This corresponds to a decline of about 32 establishments per 100,000 MSA inhabitants.

Column 2 adds three controls to the specification in column 1. The amount of land available for homesteading depended on whether there was a railroad or a canal in the area. When there was, railroad and canal land would have been given by the federal government to the respective companies managing the lines and the waterworks (Gates 1936). We control for two dummy variables that indicate whether a canal was present in or a railroad passing through the area in 1860. Also, the amount of land to be homesteaded may have depended on the presence of Native American tribes in the area, as it has been reported that the homesteading policy was also used to mobilize settlers against Native Americans (Allen 1991). We control for a dummy variable that equals one if there were recorded Native American tribes in the area in 1860 and zero otherwise. Again, the coefficient on Financial Wealth Inequality is negative and statistically significant at the 1% level.

The specifications in Columns 3, 4, and 5 test whether homesteading may directly influence entrepreneurship today by reflecting or affecting the value of farmland. In column 3, we add the average value of the land in the MSA between 1870 and 1950 as an additional control. In column 4, we consider a denser set of fixed effects, and control for state-year fixed effects. In column 5, we repeat the specification of column 4, but we exclude the states that were affected by the Dust Bowl phenomenon of the 1930s. Also, in these cases, the coefficient on Financial Wealth Inequality remains negative and statistically significant at the 1% level and maintains a similar economic significance compared to the previous specifications.[33]

[33] More generally, initial conditions related to the quality of arable land and the availability of natural resources may explain why the two-stage least squares point estimates are larger than the OLS estimates. Hughes and Cain (2010, pp. 31–35) describe the regions that in our sample display high levels of wealth inequality as areas with particularly fertile land and a good climate for agriculture during the colonial period. Good land quality could set unequal regions on a good path to economic development from the outset. To the extent that the OLS regressions do not control for them, favorable initial conditions would be jointly correlated with more inequality and more business formation, generating a type of endogeneity that would work against us finding a result. As

The specifications in Columns 6 and 7 tackle concerns related to the quality of human capital. Column 6 reconsiders the full sample and adds an extra control: the proportion of individuals in the MSA in 1860 that could not write. Column 7 excludes the states where no land was assigned via the Homestead Act to control whether there was anything special about people migrating to the West that may affect our results.[34]

We also verify our identification assumptions in the data. The quasi-random allocation of land originated by the Homestead movement would imply no relationship between the future value of farmland and the proportion of land allotted via the Homestead Act. We ascertain whether this is true by relating the value of farmland in various years between 1870 and 1959 to our instrument, MSA homesteading intensity. Table A.10 in the appendix presents the results. In none of the specifications do we find a statistically significant relationship between the value of the land and the instrument. The sign of the relationships also changes depending on the type of controls we use.

Last, but not least, we control for possible spatial correlation of the residuals, which may affect the statistical inference in a long-run analysis similar to that which we undertake in the IV regressions (see Kelly 2019). We present these results in the appendix, where we allow for spatial correlation using a variety of distance cutoffs (see Table A.11 and Section A2 of the appendix). In every specification, the F-statistic is always above 10 and the results of the second stage are basically unaltered.[35]

3.2.3 Type of establishment and income growth. Table 4 studies in greater depth the characteristics of new establishment creation and relates them to local wealth inequality. A recent stream of literature has looked at the quality and type of entrepreneurs as an important source of economic growth. In principle, businesses dedicated to the development of new technologies and product/service innovation should generate higher levels of economic growth (La Porta and Shleifer 2008, 2014; Levine and Rubinstein 2016, 2018).

Columns 1 and 2 look at the proportion of high-tech establishments operating in the MSA and find that in more unequal MSAs the proportion of high-tech establishments is lower. A one-standard-deviation increase in wealth inequality leads to a reduction of about 30% in the proportion of high-tech establishments

our instrument does not appear to be correlated with initial soil conditions and future farm values, it alleviates this concern and generates estimates that are in absolute value larger than the OLS coefficients.

34 Table A.10 in the appendix presents the regressions on the Georgia land lottery, a setting that enables us to control for the possible differences in the quality of human capital of the settlers. The results confirm our previous findings and indicate a negative relationship between wealth inequality and new establishments per capita.

35 In an additional test, we also field past local weather conditions to form an alternative set of six instruments (Engerman and Sokoloff 1997, 2002). In particular, we use the average and standard deviation for 1895–2003 of both precipitation and temperature in the MSA, and the soil salinity and depth in the MSA. Estimates for these IV regressions can be found in Table A.12 in the appendix and confirm our main findings so far. In unreported analyses, we also rerun all of the above specifications but restrict our sample to end in 2006. That the results remain largely unaffected indicates that the financial crisis does not drive our findings.

Table 4
Inequality and other outcomes

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent variable	*Proportion of high-tech establishments*		*Proportion of employment in high-tech*		*Venture capital investment divided by MSA population*		*Proportion of new establishments created by existing firms*		*Proportion of new establishments created by newly founded firms*		*Income per capita frowth*	
Model	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Financial wealth Gini	−0.269***	−0.251**	−0.307***	−0.566**	−10.428***	−15.021*	0.168***	0.668***	−0.089*	−0.660**	−0.009	−0.082*
	(0.029)	(0.123)	(0.069)	(0.249)	(3.478)	(8.653)	(0.058)	(0.242)	(0.052)	(0.272)	(0.015)	(0.047)
State × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic first stage	–	15.171	–	15.171	–	15.14	–	15.374	–	15.374	–	15.157
Number of observations	3,734	3,701	3,734	3,701	3,740	3,707	3,060	3,033	3,060	3,033	3,400	3,370
R-squared	0.714	0.715	0.439	0.414	0.573	0.574	0.678	0.650	0.650	0.612	0.690	0.694
Economic relevancy (Gini standard deviation)	−32.7%	−30.6%	−40.0%	−40.0%	−39.6%	−57.0%	2.1%	8.4%	−0.5%	−3.9%	−1.2%	−9.9%

The dependent variables are defined as: the number of high-tech establishments in the MSA divided by the total number of establishments; the proportion of employment in high-tech firms in the MSA; the total amount of venture capital invested in the MSA between 2004 and 2014 divided by total population in the MSA; the number of new establishments created in the MSA by firms that have been active for at least 10 years divided by the total number of new establishments; the number of new establishments created in the MSA by newly created firms divided by the total number of new establishments and the percent change of MSA total income per capita. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. All other variables are defined in Appendix Table A.1. Uneven numbered models are estimated with a linear regression (OLS) model while even numbered models are estimated with a 2SLS IV model. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic-to-Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of 1 if the MSA population density is above the sample median. The IV estimates also include Railroad Line in the MSA in 1860, Water Canal in the MSA in 1860, Native American Tribes in the MSA in 1860, and Historical Farm Value as additional controls. “Yes” indicates that the set of fixed effects or controls is included. “No” indicates that the set of fixed effects or controls is not included. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the State level. *$p<.1$; **$p<.05$; ***$p<.01$.

in the area. Columns 3 and 4 study the proportion of employment in high-tech sectors and again find that wealth inequality is negatively associated with the share of labor employed in high-tech sectors. Taken together these results suggests that more equal areas are more engaged in high tech and product/service innovation.

We further delve into this issue in Columns 5 and 6, where we field as a dependent variable the Amount of Venture Capital Investment per Capita. Venture capital investment can be considered another proxy for engagement in high tech, as it is usually employed in the financing of ventures involving the developments of new products, services, and technologies. In both specifications, we find a negative coefficient that is also economically large. A one-standard-deviation increase in wealth inequality leads to a decline of between 39% and 57% in venture capital investment per capita.

Columns 7 and 8 look at the type of establishments created and in particular at the proportion of establishments, in the MSA, founded by already existing firms.[36] Existing firms may have more financial capacity to support new businesses and more know-how with which to overcome institutional constraints just because they build up experience and social networks. Our results reveal a higher proportion of new establishment creation by firms already in existence in more unequal areas. A one-standard-deviation increase in wealth inequality leads to an increase of between 2.1% and 8.4% in the proportion of new establishments set up by existing firms. Mirroring these results, Columns 9 and 10 show a lower proportion of establishments created by newly founded firms in more unequal MSAs.

Columns 11 and 12 assess whether more unequal areas also display a different pattern of income per capita growth. In the IV regression, we find that unequal MSAs have lower growth of income per capita. A one-standard-deviation increase in the Financial Wealth Gini reduces the growth rate by about 10%. This result suggests that lower business formation levels also translate into lower income-per-capita growth rates.

4. An Investigation of the Channels: Wealth Inequality and Local Public Goods

Our analysis so far has identified a reduced-form relationship between wealth inequality and business formation. The mechanism underlying our results is important to understand. As we outlined in Section 1, a number of explanations can account for the negative relationship between wealth inequality and the quality and quantity of entrepreneurship in an area. We focus in particular on the channels based on redistribution and public good provision.[37] These

36 We define a firm as “existing” if it has been in operation for at least 5 years, according to the U.S. census data.

37 See Tables A.13–A.16 in the appendix for evidence on the credit channel.

explanations clearly predict the number of redistributive policies that a society intends to undertake. However, the sign remains unclear: On the one hand, Alesina and Rodrik (1994) and Persson and Tabellini (1994) state that more wealth inequality increases the probability of future redistribution and income tax rises. On the other hand, Alesina, Baqir, and Easterly (1999) and Glomm and Lagunoff (1999) suggest that more inequality comes with less redistribution and a lower provision of public goods.

4.1 Local campaign contributions

We first tackle this issue by looking at how societies express their preferences for redistributive policies and in particular we study campaign contributions between 2004 and 2014 in local and state elections as a way to understand the attitudes of a community toward such policies. If more inequality leads to fewer redistributive policies, we may expect candidates with platforms involving lower taxes and less provision of public goods to receive more/larger campaign contributions in more unequal MSAs.[38] We obtain political campaign contributions data from Bonica (2014). The data covers campaign contributions in the context of federal, state, and local elections. As we analyze the effects of local wealth inequality on local outcomes, we consider only contributions involving state and local elections. The recipients of contributions in the data are individuals (or related committees) that ran for election (both primary and main elections) but did not necessarily win.

The data includes the total amount in contributions received by each candidate and a measure of her or his ideology that, among other elements, measures her or his preference for redistribution. For state elections we consider individuals running for governor or for state congress. For local elections we consider any election for the governing councils of local areas, such as city mayor, city council, county commissioners, supervisors, and executives, as well as school and academic boards.

The ideology score encompasses many issues relevant in politics, including the environment, moral concerns, and social policies.[39] Politicians with high ideology scores tend to prefer lower taxes and less provision of public goods, while politicians with lower scores have stronger preferences for higher taxes and greater provision of public goods.

[38] An obvious alternative is to relate the number of votes received by candidates and to their political platforms on redistribution and local wealth inequality. Unfortunately, detailed voting data for state and local elections are not available. Several studies, however, document a positive relationship between the amount in contributions received and the probability of victory (Hogan 2004; Holbrook and Weinschenk 2014).

[39] We construct the ideology score by observing the pattern of donations of each contributor following the methodology pioneered by Poole and Rosenthal (1991) (for details, see Bonica 2014). The advantage of the data provided by Professor Bonica is that, unlike any other data set on political ideology, these data also cover state and local elections.

We estimate the following equation:

$$\ln\left(contribution_{i,j,t}\right)=\alpha+\alpha_i+\alpha_j\times\alpha_t+\beta Wealth\ Inequality_j \times High\ Ideology\ Score_{i,t}+\varepsilon_{i,j,t}, \tag{2}$$

where $contribution_{i,j,t}$ denotes the total amount in campaign contributions received by candidate i in MSA j in year t, α_i is candidate fixed effects, and $\alpha_j \times \alpha_t$ is MSA×year fixed effects.[40] The last of these is an important set of fixed effects as they control for any time-varying omitted variable at the MSA level. $Wealth\ Inequality_j \times High\ Ideology\ Score_{i,t}$ identifies the interaction term between local wealth inequality and the ideology score.[41] Wealth inequality is defined as the Financial Wealth Gini. High ideology score is a dummy variable that equals one if the ideology score is above the sample median and zero otherwise.[42] A positive (negative) β would suggest that candidates with high ideology scores receive larger (smaller) contributions in more unequal MSAs.

Table 5 presents the estimates. In columns 1 and 2, we show the results for local elections. Both in the OLS and the IV specification, the coefficient on the interaction term is positive and statistically significant indicating that, in more unequal MSAs, candidates with an above-median ideology score receive larger campaign contributions.[43] The economic effect is large: a one-standard-deviation increase in wealth inequality increases the total of contributions by 16%–33%. This result suggests that in more unequal areas, candidates proposing less redistribution received more/larger campaign contributions.

In a second step, we study whether this finding is driven by contributors located in either rich or poor neighborhoods of the MSA. As we outlined in Section 1, if more inequality leads to less redistributive policies, we may expect the wealthy to be especially unwilling to share their resources. We carry out this study by reestimating Equation (4), splitting the sample between contributions made in the poor (i.e., below-median wealth) and rich (i.e., above-median wealth) ZIP codes in the MSA.[44] Column 3 presents the results for

[40] Relevant for the analysis, MSA-year fixed effects also control for differences in political campaign regulations in different locations.

[41] For ease of exposition, we omit the variable $High\ Ideology\ Score_{i,t}$ in Equation (4). We always control for this variable, although it is very persistent throughout time and hence almost perfectly collinear with candidate fixed effects.

[42] We construct this dummy variable to simplify the interpretation of the results by providing a straightforward estimate of its economic relevancy. Results would be the same if we used the original ideology score.

[43] We follow the procedure used by Tsoutsoura (2015) and instrument *Wealth Inequality x High Ideology Score* with *Proportion of Land Homesteaded in the MSA x High Ideology Score*. In the political campaign regressions in Table 9, we always include MSA×year fixed effects. As a result, in the first stage of the IV, the *Proportion of Land Homesteaded in the MSA* alone is fully absorbed by the fixed effects: the F-statistics in Table 9 refer to *Proportion of Land Homesteaded in the MSA x High Ideology Score*.

[44] Splitting the sample between contributions received in below- or above-median ZIP codes does not guarantee that (1) the number of observations in each subsample will be the same and (2) the sum of the observations

Table 5
Inequality and campaign contributions

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	*Campaign contributions*					
	For local elections				*For state elections*	
Sample	Full		Below-median wealth	Above-median wealth	Full	
Model	OLS	IV	OLS	OLS	OLS	IV
Financial wealth Gini * High score politician	4.723***	9.587***	1.770	4.231***	2.738***	4.974***
	(1.004)	(2.355)	(3.247)	(1.121)	(0.513)	(1.648)
CBSA × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Office × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Politician fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic first stage	–	14.709	–	–	–	11.156
Number of Observations	80,801	80,801	19,648	66,514	737,211	737,211
R-squared	0.401	0.399	0.651	0.431	0.394	0.394
Economic relevancy (Gini standard deviation)	16.1%	33.0%	0.6%	14.4%	9.9%	18.0%

The dependent variables are defined as the natural logarithm of the $ amount of campaign contributions received by election candidates for local and state elections. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. A High Score Politician has stronger preferences for lower taxes and a lower provision of public goods and is defined based on the scores provided by Bonica (2014). All other variables are defined in Appendix Table A.1. Model (1) is estimated with a linear regression (OLS) model while Model (2) is estimated with a 2SLS IV model. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic to Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, and MSA Population Growth and a dummy variable (Urban MSA) that takes the value of one if the MSA population density is above the sample median. The IV estimates also include Railroad Line in the MSA in 1860, Water Canal in the MSA in 1860, Native American Tribes in the MSA in 1860, and Historical Farm Value as additional controls. "Yes" indicates that the set of fixed effects is included. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the State level. *$p<.1$; **$p<.05$; ***$p<.01$.

ZIP codes below-median wealth. The interaction term between the ideology score and wealth inequality is positive and not statistically significant. The size of the coefficient is also small: local wealth inequality has virtually no impact on the total amount in contributions made in poor neighborhoods. Column 4 considers above-median wealth ZIP codes. The interaction term between ideology score and MSA wealth inequality is positive and statistically significant. This indicates that wealth inequality leads to larger contributions to candidates proposing less redistribution, especially in wealthy neighborhoods. The economic importance is large: a one-standard-deviation increase in MSA wealth inequality leads to an increase of about 15% of campaign contributions to candidates with high ideology scores.

Columns 5 and 6 check whether inequality leads to candidates with high ideology scores receiving larger campaign contributions also in state elections.[45] In line with the previous findings, also these results suggest that wealth inequality is positively related to preferences for candidates who propose less redistributive policies. A one-standard-deviation increase in wealth inequality increases campaign contributions to candidates with an above-median ideology score by about 10%–18%.[46]

4.2 Provision of public goods

We further explore the idea that wealth inequality leads to less redistribution in Table 6, where we relate local wealth inequality to various measures of local public goods provision. Public goods can be seen as a form of redistribution from the wealthier to the poorer: if wealth inequality is associated with less redistribution, in more unequal areas the quality and quantity of public goods supplied should be lower. In particular, we focus on education and on the efficiency of the civil justice system, both local public goods that have been shown to support entrepreneurship and be important for economic growth.[47]

---

in the two subsamples will be equal to the number of observations in the full sample. This happens because local elections include elections in geographical units that can be either “large” (like counties) or “small” (like cities and school districts). Election candidates in large administrative units can receive contributions from both above- and below-median wealth ZIP codes and will appear in both subsamples, thus explaining (2). Candidates in administrative units that are “small” receive contributions only in a subsection of ZIP codes within the MSA, possibly only the very poor or the very rich, thus explaining (1).

[45] We use state elections to validate the results on local elections as states are often coresponsible with the local authorities for organizing local public life.

[46] A possible concern is that the data on contributions do not fully reveal the total amount of financial support received by a candidate. Since 2010, “Super PACs,” a particular form of political action committee, have been allowed to financially support political candidates without making direct contributions to their campaigns. Our data do not include Super PACs’ contributions and may not reveal the full pattern of financial support received by candidates in each MSA. Note, however, that (1) Super PACs tend to be more influential in federal rather than local elections; (2) we rerun our regression considering only pre-2010 election cycles, when Super PACs did not exist yet, and the results still hold.

[47] The quality of schooling has been shown to positively affect standard school testing outcomes (Fryer and Levitt 2004). Levine and Rubinstein (2016) shows that individuals with better school testing outcomes are more likely to become entrepreneurs, that is, individuals engaged in growth-enhancing entrepreneurial activities.

Table 6
Inequality and local public goods

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Proportion of public school revenue coming from local sources		*MSA school expenditure per pupil*		*% of individuals with at least some years of college*		*Judicial efficiency*		*Violent crimes per capita*	
Dependent variable	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Financial wealth Gini	−1.627***	−2.003***	−0.977**	−1.697	−86.706***	−74.554*	5.859***	5.625*	16.950***	40.936**
	(0.222)	(0.721)	(0.399)	(1.550)	(13.107)	(41.437)	(1.869)	(3.123)	(4.167)	(18.659)
State × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic first stage	–	15.287	–	15.287	–	15.534	–	11.002	–	14.918
Number of observations	3,707	3,674	3,707	3,674	2,040	2,022	2,417	2,417	2,329	2,308
R-squared	0.792	0.798	0.727	0.725	0.756	0.758	0.309	0.309	0.593	0.562
Economic relevancy (Gini standard deviation)	−15.0%	−18.8%	−4%	−6.4%	−6.2%	−5.8%	22.2%	21.3%	12.0%	32.0%

The dependent variables are defined as: the ratio of yearly public school revenues obtained from local sources in the MSA divided by total public school revenues between 2004 and 2014; the total amount of yearly public school expenditures per pupil in the MSA between 2004 and 2014; the proportion of individuals who had at least some years of college in the MSA between 2009 and 2014. Judicial efficiency is the length of time needed to obtain the verdict for a first-degree civil trial. Violent crimes per capita is the sum of murders, rapes, robberies and assaults divided by the population in the MSA (in 100,000). The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. All other variables are defined in Appendix Table A.1. Models (2), (4), (6), and (8) are estimated with a 2SLS estimator; all other Models are estimated using OLS. The definition of the variables can be found in Appendix Table A.1. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic to Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), and Change in MSA Age Composition and a dummy variable (Urban MSA) that takes the value of 1 if the MSA population density is above the sample median. The IV estimates in Columns (2), (4), (6), and (8) also include Railroad Line in the MSA in 1860, Water Canal in the MSA in 1860, Native American Tribes in the MSA in 1860, and Historical Farm Value as additional controls. "Yes" indicates that the set of fixed effects is included. "–" indicates that the indicated set of characteristics or fixed effects are comprised in the wider included set of fixed effects. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

First, we relate wealth inequality to education and schooling, presenting the results in columns 1 to 6. The analysis shows that, in more unequal MSAs, both the proportion of public school revenue coming from local sources and the school expenditure per pupil are lower. A one-standard-deviation increase in inequality decreases the proportion by 15%–20% and the expenditure per pupil by 4%–7%. Moreover, the proportion of individuals with some years of college is also lower in more unequal MSAs (see Columns 5 and 6). A one-standard-deviation increase in inequality results in an approximately 6% lower proportion of educated individuals.

Finally, from Columns 7 to 10, we assess the effect of local wealth inequality on the judiciary. We investigate how local inequality affects judicial efficiency as measured by the length of time to verdict for a first-degree civil trial. We obtain data on individual civil cases from the Bureau of Justice Statistics (BJS) for 2005. The BJS reports data on civil litigations for the 75 most-populous U.S. counties, which we aggregate at the MSA level. At the individual case level, we observe whether the number of days (its natural logarithm) it takes to come to a verdict in a case is affected by local inequality.[48] We observe that local inequality matters for judicial decision-making. When controlling for both state fixed effects and case controls (in the form of a set of dummy variables that controls for the nature of the case, such as breaching of a contract, intentional tort, or partnership dispute), the findings suggest that, in more unequal MSAs, court rulings take more time (about 22% more in our usual standard deviation assessment) and therefore are less efficient.[49] In the same spirit, we use violent crimes per capita as a rough proxy of the efficiency of the judicial and police system of the MSA and see whether more unequal MSAs also have a higher number of violent crimes per capita. We present these results in Columns 9 and 10 and find that this is indeed the case. A one-standard-deviation increase in inequality raises violent crimes per capita by about 32% in the IV specification.

4.3 Horse race analysis

Although a vast literature exists on the effects of education and the efficiency of the judicial system on economic growth and business formation, in the last piece of our analysis, Table 7 investigates whether the impact that public goods have on local entrepreneurship is related to wealth inequality. We take the version of our equation from Table 2, column 2, controlling for state $\times$ year fixed effects, and we include the schooling and violent crimes measures together with wealth inequality. If the mediated effect is important, we would expect the coefficient relating wealth inequality to business formation to become smaller once we control for local public goods.

48 We compute the number of days starting from the day when the case is filed until the day when a verdict is pronounced.

49 Together with the dummies indicating the typology of the case and our standard sets of controls, the judicial efficiency regressions also include the number of plaintiffs and the number of defendants, whether the trial was a jury or a bench trial, and whether the case was resolved via arbitration, as additional controls.

Table 7
Horse race of inequality with local public goods

Model	(1)	(2)	(3)	(4)
	MSA total establishment entries per capita			
Financial wealth Gini	−2.453***	−1.562**	−2.376***	−2.205***
	(0.536)	(0.617)	(0.660)	(0.622)
Proportion of public school revenue coming from local sources		0.379*	0.310***	0.604***
		(0.199)	(0.073)	(0.085)
MSA school expenditure per pupil		−0.020	−0.011	−0.004
		(0.042)	(0.043)	(0.049)
% of individuals with at least some years of college		0.004*		
		(0.002)		
Violent crimes per capita				−0.010*
				(0.005)
State × Year fixed effects	Yes	Yes	Yes	Yes
MSA controls	Yes	Yes	Yes	Yes
Number of observations	1,700	1,700	2,329	2,329
R-squared	0.676	0.687	0.73	0.733
Economic relevancy (Gini standard deviation)	−9.3%	−5.9%	−9.0%	−8.4%

All models are estimated with an OLS model. The dependent variables are expressed in natural logarithms. The definition of the variables can be found in Appendix Table A.1. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic to Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), and Change in MSA Age Composition and a dummy variable (Urban MSA) that takes the value of one if the MSA population density is above the sample median. "Yes" indicates that the set of fixed effects is included. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. *$p<.1$; **$p<.05$; ***$p<.01$.

Column 1 presents the regression relating inequality to business formation without controlling for any public goods measure, but restricting it to the years in which every schooling variable is available. In column 2, we repeat the estimation of column 1 but include the schooling variables. We see that the coefficient on the Financial Wealth Gini declines by around 40% once we include schooling, a result in line with the notion that wealth inequality has a mediating effect on entrepreneurship via public goods, and especially schooling. Columns 3 and 4 repeat the same exercise with violent crimes. In this case the coefficient declines little, by about 8%, suggesting that schooling is a more relevant mediating factor for our results.

In the appendix, we also check the relative importance of wealth inequality on credit markets and compare it with the importance of the public goods channel. In the horse race analysis we present in Table A.15, we find that the credit channel and the public goods channel have roughly the same importance in explaining business formation. The coefficient on wealth inequality declines about 35% once we include credit market controls, a similar decline of what we find for local public goods.

Overall, we conclude that the results point in the direction that inequality also affects the quality of local institutions in the form of financial sector development, education, and judicial efficiency.

We also run a formal analysis of the causality chain using the methodology developed by Becker and Woessmann (2009). We estimate a simultaneous equation model where in the first stage we relate wealth inequality to the proportion of land assigned via the Homestead Act in the MSA. In a second stage, we regress the quality of local institutions on wealth inequality. In the third stage, we relate institutional quality to entrepreneurship. See Appendix B for details and Table A.16 for the results.

5. Conclusions

We empirically test hypotheses showing how household wealth inequality may determine entrepreneurial dynamism. The relationship between wealth inequality and entrepreneurship is in theory ambiguous. On the one hand, local wealth inequality may be associated with low levels of entrepreneurship. On the other hand, inequality may help entrepreneurship by supporting local demand and generating more savings and investment.

In our empirical analysis, we employ multiple measures of wealth inequality at the U.S. MSA level. To alleviate endogeneity problems, we saturate specifications with comprehensive sets of local fixed effects and characteristics and estimate instrumental variable models relying on the 1862 Homestead Act.

The estimated coefficients suggest that local-level wealth inequality robustly decreases firm creation and exit and is associated with less high-tech activity. The relationship between wealth inequality and entrepreneurial dynamism can be mediated by various factors. We find that more unequal communities prefer less redistributive policies and consequently have lower local expenditures on schooling and a lower proportion of educated people in the area. Our findings also suggest the presence of a less effective judiciary in more unequal MSAs. Additionally, we observe that higher inequality MSAs display lower personal income per capita growth.

Overall, our findings suggest that inequality in local household wealth impedes entrepreneurial dynamism through a deterioration in schooling and justice, and that to address rising inequality (while preserving entrepreneurial dynamism), policy actions focusing on wealth rather than income should be considered.

Appendix

A1. Measuring Wealth Inequality

Readily obtaining representative measures of wealth inequality at the local level is difficult to do. Thus, we construct our own two proxies for local wealth inequality. The first, which we label *Financial Wealth Gini*, is based on current levels of financial wealth and broadly based on a methodology introduced by Mian, Rao, and Sufi (2013) and Saez and Zucman (2016), and it intends to construct local-level measures of household net worth; the second measure is based on U.S. house ownership data.

The contemporary measure of wealth inequality looks at amounts in dividends and interest earned by U.S. households in 2004, the first year of our sample period, as reported in Internal Revenue Service (IRS) Statistics of Income (SOI) data. The IRS/SOI data report the total amount of income in dividends and interest received by U.S. households in a given ZIP code. The information is reported as a total per ZIP code and is divided into five household income groups, ranging from low income to high income. Under the assumption that a typical household owns the market index for stocks and bonds, the amount in financial rents it receives depends on the quantity of stocks and bonds it holds. IRS/SOI provides three pieces of information important for the construction of our proxy:

1. The total number of households belonging to each income group;
2. For each income group, the number of households who declared nonzero dividend and nonzero interest income (we will call these nonzero households); and,
3. For each income group, the total amount in dividends and interest earned by all households.

We now report the procedure we adopted to construct our inequality proxy. For simplicity, we just describe the case where we consider only dividends as a financial rent. The procedure, which comprises six steps, is exactly the same when we also include interest income.

1. We aggregate the IRS/SOI figures at the MSA level.
2. For each MSA, we compute the number of households who declared zero dividend income and we place them into a separate category.
3. For each MSA and each income group, we compute the average dividend earned by nonzero households. We do this by dividing the total amount in dividends for each income group by the respective number of nonzero households.
4. We assume that each household in the same income group earned the average dividend computed in (2).
5. We assume that each household owns the same type and composition of stock: the equity index. As a result, the amount in dividends received depends on the quantity of stock owned.
6. We use the number of nonzero households belonging to each income group, the number of households declaring zero dividends, and the average earned in interest and dividends to compute a Gini coefficient that measures the distribution of dividend earnings within each MSA. Recall that the Gini coefficient is a measure of inequality that ranges between zero and one, where a coefficient close to zero can be interpreted as full equality, whereas a coefficient of one indicates perfect inequality. We perform the same procedure with the interest income data.[50]

Table A.3 provides an example of this computation. Column 1 lists the five income groups; column 2 provides the number of households belonging to each income group; column 3 the number of households declaring a nonzero dividend; and column 4 the sum of all dividends received by all households in each group.

First, we compute the number of zero-dividend earners by taking the difference between the totals of columns 1 and 2 and reporting it in column 4 in the row “Households with No Dividend Income.” In the remainder of column 4, we place, per income group, the number of households that declared dividends—the figures being the same as those in column 2. We then compute the average dividend earned by nonzero households by dividing the figures in column 3 by those in column 4; we report this ratio in column 5. We then compute the Gini coefficient, using the six dividend income groups. The first consists of the 1,576,927 households that earned zero dividends, the second contains the 31,604 households that earn \$1,181, and so on, up to the seventh group, which comprises 73,620 households that earn about \$11,800 in dividends. In this example, the Gini coefficient is equal to 0.91.

50 As we do not know the amount of income in dividends and interest each individual household declares, we cannot compute a unique Gini coefficient based on the sum of these amounts.

Naturally, this is a proxy, and it may be subject to measurement error. It performs well in identifying perfect equality and perfect inequality. In the former case, we would observe each household earning the same financial rents independently of the income group it belongs to, and our Gini coefficient would correctly have the value of zero. In the latter situation, our data would reveal all households but one receiving a financial rent and the Gini coefficient would correctly receive the score of one. The proxy does not work very well in every situation where in each income group the distribution of dividends is very dispersed around the mean. In these situations we underestimate the degree of inequality. Measurement error may produce biased estimates of the coefficients when relating wealth inequality to financial outcomes. We will be able to alleviate this problem by instrumenting this wealth inequality measure in various specifications.[51] We will present our main results using a Gini coefficient based on dividends, which we label Financial Wealth Gini. Results are basically the same if we use a Gini coefficient based on interest income.

The second group of contemporary measures of wealth inequality are based on homeownership data obtained from both the U.S. census and the Fannie Mae mortgage databases. From the U.S. census database, we obtain information on the number of owner-occupied housing units (i.e., households) that fully own their house as well as the corresponding value of the house. Housing values fall into one of the following categories: below \$50,000, between \$50,000 and \$100,000, between \$100,000 and \$150,000, between \$150,000 and 200,000, between \$200,000 and \$300,000, between \$300,000 and \$500,000, and more than \$500,000. Given that these properties are free of mortgages, the value of the houses corresponds to the household's housing equity. We use the lowest value of each bin to calculate a Gini index of housing inequality, which we label the No Mortgage Housing Wealth Gini. In the Gini index, we also include the number of households that rent a house, and we assume they have zero housing wealth.

Note that from the U.S. census database we only have homeowners that fully own their houses. To include homeowners who do have an outstanding mortgage, we rely on the Fannie Mae mortgage database ("Single Family Loan Level data set"), which provides us with information on mortgage originations between 1999 and 2004. The Fannie Mae data contains information on the MSA where the property is located and the loan-to-value ratio at origination, as well as the house value (again at origination). This information enables us to compute housing equity, defined as the value of the house in 2004 minus the outstanding mortgage. Based on these data, we calculate a Gini coefficient in a similar spirit as the Gini measures based on the housing information from the U.S. census ACS database. We compute our main housing Gini measure by aggregating the information on housing equity obtained from both the U.S. census and the Freddie Mac databases and label this variable the Housing Wealth Gini.

To construct our historical measure of wealth inequality, we obtain information on historical farmland plot sizes at the county level from the 1880 U.S. census. More precisely, for each county we have information on the total number of farms that—based on their total acres of farmland—fall into a certain size bin. Farms are assigned to one of the following seven bins: under 10 acres, from 10 to 19 acres, 20 to 49 acres, 50 to 99 acres, 100 to 499 acres, 500 to 999 acres, and 1,000 or more acres.

First, we assume that the lower bound farm size of each bin is the average farm size of all the farms in that bin (for the first bin we set the lower bound equal to 0.001). We then aggregate the information at the MSA level. Next, we use these lower bounds to calculate an MSA Gini coefficient in a similar way to that used in Rajan and Ramcharan (2011).

We also compute proxies of the amount of financial and housing wealth held by the top 10% of owners ranked according to financial or housing wealth, respectively, which we label the Proportion

[51] Another possible source of measurement error may come from tax evasion. U.S. financial institutions automatically report to the IRS dividend and interest income earned by their clients, making tax evasion through U.S. banks virtually impossible. But taxpayers can avoid taxes by holding wealth at foreign banks.

of Financial Wealth Owned by the Top 10% and the Proportion of Housing Wealth Owned by the Top 10%. The data are not granular enough to precisely compute the exact top 10%: in principle, we would like to have individual tax records, but our data are aggregate at the income or house value level. The problem is slightly more complicated with the IRS/SOI data as it splits households into bins based on their total income and not on income coming just from dividends. While it is likely that households with high incomes may also have larger dividend earnings, the information we have does not allow us to see that precisely.

Keeping these caveats in mind, we now present the procedure we adopt to construct our top 10% financial wealth measure. We use the IRS/SOI data as discussed above. For each income bin, we compute the average dividend received by the corresponding number of households, which we define as the total amount of dividends reported in a certain income bin divided by the total number of returns reported in the same bin. We then rank bins in respect to the average dividend. We start with the bin with the highest average dividend figure and compute the proportion of returns in this bin in respect to the total number of tax returns in the MSA. Once we have completed this computation, we face three possibilities:

1. If the proportion of returns is exactly equal to 10%, our procedure concludes here and we just compute the proportion of total dividends received by the households in the bin in respect to the total dividend amount reported in the entire MSA. This precisely corresponds to the amount in dividends/financial wealth held by the top 10% of owners when ranked according to their wealth.
2. If the proportion of returns is more than 10%, we compute the subset of tax returns within the bin that corresponds to exactly 10% of the number of returns. We then multiply it by the average dividend earned in this bin and divide it by the total amount in dividends received in the MSA. This again provides us with the amount in dividends/financial wealth held by the top 10% of owners.
3. If the proportion of returns is less than 10%, we move to the bin with the second highest average dividend, and we repeat the procedure again, until we reach the 10% of total tax returns delivered in the MSA.

The procedure we use to compute our top 10% housing wealth measure is very similar to what we just described; the only difference is that we rank each bin in relation to the corresponding housing value.

A.2 Some Evidence on the Direction of the Causality Chain: A Three-Stage Model

To test more formally the causality chain implied by our argument, and again following the methodology of Becker and Woessmann (2009), we write a system of equations intended to spell out each component of the causality argument. In particular, we model business formation, local institutional quality, and wealth inequality, as follows:

$$MSA\ Business\ Entry_{j,t} = \alpha + \alpha_s + \alpha_t + \beta \widehat{Institutions}_{j,t} + Controls_{j,t-1} + \varepsilon_{1,j,t}.$$

$$Public\ Goods_{j,t} = \alpha + \alpha_s + \alpha_t + \beta Wealth\ \widehat{Inequality}_j + Controls_{j,t-1} + \varepsilon_{2,j,t}.$$

$$Wealth\ Inequality_j = \alpha + \alpha_s + \alpha_t + \beta Homestead_j + Controls_{j,t-1} + \varepsilon_{3,j,t}.$$

In the system of equations, the first stage predicts MSA wealth inequality based on the proportion of land assigned via the Homestead Act. The part of the variation of wealth inequality related to the Homestead Act is then employed in the second stage to predict the level of local institutional quality. In the third stage, the fitted value of institutions is related to the MSA's number of new establishments per capita. This model is estimated using three-stage least squares and the results are reported in Table A.16. As in the previous analysis, we restrict the analysis to the measures of institutional quality for which we have access to data for a longer time series. We see that in the first stage, the proportion of land assigned via the Homestead Act is negatively associated with each measure of wealth inequality. In the second stage, with the exception of total expenditure, the fitted value of inequality is negatively associated with institutional quality. Lastly, in the third stage, institutional quality has a positive association with local business formation.[52]

52 In principle, one could imagine a more complex system of equations that relates wealth inequality to campaign contributions and campaign contributions to public goods. Such a system is, however, beyond the scope of our study.

Table A.1
Variable names, definitions, and data sources for the empirical analysis of business formation, type, financing, and public goods

Variable name	Variable definition	Source
Dependent variables		
Total establishments entry per capita	Total number of new establishments in the MSA between 2005 and 2014 divided by total MSA population expressed in thousands of inhabitants	USC
Total establishments exit per capita	Total number of establishments that became inactive in the MSA between 2005 and 2014 divided by total MSA population expressed in thousands of inhabitants	USC
Total establishments exit per capita (Churning)	Total number of establishments that became inactive in the MSA within 36 months of opening between 2005 and 2012 divided by total MSA population expressed in thousands of inhabitants	USC
Total establishments exit per capita (Shumpeterian)	Total number of establishments that became inactive in the MSA beyond 36 months of opening between 2005 and 2012 divided by total MSA population expressed in thousands if inhabitants	USC
Entrepreneurs‘ net worth	Net worth of of individuals who are about to start a business. Net Worth is defined by the total amount of assets belonging to an individual minus the total amount of her personal debts	PSED
Proportion of new establishments created by existing firms	Number of new establishments created in the MSA by firms that have been active for at least 10 years divided by the total number of new establishments	USC
Proportion of new establishments created by newly founded firms	Number of new establishments created in the MSA by firms that have been active for less than 10 years divided by the total number of new establishments	USC
Proportion of new establishments created by firms that are less than 36 months old	Number of new establishments created in the MSA by firms that have been active for less than 36 months divided by the total number of new establishments	USC
Venture capital investment divided by MSA population	Total Venture Capital Investment in the MSA divided by MSA population (in million $)	Thompson 1
Proportion of high-tech establishments	Number of new high-tech establishments created in the MSA divided by the total number of establishments in the MSA	USC
Proportion of employment in high-tech	The proportion of employment in high-tech firms in the MSA	USC
Income per capita growth	The percent change of MSA total income per capita	USC
Amount of campaign contribution (local)	Total Amount of Campaign contributions received by a candidate in a local election. Local elections include election in city councils, mayoral elections, counties, and school boards	Bonica 2014
Amount of campaign contribution (state)	Total Amount of Campaign contributions received by a candidate in a state election. State elections include gubernatorial and state council elections	Bonica 2014
Proportion of public school revenue coming from local sources	Total School Revenue coming from local sources divided by Total Schools Revenue	USC
MSA school expenditure per pupil	Total School Expenditure in the MSA divided by the number of pupils in the MSA	USC
% of individuals with at least some years of college	Percentage of Individuals with at least some years of college or more in the MSA	USC
Judicial efficiency	The length of time to the verdict for a first-degree civil trial	BJS
Violent crimes per capita	Total Number of Violent Crimes divided by population (in 1,000,000) of the MSA. Violent crimes are the sum of murders, rapes, robberies and assaults	FBI
Number of bank branches per capita	The number of yearly bank branches in the MSA divided by population	FDIC
Total amount of loans to small businesses per capita	The total amount of loans made to small businesses in US$ in the MSA between 2005 and 2014	FFIEC
Wealth inequality at the state level	The Gini coefficient of the distribution of wealth as measured by households' wealth in the state	PSID

(*Continued*)

Table A.1
(Continued)

Variable name	Variable definition	Source
Main independent variables		
Financial wealth Gini	The Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA	IRS
Housing wealth Gini	The Gini coefficient of the distribution of wealth measured as home equity of houses purchased between 1999 and 2004 that do have a mortgage and of houses that do not have a mortgage	FRED/USC
Mortgaged housing wealth Gini	The Gini coefficient of the distribution of wealth measured as home equity of houses purchased between 1999 and 2004 that do have a mortgage	USC
No mortgage housing wealth Gini	The Gini coefficient of the distribution of wealth measured as the home equity related to houses that do not have a mortgage	FRED
Land 1880 Gini	The Gini coefficient of the distribution of farmland in 1880 in the MSA	USC
Proportion of financial wealth owned by the top 10%	A proxy of the proportion of dividends held by the top 10% dividend earners based on the average dividends earned by each income class in the IRS tax returns	IRS
Proportion of housing wealth owned by the top 10%	The proportion of housing wealth owned by the top 10% house owners in the MSA	FRED/USC
Instrumental variables		
Homesteading in the first 10 years of incorporation of the MSA area	Amount of land distributed via Homestead Acts between 1862 and 1872, if the MSA was already settled at the enactment of the Act, or in the first 10 years since first settlement, divided by the total amount of land distributed in the MSA	
MSA average rainfall between 1895 and 2003	The average district precipitation between 1895 and 2003, where a district is defined as a group of clustered counties with similar climatic conditions	NCDC
MSA standard deviation of rainfall between 1895 and 2003	The standard deviation of district precipitation between 1895 and 2003, where a district is defined as a group of clustered counties with similar climatic conditions	NCDC
MSA standard deviation of temperature between 1895 and 2003	The standard deviation of district temperature in degrees between 1895 and 2003, where a district is defined as a group of clustered counties with similar climatic conditions	NCDC
MSA average temperature between 1895 and 2003	The average district temperature in degrees between 1895 and 2003, where a district is defined as a group of clustered counties with similar climatic conditions	NCDC
MSA soil salinity	The salt concentration of the soil solution in terms electric conductivity (ECe) in dS/m in the MSA	USDA
MSA soil depth	The distance from the top of the soil to the base of the soil horizon in centimeters in the MSA	USDA
Control variables		
MSA characteristics		BLM
Total MSA estimated financial wealth	Average dividend received by an individual in the MSA multiplied by the MSA population	IRS
Total MSA estimated housing wealth	The sum of the housing equity of the individuals living in the MSA	FRED
MSA real GDP per capita	The logarithm of one plus the real gross domestic product of the MSA during the year	BEA
MSA house price index (level)	The yearly level of the House Price Index in the MSA based on the movement of single-family house prices	FHFA
Urban MSA	= 1 if the population density of the MSA is above the sample median, = 0 otherwise	USC
MSA house price index (change)	The yearly change in the House Price Index in the MSA based on the movement of single-family house prices compared to the previous year	FHFA

(*Continued*)

Table A.1
(Continued)

Variable name	Variable definition	Source
Water canal in the MSA in 1860	= 1 if a water canal was present in the MSA in 1860, = 0 otherwise	USC
Native American tribes in the MSA in 1860	= 1 if there was at least one Native American tribe in the MSA in 1860, = 0 otherwise	USC
Railway line in the MSA in 1860	= 1 if a railway was present in the MSA in 1860, = 0 otherwise	USC
Historical average farm value in the MSA	Average value of farmland and buildings per acre	USC
Change in MSA age composition	Yearly % Change of the age composition in the MSA	USC
MSA population growth	Yearly % Change of the MSA population	USC
MSA land size	The logarithm of one plus the total MSA area in square miles at year-end 2000	USC
MSA percentage of inhabitants living in poverty	% of individuals in the MSA living below the povertry line	USC
White-to-total-population ratio	Ratio of the total MSA population of white race at divided by the MSA total population	USC
Earned income Gini	The Gini coefficient of the distribution of income measured as the sum total wages and total profit and losses coming from entrepreneurial activities as reported in the IRS-SOI data	IRS
MSA Catholic-to-Protestant ratio	Ratio of the total number of Catholics divided by the total number of Evangelicals in the MSA at year-end 2000	ARDA

The table defines the variables used in the analysis. For the sake of brevity, we do not report the MSA characteristics separately. Data sources include ARDA = Association of Religion Data Archives; BJS = Bureau for Justice Statistics; BEA = Bureau of Economic Analysis; FDIC = Federal Deposit Insurance Corporation; FFIEC = Federal Financial Institutions Examination Council; PSID = Panel Study of Income Dynamics; FHFA = Federal Housing Finance Agency; FRED = Freddie Mac; IRS = Internal Revenue Service; NCDC = National Climatic Data Center; PSED = Panel Study of Entrepreneurial Dynamics survey; USDA = United States Department of Agriculture; USC = US Census; FBI = Federal Bureau of Investigation; Bonica 2014 = Database on Ideology, Money in Politics, and Elections; BLM = Bureau of Land Management.

Table A.2
Descriptive statistics MSA controls for the empirical analysis of business formation, type, financing, and public goods

Variable name	Number of observations	Units	Mean	Median	Standard deviation	1st percentile	99th percentile
Soil instrumental variables							
MSA average rainfall between 1895 and 2003	3,982	Inch	3.16	3.30	1.11	0.72	5.25
MSA standard deviation of rainfall between 1895 and 2003	3,982	–	1.92	1.83	0.68	0.66	4.10
MSA standard deviation of temperature between 1895 and 2003	3,982	–	14.72	15.07	3.23	6.58	21.60
MSA average temperature between 1895 and 2003	3,982	°F	55.12	53.99	8.30	40.24	73.75
MSA soil salinity	3,177	ECe (dS/m)	0.53	0.01	1.04	0.00	4.86
MSA soil depth	3,186	cm	88.92	89.14	11.84	60.87	112.10
Control variables							
MSA characteristics							
Total MSA estimated financial wealth	3,186	USD million	505	41	2,751	6	8,116
Total MSA estimated housing wealth			1,104	186	3,531	19	16,582
MSA real GDP per capita	3,883	USD	41,708	39,786	12,039	20,886	78,708
MSA House Price index (level)	3,927	–	175.79	167.41	34.86	124.70	300.98
MSA House Price index (change)	3,927	–	0.02	0.01	0.07	−0.20	0.23
Water canal in the MSA in 1860	4,026	0/1	0.52	1	0.50	0	1
Native American tribes in the MSA in 1860	4,026	0/1	0.49	0	0.50	0	1
Railway line in the MSA in 1860	4,026	0/1	0.46	0	0.50	0	1
Historical farm values	3,982	USD	67.90	51.19	67.17	8.50	372.72
Change in MSA age composition	4,026	–	0.27	0.28	0.25	−0.50	0.84
Urban MSA	4,026	–	0.00	1.00	0.50	0.00	1.00
MSA population growth	4,026	–	0.01	0.01	0.04	−0.02	0.07
MSA land size	4,026	sq mi	2,495	1,637	2,934	252	14,566
MSA percentage of inhabitants living in poverty	4,022	Percentage	14.94	14.30	4.84	6.60	31.70
White-to-total population ratio	4,005	Percentage	72.86	75.33	15.78	30.22	96.02
MSA Catholic-to-Protestant ratio	4,026	–	0.68	0.34	0.92	0.01	4.35

The table provides the number of observations, mean, standard deviation, 1st percentile, the median (50th percentile), and the 99th percentile of all MSA control variables used in the empirical analysis as well as the instrumental variables based on MSA soil and weather characteristics. The definition of the variables can be found in Appendix Table A.1.

Table A.3
Example MSA inequality 2004: Financial wealth Gini construction

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income group category by size of adjusted gross income and ZIP code	Total number of returns	Taxable dividends number of returns	Taxable dividends total amount reported	Number of returns	Average dividend per household	Taxable interest number of returns	Taxable interest total amount reported
Total	1,882,964	306,037	1,287,291			651,013	1,189,469
Households with no dividend income			0	1,576,927	0.00		
Under $10,000	387,555	31,604	37,351	31,604	1.18	67,710	70,567
$10,000 under $25,000	553,957	42,503	64,756	42,503	1.52	114,076	150,519
$25,000 under $50,000	454,236	60,982	104,993	60,982	1.72	152,215	184,266
$50,000 under $75,000	231,139	55,051	113,500	55,051	2.06	123,892	156,056
$75,000 under 100,000	124,646	42,277	98,721	42,277	2.34	84,098	116,665
$100,000 or more	131,431	73,620	867,970	73,620	11.79	109,022	511,396

The table provides an example of the data used to construct our *Financial Wealth Gini* measure from 2004. We obtain data from the SOI (Statement of Income) database from the IRS on the total number of tax returns in thousands (one per household) filed in 2004 classified by zipcode and the adjusted gross income as shown in column (1). In addition we obtain information on the number of returns that declared to have obtained a dividend and the accompanying total dividend amounts reported (reported in thousands and thousands of US$), again classified by zipcode and the adjusted gross income of the household (shown in columns (2) and (3), respectively). Based on this data we calculate the average dividend amount per household reported for each income group in Column (5). The average dividend amount is reported in thousands of US$. We create an extra category of the number of households that did not declare any dividend (which is the total reports filed minus all reports that declared a dividend) which we report in the row 'Total', columns (4) and (5), respectively. We use these average dividends as well as the income group classification to construct a Gini index, aggregating the zip code information to the corresponding MSA, in line with Rajan (2011). We create a second Gini coefficient in the same way, only now based upon the amount of interests received by households in 2004, as reported in columns (6) and (7). Again, we obtain this information from the SOI database. The correlations between the Gini's based upon dividends and interest income received by households is very large and we therefore only report the results from our analysis in which we introduce the county inequality measure based upon dividends received.

Table A.4
Correlation matrix: Main measures of inequality

		(1)	(2)	(3)	(4)	(5)
Financial wealth Gini	(1)	1				
Housing wealth Gini	(2)	0.476***	1			
Proportion of financial wealth owned by the top 10%	(3)	0.646***	0.291***	1		
Proportion of housing wealth owned by the top 10%	(4)	0.488***	0.951***	0.266***	1	
Land Gini 1880	(5)	0.371***	0.256***	0.269***	0.374***	1

The definition of the variables can be found in Table A.1. *$p < .1$; **$p < .05$; ***$p < .01$.

Table A.5
Explaining business formation: Other inequality measures

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	*ln(MSA total establishment entries per capita)*							
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Housing wealth Gini	−0.997***	−2.814**						
	(0.284)	(1.208)						
Mortgaged housing wealth Gini			−1.246***	−9.573				
			(0.189)	(8.354)				
No mortgage housing wealth Gini					−0.791**	−2.542**		
					(0.308)	(1.102)		
Land 1880 Gini							−0.276**	−1.105**
							(0.131)	(0.435)
State × Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic first stage	–	15.424	–	1.214	–	15.420	–	14.193
Number of observations	3,740	3,707	3,740	3,707	3,707	3,685	3,399	3,399
R−squared	0.727	0.692	0.742	−0.436	0.722	0.695	0.696	0.659
Economic relevancy (Gini standard deviation)	−5.0%	−14.0%	−15.6%	−120.2%	−3.8%	−12.2%	−3.7%	−14.6%

The dependent variable is defined as the natural logarithm of the yearly total number of new establishments in the MSA between 2004 and 2014 divided by total MSA population. Housing Wealth Gini is defined as the Gini coefficient of the distribution of wealth measured as home equity of houses purchased between 1999 and 2004 that do have a mortgage and of houses that do not have a mortgage. The Mortgaged Housing Wealth Gini is defined as the Gini coefficient of the distribution of wealth measured as home equity of houses purchased between 1999 and 2004 that do have a mortgage. The No Mortgage Housing Wealth Gini is defined as the Gini coefficient of the distribution of wealth measured as the home equity related to houses that do not have a mortgage. The Land 1880 Gini is defined as the Gini coefficient of the distribution of farm land in 1880 in the MSA. All other variables are defined in Appendix Table A.1. Models (1), (3), (5), and (7) are estimated with a linear regression (OLS) model. Models (2), (4), (6), and (8) are estimated with a 2SLS IV model. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic-to-Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of one if the MSA population density is above the sample median. The IV estimates in Columns (2), (4), and (6) also include Water Canal in the MSA in 1860, Native American Tribes in the MSA in 1960, Railroad Line in the MSA in 1860 and Historical Farm Value as additional controls. “Yes” indicates that the set of fixed effects is included. “—” indicates that the indicated set of fixed effects are comprised in the wider included set of fixed effects. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. *$p<.1$; **$p<.05$; ***$p<.01$.

Table A.6
Specifications Explaining Business Formation: Wealth and Income

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	*ln(MSA Total Establishment Entries Per Capita)*			*ln(MSA Total) Establishment Exits Per Capita)*	*ln(MSA Total Establishment Exits Per Capita (Churning))*	*ln(MSA Total Establishment Exits Per Capita (Schumpeterian))*
Financial Wealth Gini	–	−2.964***	−2.841***	−2.610***	−2.719***	−2.579***
	–	(0.657)	(0.632)	(0.668)	(0.725)	(0.710)
Earned Income Gini	1.081***	0.712**	0.703**	0.387	0.585	0.333
	(0.292)	(0.331)	(0.328)	(0.343)	(0.424)	(0.325)
State Fixed Effects	Yes	Yes	–	–	–	–
Year Fixed Effects	Yes	Yes	–	–	–	–
State × Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	3,729	3,729	3,696	3,696	3,024	3,024
R-Squared	0.707	0.739	0.755	0.67	0.733	0.592
Economic Relevancy (Gini Standard Deviation) -Wealth		−10.6%	−10.2%	−9.4%	−9.8%	−9.3%
Economic Relevancy (Gini Standard Deviation) - Income	3.0%	3.4%	3.3%	1.8%	2.7%	1.6%

The dependent variables are defined as the natural logarithm of the yearly total number of new establishments in the MSA between 2004 and 2014 divided by total MSA population; the natural logarithm of the yearly total number of establishments that became inactive in the MSA between 2004 and 2014 divided by total MSA population; the natural logarithm of the yearly total number of establishments that became inactive in the MSA within 36 months of opening between 2004 and 2012 divided by total MSA population; and the natural logarithm of the yearly total number of establishments that became inactive in the MSA after 36 months of opening between 2004 and 2012 divided by total MSA population. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. The Earned Income Gini is defined as the Gini coefficient of the distribution of income measured as the sum total wages and total profit and losses coming from entrepreneurial activities as reported in the IRS-SOI data. All other variables are defined in Appendix Table A.I. All Models are estimated with a linear regression (OLS) model. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic-to-Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of one if the MSA population density is above the sample median. "Yes" indicates that the set of fixed effects or controls is included. "—" indicates that the indicated set of fixed effects are comprised in the wider included set of fixed effects. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the State level. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

Table A.7
Main specifications explaining business formation: Housing inequality - instrumented

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	*ln(MSA total establishment entries per capita)*						
Housing wealth Gini	−3.562*	−2.799*	−2.833**	−2.814**	−2.701**	−2.866**	−3.018*
	(1.781)	(1.403)	(1.209)	(1.208)	(1.299)	(1.274)	(1.473)
Railroad line in the MSA in 1860		−0.056	−0.055	−0.052	−0.040	−0.052	−0.044
		(0.040)	(0.040)	(0.039)	(0.040)	(0.039)	(0.065)
Water canal in the MSA in 1860		−0.038	−0.038	−0.040	−0.039	−0.042	−0.103***
		(0.026)	(0.025)	(0.024)	(0.026)	(0.025)	(0.034)
Native American tribes in the MSA in 1860		0.019	0.018	0.019	0.013	0.019	0.022
		(0.029)	(0.030)	(0.030)	(0.034)	(0.030)	(0.059)
Historical farm values			−0.086**	−0.078**	−0.064*	−0.077**	−0.073
			(0.035)	(0.035)	(0.037)	(0.036)	(0.057)
Illiteracy						0.077	−0.301
						(0.212)	(0.213)
Average plot size							
State fixed effects	Yes	Yes	Yes	–	–	–	–
Year fixed effects	Yes	Yes	Yes	–	–	–	–
State × Year fixed effects	No	No	No	Yes	Yes	Yes	Yes
MSA controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV coefficient first stage	−0.072***	−0.079***	−0.088***	−0.086***	−0.086***	−0.080***	−0.073***
F-statistic first stage	10.07	13.077	15.583	15.424	17.803	15.524	20.569
Number of observations	3,773	3,773	3,740	3,707	3,267	3,707	1,936
R-squared	0.634	0.679	0.677	0.692	0.696	0.691	0.720
Economic relevancy (Gini standard deviation)	−16.3%	−13.0%	−13.2%	−13.1%	−12.6%	−13.3%	−14.0%

The dependent variable is defined as the natural logarithm of the yearly total number of new establishments in the MSA between 2004 and 2014 divided by total MSA population. The Housing Wealth Gini is defined as Gini coefficient of the distribution of wealth measured as home equity of houses purchased between 1999 and 2004 that do have a mortgage and of houses that do not have a mortgage. All other variables are defined in Appendix Table A.1. All Models are estimated with a 2SLS estimator. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic-to-Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of one if the MSA population density is above the sample median. "Yes" indicates that the set of fixed effects or controls is included. "No" indicates that the set of fixed effects or controls is not included. "—" indicates that the indicated set of fixed effects or controls are comprised in the wider included set of fixed effects. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the State level. A possible concern with the Housing Wealth Gini is that it may not capture wealthy households that do not own a house but rent instead. To control for this possibility, we obtain information on the percentage of rented housing units as a proportion of total housing units in each MSA. When we control for this variable, in unreported regressions, the sign and significance of the wealth inequality coefficients remain unaltered. *$p < .1$; **$p < .05$; ***$p < .01$.

Table A.8
Main specifications explaining business formation: Land Gini 1880 - instrumented

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	*ln(MSA total establishment entries per capita)*						
Land 1880 Gini	−1.090**	−1.066**	−1.121**	−1.105**	−1.272**	−1.113**	−0.952*
	(0.452)	(0.486)	(0.473)	(0.490)	(0.536)	(0.531)	(0.545)
Railroad line in the MSA in 1860		−0.033	−0.031	−0.027	−0.025	−0.027	−0.026
		(0.031)	(0.031)	(0.032)	(0.036)	(0.033)	(0.063)
Water Canal in the MSA in 1860		0.024	0.025	0.020	0.029	0.020	-0.021
		(0.033)	(0.033)	(0.033)	(0.036)	(0.032)	(0.038)
Native American tribes in the MSA in 1860		0.010	0.009	0.013	0.022	0.013	0.032
		(0.026)	(0.026)	(0.027)	(0.031)	(0.027)	(0.046)
Historical farm values			−0.114*	−0.095	−0.030	−0.095	−0.091
			(0.061)	(0.064)	(0.073)	(0.064)	(0.072)
Illiteracy						0.026	−0.425
						(0.237)	(0.280)
Average plot size							
State fixed effects	Yes	Yes	Yes	–	–	–	–
Year fixed effects	Yes	Yes	Yes	–	–	–	–
State × Year fixed effects	No	No	No	No	No	No	
MSA controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV coefficient first stage	−0.261***	−0.242***	−0.245***	−0.248***	−0.244***	−0.227***	−0.216***
F-statistic first stage	15.06	15.24	16.30	14.19	12.58	11.49	8.88
Number of observations	3,454	3,454	3,454	3,399	3,014	3,399	1,639
R-squared	0.638	0.642	0.640	0.659	0.652	0.658	0.684
Economic relevancy (Gini standard deviation)	−13.4%	−13.2%	−13.8%	−13.6%	−15.1%	−13.7%	−13.0%

The dependent variable is defined as the natural logarithm of the yearly total number of new establishments in the MSA between 2004 and 2014 divided by total MSA population. The Land 1880 Gini is defined as the Gini coefficient of the distribution of farm land in 1880 in the MSA. All other variables are defined in Appendix Table A.1. All Models are estimated with a 2SLS estimator. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic-to-Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of 1 if the MSA population density is above the sample median. "Yes" indicates that the set of fixed effects or controls is included. "No" indicates that the set of fixed effects or controls is not included. "—" indicates that the indicated set of fixed effects or controls are comprised in the wider included set of fixed effects. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. $^{*}p<.1$; $^{**}p<.05$; $^{***}p<.01$.

Table A.9
Main specifications explaining business formation: Instrumented and Georgia land lotteries

Model	(1)	(2)	(3)
Dependent variable	*ln(MSA total establishment entries per capita)*		
Sample / specification	Georgia land lotteries		
Financial wealth Gini	−13.464*		
	(6.980)		
Housing wealth Gini		−1.204**	
		(0.597)	
Land 1880 Gini			−0.826**
			(0.412)
Railroad line in the MSA in 1860	−0.683***	−0.740***	−0.853***
	(0.086)	(0.068)	(0.081)
Water canal in the MSA in 1860	−0.227***	−0.353***	−0.315***
	(0.084)	(0.051)	(0.050)
Native American tribes in the MSA in 1860	0.233**	0.387***	0.438***
	(0.100)	(0.056)	(0.066)
Historical farm values	−3.858	1.474***	1.199**
	(2.761)	(0.460)	(0.544)
Illiteracy	−4.486***	−1.812***	−1.997***
	(1.258)	(0.473)	(0.474)
Average plot size	−0.002***	−0.002***	−0.003***
	(0.000)	(0.000)	(0.000)
State fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
State * Year fixed effects	No	No	No
MSA controls	Yes	Yes	No
IV coefficient first stage	−0.010***	−0.116***	−0.169***
F-statistic first stage	15.460	748.645	178.37
Number of observations	165	165	165
R-squared	0.733	0.788	0.773
Economic relevancy (Gini standard deviation)	−40.0%	−5.8%	−10.4%

The dependent variable is defined as the natural logarithm of the yearly total number of new establishments in the MSA between 2004 and 2014 divided by total MSA population. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. The Housing Wealth Gini is defined as Gini coefficient of the distribution of wealth measured as home equity of houses purchased between 1999 and 2004 that do have a mortgage and of houses that do not have a mortgage. The Land 1880 Gini is defined as the Gini coefficient of the distribution of farm land in 1880 in the MSA. All other variables are defined in Appendix Table A.I. All Models are estimated with a 2SLS estimator. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic to Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of one if the MSA population density is above the sample median. "Yes" indicates that the set of fixed effects or controls is included. "No" indicates that the set of fixed effects or controls is not included. Estimated coefficients are reported in each first row with Huber-White Robust standard errors below (in parentheses). *$p<.1$; **$p<.05$; ***$p<.01$.

Table A.10
Homesteading and farm values

Model	(1)	(2)	(3)
Dependent variable		*ln(Farm values)*	
Homesteading in the first 10 years of incorporation of the MSA area	0.093	−0.280	−0.255
	(0.537)	(0.849)	(0.843)
Railroad line in the MSA in 1860	0.211**	0.164**	0.166**
	(0.086)	(0.067)	(0.067)
Water canal in the MSA in 1860	0.258***	0.086	0.085
	(0.085)	(0.069)	(0.070)
Native American tribes in the MSA in 1860	0.065	0.079	0.076
	(0.085)	(0.077)	(0.076)
Illiteracy	−2.767***	−1.398***	−1.424***
	(0.275)	(0.493)	(0.492)
State fixed effects	No	Yes	–
Year fixed effects	No	Yes	–
State × Year fixed effects	No	No	Yes
Number of observations	4,531	4,531	4,437
R-squared	0.19	0.73	0.80
Economic relevancy (Gini standard deviation)	1.2%	−3.5%	−3.2%

The dependent variable is defined as the natural logarithm of farm value. Homesteading in the First 10 Years of Incorporation of the MSA Area is the amount of land distributed via Homestead Acts between 1862 and 1872, if the MSA was already settled at the enactment of the Act, or in the first 10 years since first settlement, divided by the total amount of land distributed in the MSA. All other variables are defined in Appendix Table A.I. All Models are estimated with a linear regression (OLS) model. "Yes" indicates that the set of fixed effects is included. "No" indicates that the set of fixed effects is not included. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. *$p<.1$; **$p<.05$; ***$p<.01$.

Table A.11
Main specifications explaining business formation: Accounting for spatial and time correlations

	Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Accounted for correlation up to	*Distance cutoff (in kilometers)*	*500*	*500*	*500*	*1,500*	*1,500*	*1,500*	*3,000*	*3,000*	*3,000*
	Time (in years)	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*
Dependent variable		*ln(MSA total establishment entries per capita)*								
Financial weath Gini		−3.79***	−3.79***	−3.79***	−3.79***	−3.79***	−3.79***	−3.79***	−3.79***	−3.79***
		(0.848)	(1.249)	(1.479)	(0.760)	(1.192)	(1.431)	(0.621)	(1.108)	(1.362)
F-statistic first stage		69.59	27.58	18.62	60.53	26.04	17.90	258.92	38.85	23.15
State fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Moran z-statistic		0.32	0.32	0.32	0.51	0.51	0.51	−1.88	−1.88	−1.88
Number of observations		3,740	3,740	3,740	3,740	3,740	3,740	3,740	3,740	3,740
R-squared		0.947	0.947	0.947	0.947	0.947	0.947	0.947	0.947	0.947
Housing wealth Gini		−2.83***	−2.83***	−2.83***	−2.83***	−2.83***	−2.83***	−2.83***	−2.83***	−2.83***
		(0.632)	(0.932)	(1.104)	(0.567)	(0.889)	(1.068)	(0.463)	(0.827)	(1.017)
F-Statistic First Stage		61.73	29.95	21.34	64.08	30.50	21.60	171.09	43.43	27.38
State Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Moran Z-Statistic		4.82	4.82	4.82	2.25	2.25	2.25	3.92	3.92	3.92
Number of Observations		3,740	3,740	3,740	3,740	3,740	3,740	3,740	3,740	3,740
R-Squared		0.947	0.947	0.947	0.947	0.947	0.947	0.947	0.947	0.947
Land 1880 Gini		−1.12***	−1.12***	−1.12***	−1.12***	−1.12***	−1.12***	−1.12***	−1.12***	−1.12***
		(0.226)	(0.334)	(0.395)	(0.203)	(0.318)	(0.382)	(0.166)	(0.296)	(0.364)
F-Statistic First Stage		103.33	38.04	25.37	119.12	39.99	26.22	294.05	49.97	30.17
State Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Moran Z-Statistic		1.16	1.16	1.16	−2.03	−2.03	−2.03	1.23	1.23	1.23
Number of Observations		3,454	3,454	3,454	3,454	3,454	3,454	3,454	3,454	3,454
R-Squared		0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945

The dependent variable is defined as the natural logarithm of the yearly total number of new establishments in the MSA between 2004 and 2014 divided by total MSA population. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. The Housing Wealth Gini is defined as Gini coefficient of the distribution of wealth measured as home equity of houses purchased between 1999 and 2004 that do have a mortgage and of houses that do not have a mortgage. The Land 1880 Gini is defined as the Gini coefficient of the distribution of farm land in 1880 in the MSA. All other variables are defined in Appendix Table A.1. All Models are estimated with a 2SLS estimator and include MSA controls which include MSA Estimated Total Financial and Housing Wealth, MSA Catholic to Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of 1 if the MSA population density is above the sample median. "Yes" indicates that the set of fixed effects or controls is included. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are adjusted for spatial and time-wise correlation up to the indicated cutoffs. *$p<.1$; **$p<.05$; ***$p<.01$.

Table A.12
Main specifications explaining business formation: Instrumented - soil and weather characteristics

Model		(1)	(2)	(3)	(4)
Dependent variable	*First-stage regression*	*ln(MSA Total Establishment Entries Per Capita)*		*ln(MSA Total Establishment Exits Per Capita)*	
Financial Wealth Gini		−5.904***	−5.802***	−6.054***	−5.966***
		(1.098)	(1.143)	(1.141)	(1.156)
MSA Average Rainfall between 1895 and 2003	0.011				
	(0.009)				
MSA Standard Deviation of Rainfall between 1895 and 2003	−0.018*				
	(0.01)				
MSA Average Temperature between 1895 and 2003	0.001				
	(0.001)				
MSA Standard Deviation of Temperature between 1895 and 2003	0.003**				
	(0.001)				
MSA Soil Salinity	0.006***				
	(0.001)				
MSA Soil Depth	0.0004**				
(0.0002)					
State Fixed Effects	Yes	Yes	–	–	–
Year Fixed Effects	Yes	Yes	–	–	–
State × Year Fixed Effects	No	No	Yes	No	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes
IV Coefficient First Stage					
F-Statistic First Stage	9.871	11.076	9.871	11.076	
Number of Observations	3,663	3,663	3,630	3,663	3,630
R-Squared	0.731	0.675	0.692	0.553	0.579
Economic Relevancy (Gini Standard Deviation)		−45%	−44%	−45%	−45%

The dependent variables are defined as the natural logarithm of the yearly total number of new establishments in the MSA between 2004 and 2014 divided by total MSA population; and the natural logarithm of the yearly total number of establishments that became inactive in the MSA between 2004 and 2014 divided by total MSA population. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. All other variables are defined in Appendix Table A.1. All Models are estimated with a 2SLS IV model. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic to Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of 1 if the MSA population density is above the sample median. "Yes" indicates that the set of fixed effects is included. "No" indicates that the set of fixed effects is not included. "—" indicates that the indicated set of characteristics or fixed effects are comprised in the wider included set of fixed effects. "Yes" indicates that the set of fixed effects or controls is included. "No" indicates that the set of fixed effects or controls is not included. "—" indicates that the indicated set of characteristics or fixed effects are comprised in the wider included set of fixed effects. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the State level. $*p<.1$; $**p<.05$; $***p<.01$.

Table A.13
Inequality and financing outcomes and business formation

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	*No. of Bank Branches per Capita*		*Total Amount of Loans to Small Business per Capita*		*ln(MSA Total Establishment Entries Per Capita)*	
Model	OLS	IV	OLS	IV	OLS	IV
Financial Wealth Gini	−2.481***	−2.453	−1.752**	−4.705	−2.675***	−1.582***
	(0.752)	(1.651)	(0.757)	(3.854)	(0.515)	(0.387)
No. of Bank Branches per Capita						0.340***
						(0.051)
Total Amount of Loans to Small Business per Capita						0.110***
						(0.025)
State × Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes	Yes
F-Statistic First Stage	–	15.136	–	11.848	–	–
Number of Observations	3,773	3,707	2,996	2,942	1,314	1,314
R-Squared	0.782	0.267	0.576	0.586	0.751	0.804
Economic Relevancy (Gini Standard Deviation)	−9.4%	−9.3%	−6.7%	−19.3%	−10.2%	−6.0%

In this table we investigate the effect of wealth Inequality on credit market outcomes. The quality of the local credit/banking market is a possible channel that can explain the relationship between wealth inequality and entrepreneurship: if financing is more prevalent in more equal areas this can foster entrepeneurship. The dependent variables are defined as: the number of yearly bank branches in the MSA divided by population between 2004 and 2014; the total number of loans made to small business in the MSA between 2006 and 2014; the total amount of loans made to small businesses in USD in the MSA between 2006 and 2014: and the natural logarithm of the yearly total number of new establishments in the MSA between 2005 and 2014 divided by total MSA population. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. All other variables are defined in Appendix Table A.1. Models (2), (4), and (6) are estimated with a 2SLS estimator; all other Models are estimated using OLS. The definition of the variables can be found in Appendix Table A.1. t-1. The definition of the variables can be found in Table A.1. MSA controls include MSA Estimated Total Wealth, MSA Catholic to Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of 1 if the MSA population density is above the sample median. The IV estimates in Columns (2), (4), (6), and (8) also include Railroad Line in the MSA in 1860, Water Canal in the MSA in 1860, Native American Tribes in the MSA in 1860, and Historical Farm Value as additional controls. "Yes" indicates that the set of fixed effects is included. "—" indicates that the indicated set of characteristics or fixed effects are comprised in the wider included set of fixed effects. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. *$p < .1$; **$p < .05$; ***$p < .01$.

Table A.14
Inequality and net worth

Model	(1)	(2)
Dependent variable	*ln(Entrepreneurs' net worth)*	
Model	OLS	IV
Financial Wealth Gini	36.046**	94.936*
	(16.448)	(50.044)
State Fixed Effects	Yes	Yes
MSA Controls	Yes	Yes
F-Statistic First Stage	–	8.394
Number of Observations	424	424
R-Squared	0.518	0.496
Economic Relevancy (Gini Standard Deviation)	114.5%	301.7%

In this table we investigate whether there is a relationship between wealth inequality and the net worth of individuals who are about to become entrepeneurs. Also these regressions study the importance of local credit markets in explaining the relationship between wealth inequality and entrepreneurship. In an economy with credit constraints, an individual's likelihood of starting a business increases with her net worth. This implies that, for any level of wealth, more unequal areas will have lower business formation as fewer individuals (possibly only the very rich) will have enough wealth to overcome credit market imperfections. The dependent variable is defined as the natural logarithm of the net worth of individuals who are about to become entrepreneurs. The Financial Wealth Gini is defined as the Gini coefficient of the distribution of wealth as measured by the distribution of the amount of declared dividends from household tax filings in the MSA. All other variables are defined in Appendix Table A.I. Model (1) is estimated with a linear regression (OLS) model while Model (2) is estimated with a 2SLS IV model. MSA controls include MSA Estimated Total Wealth, MSA Catholic to Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, MSA Population Growth, and a dummy variable (Urban MSA) that takes the value of 1 if the MSA population density is above the sample median. The IV estimates also include Railroad Line in the MSA in 1860, Water Canal in the MSA in 1860, Native American Tribes in the MSA in 1860, and Historical Farm Value as additional controls. "Yes" indicates that the set of fixed effects is included. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. *$p < .1$; **$p < .05$; ***$p < .01$.

Table A.15
Horse race of inequality with credit markets and local public goods

A. Credit markets and schooling

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MSA total establishment entries per capita						
Financial Wealth Gini	−2.636***	−1.722***	−2.411***	−1.527**	−2.537***	−1.681***	−1.059**
	(0.470)	(0.377)	(0.530)	(0.615)	(0.459)	(0.388)	(0.439)
No. of Bank Branches per Capita		0.138***				0.121***	0.111***
		(0.026)				(0.024)	(0.022)
Total Amount of Loans to Small Business per Capita		0.263***				0.266***	0.273***
		(0.050)				(0.054)	(0.053)
Proportion of Public School Revenue Coming from Local Sources				0.386*			0.148
				(0.196)			(0.154)
MSA School Expenditure per Pupil				−0.025			−0.022
				(0.043)			(0.030)
% of Individuals with at Least Some Years of College				0.004			0.005***
				(0.002)			(0.002)
Violent Crimes Per Capita							
State × Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	2,969	2,969	2,040	2,040	1,986	1,986	1,986
R-Squared	0.748	0.800	0.670	0.681	0.680	0.742	0.749
Economic Relevancy (Gini Standard Deviation)	−10.0%	−6.5%	−9.2%	−5.8%	−9.6%	−6.4%	−4.0%

B. Credit markets, schooling, and violent crimes

Model	(1)	(2)	(3)	(4)
	MSA total establishment entries per capita			
Financial Wealth Gini	−2.483***	−1.725***	−1.119**	−0.995**
	(0.504)	(0.430)	(0.482)	(0.419)
No. of Bank Branches per Capita		0.084***	0.077***	0.078***
		(0.025)	(0.024)	(0.024)
Total Amount of Loans to Small Business per Capita		0.274***	0.277***	0.281***
		(0.059)	(0.058)	(0.053)
Proportion of Public School Revenue Coming from Local Sources			0.108	0.066
			(0.149)	(0.145)
MSA School Expenditure per Pupil			−0.049	−0.055
			(0.066)	(0.065)
% of Individuals with at Least Some Years of College			0.005***	0.006***
			(0.002)	(0.002)
Violent Crimes Per Capita				−0.013**
				(0.006)
State × Year Fixed Effects	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes
Number of Observations	646	646	646	646
R-Squared	0.689	0.746	0.755	0.760
Economic Relevancy (Gini Standard Deviation)	−9.4%	−6.6%	−4.2%	−3.8%

All models are estimated with an OLS model. The dependent variables are expressed in natural logarithms. The definition of the variables can be found in Appendix Table A.I. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic-to-Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), and Change in MSA Age Composition and a dummy variable (Urban MSA) that takes the value of one if the MSA population density is above the sample median. "Yes" indicates that the set of fixed effects is included. "—" indicates that the indicated set of characteristics or fixed effects are comprised in the wider included set of fixed effects. t-1 indicates a one year lag. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. * $p < .1$; ** $p < .05$; *** $p < .01$.

Table A.16
3SLS

A. Credit markets and schooling

Model	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Stege 3	*ln(MSA Total Establishment Entries Per Capita)*														
No. of Bank Branches per Capita	1.149** (0.580)					1.149** (0.584)					0.629*** (0.110)				
Total Amount of Loans to Small Business per Capita		0.834 (0.604)					0.861 (0.621)					0.421*** (0.084)			
Proportion of Public School Revenue Coming from Local Sources			2.066* (1.234)					2.066* (1.235)					0.727** (0.324)		
School Expenditure per Pupil				0.045* (0.024)					0.045* (0.024)					0.006 (0.108)	
Violent Crime per Capita					−0.077** (0.036)					−0.075** (0.037)					−0.062*** (0.011)
Stage 2	*No. of Bank Branches per Capita*	*Total Amount of Loans to Small Business per Capita*	*Proportion of Public School Revenue Coming from Local Sources*	*School Expenditure per Pupil*	*Violent Crime per Capita*	*No. of Bank Branches per Capita*	*Total Amount of Loans to Small Business per Capita*	*Proportion of Public School Revenue Coming from Local Sources*	*School Expenditure per Pupil*	*Violent Crime per Capita*	*No. of Bank Branches per Capita*	*Total Amount of Loans to Small Business per Capita*	*Proportion of Public School Revenue Coming from Local Sources*	*School Expenditure per Pupil*	*Violent Crime per Capita*
Gini (Used in Stage 1)	−3.771** (1.629)	−5.325 (4.507)	−2.089** (0.853)	−1.046 (1.947)	49.726*** (19.162)	−3.102* (1.672)	−4.477 (4.001)	−1.722** (0.762)	−0.264 (1.817)	40.396** (19.112)	−1.242*** (0.417)	−1.982** (0.907)	−0.448** (0.226)	−0.665 (0.506)	11.622*** (2.883)
Stage 1	*Financial Wealth Gini*					*Housing Wealth Gini*					*Land Gini 1880*				
Proportion of Land Assigned via Homestead Act	−0.059*** (0.017)	−0.059*** (0.017)	−0.059*** (0.017)	−0.052*** (0.016)	−0.059*** (0.017)	−0.072*** (0.023)	−0.072*** (0.023)	−0.072*** (0.023)	−0.062** (0.025)	−0.072*** (0.023)	−0.262*** (0.057)	−0.260*** (0.056)	−0.253*** (0.054)	−0.251*** (0.054)	−0.261*** (0.057)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	3,773	2,996	3,707	3,707	2,329	3,773	2,996	3,707	3,707	2,329	3,773	2,996	3,707	3,707	2,329

All variables are defined in Appendix Table A.1. MSA controls include MSA Estimated Total Financial and Housing Wealth, MSA Catholic-to-Protestant Ratio, MSA Land Size, MSA Real GDP, MSA House Price Index (Level), MSA House Price Index (Change), Change in MSA Age Composition, and MSA Population Growth and a dummy variable (Urban MSA) that takes the value of one if the MSA population density is above the sample median, Railroad Line in the MSA in 1860, Water Canal in the MSA in 1860, Native American Tribes in the MSA in 1860, and Historical Farm Value. “Yes” indicates that the set of fixed effects is included. Estimated coefficients are reported in each first row with standard errors below (in parentheses) that are clustered at the state level. $*p<.1$; $**p<.05$; $***p<.01$.

References

Acemoglu, D., S. Johnson, and J. A. Robinson. 2001. The colonial origins of comparative development: An empirical investigation. *American Economic Review* 91:1369–401.

Acemoglu, D., and J. A. Robinson. 2013. *Why nations fail. the origins of power, prosperity and poverty*. New York: Random House.

Adelino, M., A. Schoar, and F. Severino. 2015. House prices, collateral, and self-employment. *Journal of Financial Economics* 117:288–306.

Aghion, P., and P. Bolton. 1997. A theory of trickle-down growth and development. *Review of Economic Studies* 64:151–72.

Aghion, P., and P. Howitt. 1992. A model of growth through creative destruction. *Econometrica* 60:323–51.

Akcigit, U., and W. R. Kerr. 2018. Growth through heterogeneous innovations. *Journal of Political Economy* 126:1374–443.

Alesina, A., R. Baqir, and W. Easterly. 1999. Public goods and ethnic divisions. *Quarterly Journal of Economics* 114:1243–84.

Alesina, A., and D. Rodrik. 1994. Distributive politics and economic growth. *Quarterly Journal of Economics* 109:465–90.

Allen, D. W. 1991. Homesteading and property rights; or, "How the West Was Really Won". *Journal of Law and Economics* 34:1–23.

Bach, L., L. E. Calvet, and P. Sodini. Forthcoming. Rich pickings? Risk, return, and skill in household wealth. *American Economic Review*. Advance Access published 10.1257/aer.20170666.

Bagchi, S., and J. Svejnar. 2015. Does wealth inequality matter for growth? The effect of billionaire wealth, income distribution, and poverty. *Journal of Comparative Economics* 43:505–30.

Banerjee, A. V., and E. Duflo. 2003. Inequality and growth: What can the data say? *Journal of Economic Growth* 8:267–99.

Banerjee, A. V., and A. F. Newman. 1993. Occupational choice and the process of development. *Journal of Political Economy* 101:274–98.

Barro, R. J. 2000. Inequality and growth in a panel of countries. *Journal of Economic Growth* 5:5–32.

Beck, T., R. Levine, and A. Levkov. 2010. Big bad banks? The winners and losers from bank deregulation in the United States. *Journal of Finance* 65:1637–67.

Becker, S. O., and L. Woessmann. 2009. Was Weber wrong? A human capital theory of protestant economic history. *Quarterly Journal of Economics* 124:531–96.

Berkowitz, D., and K. B. Clay. 2011. *The evolution of a nation: How geography and law shaped the American states*. Princeton, NJ: Princeton University Press.

Berkowitz, J., and M. J. White. 2004. Bankruptcy and small firms' access to credit. *RAND Journal of Economics* 35:69–84.

Black, S. E., and P. E. Strahan. 2002. Entrepreneurship and bank credit availability. *Journal of Finance* 57:2807–34.

Bleakley, H., and J. Ferrie. 2016. Shocking behavior: Random wealth in antebellum Georgia and human capital across generations. *Quarterly Journal of Economics* 131:1455–95.

Bonica, A., 2014. Mapping the ideological marketplace. *American Journal of Political Science* 58:367–86.

Bourguignon, F. 1981. Pareto superiority of unegalitarian equilibria in Stiglitz's model of wealth distribution with convex saving function. *Econometrica* 49:1469–75.

Braun, M., F. Parro, and P. Valenzuela. 2019. Does finance alter the relation between inequality and growth? *Economic Inquiry* 57:410–28.

Cagetti, M., and M. De Nardi. 2008. Wealth inequality: Data and models. *Macroeconomic Dynamics* 12:285–313.

Card, D., and A. A. Payne, 2002. School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics* 83:49–82.

Carlino, G., and W. R. Kerr. 2015. Agglomeration and innovation. In *Handbook of regional and urban economics*, eds. G. Duranton, J. V. Henderson, and W. C. Strange, vol. 5A. Amsterdam, the Netherlands: Elsevier.

Corradin, S., and A. Popov. 2015. House prices, home equity borrowing, and entrepreneurship. *Review of Financial Studies* 28:2399–428.

Davies, J. B., R. Lluberas, and A. F. Shorrocks. 2017. Estimating the level and distribution of global wealth, 2000–2014. *Review of Income and Wealth* 63:731–59.

Decker, R. A., J. C. Haltiwanger, R. S. Jarmin, and J. Miranda. 2018. Changing business dynamism and productivity: shocks vs. responsiveness. Working Paper, Federal Reserve Board.

Degryse, H., T. Lambert, and A. Schwienbacher. 2018. The political economy of financial systems: Evidence from suffrage reforms in the last two centuries. *Economic Journal* 128:1433–75.

Demirgüç-Kunt, A., and R. Levine. 2009. Finance and inequality: Theory and evidence. *Annual Review of Financial Economics* 1:287–318.

Donaldson, T. 1884. *The public domain: Its history, with statistics*, reprint. New York: Johnson.

Dunne, T., M. J. Roberts, and L. Samuelson. 1988. Patterns of firm entry and exit in U.S. manufacturing industries. *RAND Journal of Economics* 19:495–515.

Easterly, W. 2007. Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics* 84:755–76.

Engerman, S. L., and K. L. Sokoloff. 1997. Factor endowments, institutions, and differential paths of growth among new world economies: A view from economic historians of the United States. In *How Latin America fell behind*, ed. S. Haber. Stanford, CA: Stanford University Press.

———. 2002. Factor endowments, inequality, and paths of development among new world economies. *Journal of LACEA Economia* 3:41–109.

Forbes, K. J. 2000. A reassessment of the relationship between inequality and growth. *American Economic Review* 90:869–87.

Foster, L., C. Grim, J. C. Haltiwanger, and Z. Wolf. 2018. Innovation, productivity dispersion, and productivity growth. Working Paper, U.S. Census Bureau.

Fryer, R. G., and S. D. Levitt. 2004. Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics* 86:447–64.

Galor, O., O. Moav, and D. Vollrath. 2009. Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *Review of Economic Studies* 76:143–79.

Gates, P. W. 1936. The Homestead Law in an incongruous land system. *American Historical Review* 41:652–81.

———. 1969. *History of US public land law development*. New York: Zenger Publisher.

Glaeser, E., J. Scheinkman, and A. Shleifer. 2003. The injustice of Inequality. *Journal of Monetary Economics* 50:199–222.

Glaeser, E. L., and W. R. Kerr. 2009. Local industrial conditions and entrepreneurship: How much of the spatial distribution can we explain? *Journal of Economics and Management Strategy* 18:623–63.

Glomm, G., and R. Lagunoff. 1999. A dynamic tiebout theory of voluntary vs. involuntary provision of public goods. *Review of Economic Studies* 66:659–77.

Goldin, C., and L. F. Katz. 1998. The origins of state-level differences in the public provision of higher education: 1890 to 1940. *American Economic Review Papers and Proceedings* 88:303–8.

Hansen, Z. K., and G. D. Libecap. 2002. "Rain follows the plow" and dry farming doctrine: The climate information problem and homestead failure in the Upper Great Plains, 1890-1925. *Journal of Economic History* 62:86–120.

———. 2004. The allocation of property rights to land: US land policy and farm failure in the Northern Great Plains. *Explorations in Economic History* 41:103–29.

Hecker, D. E. 2005. High-technology employment: A NAICS-based update. *Monthly Labor Review of the U.S. Department of Labor and U.S. Bureau of Labor Statistics* 128:57–72.

Hilary, G., and K. W. Hui. 2009. Does religion matter in corporate decision making in America? *Journal of Financial Economics* 93:455–73.

Hogan, R. E. 2004. Challenger emergence, incumbent success, and electoral accountability in state legislative elections. *Journal of Politics* 66:1283–303.

Holbrook, T. M., and A. C. Weinschenk. 2014. Money, candidates, and mayoral elections. *Electoral Studies* 35:292–302.

Hughes, J., and L. Cain. 2010. *American economic history*, vol. 8. London, UK: Pearson.

Jiang, W. 2017. Have instrumental variables brought us closer to the truth. *Review of Corporate Finance Studies* 6:127–40.

Kaldor, N. 1961. Capital accumulation and economic growth. In *Theory of capital*, eds. F. Lutz and D. Hague. London: Macmillan.

Kelly, M. 2019. The standard errors of persistence. Discussion Paper, University College Dublin.

Kerr, S. P., W. R. Kerr, and R. Nanda. 2015. House money and entrepreneurship. Working Paper, Wellesley College.

Kerr, W. R., and R. Nanda. 2009. Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship. *Journal of Financial Economics* 94:124–49.

———. 2010. Banking deregulations, financing constraints, and firm entry size. *Journal of the European Economic Association* 8:582–93.

Kuhn, M., M. Schularick, and U. Steins. Forthcoming. Income and wealth inequality in America, 1949-2016. *Journal of Political Economy*. Advance Access published July 29, 2020, 10.1086/708815.

Kumhof, M., R. Rancière, and P. Winant. 2015. Inequality, leverage, and crises. *American Economic Review* 105:1217–45.

La Porta, R., and A. Shleifer. 2008. The unofficial economy and economic development. *Brookings Papers on Economic Activity,* 275–352.

———. 2014. Informality and development. *Journal of Economic Perspectives* 28:109–26.

Levine, R., and Y. Rubinstein. 2016. Smart and illicit: Who becomes an entrepreneur and do they earn more? *Quarterly Journal of Economics* 132:963–1018.

———. 2018. Selection into entrepreneurship and self employment. Mimeo.

Liptak, A., and A. Roberts. 2006. Campaign cash mirrors a high court's rulings. *New York Times*, October 1.

Magney, M. 2001 Old money, old virtues. *Forbes.* October 7. https://www.forbes.com/forbes/2001/1008/097.html#f92099f4bea2

Malthus, T. R. 1836. *Principles of political economy*. London: W. Pickering.

Matsuyama, K. 2002. The rise of mass consumption societies. *Journal of Political Economy* 110:1035–70.

Mian, A., K. Rao, and A. Sufi. 2013. Household balance sheets, consumption, and the economic slump. *Quarterly Journal of Economics* 128:1687–726.

Mirrlees, J. A., 1971. An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38:175–208.

Perotti, E., and E.-L. von Thadden. 2006. The political economy of corporate control and labor rents. *Journal of Political Economy* 114:145–74.

Persson, T., and G. Tabellini. 1994. Is inequality harmful for growth? *American Economic Review* 84:600–621.

Piketty, T. 2014. *Capital in the twenty-first century*. Cambridge, MA: Harvard University Press.

Poole, K. T., and H. Rosenthal. 1991. Patterns of congressional voting. *American Journal of Political Science* 35:228–78.

Rajan, R., and R. Ramcharan. 2011. Land and credit: A study of the political economy of banking in the United States in the early 20th century. *Journal of Finance* 66:1895–930.

Rajan, R. G. 2009. Rent preservation and the persistence of underdevelopment. *American Economic Journal: Macroeconomics* 1:178–218.

Ramcharan, R. 2010. Inequality and redistribution: Evidence from U.S. counties and states, 1890–1930. *Review of Economics and Statistics* 92:729–44.

Robb, A. M., and D. T. Robinson. 2014. The capital structure decisions of new firms. *Review of Financial Studies* 27:153–79.

Saez, E. 2017. Income and wealth inequality: Evidence and policy implications. *Contemporary Economic Policy* 35:7–25.

Saez, E., and G. Zucman. 2016. Wealth Inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics* 131:519–78.

Scheidel, W. 2017. *The great leveler: Violence and the history of inequality from the stone age to the twenty-first century*. Princeton, NJ: Princeton University Press.

Schmalz, M. C., D. A. Sraer, and D. Thesmar. 2017. Housing collateral and entrepreneurship. *Journal of Finance* 72:99–132.

Schumpeter, J. A. 1934. *The theory of economic development*. New Brunswick, NJ: Transactions Publishers.

Smith, A. 1776. *An inquiry into the nature and causes of the wealth of nations*. London: W. Strahan and T. Cadell.

Sonin, K. 2003. Why the rich may favor poor protection of property rights. *Journal of Comparative Economics* 31:715–31.

Stark, K. J., and J. Zasloff. 2003. Tiebout and tax revolt: Did Serrano really cause Proposition 13? *UCLA Law Review* 50:101–62.

Stiglitz, J. E. 1969. Distribution of income and wealth among individuals. *Econometrica* 37:382–97.

Tsoutsoura, M. 2015. The effect of succession taxes on family firm investment: Evidence from a natural experiment. *Journal of Finance* 70:649–88.

Vollrath, D. 2013. Inequality and school funding in the rural United States, 1890. *Explorations in Economic History* 50:267–84.