Household Inequality, Entrepreneurial Dynamism and Corporate Financing

Fabio Braggion
CentER - Tilburg University

Department of Finance
PO Box 90153, NL 5000 LE Tilburg, The Netherlands
Telephone: +31 13 4668209, Fax: +31 13 4662875
E-mail: f.braggion@uvt.nl

Mintra Dwarkasing *
CentER - Tilburg University

Department of Finance
PO Box 90153, NL 5000 LE Tilburg, The Netherlands
Telephone: +31 13 4668209, Fax: +31 13 4662875
E-mail: m.s.d.dwarkasing@uvt.nl

Steven Ongena
University of Zürich, *SFI* and *CEPR*

Department of Banking and Finance
University of Zürich
Plattenstrasse 32, CH-8032 Zürich, Switzerland
E-mail: steven.ongena@bf.uzh.ch

This Draft: August 2015

* Corresponding author. We thank Cédric Argenton, Fabio Castiglionesi, James Choi, Marco Da Rin, Joost Driessen, Thomas Hellmann, William Kerr, Alberto Manconi, Ramana Nanda, Fabiana Penas, Johann Reindl, David Robinson, Joacim Tåg, Florian Schuett, Janis Skrastins and Per Strömberg, participants at the 2015 ASSA-Meetings (Boston), 2015 FIRS Conference (Reykjavík), the Second European Workshop on Entrepreneurship Economics (Cagliari), 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 Rotterdam School of Management, Tilburg University and the Tilburg Center for Law and Economics for very valuable comments. We also like to thank the *Kauffman Foundation* for providing access to its Firm Survey. The *Netherlands Organisation for Scientific Research* (*NWO*) generously supported Dwarkasing through its Mosaic Grant Program during the writing of this paper. A previous draft of the paper was circulated with the title "Household Inequality, Corporate Capital Structure and Entrepreneurial Dynamism."

Household Inequality, Entrepreneurial Dynamism and Corporate Financing

Abstract

We empirically test hypotheses emanating from recent theory predicting that household wealth inequality may determine entrepreneurial dynamism and corporate financing. We construct two measures of wealth inequality at the US MSA/county level: One based on the distribution of financial rents in 2004 and another one related to the distribution of land holdings in the late Nineteenth century. Our results suggest that in more unequal areas business creation, especially of high-tech ventures, is lower and more likely to be financed via bank and family financing. Wealth inequality seemingly also affects local institutions such as banks, schools, and courts.

Keywords: wealth inequality, entrepreneurship, corporate financing.

JEL: D31, G3, L26.

I. Introduction

Households' wealth inequality is a defining societal characteristic with important implications for economics and finance. Already in his *magnum opus* "The Wealth of Nations", Adam Smith expressed concerns that an unequal distribution of the land may have had a negative impact on 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).

The growth of wealth inequality during recent decades has brought the issue back to the top of the agenda of policymakers and social leaders in many Western economies. While recent academic work has consequently revisited its defining, measuring, and analysing, (Chetty, Hendren, Kline and Saez (2014); Chetty, Hendren, Kline, Saez and Turner (2014); Piketty (2014); Saez and Zucman (2014)), policymaker and academic interest has also turned to the possible social and economic effects of inequality.

This paper sheds light on the economic consequences of households' wealth inequality, and in particular studies its impact on firm creation, technology, and financing choices made by young entrepreneurs. This is an important question because the creation of new business ventures, as well as the available means for financing them, are defining features of any economic system and likely have an important impact on economic development.

Recent economic theory directly links the degree of wealth inequality to economic and financial outcomes. Engerman and Sokoloff (1997), Glaeser, Scheinkman and Shleifer (2003), Sonin (2003) and Berkowitz and Clay (2011) for example describe how large levels of wealth inequality could impair the development of institutions that are conducive for economic growth. In unequal societies wealthy elites may prevent the sound development of basic institutions such as banks, schools, and courts, in order to maintain their grip on power. According to this view, an unequal society will be characterized by a poorly-developed financial system and by less effective schooling and law enforcement (see also Acemoglu and

Robinson (2013), pp. 152-158, Perotti and von Thadden (2006) and Rajan and Ramcharan (2011)).

All in all, these theories predict that wealth inequality may negatively influence the decision of individuals to become entrepreneurs and affect the technology, and financing choices made by firms. In more unequal societies, either because of a lack of higher education or because of poorly-developed financial markets, we expect to observe less entry of new firms, and those that enter to be more likely to operate in traditional sectors, and to come with a simpler corporate form. In more unequal societies, bank debt or family loans, as opposed to equity from institutional investors, will be the prevailing form of external finance.

This study will focus on households' wealth inequality measured at either the US metropolitan statistical area (MSA) or US county level and relate it to the choices made by firms located in the same area. Measurement at the MSA or county level substantially "shortens the distance" between local conditions and corporate outcomes and therefore allows us to obtain more precise estimates of the effects of interest. Important for our purposes is the observation that local administrations, in particular those at the county level, are often co-responsible (with state-level authorities) for many important elements of public life, such as the organization of schooling, the judiciary as well as the enforcement of the law and taxation.[1]

We will also study start-up firms: First, their creation and financing is more likely to depend upon local institutional and credit market conditions (Lerner (1995); Sorenson and Stuart (2001); Petersen and Rajan (2002); Guiso, Sapienza and Zingales (2004); Kerr and

[1] Ramcharan (2010), Rajan and Ramcharan (2011), Vollrath, Galor and Moav (2009) and Vollrath (2013) also relate county level inequality to various economic outcomes such as income redistribution, access to credit and schooling.

Nanda (2009); Chen, Gompers, Kovner and Lerner (2010); Berger, Cerqueiro and Penas (2014)); second, it allows us to more precisely identify entrepreneurial dynamics, as we can observe technology and production choices at the very beginning of the firm's life cycle.

Identifying the effect of inequality on entrepreneurial activity and corporate outcomes presents us with sizable empirical challenges. First, it is difficult to measure wealth inequality at a local level because direct and reliable households wealth data at a county level is virtually impossible to find. Additionally, corporate outcomes themselves (such as the local ease to start a *de novo* firm and the resultant distribution of profits) could easily determine local wealth inequality itself.

We address the first problem by constructing two proxies of wealth inequality: One based on contemporary sources and another one that relies on historical records. The present-day measure of wealth inequality is based on the amounts of dividends and interests earned by US households in 2004 (the first year for which this data is available) as reported by Internal Revenue Service (IRS) Statistics of Income (IRS-SOI) data. The IRS-SOI data report the total amount of dividends and interest income received by US households in each postal zip code. Under the assumption that a typical household holds 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, in other words, by the total amount of financial wealth it owns. For each zip code, the IRS provides this information for five income groups. It also reports the number of households belonging to each income group and the total amount of interests and dividends earned by each income group. We use this information to construct the distribution of financial rents by zip code, then we aggregate it at the county level and compute a county Gini coefficient as a measure of financial wealth inequality.[2]

The historical measure of wealth inequality is the distribution of land holdings at the US county level in 1890 (*sic*), a measure which given its historic nature is strappingly pre-determined. To construct the Gini-coefficient of land holdings in 1890, we access the US Census of Agriculture dataset that contains the size and number of farms in all counties recorded every ten years since 1860 (due to missing observations for counties in Oklahoma the state-level coefficient has to be used). Such a measure has already been employed by Galor, Moav and Vollrath (2009), Ramcharan (2010), Rajan and Ramcharan (2011) and Vollrath (2013) for example to study US historical developments in education, banking and redistribution.

For our historical measure to be able to deliver it has to be true that counties with a more unequal land distribution in 1890 are also those that are characterized by a higher wealth inequality today. We validate this measure by checking its correlation with present-day factors (also measured at county level) that arguably are related to the degree of wealth inequality. We find that 1890 land inequality displays a 36 and 46 percent positive correlation with our 2004 measures of dividend and interest inequality, for example. It is also positively correlated with the local poverty rates (43 percent) and the number of crimes per capita (33 percent) and it is negatively correlated with the number of white people living in a county (-53 percent).[3]

---

[2] Mian, Rao and Sufi (2013) use a similar methodology to construct local measures of US Households' Net Worth. Saez and Zucman (2014) also use this methodology to construct US wide long time series of wealth inequality.

[3] We assume here that the share of white people represents fairly well the proportion of the middle class living in a county. We take the data on poverty rates from the US Census bureau. These figures do not define poverty only based on income factors, but also to factors more related with wealth, such as the amount of money held in deposit accounts and the participation to various food/meal assistance programs.

The second empirical challenge consists of precisely identifying a causal relationship between wealth inequality and corporate outcomes, as wealth inequality itself may be correlated with unobserved factors likely to affect our estimates. We tackle this problem in various ways. First, as our measures of wealth inequality are at the local level (and since we know the precise location of the firms) we saturate our specifications with state, year, industry, state-year and/or industry-year fixed effects to account for any unobserved heterogeneity at those aforementioned levels. This allows us to control for competing explanations of the deep rooted determinants of institutions such as the individual States' type of colonization and legal traditions (see Berkowitz and Clay (2011), pp. 16-59; Acemoglu, Johnson and Robinson (2001)) as well as changes in their legislation and regulation. Our capital structure regressions also control for a large number of salient firm and main owner characteristics taken from the Kauffmann Survey of Business Formation database, as well as various measures of county demographic and socio-economic characteristics.

Second, following the literature (Easterly (2007); Galor, Moav and Vollrath (2009); Ramcharan (2010); Rajan and Ramcharan (2011); Berkowitz and Clay (2011), pp. 102-104), we instrument the contemporary measure of wealth inequality with a set of variables related to the local historical averages of rainfall and temperature that have been considered an exogenous predictor of contemporary wealth inequality.[4] This strategy relies on the historical evidence provided by Engerman and Sokoloff (2002b) that suggests that the quality of soil combined with the climate may have a persistent effect on the degree of inequality. In particular, regions whose soil and climate are best suited for large farms, with crops such as

[4] We measure average historical rainfall and temperature at the district level, where a district is defined by the National Climatic Data Center as a cluster of two or three counties sharing similar climatic conditions.

cotton or tobacco, should induce relatively high wealth inequality. The production of these crops entails high fixed costs. As a result, the market, in equilibrium, can support only a few farms owned by a few wealthy individuals.[5]

The underlying assumption of the instrumental variable analysis is that local weather conditions matter because, via inequality, they determine local institutions and entrepreneurship. To address the concern of whether the exclusion restriction is satisfied, we perform a falsification test that links local weather conditions to local entrepreneurship in France: a developed country, where local authorities have very limited power in setting up institutions like schooling and the judiciary. In principle, in France we should not find any correlation between local weather conditions and entrepreneurship.

Third, we exploit State changes in Estate, Inheritance and Gift (EIG) taxes between 1976 and 2000 and assess their impact on (new) firm entry and exit using a difference-in-differences approach. EIG taxes may be related with wealth inequality as they define the amount of wealth transferred from one generation to another. As a result, lower EIG taxes should promote or maintain a high level of wealth inequality. Starting in 1976, more than 30 states have eliminated their incremental EIG taxes imposed on top of the Federal tax, thus lowering the EIG tax burden on their citizens (Conway and Rork (2004)). If our conjecture on the relationship between wealth inequality and entrepreneurship is correct, we should find that States that lowered EIG taxes earlier experienced a significant drop in entrepreneurship activity.

---

[5] Under this perspective, the observed rainfall and temperature are good predictors of wealth inequality to the extent that inequality and institutions persist throughout time. Engerman and Sokoloff (1997), Engerman and Sokoloff (2002b) and Rajan (2009) suggest that this is the case and, ultimately, this is an empirical question that our first stage regressions will address.

Even after saturating specifications with the aforementioned dense sets of fixed effects and characteristics as well employing an instrumental variable analysis and a difference-in-differences methodology, the estimated coefficients robustly suggest that MSA level inequality decreases firm entry and exit in the MSA. Moreover, at the county-level wealth inequality increases the likelihood that a firm is a sole-proprietorship and boosts its proportion of family and bank financing (to owner or firm). Angel and venture capital financing on the other hand decreases in inequality, and so does the likelihood that newly created firms are high-tech.

Important for our identification strategy, we also notice that in salient specifications including controls in our regressions leaves the relationship between wealth inequality and entrepreneurial outcomes unaltered or, if anything, makes it stronger rather than weaker. To the extent that firms and state/county unobservable characteristics are correlated to our controls, it appears that endogeneity works against finding any link between wealth inequality, start-up capital structure and technology choices (Altonji, Elder and Taber (2005); Bellows and Miguel (2006); Bellows and Miguel (2009)).

Our estimates are not only statistically significant, but also economically relevant. A one standard deviation increase in MSA-level wealth inequality leads to a 30 percent increase of new establishments' entry and exit. These results lend support to the notion that in more equal areas, there is a more active process of creative destruction. Interestingly, more equal areas experience higher closure of both young and old establishments, suggesting that together with a genuine process of creative destruction (i.e., new establishments challenge the incumbents) there is lots of churning entry (i.e., lots of closures happens among newly formed establishments).

We also find sizable effects in our capital structure regressions. A one standard deviation increase in county-level wealth inequality for example increases the likelihood a sole proprietorship is locally present by 13 percent (of its own mean). A similarly constructed semi-elasticity for the impact (of inequality) on family and bank financing is equal to 30 percent. And, the likelihood that newly created firms are high-tech firms decreases by 8 percent in more unequal counties.

Our analysis identifies a reduced-form relationship between wealth inequality, entrepreneurial activity and capital structure choices and implicitly relates it to the quality of the local institutions. In the last of part of the paper, we study whether local institutions such as schooling, the banking market and the judiciary behave differently in counties with different levels of wealth inequality. For example, to the extent that in unequal counties education is poorer, less educated entrepreneurs may choose to work in low tech ventures which may require either own or bank financing. Our results show that in unequal counties there are a lower number of bank establishments per capita, which suggests that entrepreneurs may be rationed on the credit market.[6] Unequal counties also have a lower percentage of the population with at least a college degree and they are less likely to attract educated people from other geographical areas. More unequal counties also display a more inefficient civil justice system: everything else equal, in unequal counties, first degree civil justice trials have a longer completion time.

In sum, our findings vividly demonstrate the importance of inequality for corporate outcomes. These results therefore not only contribute to the already-cited literature that

[6] We follow Rajan and Ramcharan (2011) and use the number of bank establishments per capita as proxy of the supply of debt finance.

specifically links wealth inequality with entrepreneurial dynamism, firm ownership and financing, but also contributes directly to our understanding of all relevant factors shaping such outcomes. In his "Theory of Economic Development," Joseph Schumpeter stated that "an individual can only become an entrepreneur by previously becom[ing] a debtor" (Schumpeter (1934), p. 102).

While the Schumpeterian notion of creative destruction has received ample attention in the literature, detailed analyses of young entrepreneurs' financing, and especially the local conditions determining it, is still a novel field in economics and finance. Our estimates on this account demonstrate that local conditions are very important in determining the type and amount of financing entrepreneurs receive. In this sense our paper is related to Black and Strahan (2002), Berkowitz and White (2004), Kerr and Nanda (2009) and Kerr and Nanda (2010).[7]

Our work also contributes to a larger literature that investigates the salient long-term determinants of economic growth and development (Engerman and Sokoloff (1997); Acemoglu, Johnson and Robinson (2001); Acemoglu, Johnson and Robinson (2002); Engerman and Sokoloff (2002b); Nunn (2008); Acemoglu and Robinson (2013)). Last but not least, the 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 (see Demirgüç-Kunt and Levine (2009) for a review, and more recently Beck, Levine and Levkov (2010)), our paper studies how wealth inequality affects financial outcomes (and in this sense it is closer to Rajan (2009) and Degryse, Lambert and Schwienbacher (2013)).

[7] See also Kerr, Nanda and Rhodes-Kropf (2014), Carlino and Kerr (2015), among others, for reviews.

The rest of the paper is organized as follows. Section II discusses the testable hypotheses and introduces in more detail our measures of wealth inequality. Section III discusses the results on local wealth inequality and firm creation, technology, ownership and financing. Section IV explores the effect of wealth inequality on entrepreneurial dynamics exploiting the abandonment of state 'death', estate and gift taxes by different States at different points in time. Section V links local wealth inequality to local bank presence, education and the efficiency of the judicial system. Section VI concludes.

II. Inequality and Corporate Outcomes

A. Testable Hypotheses

Our hypotheses are based on the work of Engerman and Sokoloff (1997), Engerman and Sokoloff (2002b), Glaeser, Scheinkman and Shleifer (2003), and Sonin (2003). These papers describe how the emergence of wealthy elites may lead to the creation of institutions designed to preserve their political power and to maintain the existing level of inequality. According to these works, unequal societies will be characterized by less efficient capital markets, low levels of education, and an ineffective judiciary system. To the extent that efficient capital markets, good schooling, and a well-functioning judicial system are factors that contribute to higher rates of entrepreneurship, we expect business entries and exits to be lower in more unequal areas.

Acemoglu and Robinson (2013) also associate "bad institutions" to the presence of wealthy elites that prevent technological change and "creative destruction" (e.g., pp. 152-165). Modigliani and Perotti (2000)) provide a theoretical framework that shows that debt contracts are easier to enforce (vis-a-vis equity contracts) in areas where law enforcement is weaker. At the same time, if bad institutions reduce the amount of external finance available to entrepreneurs, families will have a more important role in financing business ventures in more unequal areas. All in all these works suggest that more inequality should be related to less technological change and a higher use of debt and family financing.

Within this framework we can therefore derive four testable hypotheses concerning the effect of wealth inequality on entrepreneurial dynamism and corporate financing:

Hypothesis 1: Greater wealth inequality will lead to, ceteris paribus, lower rates of business formation (i.e., entry and exit).

In unequal societies entrepreneurs themselves may prefer simpler technologies or they may find it difficult to finance riskier ventures. We therefore expect to see less high technology start-ups when county inequality is larger:

Hypothesis 2: The probability that a new business venture will be a high tech firm will, ceteris paribus, decrease in county inequality.

Hypothesis 3: As proprietorships are a simple corporate form often characterized by concentrated ownership and they usually are family owned, the likelihood that a start-up is a sole proprietorship as opposed to any other type of firm will be increasing in county inequality.

Hypothesis 4: Greater wealth inequality will lead, ceteris paribus, firm family and bank financing to be a larger fraction of total financing. Similarly, greater wealth inequality will lead, ceteris paribus, angel investor and venture capital financing to be a smaller fraction of total financing.

Hypotheses (2), (3) and (4) are also consistent with Perotti and von Thadden (2006). They build a model in which the median voter in an unequal society only owns her non-diversifiable human capital. As a result, she may prefer a financial system dominated by family and banks, "institutions" with whom she shares her aversion to risk. In more equal societies, the median voter may also own diversifiable financial wealth, and may prefer a system that also relies on equity financing and is characterized by risk-taking dynamism.

To summarize: We expect the probability that a new business venture will be formed and will be a high tech firm decreases, ceteris paribus, in local inequality (*H1*, *H2*). On the other hand, the probability that new ventures will take a simple corporate form will be increasing in wealth inequality (*H3*). The amount of bank and family financing will be increasing in local inequality, whereas the amount of 'outside' equity obtained from angel investors and venture capitalists will be decreasing in it (*H4*).

B. Empirical Strategy

In our analysis, we will estimate the impact of Wealth inequality on the number of firms' entry and exit using data at the MSA level as well as on capital structure and technology choices of startups.[8] In particular, for entry and exit, we will estimate the following equation:

$$Y_{j,t} = \alpha + \alpha_s + \alpha_t + \beta Wealth\ Inequality_j + Controls + \varepsilon_{j,t}$$

Where, $Y_{j,t}$ indicates the natural logarithm Number of Firms Entries and Exists in the Metropolitan Statistical Area *j* at year *t*. In line with Kerr and Nanda (2010), we will focus on gross business entry and exit. The variable *Wealth Inequality* indicates one of our measures of local wealth inequality. *Controls* stands for a set of MSA controls, such as population, income per capita and house prices which we will discuss in detail in the next sections. For the capital structure and technology choice regressions, our model is:

$$Y_{i,j,t} = \alpha + \alpha_s + \alpha_{Ind} + \alpha_t + \beta Wealth\ Inequality_j + Controls$$

Where, $Y_{i,,j,t}$ indicates different capital structure variables, the technology choice and corporate form of startup *i* located in county *j* at year *t*.

Since, in this case, we have firm level data, we additionally control for industry fixed effects together with the state and year fixed effects. *Controls* capture two sets of county and firm characteristics.

[8] In the analysis of entry and exit of new firms, we use data at the MSA level as data at the county level is not freely available.

C. Identification

Wealth inequality could capture either omitted factors or with the degree of entrepreneurship. We will address this problem in several ways. First, as we discussed, our measure of wealth inequality is constructed either at the MSA or county level. This allows us to control for State Fixed effects and State Trends in the analysis. This is a relevant feature of our identification strategy, as other main deep 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 what US-States have civil law (rather than common law) traditions and they link their legal traditions to the country of origins of early settlers.

Second, we make use of the available historical literature to generate an instrumental variable analysis. Engerman and Sokoloff (1997) and Engerman and Sokoloff (2002a) describe the factors that can be underlying causes of persistent differences in inequality: Different climates and geographical environments that may favor the production of one type of crop over another. Their argument suggests that climates that are best suited for large plantations, like sugar or tobacco, will induce relatively high economic inequality. The production of these crops comes at a high fixed cost; as a result, the market, in equilibrium, can support only a few farms. The outcome is thus a society controlled by few wealthy landowners. Conversely, climates supporting crops like wheat will result in a more equal society. The production of these crops do not require high fixed costs, hence the market can "bear" more producers. These societies will be more equal and mainly composed of small landowners. A feature of this theoretical framework is that inequality and "bad" institutions will be persistent over time and reinforce each other. Along these lines, Acemoglu and

Robinson (2013) and Rajan (2009) also provide a theoretical framework and empirical evidence of how institutions may persist through time.

We will also construct a falsification test to corroborate the validity of our instrumental variables. Our identification strategy implicitly assumes that local inequality has an impact on entrepreneurship because of the quality of local institutions. We will relate local weather conditions to entrepreneurship in France, a country with a similar level of development as the US but where, differently from the US, the local authorities have a very limited or no role in shaping local institutions such as schooling and the judiciary. Given that we do not have local inequality data for France, we will work with reduced form equations linking local weather conditions to local business entry and exit both for France and the US. Finding that local weather conditions have an impact on business formation in the US − because in the US local institution matters, but not in France − should justify the identification assumption.

D. Measuring Wealth Inequality

It is very difficult to readily obtain representative measures of wealth inequality at the local level. As a result we construct our own two proxies for local wealth inequality. The first one 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 (2014), and it intends to construct local level measures of household net worth; the second measure is based on historical records of land ownerships.

The contemporary measure of wealth inequality looks at the amounts of dividends and interests earned by US households in 2004, the first year in our sample period, as reported by Internal Revenue Service (IRS) Statistics of Income (SOI) data. The IRS-SOI data report the total amount of dividends and interest income received by US households in a certain zip code. The information is reported as a total per zip code, but also divided in five households' 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. IRS-SOI provides three pieces of information important to construct our proxy:

a) The total number of households belonging to each income group;

b) For each income group, the number of households who declared non-zero dividend and non-zero interest income (we will call these non-zero households); and,

c) For each income group, the total amount of dividends and interests 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 is exactly the same when we also include interest income and comprises six steps.

(1) We aggregate the IRS-SOI figures at a county level.[9]

(2) For each county, we compute the number of households who declared zero dividend income and we place them into a separate category.

(3) For each county and each income group, we compute the average dividend earned by non-zero households. We do this by dividing the total amount of dividends for each income group by the respective number of non-zero households.

(4) We assume that each household in the same income group earned the average dividend computed in (2).

---

[9] For all analyses at the MSA level we execute this and other instructions below (that mention execution at the county level) at the MSA level.

(5) We assume that each household owns the same type and composition of stock: the equity index. As a result, the amount of dividend received depends only on the quantity of stock owned.

(6) We use the number of non-zero households belonging to each income group, the number of households declaring zero dividends, and the average interests and dividend earned to compute a Gini coefficient that measures the distribution of dividend earnings within each county. Recall that the Gini coefficient is a measure of inequality that ranges between 0 and 1, where a coefficient close to 0 can be interpreted as full equality, whereas a coefficient of 1 indicates perfect inequality. We perform the same procedure with the interest income data.[10]

Table I 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 non-zero dividend, and Column 4 the sum of all dividends received by all households in each group.

First, we compute the number of non-dividend earners by taking the difference between the total of Columns 1 and 2, which we report in first row of Column 5. In the remainder of Column 5, we place the number of households that declared dividends, the same as in Column 3. We then compute the average dividend earned by non-zero households by dividing Column 4 by Column 3; we report this ratio in Column 6. We then compute the Gini coefficient, using the six dividend income groups. The first one consists of the 1,576,927 households that earned zero dividends, the second contains the 31,604 households that earn 1,181 dollars and up to

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

the sixth group composed by 73,620 households that earn about 11,800 dollars in dividends. In this example, the Gini coefficient is equal to 0.91.

[Table I around here]

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 all 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.[11] We will present our main results using a Gini coefficient based on dividends. Results are basically the same if we use a Gini coefficient based on interest income.

To construct our historical measure of wealth inequality we obtain information on historical farm land sizes at the county level from the 1890 US Census. More precisely, for each county we have information on the total number of farms that – based upon their total

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

acres of farm land – fall in 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 this bin (for the first bin we set the lower bound equal to 0.001). Next, we use these lower bounds to calculate a county Gini coefficient in a similar way as in Rajan and Ramcharan (2011).

Notice that we are unable to calculate a Gini coefficient for those counties that became incorporated after 1890, as the information on 1890 farm size distribution is unavailable. For these counties we manually look up the 1890 counties which these missing counties were then still part of before their own incorporation and take (simple) averages of the corresponding Gini coefficients. As the entire State of Oklahoma was incorporated well after 1890 (in 1907) we only have information on those eight counties that existed when it was still a territory. Based upon the information from these counties we construct a State Gini coefficient which we use for all counties in Oklahoma. To calculate this State Gini coefficient we sum the number of farms in each bin across the counties.[12]

In our dataset the average Land Gini coefficient is 0.44 and its standard deviation is 0.14. This is slightly lower compared to other contemporary measures of household wealth inequality at the aggregated level (contemporary measures of household wealth inequality at the county level do not exist). For example, De Nardi (2004) shows that the Gini coefficient for the entire US is 0.78 based upon household wealth data from the Survey of Consumer

[12] Results are unaffected if we exclude Oklahoma from the analysis.

Finances from 1989. Relying on the same survey, Wolff (2010) finds that the Gini coefficient is 0.83 in 2007.

We also find that 1890 land inequality display a 36 and 46 percent positive correlation with our measures of dividend and interest inequality. The historical measure is also correlated with other contemporary socio economic measures that may reflect the degree of wealth inequality. It displays a positive correlation with local poverty rates (43 percent) and the number of crimes per capita (33 percent) and it is negatively correlated with the number of white people (a rough proxy of the size of the middle class) living in a county (-53 percent).

III. Firm Creation, Type, Ownership and Financing

A. Data Sources with Information on Firm Creation, Type and Financing

Data on establishment entry and exit are obtained from the Business Dynamics Statistics (BDS) a database set up by the US census that provides annual measures of, amongst other things, establishment births and deaths, and firm startups and shutdowns. The BDS data is available only at the US MSA level and provides information from 1976 until 2011. The BDS also lists the physical location of establishments instead of States of incorporation, circumventing issues such as higher incorporation rates in Delaware. Like in Kerr and Nanda (2009) entrepreneurship is defined as entry of new, stand-alone firms.

From the Kauffman Firm Survey (KFS) panel dataset we obtain the financial information for a five-year period from 2004 to (and including) 2008 on 4,928 individual US start-ups during their early years of operation (see Robb and Robinson (2014) for a comprehensive discussion of the capital structure choices of firms covered by this survey). This information is particularly useful to reconstruct the sources of financing of these young firms and allows us to distinguish between family, bank and venture capital financing. Data on our main dependent variables for our capital structure regressions is obtained from a restricted-access-only database, which is the so-called "Fourth Follow-Up Database" and which is a longitudinal survey. We analyze the 3,419 firms of the baseline survey that either survived over the entire 2005-2008 period or were specifically identified as going out of business during the same period. Hence, firms that dropped out in a specific year because their owners cannot be located or refuse to respond to the follow-up survey are not included in our analysis.[13] The dataset contains response-adjusted weights (which we use) to minimize the potential non-response bias in the estimates. From this database, we construct several crucial financial outcome variables, as well as control variables in the form of firm and main owner characteristics.

B. Dependent Variables

We study the impact of county and MSA inequality on, in total, six corporate outcome variables. Table II collects all definitions of all these dependent variables, and also of all controls, and indicates the relevant data sources. Table III provides summary statistics.

[Tables II and III around here]

For the purpose of summarizing we categorize the dependent variables into business formation, firm type, firm ownership and firm financing, though we recognize this categorization is not entirely descriptive. In terms of business formation we capture *MSA*

---

[13] This is common in the literature see e.g. Berger, Cerqueiro and Penas (2014) and Robb and Robinson (2014).

Total Establishment Entries, the yearly total number of new establishments in an MSA and *MSA Total Establishments Exits*, which is the yearly total number of establishments that became inactive in an MSA (the estimates for these specifications will be reported in Tables IV and VI respectively). On average around 459 new establishments are set up each year, whereas a slightly higher number of establishments exits the market yearly, 539 on average.

As the sole firm type variable we feature: *Firm is High Tech* which equals one if the firm operates in a high technology industry, and equals zero otherwise. We classify firms as being high technology intensive (*High Tech*) based upon the High Technology Industries NAICS list from the Science and Engineering Indicators 2006 from the NSF. Based upon this classification 31 percent of the firms are considered as being high tech at inception.

Concerning firm ownership, we feature the following variable (the estimates for this specification will be reported in Table V and VII): *Firm Is Proprietorship* equals one if the firm is a proprietorship, and equals zero otherwise;[14] From the 3,419 firms in our dataset 1,294 firms (38 percent) are proprietorships at the start of the sample period.

As firm financing variables we feature (the estimates will be in Table V and VII as well): *Firm Family and Bank Financing* is the total amount of business and owners' personal bank financing plus family financing divided by total firm financing; *Firm Angel and Venture Capital Financing* is the amount of equity obtained from angels and venture capitalists divided by total firm financing.[15]

---

[14] We assign the value of one to the variable *Firm is Proprietorship* if a firm is either a sole proprietorship or a limited partnership. We assign all other firms the value of zero, i.e., these are firms that are classified as a limited liability firm, subchapter S-Corporations, C-Corporations, general partnerships or any other legal form. Alternatively defining a general partnership as a form of proprietorship does not change the results.

[15] Family Financing is the amount of equity invested by parents and/or spouse. In the main analysis, we add bank and family financing in the same dependent variable. In additional tests, we consider the two sources of financing as separate entities. It is evident from the descriptive statistics that a small percentage of firms rely on

Firms in our sample tend to rely little on equity obtained from angels and venture capitalists; only 4 percent of the firms make use of this type of financing at start-up. On average only 2 percent of total funding is obtained through these sources. The mean amount of equity from angels and venture capitalists at inception equals around $37,500. Contrary to equity obtained from angels and venture capitalists, start-ups tend to rely more on debt financing in the form of bank and family financing (on average 11 percent of total financing comes from both business and owner's personal bank and family loans).[16]

C. Control Variables

We have three sets of control variables, i.e., firm characteristics, main owner characteristics, and state and county characteristics. We discuss each of these sets of control variables now in turn. For the analysis at the MSA level (for business formation dynamics) in tables IV, VI and VIII we include a set of MSA characteristics similar to the county characteristics included in the remaining analyses.[17] For the sake of brevity we do not discuss them.

---

venture capital financing as well as bank financing. This generates many "zeros" in some of our dependent variable. To verify the robustness of our results in respect to this issue, we run two additional tests. First, we restrict our sample only to companies that rely on some form of external finance; this almost mechanically reduces the number of zeros by dropping from the sample firms that only rely on internal equity. Second, we keep the whole sample, but we transform the variables "bank and family financing" and "angels and venture capital financing" in (0/1) binary variables. We then run a rare event logit model (as in King and Zeng (2001)) and see if the explanatory power of our inequality measures still holds. In both robustness checks, inequality maintains the statistically and economic significance we observe in our main results.

[16] The largest part, i.e., 7 percent, comes from personal loans obtained by the owner(s). The average amount of bank financing that firms rely upon in our sample is $28,277 at inception.

[17] More specifically, the MSA characteristics included are: *Population, Wage Inequality, Catholic to Protestant ratio, Personal Income per Capita, Federal Government Expenditures per Capita, Land Area and the number of Whites to Total Population.*

1. *Firm Characteristics*

Total Assets is the logarithm of one plus total assets, which is the sum of cash, accounts receivable, product inventory, equipment or machinery, land and buildings, vehicles, other business owned property and other assets; *ROA* is the Return on Assets, i.e., the amount of net profit divided by total assets, which we winsorize at the 1 percent level; *Tangibility* is the amount of property, plant and equipment divided by total assets; and, *Number of Owners* is the logarithm of one plus the total number of owners.

In their first year of operations the start-ups in our dataset have total assets worth of, on average, $172,709. Some firms have a negative return on assets. For example, in the first year of operations 54 percent of all start-ups have a *ROA* below zero percent. Tangible assets make up around 56 percent of total assets on average. However, 26 percent of the businesses report no tangible assets at all in the first year of business. The majority of firms (61 percent) have one owner, whereas the remaining 39 percent of businesses is owned by multiple owners.

2. *Main Owner Characteristics*

The main owner characteristics comprise five dummies that equal one if the condition embedded in its label is fulfilled, and equals zero otherwise. These main owner dummies are: *Is Female*, *Is African-American*, *Is Hispanic*, *Is Asian* and *Is Born in the US*. *Work Experience* is the number of years of work experience of the main owner in the firm's industry.

We identify the main owner in the same way as Robb and Robinson (2014), who consider the owner with the largest amount of equity invested to be the primary owner.[18] Overall, entrepreneurs of the nascent businesses in our sample are mostly white (only 15 percent is

[18] See their paper for the exact methodology on how to define the primary owner in case multiple owners invest an equal amount of equity into the firm.

either African-American, Hispanic or Asian), male and born in the US. This is consistent with the owner characteristics of firm owners from the Survey of Consumer Finances (SCF) (see Puri and Robinson (2007)). Less than a third (27 percent) of the main owners is female, and only 9 percent is born somewhere outside of the United States. In addition, primary owners tend to have quite some work experience in the same industry as their new business is operating in, the average years of experience is a little less than 14 years (median of 11 years).

3. State and County Characteristics

As State characteristics we include its *GDP* which is the logarithm of one plus the gross domestic product of the State during the year.

As county characteristics we have: *Population* which is the total county population at year-end (one year lag); the *Catholic to Protestant Ratio* which is the (natural logarithm of the) ratio of the total number of Catholics divided by the total number of Evangelicals in the county at year-end 2000; the *Personal Income Per Capita* which is the logarithm of one plus the per capita county personal income at year-end; the *Nonfarm Establishments Per Capita* which is the total number of nonfarm establishments divided by the total population in the county at year-end; *Wage Inequality* which is the Gini coefficient of wages earned in each US-county coming from the IRS-SOI data; the *Federal Government Expenditures Per Capita* which is the total Federal government expenditures in thousands of US Dollars during the year in the county divided by the total population in the county; and the *Land Area* which is the logarithm of one plus the total county area in square miles at year-end 2000.[19]

---

[19] As we control both for population and land, we implicitly control for population density. Controlling directly for population density does not change our results.

We include *GDP* to control for the state of the local economy and in the same line for per capita county personal income; their means are respectively 10.65 and 10.48. With respect to county demographics, between counties there is considerable heterogeneity when it comes to religion, although on average Catholics outweigh Protestants by a factor of four. County federal government expenditures per capita differ quite a lot between counties as well; the 10th percentile (of its logarithm) is 4 whereas the 90th percentile is almost three times higher. Additionally, counties' Wage Inequality, as measured by the Gini coefficient based on incomes, is quite high, with a mean value of 0.55, but differs less substantially between counties: Its standard deviation is 0.04.

D. Results

1. Business Formation

We start by testing how local wealth inequality affects business formation dynamics in US *Metropolitan Statistical Areas* (MSA's). Table IV provides the first estimation results. We relate local inequality to the yearly number of new establishments as well as the total number of establishments that become inactive in a given year. We measure local inequality at the MSA level by creating a contemporary Gini coefficient using information on households' wealth in the form of dividends obtained from the IRS tax filings from 2004. Secondly, we introduce an inequality measure that is based upon historical farm land data from 1890 in robustness analysis.

To be able to make causal statements we account for possible endogeneity in several ways. To account for omitted variables because of unobserved heterogeneity at the state, year and state -year level that could affect our estimates, we introduce correspondingly broad sets of fixed effects. Moreover, we instrument our contemporary inequality measure using information on average rainfall and temperatures in the spirit of Easterly (2007) and Rajan and Ramcharan (2011).

[Table IV around here]

We report the results in Table IV, using our contemporary measure of inequality based on households' financial wealth. In Column (1) in Table IV we start with our baseline estimation using a standard Ordinary Least Squares model. We include state and year fixed effects and in addition control for MSA population. In Column (2) we repeat the estimation but additionally include an extra set of MSA characteristics. Alternatively we include in specifications (3) and (4), next to the MSA characteristics also the percentage of white individuals in a given MSA, a measure of Wage Inequality, Federal Government Expenditures per Capita and state-year fixed effects, respectively. In all these specifications we want to know if local inequality matters for new establishment entries.

In all specifications we find, in line with Hypothesis 1 (H1), that the number of new establishment entries in an MSA decreases in inequality. The point estimates decrease slightly when adding a full set of MSA control variables in Column (3) but remain quite similar in Columns (2) and (4). The effect we find is also economically significant and stable across specifications: A one standard deviation increase in MSA wealth inequality decreases the number of new establishment entries by around 22 percent.

That local inequality matters for business formation dynamics is also confirmed in Columns (5) to (8), where we relate local inequality to the yearly number of establishments that become inactive (natural logarithm). Similar to the new establishment entries regressions we include in Column 5 state and year fixed effects and control for MSA population. In Column 6 we add an additional set of MSA characteristics, whereas in Columns 7 and 8 we include the percentage of white individuals in an MSA, a measure of Wage Inequality, Federal Government Expenditures per Capita and state-year fixed effects, respectively. The economic relevancy is practically similar: A one standard deviation increase in MSA wealth inequality decreases the number of establishments that become inactive by around 37 percent, indicating that business formation (i.e. establishments' entries and exits) is less dynamic in more unequal metropolitan areas. In unreported results, we also check whether business exit has a different pattern depending on the age of the firm.

2. *Firm Type*

In the previous tables we find evidence in support of less creative destruction activity, in the form of less new establishment entries on one hand but also a lower amount of establishments that become inactive on the other hand, in more unequal MSA's. But does inequality also affect the type of firm formation? We present results on this possibility in Table V, where we examine the effect of wealth inequality on firm type (H2), firm ownership (H3) and firm financing (H4). To examine whether local county inequality affects firm type choices, we construct a dummy variable that indicates whether a Firm is High Tech or not and run a Linear Probability model in Columns (1) to (3). As the definition of High Tech is based on the NAICS industry classification, we do not include industry fixed effects and trends in these regressions. We do, however, in the different specifications include firm, owner and county controls and broad sets of fixed effects, i.e., state, year and state*year fixed effects, respectively. These regressions, as well as the following capital structure regressions, use measures of inequality at the US county level. Results are similar when we use MSA-measures of inequality.[20]

[Table V around here]

Interestingly, the coefficient on county inequality is always negative and statistically significant at the 5 percent level, indicating that the likelihood that newly created firms are high-tech decreases in local inequality. The effects are also economically relevant. Depending upon specification, a one standard deviation increase in inequality decreases the likelihood that a Firm is High Tech between 8 and 9 percent of its mean when introducing a contemporary inequality measure. In unequal societies entrepreneurs themselves may prefer simpler technologies or they may find it difficult to finance riskier ventures.

The results are suggestive of the possibility that banks are more willing to extend financing to conservative industries, making it difficult for new start-ups in more unconventional industries (such as high tech industries) to obtain bank financing in order to start up their business. Another explanation may be demand driven: In more unequal counties entrepreneurs may prefer simpler technologies themselves. We will explore possible institutional mechanisms in more detail in Section V.

It is also important to notice that including controls in our regressions leaves the relationship between wealth inequality and entrepreneurial outcomes unaltered or, in many

[20] When possible, we prefer using county measures of inequality. This is for two reasons: (1) Many of the institutional activities taking place in the local areas such as schooling and the judiciary are organized at the county level; (2) By estimating at a more local level and "closer to the firms", county level measures of inequality allow us to obtain a better control of omitted factors.

cases, makes it ultimately stronger rather than weaker. To the extent that firms and state/county unobservable characteristics are correlated to our controls, it appears that endogeneity works against finding any link between wealth inequality, start-up capital structure and technology choices (Altonji, Elder and Taber (2005); Bellows and Miguel (2006); Bellows and Miguel (2009)).

3. *Firm Ownership*

We continue by testing how local wealth inequality affects firm ownership as measured by the corporate form in which the firm is set up. Table V, Columns (4) to (6) provide the first estimation results for the dependent variable *Firm Is Proprietorship*. We relate local inequality to the probability that a start-up is a sole proprietorship.

In Column (4) we start with our baseline estimation using a standard Linear Probability model.[21] We include state and year fixed effects and control for county population. In Column (5) we repeat the estimation but additionally include extra sets of firm and county characteristics. Alternatively we include in specification (6) industry, state-year and industry-year fixed effects, respectively. We see that local inequality matters. In all specifications we find, in line with Hypothesis 3 (H3), that the probability for a start-up to be a sole proprietorship increases in inequality. The point estimates increase slightly across Columns (5) and (6). The effect we find is also economically significant and stable across specifications: A one standard deviation increase in county wealth inequality increases the

[21] We prefer a linear probability model because the large number of fixed effects we eventually want to introduce may bias the maximum likelihood estimates due to the incidental parameter problem (for a review see e.g. Lancaster (2000)). In robustness we also run specifications using logit and probit models but the point estimates of our main independent variables as well as their statistical significance actually do remain mostly unaltered.

probability for a start-up to be a sole proprietorship by around 14 percent (evaluated at the mean of the Proprietorship indicator variable).

4. Firm Financing

So far the results overall demonstrate that local wealth inequality matters for business formation dynamics, firm type and firm ownership as measured by its corporate form. In this section and based on the models in recent papers by Modigliani and Perotti (2000) and Perotti and von Thadden (2006), we ask whether it also matters for the way in which start-ups are financed. To the extent that more unequal counties are characterized by less efficient capital markets, poorer institutions and law enforcement entrepreneurs may rely more on internal financing from family and friends as well as bank financing, given that bank debt contracts are relatively easy to enforce, but less on outside equity obtained from angel investors and venture capitalists (see Modigliani and Perotti (2000) for the theoretical background). Columns (7) to (12) in Table V present the results of our OLS analyses.[22] We focus on the proportion of bank and family financing in Columns (7) to (9), controlling for the usual firm, owner and county controls as well as dense sets of fixed effects to capture unobserved state, year, industry, state-year and industry-year heterogeneity, respectively.

The results in Column (7) show that, in line with H4, the coefficient on county inequality is positive although not statistically significant. The result indicates that a one standard deviation increase in county inequality increases the proportion of bank and family financing with 7

---

[22] We also re-estimate all specifications using a Tobit model. Maximum likelihood estimators of marginal effects in Tobit models are found to be overall much less biased due to the incidental parameter problem (than those in binary dependent variable models); but when many fixed effects are introduced, and expected biases in the slope estimators (in terms of marginal effects) do emerge, it is away from zero; at the same time the estimated standard errors may be biased towards zero (Greene (2004)). We therefore report the results from using OLS.

percent (of its own mean) when using a contemporary inequality measure. When including a set of county controls, firm characteristics as well as industry fixed effects in Column (8) the effect doubles to 14 percent and the coefficient is statistically significant at conventional levels. When we add state-year and industry-year effects in Column (9) the estimated coefficient on inequality is again statistically significant (at the 5 percent level) and implies that a one standard deviation increase in county inequality increases the proportion of bank and family financing again with around 14 percent.

Overall these findings are in line with both supply and demand sides of finance hypotheses. In more unequal societies, bank debt may be easier to enforce in court and, as a result, more commonly supplied. At the same time, in more unequal societies entrepreneurs may prefer to undertake simpler technologies requiring simpler forms of external financing for their businesses.

In the remaining columns of Table V we explore, again using OLS estimation, how county inequality affects the proportion of angel and venture capital financing. The proportion of angel and venture capital financing decreases in county inequality, as we can see from the results in Columns (13) to (15). The coefficient on inequality is negative and statistically significant and robust across specifications, however the economic significance is somewhat larger in the specifications in Columns (10) and (12), when including sets of firm characteristics, county controls, industry, state-year and industry-year effects respectively. A one standard deviation increase in county inequality decreases the proportion of angel and venture capital financing between 26 and 44 percent, depending upon specification. External equity financing is decreasing in county inequality, either because of a lower supply, in line with what described in Perotti and von Thadden (2006), or a lower demand because entrepreneurs prefer more traditional technologies. The results show that local inequality matters for firm financing. The findings confirm (i.e., cannot reject) H4 and are therefore consistent with the papers by Perotti and von Thadden (2006) as well as Modigliani and Perotti (2000).

Chen, Gompers, Kovner and Lerner (2010) show that the distribution of venture capitalists in the US is concentrated in three areas: San Francisco, Boston and New York.[23] We therefore verify whether our results are not simply driven by firms located in these areas by excluding all firms located in the States of California, Massachusetts and New York. We find that results are unaffected (and therefore not tabulated).

5. *Instrumental Variable Analysis*

Even though we take several measures to prevent our estimates to be biased because of unobserved variables that are correlated with both our independent variable of interest, i.e., county wealth inequality, and the dependent variables of interest,[24] we now rule out further any possible bias in our estimates by instrumenting our contemporary 2004 county inequality measure as well.

We use instruments for county wealth inequality in the spirit of Easterly (2007). The instrumental variables are based upon past local weather conditions, i.e., there are based on the historical rainfall and temperature between 1895 and 2003 and their corresponding

---

[23] Chen, Gompers, Kovner and Lerner (2010) reports that more than 49 percent of US based companies financed by venture capital firms are located in one of these three areas. The Kauffmann survey data we use combines venture capital with angel financing. The number of firms with non-zero angel and venture capital financing located in California, Massachusetts and New York corresponds to only 19 percent of the total number of firms, suggesting angel financing may be less concentrated.

[24] Recall that we saturate specifications with state, industry, state-year and/or industry-year fixed effects. In robustness analysis in Section 6 we additionally use an historical measure of county inequality to resolve reverse causality as well.

standard deviations. More precisely, we obtain information from the National Climatic Data Center (NCDC) on local monthly precipitation and temperature (measured in inches and degrees Fahrenheit, respectively) and their corresponding standard deviations for the entire period between 1895 and 2003. We then construct simple averages of these series. The NCDC provides this weather information at the so-called "divisional" level, i.e., each state is subdivided in at most 10 divisions that are comprised of areas that are known to have similar climatic conditions. We assign each county to the state division it belongs to. We proceed in a similar way when constructing the instrumental variables at the MSA level.

Our instrumental variable strategy is also inspired by Engerman and Sokoloff (2002b) who already documented that the degree of inequality is partly determined by the local soil quality. Their findings make the historical local weather conditions suitable instruments for inequality (of course to the extent that inequality is persistent through time) becomes local weather patterns likely fixed crop yields for cotton and tobacco which in turn determined the well-being of a small group of wealthy owners (in the US South). As already shown by Vollrath (2013) and Rajan and Ramcharan (2011), even within States there is a significant amount of diversity in term of temperature and rainfall. In Kansas and Texas for instance some counties experience a yearly rainfall average of 20 inches while others go beyond 40 inches. A bit less extreme but still important are the differences in Illinois, where some counties have an average rainfall of 28 inches while others have 30 percent more (about 36 inches). Similarly, in California some counties had an average temperature of 50 F while others have 64 F.

[Table VI around here]

The *'First Stage'* column in Table VI provides the results of the first stage regression from 2SLS regressions and indicates that indeed rain and temperature are significant determinants of current MSA inequality for the dependent variable *Total Establishment Entries*. All climate variable coefficients, except for the standard deviation of temperature, are statistically significant at the 1 percent level and all enter with the expected sign: Higher rainfall levels and temperatures are associated with higher current MSA inequality, but at a decreasing rate as indicated by the negative signs on the coefficients of their respective standard deviations. In all the specifications the F-statistic of the first stage (not reported) is well above 20, confirming that we have a powerful first stage.

The following columns in Table VI report the second stage regressions for the dependent variables from Hypothesis 1. The results confirm our previous findings: In more unequal MSA's business formation is less dynamic, both the number of new establishment entries as well as the number of establishments that becomes inactive is significantly lower in more unequal MSA's. Again, the results are not only statistically significant but also very economically relevant: A one standard deviation increase in MSA inequality decreases the total number of new establishments that are set up in a given year by 12 percent. Additionally, it also results in a lower number of establishments that become inactive; the number of establishments that become inactive is around 20 percent lower in more unequal MSA's. The results hence indicate that in more unequal areas business formation dynamics are less pronounced.

We turn to the IV results for the dependent variables capturing firm type, ownership and financing in Table VII.

[Table VII around here]

The *'First Stage'* Column in Table VII again shows the results of the first stage regression, indicating that also for US counties in addition to MSA's both rain and temperature are suitable instruments for local wealth inequality. The following columns in Table VII report the second stage regressions for the dependent variables from Hypotheses 2-4 respectively. The results in Columns (1) and (2) indicate that county inequality indeed lowers the probability that a start-up firm is of High Tech nature. In Column (1) the effect is not only statistically significant but also very much economically meaningful: A one standard deviation increase in county inequality lowers the probability that a firm is of high tech nature by 21 percent. When including sets of firm and county characteristics and industry*year and state*year fixed effects in Column (2) however the coefficient is not statistically significant at conventional levels, although the sign remains negative as expected. In the same line the results mostly confirm our previous findings concerning firms' financing: In more unequal counties start-up firms rely more on bank and family financing.

Moreover, start-up firms are more likely to be of a simpler business form, i.e., a proprietorship, and, although the coefficients are not statistically significant at conventional levels when adding sets of county and firm characteristics, they also rely less on angel and venture capital financing and are less likely to be of a more complex high-tech nature. Again, the results are not only statistically significant but also very economically relevant: A one standard deviation increase in county inequality increases the probability that the start-up firm is of a simpler form (proprietorship) by more than half (53 percent increase of its mean) and increases the reliance on family and bank financing by 24 percent (again evaluated at its mean).

6. *Falsification Test*

Our identification strategy implicitly assumes that local inequality has an impact on entrepreneurship via the quality of local institutions. To address the concern of whether the exclusion restriction is satisfied, we perform a falsification test that links local weather conditions to local entrepreneurship in France.

In France local authorities have much more limited power in organizing public life than in the US. In our test, we will take Departments as the French version of the US Counties and the French Administrative Regions as equivalents of the US States. French departments constitute the second of three levels of government below the national government. They are smaller than the 27 administrative regions but larger than towns. The main areas of responsibility of a department include the organization of various welfare allowances, the maintenance of high school buildings and local roads as well as the contribution to municipal infrastructures. Importantly, and differently from the US, French Departments do not have any role in organizing the judiciary, and their administrative activities are supervised by a Prefect, the high representative of the national government.

As we do not have local inequality data for France, we will work with reduced form equations linking local weather conditions to local business entry and exit both for France and the US. We obtain data on establishments' entry and exit between 2006 and 2014 from INSEE, the national statistics office, and data on historical temperature and rainfall from METEO-France, the national meteorology institute.

We run reduced form regressions linking the local degree of firms' entry and exit to local weather patterns and control for Regional and Year Fixed effects. Naturally, weather conditions are more extreme and likely to be more diverse within the same State in the US compared to the French Departments. To control for this problem, we also perform the regression for the US restricting the sample to US MSAs whose average and standard deviation of rainfall and temperature fall within the respective French minimum and maximum averages and standard deviations. We cluster standard errors at the State/Regional level.[25] The results are reported in Table VIII.

[Table VIII around here]

As can be seen from the results in Table VIII, our instrumental variables, the average historical rainfall and temperature as well as their respective standard deviations significantly affect establishment entries and exit levels in the US. Especially when restricting the sample of observation to fall within the minimum and maximum of French rain and temperature means and standard deviations in Columns (3) and (4) this significant effect is very pronounced. On the contrary, there is no significant relationship between our weather variables and business dynamics in France, as can be seen from Columns (5) and (6). All in all, these findings provide support for the validity of the exclusion restriction when using rain and temperature as instrumental variables for local inequality.

[25] Since France has only 21 regions, the low number of clusters may lead to bias towards finding results that are not statistically significant. Results remain the same when we cluster at the department level: In this case we have 83 clusters.

7. *Inequality from 1890*

To rule out any problems of reverse causality we resort to history to obtain a historical measure of local wealth inequality based upon historical farm land data from 1890, which we introduce as a second measure of local inequality in Table IX. In addition, to account for omitted variables because of unobserved heterogeneity at the state, year, industry, state-year and industry-year level that could affect our estimates, we introduce correspondingly a broad sets of fixed effects.

[Table IX around here]

The results confirm our previous findings: Columns (1) and (2) show that business formation is less dynamic in more unequal MSA's, i.e., in more unequal MSA's there is less business formation but also fewer establishments exit. The coefficients are also statistically significant at conventional levels. In line with H2 and H3 the probability that a start-up firm is of high technology nature decreases in county inequality whereas the probability that a new venture is a proprietorship increases in county inequality as can be seen in Columns (3) and (4). Both coefficients are statistically significant and a one standard deviation increase in county inequality alters both probabilities with around 3 percent (evaluated at their respective means). For *Firm Bank and Family Financing* the coefficient on historical county inequality enters significant and with a positive sign, in line with previous findings, in Column (5): A one standard deviation increase in inequality increases financing obtained from banks and family members with a bit less than 5 percent (evaluated at its mean). Remarkably the coefficient on county inequality enters with a positive sign in Column (6), for the dependent variable *Angel and Venture Capital Financing*, contrary to previous findings, but it is not statistically significant.

IV. The Removal of State EIG Taxes and Entrepreneurship Dynamics

To further address endogeneity concerns we exploit changes in Estate Inheritance and Gift taxes that took place in various States between the 1970s and 2000. Since 1976, and at different points in time, 31 States repealed their "Death, Estate and Gift" (EIG) taxes. In particular, States switched from a system where state EIG taxes were a percentage computed on top of the corresponding federal EIG tax to a "pick up" system. In a "pick up" system the State takes on a proportion (hence *picks up*) of the Federal EIG tax applied to its citizens without increasing the total tax burden.

EIG taxes may matter for wealth inequality as they define the amount of wealth that is transferred from one generation to another. In principle, systems with very high EIG taxes should promote more equality, as wealthy parents will not be able to fully transfer their wealth to their children. Conversely, low EIG taxes should promote more inequality as it would be easier to pass wealth from one generation to the next.

These widespread changes in state EIG taxes provide substantial cross-sectional and time series variation which we will exploit in our analysis. In this analysis, we focus only on entry and exit of new firms, as the Kauffman data is available from 2004 onwards, and we estimate the following equation:

$$Y_{j,t} = \alpha + \alpha_{msa} + \alpha_t + \beta Post_{jt} + Controls + \varepsilon_{j,t}$$

Where, $Y_{j,t}$ indicates the natural logarithm Number of Firms Entries and Exists in the Metropolitan Statistical Area *j* in year *t*. The variable of interest is $Post_{jt}$. Post is a dummy variable that takes the value of 1 for the years following the introduction of the so called 'Pick Up' system in the State where the MSA belongs to. Like Kerr and Nanda (2010), we will also study specifications where we substitute $Post_{jt}$, with the number of years since the reform was introduced.[26] Since the data has a panel dimension, we control for MSA and year fixed effects. As the EIG tax reforms are defined at the State level, we cluster the standard errors at the corresponding State level. We consider data between 1976 and 2000. In 2001, the federal government introduced legislation that phased out the pick-up system, hence generating an important confounding event.[27]

We present the results in Table X. In columns (1)-(3) and (7) – (9) the dependent variable is log of the total number of firm entries. The coefficient on the variable post is negative and statistically significant at the 10 percent level, indicating that switching to a pick up system reduces the number of new firms that enter in the MSA. To the extent, that lower EIG taxes promote more wealth inequality, this result is in line with our baseline results that show a negative relationship between local wealth inequality and entrepreneurship dynamics. The economic significance is also sizable, after the introduction of the new pick up system, new business formation is reduced by around 4 percent. This is a sizable effect, considering that the mechanism we attempt to identify relies on economic institutions, which take a long time to change.

---

[26] And like Kerr and Nanda (2010) we normalize the long run to four years.

[27] Most States responded to the Federal legislation re-initiating a State EIG tax equal to the amount of the federal credit as determined by the IRS code as of January 2001.

Similarly, the introduction of the pick-up system tends to decrease the number of firms' exists. The economic significance is slightly lower compared to the economic significance for firms' entry. After the enactment of the pick-up system, the number of exists declines of about 2.5 percent.

V. Wealth Inequality and Local Institutions

Our analysis so far has identified a reduced form relationship between wealth inequality, business formation, firms' technology choices and capital structure. For instance, wealth inequality may result in inefficient financial markets and yield restrictions to the supply of external finance. At the same time, inequality could be associated with a lower demand for external finance because of the entrepreneurs' education level. Less educated entrepreneurs may prefer to work with simpler technologies and more traditional productions that require a lower amount of external finance, and in particular equity finance.

We use data from the US Census to evaluate the relative importance of supply and demand factors in explaining our results. We focus in particular on banking development, on education, and efficiency of the civil justice system. Particularly, we study whether average county education levels depend on the local level of wealth inequality.

As a measure of banking development we follow Rajan and Ramcharan (2011) and use the number of bank establishments per capita. Rajan and Ramcharan (2011) show that in the 1930s US counties displaying more wealth inequality had a significant lower number of bank establishments per capita.

In Table XI we first examine the effect of contemporary county inequality on banking development, as measured by the number of banks per 1,000 Capita (as in Rajan and

Ramcharan (2011)). In Column (1) we include state and year fixed effects, whereas in Column (2) we add a comprehensive set of county characteristics. Column (3) presents the results from the second stage from a 2SLS IV regression where we instrument county inequality with the average historical rainfall, temperature and their standard deviations. The results indicate that inequality indeed hampers banking development: The county inequality coefficient is negative and statistically significant throughout specifications.

The results are also economically meaningful: A one standard deviation increase in inequality decreases the number of banks per 1,000 capita for example by around 9 percent of its mean in Columns (1) and (2) respectively.

[Table XI around here]

Turning to education as another institutional feature of the local environment, we present the results in Columns (4) to (9), again using a county inequality measure based upon households' financial wealth from 2004. In Columns (4) to (6) the analysis shows that in more unequal counties the percentage of adults with a college degree or higher is lower. In fact, a one standard deviation increase in county inequality decreases the percentage of adults with a college degree or more between 16 to 45 percent of its mean, depending upon specification. Moreover, the population inflow of educated individuals (i.e., those with at least a college degree) is also lower in more unequal counties, as can be seen from Columns (7) to (9). A one standard deviation increase in inequality results in a significant lower inflow of educated individuals. The economic effect is also sizable however: A decrease between 30 percent and 50 percent evaluated at the mean of county inflow. In Columns (10) – (12) we assess the effect local wealth inequality has on another local institution: The judiciary. In specific we investigate how local inequality affects the judicial efficiency as measured by the length of time to verdict for a first degree civil trial. For this aim we obtain data on individual civil cases from the Bureau of Justice Statistics (BJS) from 2005. At the individual case level we see whether the length in terms of the number of days (its natural logarithm) it takes to come to a verdict in a case is affected by local inequality. We see that local inequality matters for judicial decision making. When controlling for both state fixed effects and case controls (in the form of the number of plaintiffs involved as a proxy for case complexity and whether or not an appeal was granted as well as the total number of cases in a certain court) the findings suggest that in more unequal counties court rulings take more time and therefore are less efficient.

Overall, it can be concluded that the results point in the direction that inequality also affects the quality of local institutions in the form of banking development, education and judicial efficiency.

VI. Conclusions

We empirically test hypotheses emanating from recent theory showing how household wealth inequality may determine both entrepreneurial dynamism and corporate financing. Local wealth inequality may be associated with poorer institutions leading entrepreneurs to choose simpler corporate forms for their businesses and to rely on bank and family finance.

To test the hypotheses emanating from these frameworks we employ two measures of wealth inequality: One based on the current distributions of dividends and another one that relies on the distribution of land holdings within US Counties in 1890. To overcome endogeneity problems, we saturate specifications with comprehensive sets of fixed effects and characteristics and we estimate instrumental variable models. Additionally we exploit the removal of Estate, Inheritance and Gift (EIG) taxes in various States between the 1970s and 2000 in a difference-in-differences framework.

The estimated coefficients suggest that local-level wealth inequality robustly decreases firm creation, in particular of the high-tech type, and it also decreases firm exit. At the same time inequality increases sole-ownership and the proportion of equity, family and bank financing, yet decreases angel and venture capital financing.

The effects of wealth inequality on entrepreneurs' financing and technology could be mediated by various factors. We find that in more unequal counties, there are fewer bank establishments per capita consistent with the existence of local credit rationing. But we also find that in more unequal counties entrepreneurs are less likely to have a college degree (or higher). In principle, these are the entrepreneurs that could be more likely to work with traditional technologies that require simpler forms of financing. Moreover, our findings also suggest the presence of a less effective judiciary in more unequal counties.

References

Acemoglu, D., S. Johnson, and J. A. Robinson, 2001, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91, 1369-1401.

—, 2002, "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics* 117, 1231-1294.

Acemoglu, D., and J. A. Robinson, 2013, *Why Nations Fail. The Origins of Power, Prosperity and Poverty*, Random House, New York NY.

Altonji, J., T. Elder, and C. Taber, 2005, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy* 113, 151-184.

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, 1637-1667.

Bellows, J., and E. Miguel, 2006, "War and Institutions: New Evidence from Sierra Leone," *American Economic Review* 96, 394-399.

—, 2009, "War and Local Collective Action in Sierra Leone," *Journal of Public Economics* 93, 1144-1157.

Berger, A. N., G. Cerqueiro, and M. F. Penas, 2014, Market Size Structure and Small Business Lending: Are Crisis Times Different from Normal Times?, University of South Carolina, Columbia SC.

Berkowitz, D., and K. B. Clay, 2011, *The Evolution of a Nation: How Geography and Law Shaped the American States*, Princeton University Press, Princeton NJ.

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, 2807-2834.

Carlino, G., and W. R. Kerr, 2015, "Agglomeration and Innovation," in G. Duranton, J. V. Henderson, and W. C. Strange, eds, *Handbook of Regional and Urban Economics Volume 5A*, Elsevier, Amsterdam.

Chen, H., P. A. Gompers, A. Kovner, and J. Lerner, 2010, "Buy Local? The Geography of Successful Venture Capital Expansion," *Journal of Urban Economics* 67, 90-102.

Chetty, R., N. Hendren, P. Kline, and E. Saez, 2014, Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States, National Bureau for Economic Research, Cambridge MA.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner, 2014, Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility?, National Bureau for Economic Research, Cambridge MA.

Conway, K. S., and J. C. Rork, 2004, "Diagnosis Murder: The Death of State Death Taxes," *Economic Inquiry* 42, 537-559.

De Nardi, M., 2004, "Wealth Inequality and Intergenerational Links," *Review of Economic Studies* 71, 743-768.

Degryse, H., T. Lambert, and A. Schwienbacher, 2013, 2013, The Political Economy of Financial Systems: Evidence from Suffrage Reforms in the Last Two Centuries, CEPR, London.

Demirgüç-Kunt, A., and R. Levine, 2009, "Finance and Inequality: Theory and Evidence," *Annual Review of Financial Economics* 1, 287-318.

Easterly, W., 2007, "Inequality Does Cause Underdevelopment: Insights from a New Instrument," *Journal of Development Economics* 84, 755-776.

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 S. Haber, ed., *How Latin America Fell Behind*, Stanford University Press, Stanford.

Engerman, S. L., and K. L. Sokoloff, 2002a, Factor Endowments, Inequality, and Paths of Development Among New World Economies, National Bureau of Economic Research, Cambridge MA.

Engerman, S. L., and K. L. Sokoloff, 2002b, "Factor Endowments, Inequality, and Paths of Development among New World Economies," *Journal of LACEA Economia* 3, 41-109.

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, 143-179.

Glaeser, E., J. Scheinkman, and A. Shleifer, 2003, "The Injustice of Inequality," *Journal of Monetary Economics* 50, 199-222.

Greene, W., 2004, "The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects," *Econometrics Journal* 7, 98-119.

Guiso, L., P. Sapienza, and L. Zingales, 2004, "Does Local Financial Development Matter?," *Quarterly Journal of Economics* 119, 929-970.

Kerr, W. R., and R. Nanda, 2009, "Democratizing Entry: Banking Deregulations, Financing Constraints, and Entrepreneurship," *Journal of Financial Economics* 94, 124-149.

—, 2010, "Banking Deregulations, Financing Constraints, and Firm Entry Size," *Journal of the European Economic Association* 8, 582-593.

Kerr, W. R., R. Nanda, and M. Rhodes-Kropf, 2014, "Entrepreneurship as Experimentation," *Journal of Economic Perspectives* 28, 25-48.

King, G., and L. Zeng, 2001, "Logistic Regression in Rare Events Data," *Political Analysis* 9, 137-163.

Lancaster, T., 2000, "The Incidental Parameter Problem since 1948," *Journal of Econometrics* 95, 391-413.

Lerner, J., 1995, "Venture Capitalists and the Oversight of Private Firms," *Journal of Finance* 50, 301-318.

Mian, A., K. Rao, and A. Sufi, 2013, "Household Balance Sheets, Consumption, and the Economic Slump," *The Quarterly Journal of Economics* 128, 1687-1726.

Modigliani, F., and E. Perotti, 2000, "Security Markets versus Bank Finance: Legal Enforcement and Investors' Protection," *International Review of Finance* 1, 81-96.

Nunn, N., 2008, "The Long-Term Effects of Africa's Slave Trades," *Quarterly Journal of Economics* 123, 139-176.

Perotti, E., and E.-L. von Thadden, 2006, "The Political Economy of Corporate Control and Labor Rents," *Journal of Political Economy* 114, 145-174.

Petersen, M. A., and R. G. Rajan, 2002, "Does Distance Still Matter? The Information Revolution in Small Business Lending," *Journal of Finance* 57, 2533-2570.

Piketty, T., 2014, *Capital in the Twenty-First Century*, Harvard University Press, Cambridge MA.

Puri, M., and D. T. Robinson, 2007, "Optimism and Economic Choice," *Journal of Financial Economics* 86, 71-99.

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, 1895-1930.

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, 729-744.

Robb, A. M., and D. T. Robinson, 2014, "The Capital Structure Decisions of New Firms," *Review of Financial Studies* 27, 153-179.

Saez, E., and G. Zucman, 2014, Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data, National Bureau of Economic Research, Cambridge MA.

Schumpeter, J., 1934, *The Theory of Economic Development*, Transactions Publishers, New Brunswick.

Smith, A., 1776, *An Inquiry into the Nature and Causes of the Wealth of Nations*, W. Strahan and T. Cadell, London.

Sonin, K., 2003, "Why the Rich May Favor Poor Protection of Property Rights," *Journal of Comparative Economics* 31, 715-731.

Sorenson, O., and T. E. Stuart, 2001, "Syndication Networks and the Spatial Distribution of Venture Capital Investments," *American Journal of Sociology* 106, 1546-1588.

Vollrath, D., 2013, "Inequality and School Funding in the Rural United States, 1890," *Explorations in Economic History* 50, 267-284.

Vollrath, D., O. Galor, and O. Moav, 2009, "Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence," *Review of Economic Studies* 76, 143-179.

Wolff, E. N., 2010, Recent Trends in Household Wealth in the United States: Rising Debt and the Middle-Class Squeeze - An Update to 2007, Levy Economics Institute.

TABLE I

EXAMPLE COUNTY INEQUALITY 2004 CONSTRUCTION

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income Group Category by Size of Adjusted Gross Income and Zip-Code	Total No. of Returns	Taxable Dividends: No. Of returns	Taxable Dividends: Total Amount Reported	No. of Returns	Average Dividend per Household	Taxable Interest: No. of Returns	Taxable Interest: Total Amount Reported
Total	1,882,964	306,037	1,287,291	1,576,927	0.00	651,013	1,189,469
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

NOTES. The table provides an example of the data used to construct our *County Inequality* 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 upon 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', column (4) and (5) respectively. We use these average dividends as well as the income group classification to construct a Gini index 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 anlysis in which we introduce the county inequality measure based upon dividends received. To construct our *MSA Inequality* measure from 2004 we follow the same procedure but aggregate the zipcode data to the corresponding metropolitan statistical area (MSA) as opposed to the county.

TABLE II

VARIABLE NAMES, DEFINITIONS, AND DATA SOURCES FOR THE EMPIRICAL ANALYSIS OF FIRM OWNERSHIP, FINANCING AND TYPE

Variable Name	Variable Definition	*Source*
Dependent Variables		
MSA Total Establishment Entries	The logarithm of the yearly total number of new establishments in the MSA between 2005-2012	*USC*
MSA Total Establishment Exits	The logarithm of the yearly total number of establishments that became inactive in the MSA between 2005-2012	*USC*
Firm is High Tech	= 1 if firm operates in a high technology industry, = 0 otherwise	*NSF*
Firm Is Proprietorship	= 1 if firm is a proprietorship, = 0 otherwise	*KFS*
Firm Bank and Family Financing	The amount of business and owners' personal bank financing and the amount of equity invested by parents and/or spouse divided by total firm financing	*KFS*
Firm Angel and Venture Capital Financing	The amount of equity obtained from angels and venture capitalists divided by total firm financing	*KFS*
Main Independent Variables		
County Inequality in 2004	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 county	*IRS*
MSA Inequality in 2004	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 (MSA)	*IRS*
County Inequality in 1890	The Gini coefficient of the distribution of farm land in 1890 in the county (for counties in Oklahoma the state-level coefficient is used)	*USC*
MSA Inequality in 1890	The Gini coefficient of the distribution of farm land in 1890 in the metropolitan statistical area (MSA)	*USC*
Instrumental Variables		
Rain	The average district precipitation between 1895-2003, where a district is defined as a group of clustered counties with similar climatic conditions	*NCDC*
Temperature	The average district temperature in degrees between 1895-2003, where a district is defined as a group of clustered counties with similar climatic conditions	*NCDC*
Control Variables		
Firm Characteristics		
Firm Total Assets$_{t-1}$	The logarithm of one plus total assets, which is the sum of cash, accounts receivable, product inventory, equipment or machinery, land and buildings, vehicles, other business owned property and other assets	*KFS*
Firm ROA$_{t-1}$	Return on Assets, i.e., the amount of net profit divided by total assets winsorized at the 1% level	*KFS*
Firm Tangibility$_{t-1}$	The amount of property, plant and equipment divided by total assets	*KFS*
Firm Number of Owners$_{t-1}$	The logarithm of one plus the total number of owners	*KFS*
Main Owner Characteristics		
Main Owner Is Female	= 1 if main owner is a female, = 0 otherwise	*KFS*
Main Owner Is African-American	= 1 if main owner is African-American, = 0 otherwise	*KFS*
Main Owner Is Hispanic	= 1 if main owner is Hispanic, = 0 otherwise	*KFS*
Main Owner Is Asian	= 1 if main owner is Asian, = 0 otherwise	*KFS*
Main Owner Is Born in the US	= 1 if main owner was born in the US, = 0 otherwise	*KFS*
Main Owner's Work Experience	Number of years of work experience of the main owner in the firm's industry	*KFS*
State and County Characteristics		
State GDP$_{t-1}$	The logarithm of one plus the gross domestic product of the state during the year	*USC*
County Population	Total county population at year-end	*USC*
County Catholic to Protestant Ratio	Ratio of the total number of Catholics divided by the total number of Evangelicals in the county at year-end 2000	*ARDA*
County Personal Income Per Capita	The logarithm of one plus the per capita county personal income at year-end	*BEA*
County Nonfarm Establishments Per Capita	Total number of nonfarm establishments divided by the total population in the county at year-end	*USC*
County Wage Inequality	The Gini coefficient of the distribution of wages as measured by the distribution of the amount of labor from household tax filings in the county	*IRS*
County Federal Government Expenditures Per Capita	Total Federal government expenditures in thousands of US Dollars during the year in the county divided by the total population in the county	*USC*
County Land Area	The logarithm of one plus the total county area in square miles at year-end 2000	*USC*

NOTES. The table defines the variables used in the empirical analysis of firm ownership, financing and type, as well as the corresponding data sources used. Total firm financing is the sum of total debt and equity financing. *t-1* indicates a one year lag is used in the empirical analysis. For the sake of brevity we do not report the MSA characteristics separately. Data sources include: *ARDA* = Association of Religion Data Archives; *BEA* = Bureau of Economic Analysis; *IRS* = Internal Revenue Service; *KFS* = Kauffman Firm Survey; *NCDC* = National Climatic Data Center; *NSF* = National Science Foundation; *USC* = US Census.

TABLE III

DESCRIPTIVE STATISTICS FOR THE EMPIRICAL ANALYSIS OF BUSINESS FORMATION, FIRM OWNERSHIP, FINANCING AND TYPE

Variable Name	Number of Observations	Mean	Standard Deviation	10%	Median (50%)	90%
Dependent Variables						
MSA Total Establishment Entries	3,296	6.13	1.18	4.79	5.92	7.72
MSA Total Establishment Exits	3,296	6.29	1.18	5.13	6.05	7.95
Firm is High Tech	15,328	0.31	0.46	0	0	1
Firm Is Proprietorship	14,051	0.35	0.48	0	0	1
Firm Bank and Family Financing	10,540	0.11	0.25	0.00	0.00	0.50
Firm Angel and Venture Capital Financing	7,229	0.02	0.11	0.00	0.00	0.00
Main Independent Variable						
County Inequality in 2004	13,875	0.85	0.05	0.79	0.85	0.90
MSA Inequality in 2004	3,128	0.81	0.16	0.43	0.86	0.92
County Inequality in 1890	13,908	0.44	0.14	0.28	0.42	0.64
MSA Inequality in 1890	3,080	0.43	0.13	0.27	0.41	0.61
Instrumental Variables						
Rain	12,757	3.04	1.08	1.35	3.16	4.35
Temperature	11,787	54.73	8.45	45.12	52.89	68.92
Control Variables						
Firm Characteristics						
Firm Total Assets	14,015	9.41	3.71	1.79	10.23	12.91
Firm ROA	12,016	0.26	2.26	-0.91	0.04	1.67
Firm Tangibility	12,602	0.56	0.37	0.00	0.64	1.00
Firm Number of Owners	14,039	0.91	0.40	0.69	0.69	1.39
Main Owner Characteristics						
Main Owner Is Female	14,006	0.27	0.44	0	0	1
Main Owner Is African-American	14,050	0.07	0.25	0	0	0
Main Owner Is Hispanic	14,050	0.04	0.20	0	0	0
Main Owner Is Asian	14,050	0.04	0.20	0	0	0
Main Owner Is Born in the US	13,997	0.91	0.29	1	1	1
Main Owner's Work Experience	14,002	13.49	10.96	1	11	30
State and County Characteristics						
State GDP	13,875	10.65	0.14	10.51	10.64	10.80
County Population	13,875	905,644	1,557,066	42,269	405,142	2,015,355
County Catholic to Protestant Ratio	13,870	4.14	6.29	0.18	1.84	11.52
County Personal Income Per Capita	13,875	10.48	0.54	10.17	10.47	10.85
County Nonfarm Establishments Per Capita	13,875	0.03	0.01	0.02	0.03	0.03
County Wage Inequality	13,875	0.55	0.04	0.50	0.54	0.60
County Federal Government Expenditures Per Capita	13,875	7.46	6.62	3.99	6.34	11.07
County Land Area	13,875	14.41	0.64	13.78	14.46	15.06

NOTES. The table provides the number of observations, mean, standard deviation, 10th percentile, the median (50th percentile) and the 90th percentile of all variables used in the empirical analysis. Due to confidentiality the minimum and maximum are not reported. For the sake of brevity we do not include MSA characteristics separately.

TABLE IV

MAIN SPECIFICATIONS EXPLAINING BUSINESS FORMATION

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	*MSA Total Establishment Entries*				*MSA Total Establishment Exits*			
MSA Inequality in 2004	-3.540***	-2.635***	-1.554***	-2.463***	-3.426***	-2.952***	-1.668***	-2.702***
	(0.723)	(0.725)	(0.533)	(0.746)	(0.803)	(0.688)	(0.490)	(0.711)
MSA Population$_{t-1}$	1.060***	0.997***	1.006***	1.001***	1.029***	0.960***	0.974***	0.962***
	(0.0170)	(0.0155)	(0.0169)	(0.0169)	(0.0190)	(0.0155)	(0.0198)	(0.0172)
MSA Catholic to Protestant Ratio	--	0.0206	-0.0130	0.0165	--	0.0212	-0.0100	0.0175
	--	(0.0237)	(0.0216)	(0.0255)	--	(0.0185)	(0.0232)	(0.0208)
MSA Personal Income Per Capita	--	0.338***	0.334***	0.297***	--	0.335***	0.329***	0.314***
	--	(0.0988)	(0.0903)	(0.104)	--	(0.101)	(0.0798)	(0.103)
MSA Land Area	--	0.0452**	0.0443**	0.0412*	--	0.0466**	0.0359*	0.0442**
	--	(0.0216)	(0.0190)	(0.0231)	--	(0.0180)	(0.0185)	(0.0196)
MSA House Price Index	--	0.00167***	0.00159***	0.00282***	--	0.000309	-0.000135	0.00163***
	--	(0.000376)	(0.000379)	(0.000389)	--	(0.000320)	(0.000303)	(0.000296)
MSA Whites to Total Population Ratio	--	--	0.424***	--	--	--	0.349**	--
	--	--	(0.132)	--	--	--	(0.150)	--
MSA Wage Inequality in 2004	--	--	2.047***	--	--	--	2.966***	--
	--	--	(0.594)	--	--	--	(0.758)	--
MSA Federal Government Expenditures Per Capita	--	--	-0.00416**	--	--	--	-0.00313	--
	--	--	(0.00193)	--	--	--	(0.00262)	--
State Fixed Effects	Yes	Yes	Yes	--	Yes	Yes	Yes	--
Year Fixed Effects	Yes	Yes	Yes	--	Yes	Yes	Yes	--
State*Year Fixed Effects	No	No	No	Yes	No	No	No	Yes
Number of Observations	2,888	2,704	2,028	2,704	2,888	2,704	2,028	2,704

NOTES. All Models are estimated with a linear regression (OLS) model. The definition of the variables can be found in Table II.*t-1* indicates a one year lag. "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. Standard errors are clustered at the state level. Standard errors are given in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

TABLE V

MAIN SPECIFICATIONS COUNTY INEQUALITY 2004: EXPLAINING FIRM TECHNOLOGY, OWNERSHIP AND FINANCING

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable	*Firm Is High Tech*			*Firm Is Proprietorship*			*Firm Bank and Family Financing*			*Firm Angel and Venture Capital Financing*		
County Inequality in 2004	-0.670***	-0.555***	-0.550***	1.159***	1.058***	1.060***	0.185	0.347*	0.359**	-0.113*	-0.189***	-0.193***
	(0.171)	(0.165)	(0.168)	(0.213)	(0.257)	(0.259)	(0.128)	(0.177)	(0.177)	(0.0635)	(0.0521)	(0.0557)
County Population	0.0208***	0.00836**	0.00828**	-0.0462***	-0.0451***	-0.0452***	-0.0137***	0.359**	-0.00961*	0.00326***	0.00311**	0.00349**
	(0.00322)	(0.00327)	(0.00335)	(0.00888)	(0.00749)	(0.00764)	(0.00326)	(0.177)	(0.00483)	(0.000972)	(0.00136)	(0.00167)
Firm Total Assets$_{t-1}$	--	-0.0137***	-0.0140***	--	-0.0523***	-0.0531***	--	0.0235***	0.0239***	--	0.00246	0.00266
	--	(0.00258)	(0.00257)	--	(0.00544)	(0.00550)	--	(0.00264)	(0.00265)	--	(0.00155)	(0.00160)
Firm ROA$_{t-1}$	--	0.00538**	0.00521**	--	0.00906***	0.00924***	--	-0.00295	-0.00292	--	-0.00324***	-0.00330***
	--	(0.00242)	(0.00249)	--	(0.00256)	(0.00254)	--	(0.00205)	(0.00217)	--	(0.00112)	(0.00118)
Firm Tangibility$_{t-1}$	--	-0.182***	-0.184***	--	0.195***	0.197***	--	0.0548***	0.0548***	--	-0.00477	-0.00497
	--	(0.0111)	(0.0111)	--	(0.0224)	(0.0223)	--	(0.0110)	(0.00973)	--	(0.00439)	(0.00445)
Firm Number of Owners$_{t-1}$	--	0.0248	0.0252	--	-0.367***	-0.366***	--	-0.00722	-0.00584	--	0.0412***	0.0421***
	--	(0.0192)	(0.0195)	--	(0.0422)	(0.0436)	--	(0.0133)	(0.0130)	--	(0.00948)	(0.00970)
Main Owner Is Female	--	-0.0487***	-0.0491***	--	0.0517**	0.0502**	--	0.00412	0.00386	--	-0.00987***	-0.00981***
	--	(0.0146)	(0.0147)	--	(0.0233)	(0.0238)	--	(0.0116)	(0.0119)	--	(0.00287)	(0.00308)
Main Owner Is African-American	--	0.0361	0.0364	--	0.00421	-0.000277	--	-0.0331*	-0.0326*	--	-0.00337	-0.00324
	--	(0.0270)	(0.0271)	--	(0.0420)	(0.0429)	--	(0.0184)	(0.0187)	--	(0.00284)	(0.00336)
Main Owner Is Hispanic	--	-0.0123	-0.0121	--	-0.00413	-0.00679	--	-0.00478	-0.00606	--	0.00520	0.00587
	--	(0.0263)	(0.0265)	--	(0.0769)	(0.0793)	--	(0.0134)	(0.0136)	--	(0.00723)	(0.00847)
Main Owner Is Asian	--	-0.00323	-0.00184	--	-0.0108	-0.0105	--	-0.00279	-0.00503	--	0.00489	0.00576
	--	(0.042)	(0.0422)	--	(0.0443)	(0.0456)	--	(0.0248)	(0.0242)	--	(0.0118)	(0.0119)
Main Owner Is Born in the US	--	-0.0951***	-0.0945***	--	0.0775**	0.0778**	--	-0.00165	-0.00177	--	-0.00112	-0.00124
	--	(0.0278)	(0.0280)	--	(0.0338)	(0.0343)	--	(0.0158)	(0.0163)	--	(0.00465)	(0.00458)
Main Owner's Work Experience	--	0.00388***	0.00388***	--	0.000401	0.000449	--	-0.000480	-0.000511	--	0.0000576	0.0000681
	--	(0.000404)	(0.000407)	--	(0.000718)	(0.000734)	--	(0.000433)	(0.000442)	--	(0.000150)	(0.000146)
State GDP$_{t-1}$	--	0.0820	--	--	0.0481	--	--	-0.0248	--	--	-0.0435	--
	--	(0.132)	--	--	(0.139)	--	--	(0.128)	--	--	(0.0485)	--
County Control Variables	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
State Fixed Effects	Yes	Yes	--	Yes	Yes	--	Yes	Yes	--	Yes	Yes	--
Year Fixed Effects	Yes	Yes	--	Yes	Yes	--	Yes	Yes	--	Yes	Yes	--
2-digit Industry Fixed Effects	No	No	No	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
State*Year Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Industry*Year Fixed Effects	No	No	No	No	No	Yes	No	No	Yes	No	No	Yes
Number of Observations Panel A	13,875	8,830	8,837	13,865	8,826	8,833	10,404	6,456	6,460	7,122	4,505	4,509

NOTES. All Models are estimated with a linear regression model (OLS) and take into account cross-sectional Kauffman Firm Survey weights. The definition of the variables can be found in Table II*t-1* indicates a one year lag. "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. Standard errors are clustered at the state level. Standard errors are given in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

TABLE VI

MAIN SPECIFICATIONS EXPLAINING BUSINESS FORMATION: MSA INEQUALITY 2004 - INSTRUMENTED

Model	First Stage	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	*MSA Inequality in 2004*	*Total Establishment Entries*			*Total Establishment Exits*		
MSA Inequality in 2004	--	-5.295***	-2.248*	-3.608***	-6.765***	-6.063***	-6.047***
	--	(0.853)	(1.240)	(1.090)	(1.083)	(1.685)	(1.414)
MSA Population$_{t-1}$	-0.0100***	1.053***	0.998***	0.989***	1.013***	0.937***	0.931***
	(0.000389)	(0.0168)	(0.0158)	(0.0137)	(0.0184)	(0.0214)	(0.0182)
MSA Catholic to Protestant Ratio	-0.00382***	--	-0.00759	0.0190	--	0.0127	0.0220
	(0.000569)	--	(0.0229)	(0.0267)	--	(0.0291)	(0.0292)
MSA Personal Income Per Capita	-0.00441***	--	0.318***	0.236**	--	0.198**	0.134
	(0.00140)	--	(0.0862)	(0.0978)	--	(0.0933)	(0.107)
MSA Land Area	0.00990***	--	0.0624***	0.0624***	--	0.0878***	0.0868***
	(0.000454)	--	(0.0158)	(0.0164)	--	(0.0228)	(0.0211)
MSA House Price Index	0.0000645***	--	0.00155***	0.00257***	--	-0.000389	0.000947**
	(0.0000101)	--	(0.000395)	(0.000470)	--	(0.000347)	(0.000457)
MSA Whites to Total Population Ratio	-0.0230***	--	0.403***	--	--	0.272	--
	(0.00324)	--	(0.133)	--	--	(0.177)	--
MSA Wage Inequality in 2004	-0.675***	--	1.617*	--	--	0.528	--
	(0.0142)	--	(0.960)	--	--	(1.193)	--
MSA Federal Government Expenditures Per Capita	-0.000817***	--	-0.00555***	--	--	-0.00604**	--
	(0.0000976)	--	(0.00164)	--	--	(0.00257)	--
Rain	0.00286***	--	--	--	--	--	--
	(0.000377)	--	--	--	--	--	--
Rain Standard Deviation	0.00285***	--	--	--	--	--	--
	(0.0000650)	--	--	--	--	--	--
Temperature	-0.0126***	--	--	--	--	--	--
	(0.000543)	--	--	--	--	--	--
Temperature Standard Deviation	-0.000151	--	--	--	--	--	--
	(0.000136)	--	--	--	--	--	--
State Fixed Effects	Yes	Yes	Yes	--	Yes	Yes	--
Year Fixed Effects	Yes	Yes	Yes	--	Yes	Yes	--
State*Year Fixed Effects	No	No	No	Yes	No	No	Yes
Number of Observations	8,654	2,880	2,022	2,696	2,880	2,022	2,696

NOTES. All models are estimated with a 2SLS IV model. The first column contains the results of the first stage regression. MSA Inequality in 2004 is instrumented with average division rain fall and temperature between 1895 - 2003 and their corresponding standard deviations. The definition of the variables can be found in Table II. *t-1* indicates a one year lag. "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. Standard errors are clustered at the state level. Standard errors are given in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

TABLE VII

MAIN SPECIFICATIONS COUNTY INEQUALITY 2004 - INSTRUMENTED: FIRM TECHNOLOGY, OWNERSHIP AND FINANCING

Model	First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	*County Inequality in 2004*	*Firm Is High Tech*		*Firm Is Proprietorship*		*Firm Bank and Family Financing*		*Firm Angel and Venture Capital Financing*	
County Inequality in 2004	--	-1.520***	-0.285	0.724	4.175**	0.770	1.374*	-0.366***	-0.284
	--	(0.442)	(1.290)	(1.141)	(2.017)	(0.642)	(0.812)	(0.109)	(0.395)
County Population	0.0105***	0.0201***	0.000575	-0.0398***	-0.0730***	-0.0164***	-0.0195*	0.00296**	0.00239
	(0.000542)	(0.00343)	(0.0189)	(0.00729)	(0.0241)	(0.00291)	(0.0105)	(0.00119)	(0.00485)
Firm Total Assets$_{t-1}$	0.000113	--	-0.0122***	--	-0.0540***	--	0.0237***	--	0.00175
	(0.000243)	--	(0.00299)	--	(0.00621)	--	(0.00276)	--	(0.00153)
Firm ROA$_{t-1}$	0.000265	--	0.00487*	--	0.0105***	--	-0.00149	--	-0.00336***
	(0.000220)	--	(0.00264)	--	(0.00254)	--	(0.00210)	--	(0.00129)
Firm Tangibility$_{t-1}$	0.000851	--	-0.190***	--	0.208***	--	0.0533***	--	-0.00616
	(0.00129)	--	(0.0120)	--	(0.0296)	--	(0.00918)	--	(0.00526)
Firm Number of Owners$_{t-1}$	-0.000745	--	0.0238	--	-0.355***	--	0.00170	--	0.0496***
	(0.00113)	--	(0.0219)	--	(0.0476)	--	(0.0126)	--	(0.00897)
Main Owner Is Female	0.000698	--	-0.0518***	--	0.0407	--	0.0146	--	-0.0119***
	(0.00101)	--	(0.0142)	--	(0.0253)	--	(0.0138)	--	(0.00358)
Main Owner Is African-American	0.0109***	--	0.0391	--	-0.0138	--	-0.0443**	--	-0.00495
	(0.00202)	--	(0.0346)	--	(0.0424)	--	(0.0224)	--	(0.00367)
Main Owner Is Hispanic	0.00533***	--	-0.0168	--	-0.0175	--	-0.0146	--	0.00660
	(0.00206)	--	(0.0259)	--	(0.0728)	--	(0.0137)	--	(0.0106)
Main Owner Is Asian	-0.000239	--	0.00102	--	-0.00654	--	-0.0393*	--	0.00921
	(0.00229)	--	(0.0459)	--	(0.0525)	--	(0.0201)	--	(0.0130)
Main Owner Is Born in the US	-0.00554***	--	-0.0882***	--	0.0729*	--	-0.0126	--	-0.00453
	(0.00163)	--	(0.0283)	--	(0.0411)	--	(0.0199)	--	(0.00509)
Main Owner's Work Experience	-0.0000441	--	0.00386***	--	0.00171**	--	-0.000397	--	0.0000155
	(0.0000428)	--	(0.000387)	--	(0.000767)	--	(0.000527)	--	(0.000152)
State GDP$_{t-1}$	0.0100	--	--	--	--	--	--	--	--
	(0.0214)	--	--	--	--	--	--	--	--
Rain	0.00752***	--	--	--	--	--	--	--	--
	(0.00262)	--	--	--	--	--	--	--	--
Rain Standard Deviation	-0.0156***	--	--	--	--	--	--	--	--
	(0.00310)	--	--	--	--	--	--	--	--
Temperature	0.000982***	--	--	--	--	--	--	--	--
	(0.000265)	--	--	--	--	--	--	--	--
Temperature Standard Deviation	-0.00444***	--	--	--	--	--	--	--	--
	(0.000611)	--	--	--	--	--	--	--	--
County Control Variables	Yes	No	Yes	No	Yes	No	Yes	No	Yes
State Fixed Effects	Yes	Yes	--	Yes	--	Yes	--	Yes	--
Year Fixed Effects	Yes	Yes	--	Yes	--	Yes	--	Yes	--
2-digit Industry Fixed Effects	Yes	No	No	No	--	No	--	No	--
State*Year Fixed Effects	No	No	Yes	No	Yes	No	Yes	No	Yes
Industry*Year Fixed Effects	No	No	No	No	Yes	No	Yes	No	Yes
Number of Observations	5,442	11,696	7,467	11,687	7,463	8,796	5,469	5,983	3,787
R-Squared	0.675	0.015	0.107	0.094	0.291	0.026	0.125	0.017	0.185

NOTES. All models are estimated with a 2SLS IV model. The first column contains the results of the first stage regression. County Inequality in 2004 is instrumented with average division rain fall and temperature between 1895 - 2003 and their corresponding standard deviations. All models take into account cross-sectional Kauffman Firm Survey weights. The definition of the variables can be found in Table II. t-1 indicates a one year lag. "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. Standard errors are clustered at the state level. Standard errors are given in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

TABLE VIII

FALSIFICATION TEST INSTRUMENTAL VARIABLES: US VERSUS FRANCE

Model	(1)	(2)	(3)	(4)	(5)	(6)
	US Full Sample		US Restricted Sample		France Full Sample	
Dependent Variable	*Total Establishment Entries*	*Total Establishment Exits*	*Total Establishment Entries*	*Total Establishment Exits*	*Total Establishment Entries*	*Total Establishment Exits*
Rain	-0.075**	-0.054	-0.140**	-0.102*	-0.002	0.001
	(0.036)	(0.033)	(0.066)	(0.055)	(0.003)	(0.010)
Rain Standard Deviation	0.117*	0.096**	0.332**	0.206*	0.003	0.014
	(0.058)	(0.044)	(0.126)	(0.118)	(0.002)	(0.016)
Temperature	-0.017**	-0.019***	-0.037***	-0.030***	0.027	0.006
	(0.007)	(0.007)	(0.006)	(0.006)	(0.024)	(0.054)
Temperature Standard Deviation	-0.034**	-0.022	-0.049***	-0.043**	-0.009	0.355
	(0.015)	(0.016)	(0.012)	(0.016)	(0.079)	(0.280)
MSA/ Department Population$_{t-1}$	1.049***	1.051***	1.062***	1.061***	1.127***	0.890***
	(0.024)	(0.023)	(0.022)	(0.022)	(0.032)	(0.082)
MSA/ Department Land Area	0.015	-0.001	0.011	-0.001	-0.212***	-0.199***
	(0.025)	(0.023)	(0.039)	(0.035)	(0.035)	(0.058)
State/Region Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	12,600	12,600	5,250	5,250	903	903

NOTES. All models are estimated with a linear regression model (OLS). Column (1) and (2) contain the results of the regressions including all observations from the US sample. Columns (3) and (4) provide the results when restricting the US sample to all those observations within the minimum and the maximum of the French averages and standard deviations. Column (5) and (6) present the results for the entire French data sample. The unit of analysis is the MSA for the US and the department for France. The definition of the variables can be found in Table II. *t-1* indicates a one year lag. "Yes" indicates that the set of fixed effects is included. "No" indicates that the set of fixed effects is not included. Standard errors are clustered at the state/ regional level. Standard errors are given in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

TABLE IX

ROBUSTNESS ANALYSIS - INEQUALITY 1890: EXPLAINING BUSINESS FORMATION, FIRM TECHNOLOGY, OWNERSHIP AND FINANCING

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	*MSA Total Establishment Entries*	*MSA Total Establishment Exits*	*Firm Is High Tech*	*Firm Is Proprietorship*	*Firm Bank and Family Financing*	*Firm Angel and Venture Capital Financing*
MSA Inequality in 1890	-0.253**	-0.206*	--	--	--	--
	(0.121)	(0.112)	--	--	--	--
County Inequality in 1890	--	--	-0.182**	0.243***	0.120***	0.00415
	--	--	(0.0710)	(0.0773)	(0.0443)	(0.0200)
County Control Variables	No	No	Yes	Yes	Yes	Yes
MSA Control Variables	Yes	Yes	No	No	No	No
State Fixed Effects	Yes	Yes	--	--	--	--
Year Fixed Effects	Yes	Yes	--	--	--	--
2-digit Industry Fixed Effects	No	No	Yes	No	Yes	Yes
State*Year Fixed Effects	No	No	Yes	Yes	Yes	Yes
Industry*Year Fixed Effects	No	No	Yes	No	Yes	Yes
Number of Observations	9,909	9,909	8,777	8,773	6,420	4,487
R-Squared	0.961	0.963	0.101	0.316	0.138	0.147

NOTES. All Models are estimated with a linear regression model (OLS). Models (3) - (7) take into account cross-sectional Kauffman Firm Survey weights. The definition of the variables can be found in Table II. t-1 indicates a one year lag. "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. Standard errors are clustered at the state level. Standard errors are given in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

TABLE X

DIFFERENCE-IN-DIFFERENCES ANALYSIS: STATE ELIMINATIONS OF ESTATE, INHERITANCE AND GIFT TAXES

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable	*Total Establishment Entries*			*Total Establishment Exits*			*Total Establishment Entries*			*Total Establishment Exits*		
Post	-0.040*	-0.038*	-0.026*	-0.014	-0.013	0.003	--	--	--	--	--	--
	(0.025)	(0.023)	(0.015)	(0.014)	(0.014)	(0.017)	--	--	--	--	--	--
Number of Years since Reform	--	--	--	--	--	--	-0.013*	-0.013*	-0.008	-0.008**	-0.008**	-0.004
	--	--	--	--	--	--	(0.008)	(0.008)	(0.006)	(0.004)	(0.004)	(0.006)
MSA Population$_{t-1}$	0.681***	0.686***	0.600***	0.969***	0.971***	1.050***	0.692***	0.696***	0.606***	0.977***	0.978***	1.053***
	(0.130)	(0.117)	(0.101)	(0.060)	(0.056)	(0.080)	(0.128)	(0.119)	(0.099)	(0.059)	(0.056)	(0.079)
MSA Catholic to Protestant Ratio	--	0.034	-0.002	--	-0.002	-0.021	--	0.038	0.000	--	0.000	-0.020
	--	(0.030)	(0.026)	--	(0.029)	(0.022)	--	(0.031)	(0.025)	--	(0.030)	(0.023)
MSA Whites to Total Population Ratio	--	1.610**	1.497**	--	0.509	0.389	--	1.582**	1.496**	--	0.477	0.361
	--	(0.713)	(0.654)	--	(0.517)	(0.484)	--	(0.671)	(0.636)	--	(0.494)	(0.467)
MSA House Price Index	--	--	0.003***	--	--	-0.000	--	--	0.003***	--	--	0.000
	--	--	(0.001)	--	--	(0.001)	--	--	(0.001)	--	--	(0.001)
MSA Personal Income Per Capita	--	--	0.636***	--	--	0.344*	--	--	0.613***	--	--	0.319*
	--	--	(0.232)	--	--	(0.182)	--	--	(0.226)	--	--	(0.187)
MSA Fixed Effects	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
Number of Observations	8,303	8,303	5,912	8,303	8,303	5,912	8,303	8,303	5,912	8,303	8,303	5,912

NOTES. All models are estimated with a linear regression model (OLS). *Post* is an indicator variable that takes on the value of 1in the years after the state in which an MSA is located abandons the EIG taxes and 0 otherwise. *Number of Years since Reform* indicates the total number of years that have passed after the EIG Reform was introduced in the state in which the MSA is located, with a maximum of 4 years. The definition of the other variables can be found in Table II. t-1 indicates a one year lag. "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. Standard errors are clustered at the state level. Standard errors are given in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.

TABLE XI

EXPLORING DEMAND AND SUPPLY: MAIN SPECIFICATIONS EXPLAINING COUNTY BANK FINANCING SUPPLY, COUNTY EDUCATION, EDUCATED COUNTY POPULATION INFLOW AND CIVIL JUSTICE EFFICIENCY

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent Variable	*No. of Bank Establishments per 1,000 Capita*			*County Percentage of Adults with College Degree or More*			*Population Inflow with at Least College degree*			*Judicial Efficiency*		
County Inequality in 2004	-0.806***	-0.971***	-6.579***	-95.993***	-61.234***	-171.166**	-17.121***	-7.237***	-19.955**	1.172**	0.525**	-1.177
	(0.296)	(0.315)	(1.803)	(14.748)	(11.156)	(83.163)	(3.844)	(1.300)	(8.602)	(0.442)	(0.236)	(1.15)
County Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	No	No
Case Controls	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	No	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Number of Observations	15,690	12,376	11,416	15,690	3,094	2,854	2,515	2,487	2,271	8,300	8,296	8,234

NOTES. Models (3), (6), (9) and (12) are estimated with a 2SLS IV model. All other Models are estimated using OLS. *County Inequality in 2004* is instrumented with average division rain fall and temperature between 1895 - 2003 and their corresponding standard deviations in Model (3), (6), (9) and (12) respectively. The definition of the variables can be found in Table II. *t-1* indicates a one year lag. "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. Standard errors are clustered at the state level. Standard errors are given in parentheses. ***, ** and * indicate significance at the 1%, 5% and 10% level respectively.