Essays in History & Finance

by

Francesco D'Acunto

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ulrike Malmendier, Chair
Professor Stefano DellaVigna
Professor Barry Eichengreen
Professor Ross Levine
Professor Gustavo Manso

Spring 2015

Essays in History & Finance

Abstract

Essays in History & Finance

by

Francesco D'Acunto

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Ulrike Malmendier, Chair

To which extent may history help us understand current financial outcomes? Financial History typically tests for financial theories using historical data, or studies past financial outcomes to understand the present through analogy. This dissertation focuses on an alternative approach to the use of history in finance, which I label "History & Finance." This approach exploits facts and institutions of the past, and the persistence of the economic and social phenomena they determine, to *directly* explain current financial outcomes.

In the first chapter of the dissertation, I define the scope of History & Finance based on its differences with Economic and Financial History. History & Finance has its roots in the Political Economy approach that exploits the long-run effects of historical shocks and institutions to understand current economic outcomes. I survey the research in History & Finance, organizing it across the subfields of Finance which have used this approach so far. I discuss the challenges that this method poses to researchers, and in particular the difficulty of making plausible causal statements in long-run settings, and of determining the channels of transmission of the effects of historical phenomena on current financial outcomes. I then propose five directions that scholars may follow to enlarge the scope of this approach to research in finance.

In the second chapter of the dissertation, I use the History & Finance approach to understand the current spatial distribution of innovation. I focus on the innovation of traditional industries, a margin of innovation which is largely neglected in the literature, although it includes the vast majority of the innovations produced every year in Europe and the United States. I exploit newly assembled data on innovation and education at the level of European regions, as well as a unique firm-level data set on the innovation of traditional industries, to show that the amount of formal education of blue-collar workers is an important determinant of the innovation produced in traditional industries. The relevance of the formal education of blue-collar workers for innovation is alternative to the learning-by-doing hypothesis and the directed technical change literature, none of which recognizes a role for

formal education in the productivity of jobs on the line. Moreover, I document for the first time that the variation in the amount of basic education across European regions is highly persistent over the decades, despite the major institutional and economic shocks that have differentially affected European regions over time. Hence, I use the historical variation of basic education across regions to address the issue of the reverse causality between current innovation and current basic education. I also propose an historical natural experiment, the quasi-exogenous diffusion of the printing press across Europe after 1450 AD, as a source of variation for cross-regional basic education in the past. Higher basic education increases the amount of innovation produced by firms in traditional industries, whose workers are in large part basically educated individuals. The effect is indeed stronger for firms that employ more blue-collar workers. I also document that higher basic education determines higher capital expenditures and lower financial constraints for those firms that innovate more.

In the third chapter of the dissertation, which is coauthored with Marcel Prokopczuk and Michael Weber, I use the History & Finance approach to understand the limited stock market participation of households. Because the Jewish population has been associated with the provision of financial services in Europe for more than eight centuries, we test whether those German counties with a higher persecution of local Jewish communities over time tend to invest less in stocks. Indeed, in those counties where historical persecution was higher as far back as in the Middle Ages, present-day households invest less in stocks. The size of the effect is similar across cohorts, and it does not fade away over time. The evidence is consistent with a cultural norm of distrust in finance that has started in those areas where Jews were persecuted more in the past, and has transmitted across generations over the centuries. To obtain quasi-exogenous variation in the extent of Jewish persecution across German counties in the distant past, we exploit the forced migrations of the first German Jewish communities from the Rhine Valley to current Poland to instrument for the existence of Jewish communities in counties in the Middle Ages, and hence the likelihood that persecution against the communities was documented.

This dissertation defines History & Finance as a rising approach to the study of financial phenomena, and it proposes several promising routes that researchers across subfields of finance may take to bring this approach from its infancy to a mature stage.

Contents

List of Figures

List of Tables

Chapter 1

History & Finance: A Nascent Field

1.1 Introduction

Economic history has developed as an autonomous field of research at least since 1933, when the first chair was established at the London School of Economics. History is the cornerstone of economic history research: either historical facts and institutions are interesting to study by themselves, or economic historians study them to infer lessons about the present, based on the analogies and differences between the historical and current settings. A turning point for economic history was the cliometric revolution initiated in the 1960s by Douglass North and Robert Fogel, among others, which injected the analytical and econometric advances of economics into the study of facts and institutions of the past. With cliometrics, history also became a fertile source of natural experiments, which allow the testing of economic theories that would otherwise be hard to study based on current data. Because experiments and outcomes are confined to the past, history maintains the center of the stage in cliometrics, whereas economics provides the research tools.

In this essay, I describe another approach to the use of historical facts and institutions in finance, which I propose labeling "History & Finance." This approach echoes the movement in the economic growth literature that uses facts and institutions of the past and the long-run persistence of economic and social phenomena to *directly* explain current economic outcomes, initiated by the seminal works of La Porta, Lopez-de Silanes, Shleifer, and Vishny (1998), Levine (1998), Acemoglu, Johnson, and Robinson (2001), and Easterly and Levine (2003), and recently surveyed by Spolaore and Wacziarg (2013) and Nunn (2014).[1]

First, I describe the scope of History & Finance and how it relates to the fields of Economic and Financial History. Second, I survey the body of work in History & Finance. I discuss the challenges History & Finance pose for finance scholars, and how existing work

[1] To the best of my knowledge, this approach to economic research has not been labeled yet. Based on the arguments below, the approach could be labeled "History & Economics."

has attempted to address these challenges. Finally, I comment on the questions and avenues for future research that History & Finance opens to finance scholars and economic historians alike.

1.2 History & Finance vs. Economic and Financial History

In the fields of Economic and Financial History, the historical component is the cornerstone on which the research investigation is built. One could roughly think about two approaches, which I present with stylized diagrams in Panel A and Panel B of Figure 1.1. The first approach studies historical events or institutions as objects worthy of interest by themselves, for instance, because they represent milestones in the history of humanity.[2] Two monumental contributions in this vein are Fogel (1964) on the role of railroads in the development of the U.S. economy,[3] and Fogel and Engerman (1974), who demonstrated the profitability of slavery in the U.S. South and challenged the common view that slavery would have disappeared in the United States regardless of the outcome of the Civil War. The aim of studying the effects of railroads in the developing phase of the U.S. economy, or an abolished labor market institution, was not to provide insights on the U.S. economy of the 1960s and 1970s. The topics were important objects of study for their own sake. Recent examples in finance are Calomiris and Schweikart (1991) on the origins and consequences of the Panic of 1857, and the interdisciplinary study of Goetzmann, Labio, Rouwenhorst, and Young (2013) on the financial bubble of 1720.

The second approach exploits facts, institutions, and natural experiments of the past and their immediate or longer-term effects on *past* outcomes to test economic theories that would be hard to test based on current data.[4] Historical events and institutions are worth studying because they may inform us about current outcomes by analogy, based on the similarities and differences between the historical and current settings. This approach was used well before the start of modern economic thought. For instance, Xenophon (60BC) introduced economics as the set of "norms to run the household", and discussed these norms based on Isomachus's fictitious description to Socrates two generations earlier. The second book of Aristotle (20BC) enlarged the scope of economics to the norms that regulate four different economic entities based on the earlier history of Greek cities (*poleis*).[5] Classic

[2]McCloskey (1976) puts it as follows: "For the professional economic historian the worth of economic history is that of general history, to which it contributes, and it is because he puts high value on history, economic or not, that he chooses to study it."

[3]See Goldin (1995) on how this work introduced the notion of a *counterfactual* in economic history.

[4]Greif (1997) discusses the early critique of this cliometric approach with regards to the underlying tenet that economic theory and its assumptions are ahistorical.

[5]No consensus exists on the authors and issue date of the *Oikonomika*, although historians and philologists agree members of the Aristotelian peripatetic school wrote it.

economics has also built on this approach, for instance, to discuss the historical origins of the division of labor (Smith (1776)). As for finance, this second approach has recently tested financial theory claims using unique historical micro-data and natural experiments in the past. Koudijs (ming) exploits exogenous shocks to the flow of information in the 18^{th}-century Amsterdam stock market to test for the effects of information on the reaction of asset prices on top of non-information-driven price volatility. Frehen, Goetzmann, and Rouwenhorst (2013) use hand-collected data on stock prices and investor accounts in 1720 to test for theories of asset price bubbles, whereas Giusti, Noussair, and Voth (2013) reconstruct the institutional characteristics of the South Sea Bubble of 1720 in a controlled environment to study which institutional aspects determine the start of asset price bubbles.

The two approaches described above can be interconnected. For instance, in corporate governance, Malmendier (2002) introduces the *societas publicanorum* as a historical object of interest by itself. It represented a "prototypical" form of corporation, its shares were traded, and its ownership was separated from control. It challenges the commonly held view that no institutions close to limited liability had existed before the 17^{th} century. At the same time, the *societas publicanorum* existed under the Roman Republic, a time when the Roman legal system had not yet reached its full maturity, but it disappeared afterwards. Hence, it helps disentangle the legal system from institutions, which cannot be done in current settings.[6]

The History & Finance approach

History & Finance differs from the previous approaches because it aims to *directly* understand current financial outcomes and institutions exploiting facts and institutions from the past, and the long-run persistence of economic and social phenomena. Thus, History & Finance reverses the role and status of history as an object of research compared to Economic and Financial History. The approach of History & Finance is illustrated in Panel C of Figure 1.1. The main object of interest of History & Finance is a current phenomenon, and the baseline analysis is run at current times. The historical component enters the analysis as a means to obtaining quasi-exogenous variation in cultural or institutional characteristics that could not be disentangled from other dimensions if focusing on the current setting. The historical component is a method, and not the main object of research. Whereas the research in Economic and Financial History may have a motivating current economic phenomenon in the background, in History & Finance, the current phenomenon is the cornerstone on which the research effort is built, and history is an instrument to understand it. One may compare the methodological aims of History & Finance to those of cliometrics. The cliometric revolution aimed to bring the theoretical and empirical toolbox of economics into the study of historical phenomena. History & Finance, instead, aims to bring the institutional and factual toolbox of history into the study of current financial phenomena.

[6] Moreover, the suggestive evidence on the trading of shares of the *societates publicanorum* informs on the dynamics of the *calor licitantis* (bidding fever), for which Capen, Clapp, and Campbell (1971) offered the first evidence in economics, and Malmendier, Moretti, and Peters (2012) in corporate finance.

Defining a research approach cannot avoid the proposal of an appropriate label for the approach. This essay uses the label "History & Finance," so a legitimate question is why this label may be appropriate. The foundations of western research methodology suggest that technical labels should follow the nature of the objects they describe.[7] The labels "Economic History" and "Financial History" have only one noun, which emphasizes one object of research: history. In these labels, "economic" and "financial" are mere adjectives, and describe an attribute of the noun defining the research method. History is the focus of these fields, and the fields obtain their research aim by exploiting the methods of economics and finance. An analogous argument can be made for the label "cliometrics," proposed by Stanley Reiter and first appearing in Davis, Hughes, and Reiter (1960). This one-word label combines two nouns: "clio" refers to an object, Clio, the muse of history,[8] and "metrics" is a methodology.[9] Thus the main object of cliometric research is history, and the economics toolbox is an instrument to understand history.

The assertion that a label for the approach described in this essay should emphasize the role of current financial facts and institutions as the objects of research naturally follows the discussion in the previous section. This emphasis is the driving force behind the label "History & Finance." Finance is an object under study, and history is no longer the only focus of research. Note that the more radical label "Historical Finance" would not be appropriate, because it would shift the attention from the current financial facts and institutions into their historical dimension, whereas the historical dimension is relevant to History & Finance only to the extent that it helps explain current phenomena.

What's in it for economic and financial historians?

History & Finance impacts the research of financial historians by enlarging its scope. Because the first approach to financial history research sees historical entities as objects of interest by themselves, it may not speak to those who have no interest in the history of specific institutions, cultures, or countries. The second approach to financial history looks at historical facts and institutions to understand the present through analogy. But the extent to which the current and past settings relate is often hard to assess.[10] History & Finance

[7]For instance, the second book of the *Institutiones*, one of the first "textbooks" ever based on research on earlier sources, discusses how the labeling of a certain type of donation during marriage had to be changed to adhere to the nature of the institution it defined: *"sed tamen nomen inconveniens remanebat, cum ante nuptias quidem vocabatur post nuptias autem tale accipiebat incrementum. sed nos plenissimo fini tradere sanctiones cupientes et consequentia nomina rebus esse studentes, constituimus, ut tales donationes [...] propter nuptias vocentur."* (Justinian, 533).

[8]The name of the muse derives from the Greek noun κλεος, which means fame, glory. Also, according to Greek mythology, Clio was the daughter of Mnemosyne, or memory. Hence, Clio is the recall of past glory, of historical milestones.

[9]The alternative holistic label "New Economic History" reflects both the empirical and theoretical methods the cliometric movement incorporates into economic history.

[10]Prominent economic historians emphasize this point. For instance, Barry Eichengreen explained that one contribution of economic history is an understanding of which dimensions of the past should and should

enlarges the scope of financial history by focusing on current financial phenomena, which is the common object of interest to all subfields of finance. Whereas other fields look for contemporaneous exogenous shocks to the incentives of financial actors, History & Finance draws on exogenous shocks in the past, and their effects over time, to obtain variation in dimensions for which contemporaneous shocks are hard to find. By changing the object of research from history to current outcomes, History & Finance transforms the research by financial historians into general interest research. This change is especially important for financial history scholars who study facts or institutions outside the United States. Because the capitalistic economic system that has developed in the United States since the end of the 18th century is the backbone of current economic institutions worldwide, understanding the development of the American economy is especially relevant to finance. Institutions and facts from other countries are instead often seen as niche topics, or intellectual curiosities. Explaining the *current* financial outcomes alongside the roots of a country's history makes the historical research effort more palatable to a general interest audience.

What's in it for finance scholars?

Finance scholars, and especially empirical scholars, search for natural experiments or other quasi-exogenous shocks to the observables whose effects on financial outcomes they aim to investigate. But several observables are entrenched in the culture, politics, and institutions that determine the outcomes to be explained. The History & Finance approach enlarges the time and cross-sectional dimensions along which finance scholars can identify appropriate exogenous shocks, by looking at shocks in the distant past and across countries. This broadening of the span of available experiments requires finance scholars to build the skills needed to understand past facts and institutions and their development over time. Building this set of skills is an intellectual enterprise similar to building econometric or economic theory knowledge. Hence History & Finance encourages cooperation between financial historians and other scholars.

1.3 Survey of Research in History & Finance

In this section, I survey the recent empirical literature in finance that exploits historical facts and institutions to directly understand current financial outcomes. I briefly describe the questions and methods researchers use across various subfields of general interest finance research.

---

not be informative to the present (Eichengreen, 2014b).

History & Asset Pricing

Understanding the behavior of assets' prices or returns over time is a crucial aim of asset pricing. Because consumption and risk attitudes vary at low frequencies, excess returns are unpredictable at high frequencies, but they may be predictable at longer horizons if risk aversion varies over time (e.g., see Cochrane (2001)). A crucial component of the toolbox of empirical asset pricing scholars is time-series econometrics, which allows the study of the characteristics of observed realizations of variables over time. The traditional approach to asset pricing does not necessarily aim to identify causal effects of cross-sectional dimensions on asset returns, which avoids the methodological challenges posed by History & Finance (section 4). Despite the vital need for long historical series to test the out-of-sample performance of asset pricing models, data availability and quality have traditionally limited the study of asset pricing to realizations since the Second World War, or at the earliest, since 1926, the starting date of the Center for Research in Security Prices (CRSP) sample. Campbell and Shiller (1998) and Campbell and Shiller (2001) are among the first to construct aggregate U.S. stock market data from 1871 until the end of the 20^{th} century, and to test for the performance of price-earning ratios and dividend-price ratios in forecasting dividend growth, earnings growth, or productivity growth, as implied by models based on the efficient-market hypothesis. Bunn and Shiller (2014) enlarge this program by constructing sector-level series from the 1870s until 2013. The historical data allows the running of an out-of-sample test for whether the value factor predicts returns. Koudijs and Golez (2014) construct for first time series of dividends and discount rates over the centuries, and use them to test for asset pricing theories. As for asset classes alternative to stocks, Ambrose et al. (2013) analyze the relationship between prices and housing rents in Amsterdam from 1650 to 2005, which allow for observation of the different responses of prices to several mispricing events.

History & International Finance

Because of the similar scope to asset pricing, international finance is also a field in which the approach of History & Finance can be naturally applied. But even in this case, the limited availability of historical data on returns and holdings across countries is a challenge. Chitu, Eichengreen, and Mehl (2014) are among the first to employ a History & Finance approach to document a positive effect of historical U.S. bilateral holdings of foreign bonds on their current geographic allocation, which also extends to the foreign holdings of other assets. The auhors interpret this "history effect" as evidence that having sunk the fixed costs of obtaining information about a country's assets in the past favors the holding of those assets in the long run.

History & Banking

Beck, Demirguc-Kunt, and Levine (2003) first test the relevance of two alternative determinants of current financial development across countries: initial endowments, and historical legal traditions that differ on the treatment of the protection of indiviual property rights. Initial endowments appear to be more robust predictors of the current cross-country variation in financial development, although both determinants are associated with it. Eichengreen (2014a) emphasizes the long-run peristence of thel features of banking and finance institutions within countries, and it discusses the importance of using historical evidence to shed light on the current connection between finance and growth. The introduction of historical natural experiments to identify causal effects in cross sections has indeed first started in the literature on banking, possibly because scholars in political economy, influenced by the seminal work in "History & Economics" by La Porta et al. (1998) and Acemoglu et al. (2001), have studied banking institutions extensively. Grosjean (2011) builds on the occupation of the Balkans and southern Europe by the Ottoman Empire. The Ottoman Empire kept the ban on lending with interest alive well into the 19^{th} century, whereas the Catholic church had started a *de facto* relaxation of the ban in the 15^{th} century. Because of the fragmentation and historical development of borders in Southeast Europe, Grosjean (2011) can design a within-country test for the effects of the ban on current financial development, measured with micro-level data on whether households are banked. The author finds the ban on lending with interest reduces the current within-country banking penetration by 4%.

Pascali (ming) studies the persistent differences in financial development within Italy. He exploits the fact that Italian cities with Jewish communities in the Middle Ages had more developed banking systems in the past as well as today. In the special case of Italy, charity-lending institutions were created to counteract the power of Jewish bankers in cities that hosted Jews. The author exploits the Jewish demographics across Italy in the 16^{th} century to instrument for banking development in the past, and finds that more developed banking in the past increases the current availability of credit to firms by at least 80%, and the current bank branch density by 24%. The author also exploits the quasi-exogenous expulsion of the Jews from the southern Italian cities under the rule of the Spanish Crown in 1541 to support the causal interpretation of the results.

History & Corporate Finance

Pierce and Snyder (2014) study the effect of deep-rooted cultural shocks on the long-run access to external financing by corporations. They exploit the variation in slave extraction across African countries from the 15th to the 19th century, based on the distance of countries from the Western coast of Africa, which increased the cost of slave extraction. First, the authors show that firms in countries with a history of higher slave extraction are more likely to report that access to finance is a major obstacle to their business activities, and are less likely to access several formal sources of external financing. Second, they find that the lower access to formal external financing is not compensated by higher access to informal financing, such as trade credit. As for the channels that transmitted the long-run effect of slave extraction on financial development, the authors find evidence consistent with a role of ethnic fractionalization, and the centralization of political power within ethnicities before the start of the slave trade. Both channels decrease the amount of generalized trust in countries with a history of high slave extraction, and lower trust reduces financial development. Contrary to access to finance, several other obstacles to business activities, which are plausibly unrelated by generalized trust, are unrelated to the historical extraction of slaves in a country. D'Acunto (2015c) studies the effects of the education level of regional blue collar workers on the innovation of low-tech firms, and their investments, capital structures, and financing conditions. The first stage of the analysis studies the current link between the education level of workers and a set of corporate finance outcomes. Because research has shown low-tech innovation to be spuriously persistent, the education level of blue-collar workers must be persistent to be a channel that affects low-tech innovation over time. Therefore, in a second step, D'Acunto (2015c) verifies that the amount of schooling for those with basic levels of education persists in locations, because literacy rates in European regions as far back as 1880 predict the current amount of schooling for inhabitants with basic education levels. But if basic education is persistent, a long-run effect of basic education on current firm-level outcomes can only be identified using some exogenous shock to the incentives to acquire literacy in the distant past. Thus, the author exploits the quasi-exogenous diffusion of the printing press out of the city of Mainz after 1450 to instrument for the historical literacy rates observed in 1880, and to test for the effects of historical literacy on current regional and firm-level outcomes.

History & Household Finance

A strand of literature at the intersection of political economy and finance has recently investigated the cultural and institutional determinants of households' investment behavior. Because culture and institutions evolve slowly over time, testing for the effects of cultural norms and institutions on households' behavior is a prime application of the History & Finance approach. D'Acunto, Prokopczuk, and Weber (2015) test for the effect of distrust in finance on the likelihood that households invest in stocks. They exploit the historical anti-Jewish sentiment to proxy for the distrust in finance across current German counties. The rationale is that Christians had been banned from lending money at interest from the 11^{th} to the 16^{th} century; hence, Jews have run the financial sector in Europe, and especially in current Germany, for centuries. In areas where the population felt more hatred toward Jews, it also probably felt more distrust toward finance was probably also distrusted more, because Jews were associated with financial services. Distrust may have transmitted across generations. The authors verify this claim by running a survey that shows that current distrust in finance, which is positively associated with Jewish persecution in a German county as far back as the Black Death of 1349. The authors also find that in counties with higher Jewish persecution, current inhabitants are less likely to invest in stocks, but not in non-financial investments.

History & Entrepreneurial Finance

The essays in Lamoreaux and Sokoloff (2007) dissect the evolution of the relationship between financing, entrepreneurship, and technological innovation in the United States from 1870 onward. Although each essay focuses on a specific period in the U.S. financial history, the sum of all essays, as well as the editor's cohesive introduction, provide a full picture to help us understand the current relationship between financing and innovation through the prism of the past.

Glaeser, Pekkala Kerr, and Kerr (2014) exploit a natural experiment in the past to study the current positive relationship between entrepreneurship and employment growth. The authors build on the thesis that cities that specialized in scale economy industries early were less likely to develop a culture of entrepreneurship over time. They use the distance of cities from coal mines, because cities that happened to be close to historical coal mines were more likely than others to specialize in scale economy industries. The crucial insight is that if an entrepreneurial culture did not develop, not only the mining industry, but also industries unrelated to mining would show lower entrepreneurship rates today than the same industries in cities far away from mines. This result is what the authors confirm in their first-stage analysis, which allows them to use the distance from historical mines to instrument for current entrepreneurship in cities. They estimate that a one-standard-deviation lower average establishment size increases employment growth in cities by about 0.8 standard deviations in the period 1982-2002.

Agrawal, Galasso, and Oettl (2014) study the relationship between knowledge flows and innovation, for which a direction of causality is hard to identify because of the ineherent endogeneity of the two dimensions. They use the historical plans of the U.S. interstate highway system to instrument for the extent of knowledge flows across cities. They find that a 10% increase in the stock of transport infrastructures in a region increases patenting in the region by 1.7%, and the increase is driven by easier knowledge flows and easier access to knowledge inputs by innovators in cities connected by highways.

1.4 Challenges and Methodology

History & Finance poses a major methodological hurdle: the outcomes to be explained and the historical shocks are realized centuries apart; hence, interpreting the effect of past shocks on current outcomes in a causal way is problematic. Persistent unobservables may be correlated with both the determinants of historical shocks and current outcomes. Supporting the exclusion restriction that the historical shocks only affected current outcomes through one specific channel of transmission is also hard. Detailed knowledge of the institutional determinants of past shocks, and of their evolution over time, is not sufficient to address the concerns that unobserved channels may be relevant. In this section, I describe three methods researchers have used to address these hurdles.

Natural Experiments and Persistence

The first method combines historical natural experiments that affected culture or institutions with evidence on the persistence of the same traits over the centuries. This method aims to obtain quasi-exogenous variation in the spatial distribution of a trait at the time it first appeared in history. By verifying that the spatial distribution caused by the historical shock has persisted until the current day, researchers reduce the concerns about the fact that historical shocks affected observables and unobservables other than the dimension of interest over time. Pascali (ming) uses this approach to explain the early development of banking systems across Italian towns. Jewish demographics drove the initial spatial distribution of banking activities in Italy: cities with more Jews were more likely to develop a more complex banking system at the time banking started in history. Pascali (ming) verifies the persistence of this spatial distribution of banking over the centuries: cities with more developed banking systems in the past have more developed banking systems today. The transmission of the shock has a peculiar twist in the Italian case: current banking is not more developed because the spatial distribution of Jewish demographics did not change, but because of the Catholic response to the early specialization of Jews in banking. Community banks were created to compete with Jewish banks. Whereas Jewish banking has disappeared over time, community banks have persisted, and they constitute the backbone of the current Italian banking system. Pascali (ming) builds on detailed historical knowledge of the first determinants of banking - Jewish demographics - and the evolution of banking - the persistence of Catholic community banks - to address the issue that historical Jewish demographics may have driven current banking development through unobserved channels. Pascali (ming) also exploits the quasi-exogenous variation in the spatial distribution of Jewish demographics resulting from the expulsion of Jews from the Italian cities under the Spanish Crown in 1541.

To understand current stock market participation, D'Acunto, Prokopczuk, and Weber (2015) build on the selection of Jews into financial services in Germany. In counties where Jews were persecuted more in the past, anything associated with Jews should have been distrusted more, including financial services. To get a shock to the likelihood that Jews were persecuted in a county in the Middle Ages, the authors use the forced migrations of the first German Jewish communities out of the Rhine Valley due to crusaders who traveled from Britain towards the Holy Land. Figure 3.9 depicts the argument. Because of transportation costs, Jewish communities slowly migrated east over the decades. The authors draw isodistance curves parallel to the Rhine Valley, that is, lines along which the likelihood of hosting Jewish communities in the past was similar, with the likelihood declining going eastward. They build on Voigtlaender and Voth (2012) to confirm that the early presence of Jewish communities is correlated with the early persecution as well as the Nazi persecution. The early persecutions of Jews also explain German households' current distrust in finance. Armed with these results, the authors design a three-stage IV strategy in which current distances from the Rhine Valley instruments for the likelihood that a German county hosted a Jewish community in the past, which instruments for the extent of early persecutions against Jews, which captures current distrust in finance.

D'Acunto (2015c) exploits the diffusion of the printing press after 1450 as a quasi-exogenous shock to the spatial distribution of literacy rates across European regions at the time mass literacy started. The argument is based on Dittmar (2011) and Barbier (2006), who show the printing press diffused concentrically out of the city of Mainz, where it was invented. European regions farther away from Mainz obtained printing presses later than regions closer to Mainz. The concentric diffusion of the printing press allows for the drawing of isodistance circles out of Mainz (Figure 1.3). Along the circles, the likelihood of having printing presses early was constant, and declined as the radius of the circles lengthens. D'Acunto (2015c) shows the early distribution of literacy implied by the diffusion of the printing press is consistent with the current spatial distribution of basic education. The author then uses the distance of a region from Mainz to instrument for the effect of basic education on current financial outcomes.[11]

High Resolution Fixed Effects Analysis

A second method is the design of fixed-effects analyses that plausibly allow controlling for time-invariant unobservables within countries or smaller partitions. Grosjean (2011) exploits the history of fragmented borders across countries in southeast Europe to test for the long-run effects of financial development. She uses the variation in the enforcement of the ban on lending across areas of a same current country that were under the Ottoman Empire or not before the 19^{th} century. This setting allows the author to keep constant institutional and cultural dimensions of current countries that only appeared after the fall of the Ottoman Empire.

D'Acunto (2015c) builds on Michalopoulos (2012) to increase the resolution at which the within-analysis is run from the country to the regional level, even if the main observables in his analysis are at the regional level. He constructs virtual regions, that is, arbitrary spatial aggregations from the intersection of a grid with the map of Europe. Because most regions enter multiple virtual regions, the design allows controlling for time-invariant regional effects. This approach is helpful when institutions and culture vary within countries; hence, averaging out country-level effects does not help.

Channels of transmission

The methods described above do not allow pinning down the channels of transmission of the long-run effects of institutions or cultural traits on current outcomes. D'Acunto, Prokopczuk, and Weber (2015) attempt to disentangle three demand-side channels that could drive the negative effect of historical Jewish persecution on current stock market participation: (i) persecution could proxy for retrograde counties, where people distrust stocks as well as any

[11] I am not aware of studies in finance that use historical regression discontinuity designs (RDD) to identify cuasal effects over time. Although unrelated to finance, Dell (2010) introduces a two-dimensional spatial RDD to identify the effect of a historical extractive institution, the *mita*, on long-run development-related outcomes.

other non-familiar things; (ii) persecution could capture counties that are more antisemitic today and consciously invest less in stocks, because they still associate Jews with stocks; (iii) persecution could capture a deep-rooted norm of distrust in finance that has transmitted autonomously from antisemitism, and is still vivid today. The three hypotheses have alternative predictions for the transmission of the effect through education. More educated households are less retrograde and less antisemitic today. But because the curricula of German schools do not include the benefits of investing in stocks, education should not moderate the effect of historical Jewish persecution on current distrust in finance. Indeed, education reduces the backwardness and antisemitism of households, but it does not reduce their distrust in finance, and it does not reduce their likelihood to invest in stocks.

1.5 Future Directions

History & Finance is a recent approach to finance research, and has only involved a few subfields of finance so far. It naturally follows that the scope and methodology of History & Finance await substantial broadening along several dimensions.

Reinterpreting known historical experiments

The first avenue is to reinterpret historical natural experiments already analyzed in the economic history literature in light of their potential to help explain current financial outcomes. This approach seems the most promising first step to advance History & Finance as mainstream finance research. Clearly, only a small subset of previously studied experiments have the characteristics to be used in History & Finance. First, the researcher needs to propose convincing arguments for why the shock has affected economic or financial observables in the past, and why the effect has persisted over the centuries. The arguments need to be quantitative in nature, but also build on the historical and institutional development of the setting under study. Because the original work that proposed the experiment only looked at its effects in the short run, the long-term arguments must be newly developed. Second, the researcher needs to provide corroborating evidence that the alternative channels plausibly affected by the shock are not likely to be relevant to the effect under study. Third, the researcher needs to build on one of the methods discussed in section 4, or on novel methods, to strenghten his/her causal argument. Collaboration with economic and financial historians seems the most promising way to ensure finance scholars are aware of the complex institutional details of the setting under study, and that historians are aware of the methodological issues to be addressed for plausible claims about causal effects over the centuries.

Proposing novel experiments and data

A harder yet more fruitful avenue is to propose novel experiments, which would move hand-in-hand with collecting novel historical data. Finding new experiments requires a detailed knowledge of the history of countries and institutions; hence, the contribution of historians is invaluable. Finance researchers should also bear in mind that most historians do not necessarily share the notion of a quasi-exogenous shock to economic or social observables.[12] An enormous amount of knowledge on historical events across countries is available that could be interpreted as natural experiments even if they have not yet been interpreted as such. According to McCloskey (1976), "History is society's laboratory." Whereas the historians contribute detailed knowledge of the facts and their dynamics,the responsiblity of those trained in economics is to judge whether a set of facts constitutes a quasi-exogenous shock to some observables. Identifying a novel natural experiment raises the issue of collecting the historical data needed to study the shock at the time it happened. Historians' archival experience is invaluable in detecting meaningful and accessible data sources.

Expanding the evidence to other countries

The peculiar facts and institutions that determined the history of a country or a territory can be transformed from intellectual curiosities into general-interest contributions to understand the current behavior of financial actors. Researchers based outside the United States are in the best position to access the information and data needed to test for the effects of historical shocks on current outcomes in their countries, mainly due to language and organizational barriers that U.S. researchers face. Whereas replicating well-known results in finance in settings different from the original one is often not perceived as a general-interest contribution, proposing historical experiments and testing for their effects on current outcomes in countries other than the United States becomes a matter of general interest because it allows understanding the current reality in a way that U.S.-based analyses cannot achieve.

Involving other fields of finance

So far, History & Finance has progressed in two directions: exploiting long financial series in asset-pricing-based analyses, and incorporating methods from political economy into the subfields of finance for which the links with political economy are easier to establish, such as banking, corporate finance, or household finance. Other subfields of finance are good candidates for enlarging the scope of History & Finance. For instance, researchers that study investment management and its effects on individual and aggregate outcomes could benefit from identifying the effects of institutional or cultural incentives of money managers through the history and tradition of money management in their countries. Moreover, the recent movement in finance that studies the effects of heterogeneities in preferences and

[12] Historians here refers to those scholars in history departments that do not specialize in economic or financial history research, and hence have no training in economics.

beliefs across investors and managers may benefit from studying the deep roots of these heterogeneities, and identifiying their sources, to better understand their consequences and policy implications. The body of research that studies the dynamics of corporate finance outcomes could also benefit from empirical tests in the long run, where initial conditions are set, and the evolution of financial decisions over time can be observed.

Proposing new methods to address the challenges of History & Finance

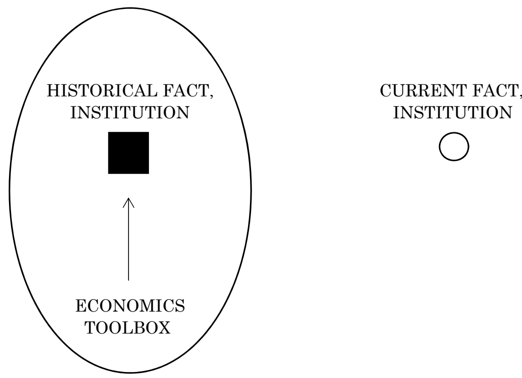
Finally, methodological contributions are necessary for the evolution of the field. Section 4 reviews the methods used so far to reduce the concerns that History & Finance poses in terms of threats to the identification of causal effects over long periods of time. More thoughts are needed to discipline and generalize this state-of-the-art methodology. Contributions are especially awaited on methods assigning bounds to the magnitudes of estimated effects based on the several dimensions that could bias the estimates over the long run. Moreover, advances in methods that exploit the channels of transmission of historical shocks are a promising avenue. For instance, the economic literature has used immigrants as a source of variation for the lifetime exposure to institutions and culture. But immigrants are selected groups of individuals in several dimensions. Identifying alternative sources of variation that do not involve selected populations would be of great help in sharpening results based on the channels of transmission.

1.6 Conclusions

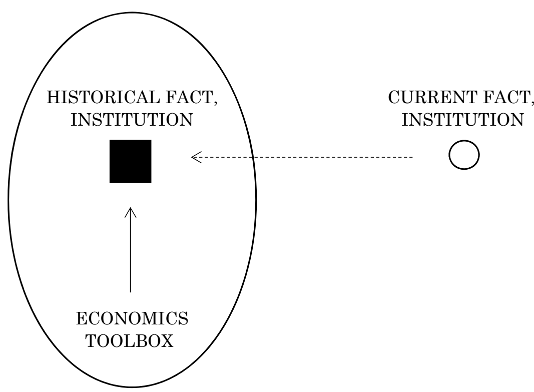
I describe the methods and results of a recent approach to finance research, which I label History & Finance. In History & Finance, current financial phenomena are the main object of interest, and the history toolbox helps investigate these phenomena. The history toolbox includes historical natural experiments and detailed institutional knowledge of historical facts and their evolution over time. This approach differs from the fields of Economic and Financial History, where the main object of research is the historical dimension, either because it is worthy of interest by itself, or because it may inform the present by analogy. I argue that History & Finance may enlarge the scope of the research of financial historians by making it a general interest research effort, as well as the scope of finance research by enlarging the span of available natural experiments over time and across countries. I survey the History & Finance work, which has mainly involved asset pricing and those fields that more closely connect to applied microeconomics and political economy. I also discuss the major challenges that History & Finance poses for empirical researchers, and some of the methods that have been used so far to reduce these concerns. I then propose dimensions along which the field of History & Finance can be enlarged, and which could trigger fruitful cooperations between financial historians and other finance scholars. First, historical natural experiments already analyzed in the economic and financial history literature may have implications over the long run. Second, the substantial body of knowledge produced on historical facts and past institutions without an eye to the incentives of financial actors represents a goldmine for research efforts in History & Finance. Third, peculiar facts and institutions across countries are a promising avenue, especially for researchers in those countries, who face lower barriers in their access to historical sources and institutional information. Fourth, other fields in finance should also be involved in the efforts of History & Finance, especially those that test for the effects of the heterogeneous preferences and beliefs of economic agents, and could benefit from identifying the deep roots of such heterogeneities. Finally, History & Finance needs methodological contributions that address the challenges posed by the identification of long-run causal effects. Economics and finance have only recently tackled these challenges. History & Finance is a field of research and a set of methods that are in their infancy: the creative and rigorous efforts of many researchers are needed to bring the field to a mature stage.

Figure 1.1: Economic and Financial History vs. History & Finance

A. First approach to research in Economic and Financial History



B. Second approach to research in Economic and Financial History



C. History & Finance

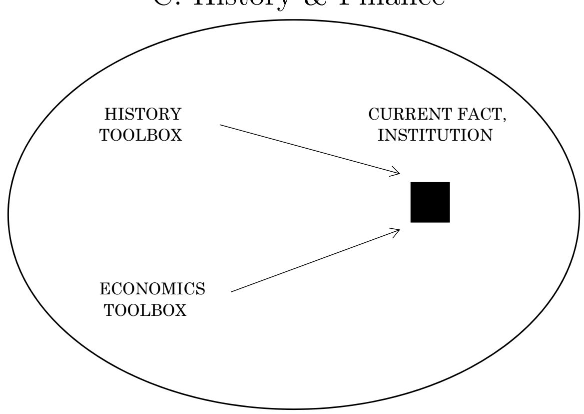
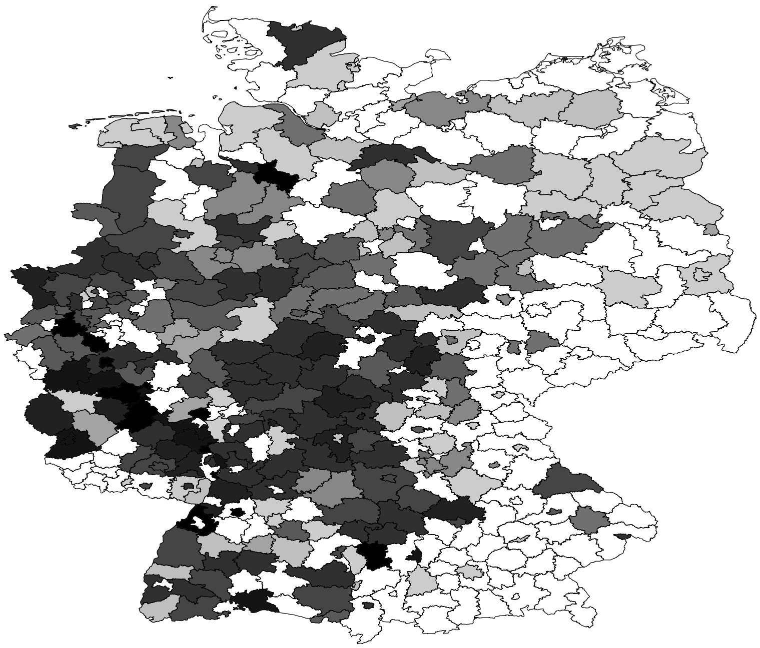
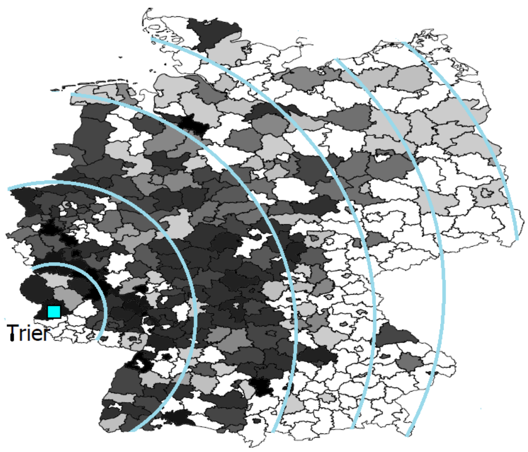


Figure 1.2: Isodistance curves 1: Distance from the Rhine Valley and Jewish Communities

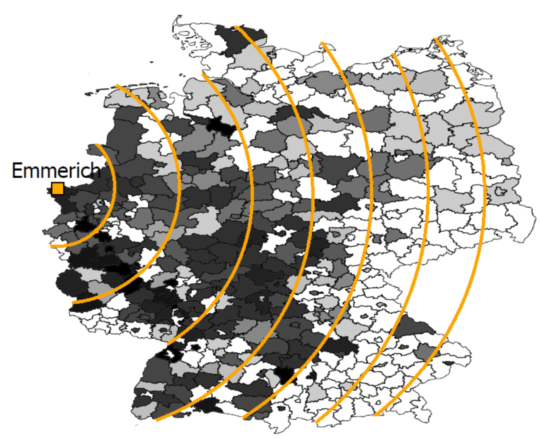
(a) Year when the first Jewish community was documented



(b) Distance from Trier

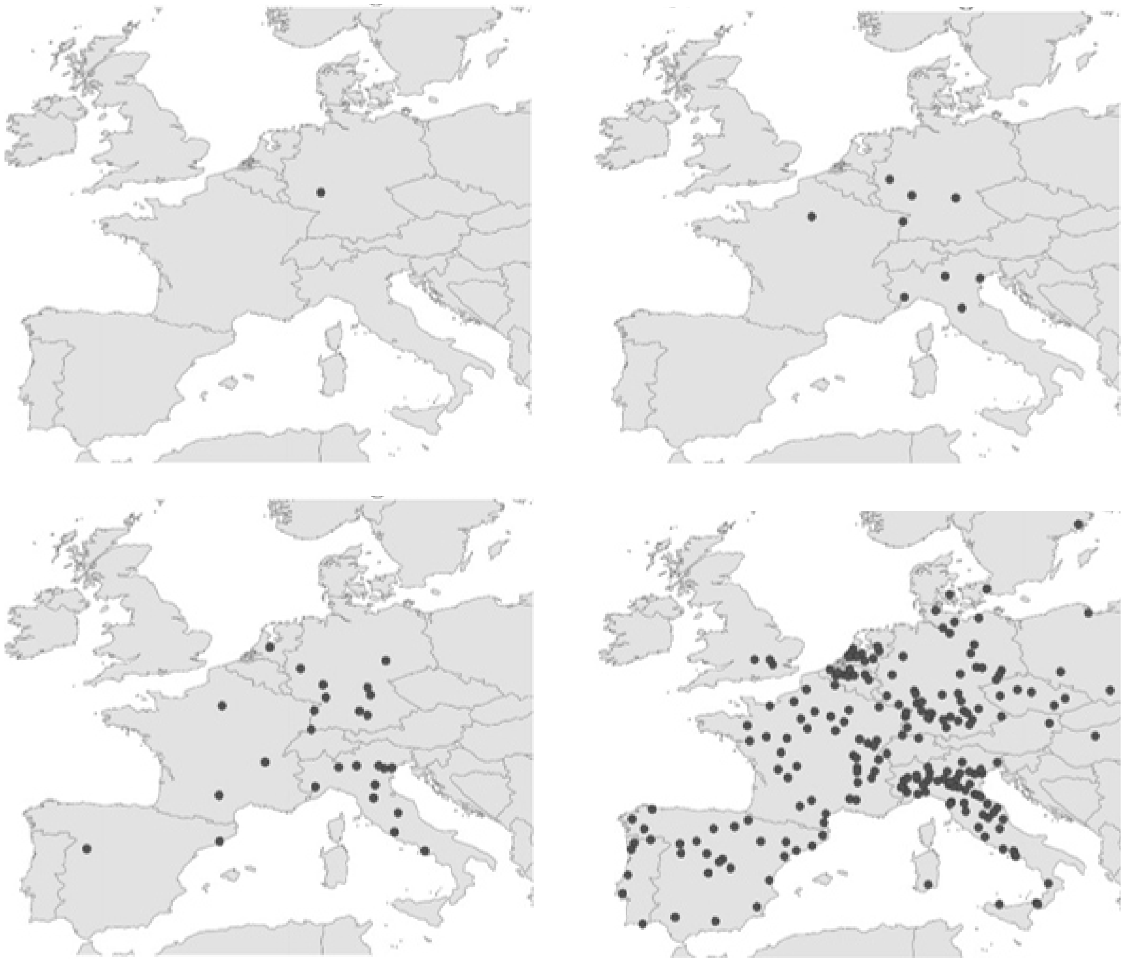


(c) Distance from Emmerich



In the maps above, the darker a county is, the earlier a Jewish community was documented in that county. Blank counties are those for which the data are not available. The bottom maps show the location of the cities of Trier, on the Mosel, and Emmerich, on the northern end of the German Rhine. The isodistance curves out of the two cities emphasize which counties are at the same distance from Trier or from Emmerich.

Figure 1.3: Isodistance curves 2: Diffusion of the Printing Press out of Mainz after 1450



The four pictures are an exact replication of Figure III in Dittmar (2011).

Chapter 2

Innovating to Invest: The Role of Basic Education

2.1 Introduction

Despite the limelight on high-tech innovation, 79% of the patents filed each year in the United States, more than 75% of those filed in Europe, and virtually all the unpatented innovation, are process and product improvements in traditional manufacturing industries (USPTO (2012), EuropeanCommission (2013)).[1] Traditional industries produce 54% of the American and 50% of the European manufacturing value added (BEA (2014)), but we know surprisingly little about the determinants and effects of their innovation.

In this paper, I study a neglected determinant of innovation. I show that the basic education of blue-collar workers increases the innovation activities of traditional manufacturing firms, and affects their investment and capital structure. Hence, firms can benefit from the formal education of blue-collar workers, and not only from their learning-by-doing within the firm. The results suggest that policies that discourage high school dropping out and increase the quality of high school education might be beneficial not only to workers, but to firms as well.

Patents in traditional industries, but not high-tech patents, increase with the amount of high school graduates (extensive margin of basic education), and with the years of schooling of inhabitants without a college degree (intensive margin of basic education) in European

[1]All manufacturing patents not produced in high-tech sectors based on the R&D-intensity-based definition of the *International Patent Classification* (IPC) are traditional manufacturing innovations (WIPO, 2006). There are eight sectors (Aviation, Computer, Communication Technology, Lasers, Micro-organism and Genetic Engineering, Semi-conductors), and several manufacturing sub-sectors that are classified as high-tech.

regions.[2] A 10% increase in the amount of high school graduates increases patents by 10%, and a 10% increase in years of schooling increases patents by 14%, after controlling for regional observables and limiting the variation within countries. The size of these associations is about 60% of the magnitude of the relationship between college education and patents in the sample. The associations are mainly driven by the regions with the lowest population density, the most rural regions, which are normally neglected by the focus of the innovation literature on high-tech clusters, cities, and areas with a large scope for agglomeration economies.

Kogan, Papanikolau, Seru, and Stoffman (2013) find that firms that innovate more have higher investment, and attract more capital.[3] Hence, basic education may affect firm-level investment and capital structure through innovation. In a unique data set on all the patented and unpatented innovation activities of 14,759 firms across 141 European regions, I confirm that higher regional basic education increases the product and process innovation of firms operating in the region. The effect is higher in firms with more basically educated employees, and is driven by traditional manufacturing firms. Consistent with Kogan, Papanikolau, Seru, and Stoffman (2013), basic education increases the capital expenditures and long-term debt of firms that innovate, but not of others.

The endogeneity of regional basic education to economic, institutional, and cultural dimensions poses a set of empirical hurdles. First, region-level policies and unobservables may determine both education and innovation. Second, unobservables such as local economic conditions, culture, and genetics, vary within regions and spill over across regions. Third, reverse causality may drive the results: regions with more innovation in traditional industries may attract better schools or individuals with higher incentives to stay in school, irrespective of an effect of schooling on innovation.

To address the issue of regional unobservables, I run the analysis at the level of European counties (NUTS 3 level), which allows averaging out any regional time-invariant effect, and I confirm the baseline results. The effects are larger in magnitude when excluding counties with at least one urban area above 200,000 inhabitants. Hence, consistent with the regional analysis, counties with a large scope for agglomeration economies do not drive the results.[4]

To address the issue of unobservables that vary at levels other than regions, I build on Michalopoulos (2012) and Michalopoulos and Papaioannou (2013) and develop a method, *Virtual Regions*, that allows me to control for unobservables flexibly within groups of neighboring counties. I impose a grid of 100 km by 100 km on the map of Europe.[5] The squares

[2]Focusing on European regions instead of US metropolitan statistical areas allows me to exploit a historical natural experiment to address the issue of reverse causality between manufacturing innovation and basic education.

[3]A growing literature in finance studies the effect of external and internal financing on innovation inputs such as R&D expenditures, including the role of LBOs (Opler and Titman (1994); Lerner, Sorensen, and Stromberg (2011)), venture capital (Gompers (1995), Kaplan and Stromberg (2003), Kortum and Lerner (2000)), and angel investing (Kerr, Lerner, and Schoar (2014)).

[4]Commutes as short as 15 minutes allow spanning multiple counties in most regions; hence, the pool of blue-collar workers that firms can hire hardly coincides with the workers in the county where they operate.

[5]I project the map of Europe with a cylindrical equal-area projection to avoid the distortions of geo-

of land the grid creates are the Virtual Regions. The Virtual Regions' borders are arbitrary, and overcome the seldom-addressed issue of the endogeneity of regional borders. To construct the Virtual-Region data set, I sum the county-level variables for the counties that fall inside each Virtual Region.[6] For counties that are split across two or more Virtual Regions, I assign the county-level values to the Virtual Region that covers the largest part of the county.[7]

Because Virtual Regions do not fully overlap with regions, I can add regional fixed effects to the Virtual-Region specifications. To define the regional fixed effects, I assign each Virtual Region to the region that comprises its largest part. Thus, adding regional fixed effects means I only exploit the variation across the Virtual Regions that are mainly covered by the same underlying region. This procedure allows me to control for any spatially correlated unobservables common to neighboring Virtual Regions. The magnitude of the effect of basic education on patents remains similar to the baseline analysis: a 10% increase in the amount of high school graduates in a Virtual Region increases the Virtual-Region patents by 12%.

The configuration of the Virtual Regions and the shape of the grid are arbitrary. To test if the results are sensitive to the configuration, I repeat the analysis using four alternative shapes of the grid. For each shape, I move the grid in 27 positions, which repeatedly translate it: (i) 10 km east, (ii) 10 km north, (iii) 10 km north east. Alternative shapes and positions modify the composition of the Virtual Regions and the associated regional fixed effects. Across 108 alternative configurations, 106 estimated coefficients (98.1%) for the effect of basic education on patents are different from zero at the 5% level of significance. The mean of the coefficients is 12%, and the standard deviation is 2.5 percentage points. The estimated coefficients are similar across subsample analyses, such as excluding the Virtual Regions with the highest number of patents, the smallest Virtual Regions,[8] the Virtual Regions with the highest population density, and each country at a time.

I then consider the issue of reverse causality between basic education and innovation. Ideally, one would need a source of exogenous variation in regional basic education at a time when manufacturing did not exist, and such variation should have persisted over time. I first establish that basic education persists in regions for decades, because literacy rates around 1880 predict the present-day amount of schooling for inhabitants with a high school degree or lower. This result is surprising, because European regions were differentially exposed to several institutional and economic shocks between the 19^{th} century and the present. Institutions and wealth are considered major determinants of human capital, but their abrupt changes across European regions have not modified significantly the pre-existing spatial distribution of basic education.

---

graphical projections (see Dell (2009)).

[6]Figure A.2 of the Appendix plots the densities of some regional observables when reconstructed at the Virtual-Region level.

[7]The pieces of land in squares that do not cover the majority of any county, typically coastal areas, do not constitute a Virtual Region.

[8]I cannot split the largest Virtual Regions, because the maximal size of a Virtual Region is a full square, and 64% of the Virtual regions have full size.

I then consider an epitomic shock to the incentives to acquire literacy in the past: the diffusion of the printing press after 1450 AD out of the city of Mainz, in current Germany (Dittmar (2011), Barbier (2006)). The printing press reduced substantially the cost of printed sources, and hence increased the incentives to acquire literacy. It was a quasi-proprietary technology, and it diffused concentrically out of the city of Mainz because of transportation costs. Being farther away from Mainz reduces regional literacy in 1880 and present-day basic education. It has no effect on college education, values, beliefs, or the quality of historical institutions.

The within-country distance from Mainz allows me to instrument for historical literacy to explain present-day innovation if it satisfies a demanding exclusion restriction: (i) the distance should not affect present-day innovation through historical channels different from historical basic education; and (ii) no unobservables should be systematically correlated with the distance and with innovation.[9] This exclusion restriction is untestable. I provide a set of results and placebo first-stage analyses that fail to document violations of the exclusion restriction. Historical literacy instrumented with the within-country distance from Mainz increases traditional manufacturing innovation in regions and firms. The results do not change if I exclude Germany, Southern countries, or Eastern post-Communist countries.

The IV analysis allows rejection of the hypothesis that manufacturing innovation itself drives the whole relationship between basic education and current innovation and investment. As for the channels that may have transmitted the effect of historical literacy on innovation, I provide correlational evidence that the quality of historical and current regional institutions, college education, the dispersion of GDP within regions, or generalized trust, do not appear to have transmitted the effect of historical literacy on innovation. Present-day basic education appears to be a channel through which historical literacy affects innovation.

The main contribution of the paper is to shed light on a neglected driver of innovation in traditional industries, namely, the education of blue-collar workers. Moreover, the paper establishes that the quality of basic education is persistent in locations, and the economic and institutional shocks of the last decades did not significantly modify its spatial distribution. Traditional industries are neglected producers of innovation, and the results in this paper raise several questions for economics, finance, and strategy scholars, such as how the financing of high-tech innovation and innovation in traditional industries differ, how managers' and workforce characteristics interact to produce innovation, and the extent to which this type of innovation might boost economic development in countries without the resources for high-tech endeavors.

Related Literature

The paper belongs to four strands of literature in finance and economics.

[9]The exclusion restriction does *not* require that present-day basic education is the only channel through which historical basic education affects outcomes, because the distance from Mainz aims to instrument for historical literacy, and not for present-day basic education.

A growing literature in finance studies the relationship between financing and innovation investment in firms, such as R&D investment. The focus has been on the effect of LBOs on innovation investment (Opler and Titman (1994); Lerner, Sorensen, and Stromberg (2011)), the role of venture capital in financing innovation (Gompers (1995), Kaplan and Stromberg (2003), Kortum and Lerner (2000)), and the effects of angel investing (Kerr, Lerner, and Schoar (2014)). In this paper, I build on Kogan, Papanikolau, Seru, and Stoffman (2013), who find that higher innovation brings firms to invest more and to attract more capital. I focus on the effect of innovation in traditional industries, which is incremental and requires little or no R&D investment, on the subsequent investment and financing of firms.

The paper contributes to the research that investigates the causes and consequences of innovation. The pioneering work of Schumpeter (1911) and Veblen (1904) shaped this area. Recent theoretical and empirical analyses include Manso (2011), Desmet and Rossi-Hansberg (2012), and Moretti and Wilson (ming). Evidence exists on the role of input sharing and financing (e.g., Gompers and Lerner (2001)), peer effects (e.g., Lerner and Malmendier (2013)), wrongful discharge legislation (Acharya et al. (2013)), patenting laws (Moser (2005), Galetovic, Haber, and Levine (2014)), and agglomeration economies (e.g., Saxenian (1994), Chatterji, Glaeser, and Kerr (2013)), in the production of innovation. I contribute by studying the role of the formal education of workers on innovation in traditional industries, as opposed to specialized education and high-tech innovation, which is the focus of most previous literature. Formal education may complement the effect of learning-by-doing inside the firm (e.g., Levitt, List, and Syverson (2013)).

The paper also fits in the extensive body of micro- and macroeconomic research on the effects of human capital on firm-level and aggregate productivity and growth. Pioneering contributions are Mincer (1958), Becker (1962), Nelson and Phelps (1966), Easterlin (1981), Lucas (1988), and Barro (1991). Goldin and Katz (2008) dissect the relationship between mass education and technological advancement over time, and their effects on labor markets. Acemoglu and Autor (2012) discuss the interactions between skills and technologies and their implications for economic growth. Other contributions include Glaeser et al. (2004), Aghion et al. (2006), Ciccone and Papaioannou (2009), Becker et al. (2011), Schoellman (2012), Hanushek and Woessmann (2012), and Gennaioli et al. (2013).[10] I contribute by documenting the persistence of basic education in regions, and its effects on innovation and investment.

I also investigate the deep roots of present-day economic outcomes (Spolaore and Wacziarg (2013), Nunn (2014)). Acemoglu et al. (2001) document the long-run effects of colonial institutions on European colonies' growth. Putterman and Weil (2010) look at historical migratory patterns and growth. Dell (2010) studies the long-run effects of the Andean mining Mita on local outcomes, and Glaeser et al. (2014) study the long-run effects of carbon mines on the entrepreneurship of cities established nearby. Recently, historical arguments have helped explain present-day financial outcomes, a nascent approach discussed by D'Acunto

[10]Whether human capital is a fundamental or proximate cause of growth is the subject of a long-term debate. The most recent discussion is in Acemoglu et al. (2014). The shock to the spatial distribution of historical basic education in section 7 is unrelated to the quality of historical regional institutions.

(2015a). In D'Acunto, Prokopczuk, and Weber (2015), past Jewish persecution reduces stock investments. In this paper, I use an historical shock to the distribution of literacy to study the effects of past basic education on innovation and investment.

2.2 Data

The unit of observation for the first part of the analysis is a European region. I assemble European data for the analysis because I cannot design a test to address the reverse causality of manufacturing innovation on education based on a historical natural experiment at a time when manufacturing did not exist in the United States. European regions are the NUTS 2 entities in the official administrative taxonomy of the European Union. The average region is a square with a side of 132 km (82 miles), and the average population is 1.8 million inhabitants. NUTS 2 entities generally coincide with the second-highest level of local government in each country, such as the *Regioni* in Italy. An exception is Germany, where NUTS 2 entities are the *Regierungsbezirke*, that is, subdivisions of the states. The baseline units of observation are the regions because commutes as short as twenty minutes allow traveling across several counties, the smaller available partitions, even in the largest regions. Hence, the pool of blue-collar workers firms can hire is not confined to the counties where firms operate.

The regional data include geographic, demographic, and institutional characteristics. The main source for the geographic and demographic characteristics is the Eurostat Regional Database. The database includes the shares of regional population with a high school and a college degree, but it does not include the average years of schooling of inhabitants by education levels. I measure them as the average years of schooling of the respondents to the World Value Survey Wave 9, covering the period from 1999-2004. This measure is not available for regions in Austria, Belgium, Denmark, Greece, and Portugal. The demographic and patenting regional data for the cross-sectional analysis of the next section thus refer to 2005 to guarantee consistent results across the measures of education. All the results in the paper are similar if I use the regional data for any other year from 1998 to 2008, for which both patenting and demographic data are available at the regional level.

I measure the index of the quality of cultivable land at the regional level by averaging for each region the underlying 1 by 1 degree raster data of Ramankutty et al. (2002). The index of the quality of cultivable land allows controlling for the productivity of the agricultural sector in each region, because agriculture is an alternative sector of employment to manufacturing for individuals with basic levels of education. I collect the first-wave results of the *EU Regional Competitiveness Index - RCI*, described in detail in Annoni and Kozovska (2010); I use the components on the quality of institutions and the quality of infrastructures. To obtain a regional measure of generalized trust, I average the individual level responses from the World Value Survey Wave 9 at the regional level.

The regional patent data are from the PATSTAT-Crios database,[11] described in detail in Coffano and Tarasconi (2014). The database provides fractional count patent data at the levels of regions and counties of Europe based on the European Patent Office PATSTAT raw database. PATSTAT is run by the European Patent Office in collaboration with the World intellectual Property Organisation, Eurostat, and the OECD. It collects raw patent data from 90 patent offices worldwide, including the United Statest Patent and Trademark Office. I use the fractional count by applicant for the priority year 2005 for the main analysis, but also establish the robustness of the results when using the integer count patent data and estimating negative binomial regressions to account for the over-dispersed count nature of the patent data.

I collect the historical regional literacy rates around 1880, the urbanization rates from 1860 to 1880, and the index of quality of historical institutions at the level of regions from Tabellini (2010) for the NUTS 2 regions included in the data set. I integrate the data set of Tabellini (2010) with a set of primary and secondary sources. For the regions of Greece, I compute the literacy rates based on the hand-collected raw population data on literate and illiterate inhabitants of Greek regions from the 1881 national census of Greece.[12] For the regions of Germany and those of Poland that were part of Prussia in the 19^{th} century, I use the raw data on the literate population by historical Prussian counties from the Ifo-iPEHD project. The database is described in detail in Becker et al. (2014). For the regions of Belgium, I compute the historical literacy rates from the province-level data on analphabetism in Ruwet and Wellemans (1978). For the regions of the United Kingdom, I test the robustness of the results to using hand-collected data on the ratio of signed marriage acts from the 43^{rd} Annual Report of the Registrar-General of 1880 at the level of counties. Overall, I can compute the measure of historical literacy rates at the regional level for 228 European regions.

Panel A of Table 2.1 reports the regional statistics. Columns (1)-(3) refer to all available observations, and columns (4)-(6) refer only to the 228 regions for which I can compute the historical literacy data from the primary and secondary sources described above. On average, 173 patents are filed in each region, but the cross-regional variation is high. As for education, in the average region, 48% of inhabitants have high school degrees, whereas 23% of inhabitants have college degrees.[13] Figure A.4 of the Appendix plots the regional distribution of historical literacy. Positive spatial correlation of literacy exists across neighboring regions, which I will take into account in the empirical analysis. Regional literacy rates range from 15% in *Calabria* to 99% in the *Stockholm* region. The standard deviation

[11] The database was called Patstat-KITES at the time when the data were assembled for the analyses of this paper. The recent denomination change follows the change in the denomination of the institute that funds the maintenance of the database in 2014, from KITES to Crios.

[12] The census can be accessed online at the following address: http://dlib.statistics.gr/portal/page/portal/ESYE/showdetails?p_id=11807138&p_derive=book.

[13] In Figure A.1 of the Appendix, I plot the densities of regional variables across four countries, because I only use within-country variation in the analyses. Panel A of Figure A.4 of the Appendix plots the regional distribution of patents per capita. The figures shows that substantial cross-regional variation exists within each country.

of historical literacy is 26 percentage points, and large variation occurs within countries.

Firm-level data are from two sources: (i) Amadeus by Bureau van Dijk, which includes the financials of public and private companies incorporated in Europe;[14] and (ii) the EFIGE/Bruegel-Unicredit Dataset, based on a unique firm-level survey of 14,759 private and public firms in 141 European regions from seven countries (Austria, Italy, Spain, UK, Germany, France, Hungary). Firms were surveyed between 2008 and 2010. Altomonte and Aquilante (2012) describe the data set and the sampling procedures in detail. The firm-level responses include a large amount of soft information absent in the most common data sets used in corporate finance. The responses are organized around several areas that include the firm's ownership structure, workforce characteristics, innovation activities, foreign operations, and financing, among others. I create three dummies that equal 1 if a firm declares it engaged in any product, process, or both product and process innovations in the year prior to the interview. The definitions of product and process innovations are from the *Oslo Manual* (OECD, 2005): a product innovation is "the introduction of a good or service that is new or significantly improved with respect to its characteristics or intended uses." A process innovation is "the implementation of a new or significantly improved production or delivery method."

Panel B of Table 2.1 describes the financials, innovation, and firm-level financial constraints for all the firms with available data in both databases. Firms in the sample have on average 15 million euros in assets. Firm-level investment is capital expenditures (property, plant, and equipment) normalized by previous end-of-year assets. Tangibility - the ratio of tangible assets to total assets - is on average 0.26.[15] Firms have on average 117 employees. About 70% of the firms are controlled by a family or an entrepreneur. In the sample, 49% of firms innovated their products, and 32% innovated their processes. Seventeen percent of the surveyed firms were subject to financial constraints, either because their loan applications were rejected, or they had not applied because they expected a rejection, or they were only able to obtain a fraction of the capital they needed. Figure A.3 of the Appendix plots the densities of the most relevant firm-level variables across four countries.

2.3 Basic Education and Innovation in Regions

In this section, I document that basic education correlates positively with regional innovation in traditional manufacturing.

[14]European private and public firms enter the Amadeus database if they fulfill at least one of three criteria: (i) operating revenues of 10 million euros or higher, (ii) total assets of 20 million euros or higher, and (iii) more than 100 employees. The sample does not include smaller firms.

[15] I do not compute Tobin's Q or profitability measures based on market value, because more than 85% of the firms in the sample are private.

I first consider the extensive margin of basic education, that is, the amount of regional inhabitants with a high school degree.[16] The following is my estimating equation[17]:

$$Ln(Patents)_{r,c} = \alpha + \beta Ln(HighSchool)_{r,c} + \gamma Ln(College)_{r,c} + \theta Ln(NoDegree)_{r,c}$$

$$+Ln(X)'_{r,c}\delta + \eta_c + \epsilon_{r,c}. \quad (2.1)$$

where $Ln(Patents)_{r,c}$ are the log of patents in region r, country c;[18] $Ln(HighSchool)_{r,c}$, $Ln(College)_{r,c}$, and $Ln(NoDegree)_{r,c}$ are the log of individuals with high school, college, or no degrees living in the region in 2005, the year I use for the cross-sectional analysis; X are geographic, demographic, and economic regional controls that include: the log of latitude, log of area, log of population density in 2005, the log of the index of the quality of cultivable land, the *EU Regional Competitiveness Index - RCI* for the institutional quality and infrastructures components (Annoni and Kozovska, 2010), the log of GDP per capita in 2005, and a dummy variable equal to 1 if the region was subject to a communist regime after the Second World War; and η_c is a set of country fixed effects. The log-log specification is implied by the following Cobb-Douglas production function, which is assumed to define the amount of patents produced in region r, country c:

$$Patents_{rc} = (\underbrace{HighSchool_{r,c}{}^{\beta}}_{\substack{\text{Inhabitants with}\\ \text{High School degree}}} * \underbrace{College_{r,c}{}^{\gamma}}_{\substack{\text{Inhabitants with}\\ \text{College degree}}} * \underbrace{NoDegree_{r,c}{}^{\theta}}_{\substack{\text{Inhabitants with}\\ \text{No Degree}}}) * X_{r,c}{}^{1-\beta-\gamma-\theta} * e^{\eta_c * u_{r,c}}. \quad (2.2)$$

In all specifications, standard errors are clustered at the level of groups of regions (NUTS 1 level), as defined by Eurostat. A group of regions normally includes four or five neighboring regions. The standard errors are similar if I cluster them at the country level and correct for the low number of clusters following Cameron and Miller (2011), or if I allow for linearly-decaying spatial correlation in longitude and latitude following Conley (1999).[19]

I report the results for estimating 2.1 in Panel A of Table 2.2. Within countries, a 10% increase in the amount of high school graduates is associated with a 14% increase in the amount of patents filed in the region (column (1)). The estimated association is about 60% as large as the association of patents with regional college graduates. Column (2) adds two controls: the index of regional competitiveness - not observed for all regions - and the regional GDP per capita. Adding GDP per capita as a right-hand side variable is likely

[16]The definition of high school degree for each country follows the level 3 classification of the *International Standard Classification of Education* (ISCED) by UNESCO. In Table A.1 of the Appendix, I report the degrees and certifications classified as high school terminal degrees for each country.

[17]I show the robustness of the results to using alternative specifications below, namely negative binomial regressions to account for the over-dispersed count nature of the patent data, and OLS regressions of the patents per capita on the ratios of regional inhabitants with a high school degree and a college degree.

[18]The number of patents is increased by 1 so that the measure is defined for all regions.

[19]I estimate the most conservative standard errors with a cutoff parameter of 2.5 degrees in longitude and latitude, and across all but one specification, the Conley standard errors are lower than those clustered at the level of groups of regions.

to overcontrol for regional economic activities, because if indeed basic education affects innovation it should also affect GDP per capita through innovation. Yet, it is important to verify how the estimated effect of basic education on innovation vary once a variable that proxies for local economic conditions enters the estimation. A 10% increase in high school graduates increases patents by 10% once I add GDP per capita to the list of controls. In untabulated results, I also add the regional index of generalized trust based on the responses to the World Value Survey (Wave 9), which is not available for 27 regions, and the size of the estimated coefficient is similar to column (2) of Table 2.2 (1.041, s.e. 0.445). In columns (3)-(6), I find the association is driven by patents in traditional industries, whereas it is indifferent from zero for high-tech patents.[20] Figure 2.1 depicts these associations. College educated inhabitants are important for high-tech patents, whereas their effect on patents in traditional industries is lower than the effect of high school graduates.

I move on to document the positive association between the intensive margin of basic education, that is, the years of schooling of those with high school degrees or below, and regional patents. I consider the following estimating equation:

$$Ln(Patents)_{r,c} = \alpha + \beta\theta Ln(s_{B,r,c}) + \beta Ln(BasicEducated)_{r,c} + \gamma\theta Ln(s_{C,r,c}) + \gamma Ln(HigherEducated)_{r,c} + Ln(X)'_{r,c}\delta + \eta_c + \epsilon_{r,c}, \quad (2.3)$$

where $s_{B,r,c}$ is the average number of years of schooling for regional inhabitants with high school degrees or lower levels of education, $BasicEducated_{r,c}$ is the amount of regional inhabitants with high school or lower education, $s_{C,r,c}$ is the average number of years of schooling of those with at least some college education, and $CollegeEducated_{r,c}$ is the amount of regional inhabitants with more than high school education.

The average number of years of schooling is computed from the respondents to the World Value Survey, and is not available for all regions. Ten percent more years of schooling for those with basic education is associated with 17% more patents (column (1) of Panel B of Table 2.2). The size of the association is 14% once I add GDP per capita as a controlling variable. The association also stays sizable and statistically significant for mid-low-tech patents (columns (5)-(6)).[21]

In Table 2.3, I show the regional results are robust to the exclusion of the regions that produce most patents, the regions with the highest GDP per capita, and the regions with the lowest GDP per capita. Moreover, I show the robustness of the results to estimating alternative specifications for the relationship between regional patents and basic education. In Panel D of Table 2.3, I estimate negative binomial regressions to account for the over-dispersed count nature of the patenting data. In Panel E, I regress the patents per capita in a region on the ratio of high school graduates, the ratio of college graduates, and the other

[20] The split between high-tech and mid-low-tech patents is available for 247 regions.

[21] In column (3), the association between average years of schooling and high-tech patents is positive, but it loses statistical significance in column (4).

regional controls.

To investigate which variation in the cross section of regions drives the results, in Panel A of Figure 2.2, I plot the coefficients estimated when sorting the regions in three groups based on the ratio of college-educated inhabitants. Regions with the lowest ratio of college-educated inhabitants drive the results. In Panel B of Figure 2.2, I sort regions by population density, which is relevant because (i) high-tech innovation clusters are mostly densely populated and urbanized areas, and (ii) the scope for agglomeration economies is larger in highly urbanized areas (e.g., see Carlino and Kerr (2014)). The least densely populated regions drive the associations.

2.4 Endogeneity Concerns: Three Approaches

I now address a set of concerns about the endogeneity of the association between basic education and regional innovation. First, region-level dimensions such as policies and institutions, may drive both basic education and innovation. Second, several unobservable and non-measurable confounding variables, such as local economic shocks, attitudes toward risk, or ideology, vary within regions and spill over across neighboring regions. Third, reverse causality may explain the results.

Regional Unobservables: County-level Analysis

To account for time-invariant regional unobservables, I exploit the variation across counties, or NUTS 3-level partitions in the official taxonomy of the European Union. The average county is a square with a side of 54 km (34 miles), and the average population is 311,000 inhabitants. On average, a region is composed of approximately 5 counties, although there is variation in the number of counties per region, as well as in the size and population of counties across countries. There are 1,254 counties in the sample, which on average filed 38 patents each in 2005. Using the county as the level of observation allows for the absorption of regional fixed effects common to all the counties in a same region[22]:

$$Ln(Patents)_{k,r} = \alpha + \beta_k Ln(HighSchool)_{k,r} + \gamma Ln(College)_{k,r} + \theta Ln(NoDegree)_{k,r}$$

$$+Ln(X)'_{k,r}\delta + \eta_r + \epsilon_{k,r}, \tag{2.4}$$

where $Ln(Patents)_{k,r}$ is the log of patents filed in county k of region r, and the other covariates are the same as in the regional analysis, but observed at the county level, with the exception of the *EU Regional Competitiveness Index - RCI* for the institutional quality and infrastructures components, which is not available at the county level. Table 2.4 reports

[22] The indices of institutional quality and of generalized trust are not available at the county level.

the estimated coefficient $\hat{\beta}_k$ from 2.4. In column (1), a 10% increase in the number of high school graduates increases county-level patents by 15%; the estimated magnitude of the effect decreases to 11% once I control for regional unobservables in column (2). In columns (3) and (4), I exclude counties that host urban conglomerates with more than 500,000 and 200,000 inhabitants, because counties that host large cities, have a high scope for agglomeration economies, or host industrial districts, may drive the results. Excluding counties that host cities with more than 200,000 inhabitants in column (4) reduces the sample by 46%. The size of the estimated coefficients is, if anything, higher once I exclude from the analysis the counties with large urban conglomerates. This result is consistent with the fact that the baseline association between basic education and patents at the regional level is driven by regions with low population density (Panel B of Figure 2.2).

Unobservables within and across Regions: Virtual Regions

The analysis at the county level with regional fixed effects controls for any source of unobserved heterogeneity that varies at the regional level. But the analysis is still subject to two shortcomings. First, commuting times as short as 15 minutes allow spanning several counties even in the largest regions (see Figure A.5 of the Appendix). The pool of blue-collar workers a firm can hire is likely dispersed across several counties, which suggests counties are not the appropriate unit of observation for the analysis, which instead should rely on larger partitions. Second, local economic shocks and unobservables that are likely to affect both the willingness to acquire education and the innovativeness of firms, such as local economic conditions, ideology, beliefs, or the average risk attitudes of the population, can be spatially correlated across counties irrespective of regional borders, because they were determined before regions existed. For instance, populations in areas across neighboring regions often are affected by the same local economic conditions, speak the same local language, have the same cultural and historical heritage, and these unobservables are plausibly important determinants of the schooling and innovating decisions of households and entrepreneurs.

To address the two concerns with the county-level analysis, I need a tool that allows me to flexibly control for local unobserved dimensions common to neighboring groups of counties irrespective of the regional borders. I build on Michalopoulos (2012) and Michalopoulos and Papaioannou (2013) to construct Virtual Regions, that is, arbitrary aggregations of neighboring counties of smaller size than regions, which pool together neighboring counties in the same or in different regions.

To create the Virtual Regions, I project the map of Europe with a cylindrical equal-area projection to avoid the distortions of geographic projections (Dell (2009)), and I impose a grid of squares of 100 km by 100 km on the map. All the intersections of the squares with land are the Virtual Regions. The shape and configuration of the grid are arbitrary; hence, they cannot be endogenous to any economic, institutional, or political dimensions that have shaped the regional borders over time.[23]

[23]I will show the results are insensitive to the shapes and configurations of the grid in the next section.

I use the Virtual Regions as a new unit of observation. Figure 2.3 describes the construction of the Virtual-Region data set for the variable *Patents* across three Italian regions. Panel A reports the number of patents at the regional level for the three regions in the example. I sum up the county-level patents for the counties that enter each Virtual Region (Panel B). On average, three to four counties enter a Virtual Region. If a county is split between two Virtual Regions, I assign the county-level value to the Virtual Region that covers the majority of the county. This rule implies that small Virtual Regions, typically coastal areas, which cover no counties, are dropped from the analysis. I then sum up the regional portions of patents inside each Virtual Region (Panel C). Virtual Regions like A and B in Panel C of Figure 2.3 become the new units of observation. Virtual Regions are arbitrary groups of counties, which may belong to the same region or to neighboring regions.

Because Virtual Regions are not composed by the same counties as regions, the distributions of variables differ across regions and Virtual Regions. Figure A.1 and Figure A.2 of the Appendix plot the distribution of variables across Virtual Regions. Also, the borders of the Virtual Regions do not coincide with those of the regions. I index each Virtual Region by the underlying region $r*$ that covers the majority of its area. The index $r*$ emphasizes that Virtual Regions are not fully included in the underlying regions r. The following is the specification at the Virtual-Region level:

$$Ln(Patents)_{v,r*} = \alpha + \beta_v Ln(HighSchool)_{v,r*} + \gamma_v Ln(College)_{v,r*} + Ln(X')_{v,r*}\delta_v + \eta_{r*} + \epsilon_{v,r*}. \tag{2.5}$$

2.5 includes a full set of fixed effects for the underlying regions $r*$, η_{r*}.[24] These fixed effects are not collinear with the Virtual-Region variables, because the borders of the Virtual Regions do not fully overlap with those of the underlying regions. Adding the regional fixed effects allows demeaning all the variables by the average across the Virtual Regions indexed by the same $r*$:[25]

$$Ln(Patents)_{v,r*} - \overline{Ln(Patents)}_{r*} =$$

$$\beta_v(Ln(HighSchool)_{v,r*} - \overline{Ln(HighSchool)}_{r*}) + (Ln(X')_{v,r*} - \overline{Ln(X')}_{r*})\delta_v + (\epsilon_{v,r*} - \bar{\epsilon}_{r*}), \tag{2.6}$$

where

$$\overline{Ln(HighSchool)}_{r*} = \frac{1}{N} \times \sum_{1}^{N} Ln(HighSchool)_{v,r*},$$

for the N neighboring Virtual Regions that are indexed by the same $r*$. In the sample, the number of Virtual Regions indexed by the same $r*$ is one, two, or three. If only one Virtual Region is indexed by a $r*$, the information included in that square effectively

[24] Of the 269 regions in the baseline analysis, 242 are associated with at least one Virtual Region.

[25] Below, the variable *College* is subsumed in the set of controls X.

drops from the analysis. When two or three Virtual Regions are indexed by the same $r*$, 2.6 absorbs any time-invariant unobservables common to the two or three Virtual Regions. These areas are geographically close, because the largest Virtual Regions have a side of 100 km (62 miles); hence, unobservables such as the quality of institutions, values, beliefs, or risk attitudes, hardly vary across neighboring Virtual Regions. Virtual Regions would allow interpreting the estimated effect $\hat{\beta}_v$ in 2.6 causally only if such unobservables were the only threat to identification.

A county-level unobservable that is not spatially correlated across neighboring counties is not absorbed in the Virtual-Region analysis. An example of such unobservables would be a set of policy interventions to increase the county-level basic education, and to increase the firm-level innovation, which only apply to individuals and firms residing in the county. These policies should be implemented in several counties across different Virtual Regions to drive the results. But note that the educational policies of European countries are mainly implemented at the national and regional levels. Policies like transfers or tax cuts to firms are mainly managed at the national level or through European Union funds that are distributed by regions.

Table 2.5 reports the estimated coefficient $\hat{\beta}_v$ in 2.6 across four specifications. In columns (1)-(2), I restrict the variation within countries. This specification allows direct comparison between the estimated magnitude of the effect in the Virtual-Region data set and the regional data set. The comparison is important, because the Virtual-Region distributions are in general different from the regional distribution. The presence of 10% more high school graduates in a Virtual Region increases the patents filed in the Virtual Region by 11.6%. The effect is similar in size to the baseline regional analysis, in which the same increase in regional high school graduates is associated with 14% more patents filed in regions. In column (3), I add the regional fixed effects of 2.6 to a specification without controls. The explained variation in the outcome is close to 1, and 10% more patents in Virtual Regions increase patents by 12.1%. In column (4), I add the covariates measured at the level of Virtual Regions to the specification, and the size of the estimated coefficient is similar. Adding the covariates allows me to address the concern that the procedure to create Virtual Regions may artificially produce systematic variation in the data. This concern does not drive the results, because adding the covariates does not reduce the magnitude of the estimated coefficient.

The last point can be formally tested using the Hausman-type decomposition discussed by Gelbach (2014), which allows derivation of a test statistic for the null that the estimated coefficients with and without covariates are the same. This null hypothesis is equivalent to the hypothesis that the covariate of interest is uncorrelated with the other covariates in the specification. The statistic is defined as follows:

$$t = \frac{\hat{\beta}_{baseline} - \hat{\beta}_{full}}{s.e.(\hat{\beta}_{baseline} - \hat{\beta}_{full})}, \tag{2.7}$$

where $\hat{\beta}_{full}$ and $\hat{\beta}_{baseline}$ are the estimated coefficients for the same OLS specification with or without the covariates of the model.[26] For the specification with country fixed effects, the null is rejected by construction, because the estimated standard errors in column (2) are higher than in column (1). For the specification with regional fixed effects, instead, we can compute the statistic in 2.7, which is 0.35. Hence, after adding the regional fixed effects in the Virtual-Region specification, we cannot reject the null of no correlation between basic education and the other covariates at any plausible level of significance.

This test is relevant to exclude alternative explanations. For instance, cities or industrial districts in Virtual Regions cannot drive the effect; otherwise, adding population density as a covariate would have decreased the coefficient on basic education.

In Table 2.6, I report a set of robustness checks and subsample analyses for the Virtual-Region specification. In columns (1)-(2), I exclude the top 5% and 10% Virtual Regions by number of patents, whereas in columns (3)-(4), I exclude the bottom 5% and 10% Virtual Regions by area.[27] The size of the estimated coefficients is similar to the baseline analysis. In columns (5)-(6), I exclude the Virtual Regions with the highest population density. Consistent with the regional baseline analysis, the size of the estimated coefficient is higher when I drop from the analysis the most urbanized areas, which have a high scope for agglomeration economies.

The grid in the analysis so far was arbitrary. The optimal grid would combine neighboring counties with most unobservables in common. By definition, I cannot define the optimal grid. Absent a criterion to define the optimal grid, I investigate how sensitive the results are to the shape of the grid and to its position. I replicate the Virtual-Region analysis across several alternative configurations, which employ four different shapes and 27 alternative positions for each shape.

The shapes include: (i) 100 km by 100 km squares, like those in the analysis so far; (ii) isosceles triangles of side 100 km, obtained by dividing each square in half along its northwest to southeast diagonal; (iii) parallelograms of side 200 km and height 100 km, obtained by merging two triangles across different squares; and (iv) squares whose sides are the diagonals of the original squares. By construction, the four shapes produce Virtual Regions that generally differ from each other in their county composition.

As for the positions of the grid, I translate each grid in 27 alternative ways to explore alternative positions with non-overlapping grids: (i) 9 times east, in increments of 10 km; (ii) 9 times north, in increments of 10 km; and (iii) 9 times northeast, in increments of 10 km * $\sqrt{2}$.

Translated grids for each shape modify the composition of Virtual Regions as long as counties mainly covered by one square move to another square after the translation. Moreover, the grids produce alternative rules of regional indexing; hence, different groups of Virtual Regions are associated with each $r*$ for each configuration of the grid.[28]

[26]Note that the statistic is derived based on the assumption that the model with covariates is correctly specified.

[27]I cannot exclude the largest Virtual Regions, because the size of Virtual Regions is lower or equal to a full square, and 64% of the Virtual Regions in the sample have full size.

[28]The Virtual Regions and fixed effects for which a 10 km translation changes none of the component counties will stay the same.

I run the specification in 2.6 for each of the 108 shapes and positions of the grid, and hence obtain 108 estimated coefficients and standard errors. Table 2.7 reports the summary of the results. Of 108 estimated coefficients, 106 (98.1%) are significantly different from zero below the 5% level of significance. The mean of the coefficients is 12%, and their standard deviation is 2.5 percentage points. The mean of the estimated standard errors is 0.35. Because the mean of the estimated coefficients is close to 11.8%, the magnitude of the coefficient in the original Virtual-Region analysis (column (4) of Table 2.5), I conclude that the results of the Virtual-Region analysis are not sensible to the shape or position of the grid I use. Table 2.7 also reports the median and mean estimated coefficients separately for each shape of the grid. The within-shape means range from 11.7% to 12.4%.

Reverse Causality: Persistence of Basic Education and Historical Instrument

The analysis so far has shown that region-level unobservables, or unobservables common to counties within and across regions, do not drive the effect of basic education on local innovation. A remaining concern is that reverse causality drives the results: regions with more innovative firms may attract the best schools, or individuals that stay in school longer irrespective of an effect of schooling on innovation. To address this issue, I exploit an exogenous shock to the incentives to acquire basic education in the distant past - when manufacturing did not exist - and whose effects have persisted over time.

Persistence of Basic Education. First, I establish the persistence of basic education in regions. I show that a measure of basic education in the past, the historical regional literacy rate, predicts the present-day amount of basic education at the regional and individual levels, but not higher levels of education. For the regional level, I estimate the following specification:

$$Ln(HighSchool)_{r,c} = \alpha + \beta Literacy1880_{r,c} + X'_{r,c}\gamma + \eta_c + \epsilon_{r,c}, \tag{2.8}$$

where $Ln(HighSchool)_{r,c}$ is the log of the regional population with a high school degree[29] but not higher education in region r and country c; $Literacy1880_{r,c}$ is the regional literacy rate around 1880; $X_{r,c}$ is a set of region-level geographic, historical, and demographic characteristics; and η_c are country fixed effects. In Table 2.8, a 10-percentage-point increase in literacy in 1880 is associated with a 2% increase in the number of current regional high school graduates (columns (1)-(2) of Panel A). Instead, historical literacy is uncorrelated with the current amount of college-educated individuals (columns (3)-(4)). One may be concerned that regions with higher past and current basic education attract more immigrants, but this

[29]The results are similar if I use the share of regional inhabitants with a high school degree as the dependent variable.

concern is inconsistent with the results in columns (5)-(6) of Table 2.8[30] The persistence of basic education from 1880 onward is surprising because European regions were differentially exposed to severe institutional and economic shocks from the end of the 19th century onwards, including two World Wars, the economic crisis of the 1930s, and the experience of totalitarian political regimes.

For the individual level, historical literacy rates in regions predict the years of schooling of current inhabitants with basic levels of education, but not of others. I consider the following:

$$YearsSchooling_{i,r,c} = \alpha + \beta Literacy1880_{r,c} + X'_{r,c}\gamma + D'_{i,r,c}\delta + \tau_{i,r,c} + \eta_c + \epsilon_{i,r,c}, \qquad (2.9)$$

where $YearsSchooling_{i,r,c}$ is the number of years of schooling for individual i in region r and country c, which take parts in the World Value Survey; $D_{i,r,c}$ is a set of respondent characteristics; and $\tau_{i,r,c}$ are five dummies for town size, which aim to eliminate the systematic variation across individuals living in rural or urban areas. Individual characteristics are available for about one half of the observations. In Panel B of Table 2.8, a 10-percentage-point increase in historical literacy is associated with 0.35 additional years of schooling in the full sample (column (1)). A major concern is that those who repeat one or more years of schooling before obtaining an academic degree drive the results. This concern is compelling because most countries impose age-based thresholds before citizens can legally withdraw from school.[31] In column (2), I exclude the variation in schooling across academic degrees,[32] and reassuringly, the effect of historical literacy on schooling disappears: those who stay in school longer because they repeat classes do not drive the persistence result. In columns (3)-(6), I find persistence holds only for individuals with basic education, and not for those with higher levels of education.

Instrument for Historical Basic Education. To address the possibility of reverse causality with the effect of basic education on innovation, one would need a source of exogenous variation in basic education at a time when manufacturing did not exist, and build on the persistence of basic education in regions. In this section, I describe the quasi-exogenous diffusion of the printing press after 1450 out of the city of Mainz, in current Germany (Dittmar, 2011), which is close to such as ideal exogenous shock.

The printing press was invented around 1450 by zu Gutenberg in Mainz (Pfalz), and it was a quasi-proprietary technology. The printing-press technology represented a positive shock to the incentives to acquire literacy for 15th-century European households, because it dramatically reduced the cost of printed sources. In modern Europe, delivering books from the cities where they were printed was as expensive as the monthly wage of a skilled craftsman (Dittmar, 2011). Printing presses started to spread around the city of Mainz through

[30] I focus on within-country immigrants, measured by Eurostat, because (i) out-of-country immigrants are often not registered and (ii) within-country immigrants are usually skilled workers attracted by job opportunities (Moretti, 2012), whereas the motivations of other immigrants may be different.

[31] Strikingly, these regulations are often not enforced. OECD (2013) describes early school leaving across European countries, a phenomenon that varies substantially within countries.

[32] The terminal years of degree cycles vary across countries. I follow the WVS taxonomy for six international levels of education conducive to degrees across different European countries.

zu Gutenberg's own collaborators. The diffusion path was slow and concentric out of the city of Mainz because of transportation costs (Barbier, 2006). The treatment of having a printing press early across regions at a similar distance from Mainz can be assumed to be as good as randomly assigned. Figure 2.4 describes the pattern of diffusion of the printing press across European regions.

I detect negative unconditional correlation between the Euclidean distance of a region from Mainz and the literacy rate in the region in 1880 (Panel A of Figure 2.5). The distance from Mainz is unconditionally negatively correlated with the present-day ratio of inhabitants with high-school degrees, but it is unrelated to the present-day spatial distribution of college education (Panel B of Figure 2.5).

The distance from Mainz can be a valid instrument for historical literacy only if it does not affect current innovation through historical channels unrelated to historical literacy, and only if there are no unobservables that are correlated with both the distance from Mainz and current innovation. The exclusion restriction cannot be formally tested. In Table 2.9, I provide a set of results that aim to inform on the plausibility that the exclusion restriction holds.

First, the minimal Euclidean distance of the centroid of a region from the city of Mainz is not associated with region-level and firm-level observables.[33] In Panel A of Table 2.9, no associations are economically or statistically different from zero, except for a negative association of the distance with the leverage of firms (columns (1) and (3)). In Figure A.8 of the Appendix, I plot the unconditional correlation between the distance from Mainz and several historical, geographic, and current observables, and I do not detect patterns similar to those for literacy rates in 1880.

Second, the distance from Mainz may have determined historical economic growth, or other historical dimensions that favored current innovation but are no more observable. To investigate this possibility, I create a purged measure of distance, which consists of the residuals from an OLS regression of the minimal Euclidean distance from Mainz on a set of geographic and historical observables that may have affected the past economic conditions of regions: the latitude, area, quality of cultivable lands (Ramankutty et al., 2002), whether the region hosted any cities in the Hanseatic League, whether the region was Catholic after the Peace of Augsburg of 1555, and whether the region was in the Communist block after the Second World War. I then regress current region-level and firm-level observables on this purged measure of distance. In columns (2) and (4) of Panel A of Table 2.9, I find no significant associations of the purged distances with any of the current region- and firm-level observables.

Third, I propose two reduced-form specifications, in which the log of the distance from Mainz enters as a covariate together with other regional characteristics, and with the log of historical literacy (columns (2)-(3) of Panel B of Table 2.9).[34] The distance from Mainz

[33]Detecting no associations is not evidence that the exclusion restriction holds, because (i) by definition, I cannot test if any unobservables are correlated with the distance, and (ii) the imprecision of the estimated effects may drive the non-results.

[34]Column (1) of Panel A of Table 2.9 coincides with column (2) of Panel A of Table A.3 of the Appendix.

is negatively associated with current regional innovation when it enters alone, but the estimated association drops in magnitude by about 30% and becomes statistically insignificant when the distance enters the same specification as historical literacy. This result is not easily compatible with an autonomous association of the distance from Mainz with present-day innovation.

Fourth, I run placebo first stages. I predict historical literacy with the distance of a region from cities other than Mainz, which the economic history literature has described as wealthy and proto-industrial in the 15^{th} and 16^{th} centuries (Prague, Amsterdam), or the cradle of nation states and national politics (London, Madrid). I also look at Florence, which was the cradle of Humanism after the 15^{th} century, and Aix-la-Chapelle (Aachen), where Holy Roman Emperors were enthroned as of the 15^{th} century, which is close to Mainz. Panel C of Table 2.9 reports the first-stage statistics for these placebo first stages, and it shows that none of the alternative distances seems to satisfy the relevance assumption for an instrument of historical literacy, although the magnitudes of first-stage test statistics increase with the correlation of each placebo distance with the distance from Mainz.

I run a two-stage least-squares analysis at the regional level, where historical literacy is instrumented with the distance from Mainz.[35] Panel A of Table 2.10 reports the outcomes of the second stage, where the dependent variable is the log of regional patents. In Panel B, I report the first-stage statistics: (i) the Cragg-Donald F-statistic, which is based on i.i.d. standard errors, but is used to compute the critical values reported in Table 5.2. of Stock and Yogo (2005); (ii) the Kleibergen-Paap F-statistic, which is computed for correcting the standard errors for correlation of unknown form at the level of groups of regions; and (iii) the Angrist-Pischke chi-square statistic, which can be used for a rank test of the matrix of the reduced-form equation coefficients and the excluded instruments.

In column (1) of Table 2.10, a one-standard-deviation increase of the instrumented log of literacy in 1880 increases the current log of regional patents by 0.57 standard deviations. The size of this estimated effect is about twice the size of the corresponding reduced-form specification (column (2) of Table A.3) of the Appendix. The tests for weak identification and underidentification that use the Kleibergen-Paap F-statistic and the Angrist-Pischke chi-square statistic reject the null hypotheses based on the Stock-Yogo critical value for a worst-case IV bias size of 15% or lower. The results survive when excluding the regions of ex-Communist countries, of southern Europe, of Germany, or all three groups.

$$Ln(Patents)_{r,c} = \alpha + \beta\theta Ln(s_{B,r,c}) + \beta Ln(BasicEducated)_{r,c} + \gamma\theta Ln(s_{C,r,c})$$
$$+\gamma Ln(HigherEducated)_{r,c} + Ln(X)'_{r,c}\delta + \eta_c + \epsilon_{r,c}, \quad (2.10)$$

where $s_{B,r,c}$ is the average number of years of schooling for regional inhabitants with high school degrees or lower levels of education, $BasicEducated_{r,c}$ is the amount of regional inhabitants with high school or lower education, $s_{C,r,c}$ is the average number of years of

[35] In the Appendix, I report the results for regressing current innovation and firm-level outcomes on historical literacy.

schooling of those with at least some college education, and $CollegeEducated_{r,c}$ is the amount of regional inhabitants with more than high school education.

The average number of years of schooling is computed from the respondents to the World Value Survey, and is not available for all regions. Ten percent more years of schooling for those with basic education is associated with 17% more patents (column (1) of Panel B of Table 2.2). The size of the association is 14% once I add GDP per capita as a controlling variable. The association also stays sizable and statistically significant for mid-low-tech patents (columns (5)-(6)).[36]

In Table 2.3, I show the regional results are robust to the exclusion of the regions that produce most patents, the regions with the highest GDP per capita, and the regions with the lowest GDP per capita. Moreover, I show the robustness of the results to estimating alternative specifications for the relationship between regional patents and basic education. In Panel D of Table 2.3, I estimate negative binomial regressions to account for the over-dispersed count nature of the patenting data. In Panel E, I regress the patents per capita in a region on the ratio of high school graduates, the ratio of college graduates, and the other regional controls.

To investigate which variation in the cross section of regions drives the results, in Panel A of Figure 2.2, I plot the coefficients estimated when sorting the regions in three groups based on the ratio of college-educated inhabitants. Regions with the lowest ratio of college-educated inhabitants drive the results. In Panel B of Figure 2.2, I sort regions by population density, which is relevant because (i) high-tech innovation clusters are mostly densely populated and urbanized areas, and (ii) the scope for agglomeration economies is larger in highly urbanized areas (e.g., see Carlino and Kerr (2014)). The least densely populated regions drive the associations.

2.5 Basic Education and the Investment and Financing of Firms

In this section, I study the implications of the effect of basic education on innovation for the investment and financing of traditional manufacturing firms. Kogan, Papanikolau, Seru, and Stoffman (2013) find that firms that innovate more have higher investment, and attract more capital. Hence, basic education should affect firm-level investment and capital structure through innovation. I use a unique data set on the patented and unpatented innovation activities of 14,759 firms to document that manufacturers in European regions with higher basic education are more likely to innovate their processes and products than manufacturers in other regions. They also invest more in capital expenditures and raise more long-term debt as a ratio of total debt.

[36] In column (3), the association between average years of schooling and high-tech patents is positive, but it loses statistical significance in column (4).

Several manufacturing innovations are never patented, because patenting requires financial and organizational resources. Observing *all* the innovation activities of firms, both patented and unpatented, is therefore crucial to studying innovation in traditional manufacturing. I use the unique EU-EFIGE/Bruegel database to estimate the following probit specification:

$$Pr(Innovation = 1)_{frc} = \Phi(\alpha + \beta Ln(HighSchool)_{rc} + X'_{rc}\gamma + F'_{frc}\delta + \eta_c + \eta_a + \eta_s + \eta_l). \quad (2.11)$$

Across three specifications, $Innovation$ is a dummy equal to 1 if the firm declares it engages in product, process, or both types of innovation. X_{rc} and F_{frc} are region- and firm-level covariates, and η_a, η_s, and η_l are firm age group-, size-, and sector-fixed effects. In column (1) of Table 2.11, a one-standard-deviation increase in the log of regional high school graduates increases the likelihood that a firm in the region engages in product innovations by 2.9 percentage points, which is a 6% increase in the average likelihood of product innovation (49%). The same increase in the log of high school graduates increases the likelihood of process and both types of innovations by 7.4 and 6.7 percentage points.[37]

To investigate which variation in the cross section of firms drives the results, I sort firms by the ratio of college-educated employees. Panel A of Figure 2.6 shows that across all three margins of innovation, the association decreases monotonically with the ratio of college-educated employees. I also run a double-interaction analysis, sorting firms by the ratio of college-educated employees and by the Pavitt sectorial taxonomy, which is based on the technological intensity of sectors (Figure A.7 of the Appendix). Firms that employ more basically-educated workers *and* operate in the least technologically intensive sectors drive the results.

Firms that innovate more invest more (Kogan et al. (2013)). For instance, they need to invest in property, plant, and equipment to buy the machines that allow producing their new products. In column (4) of Table 2.11, a one-standard-deviation increase in the log of high-school graduates is associated with a higher ratio between capital expenditures and previous end-of-year assets by 0.11, which is 4% of the average capital expenditures. I then split the sample of firms between those that innovate and those that do not innovate across the three margins (product, process, both types of innovations). Panel B of Figure 2.6 shows that firms that innovate across all margins drive the effect of basic education on firm-level investment. This result addresses the concern that basic education may capture regional determinants of investment unrelated to innovation.

In the sample, capital expenditures such as machines, property, and equipment are more likely to be financed with long-term debt (see Figure A.6 of the Appendix)). Hence, if higher basic education increases capital expenditures, it should also increase the ratio between long-term debt and total debt of firms. In column (5) of Table 2.11, a one-standard-deviation

[37] The results are similar if I estimate without restricting the variation within sectors, within firm age groups, and within firm size groups. The estimated coefficients corresponding to columns (1)-(3) of Table 2.11 are 2.6 p.p. (s.e. 1.4 p.p.), 7.4 p.p. (s.e. 1.4 p.p.), and 6.8 p.p. (s.e. 1.2 p.p.), respectively.

increase in the log of high school graduates in a region increases the ratio between long-term debt and total debt by about 16 percentage points. The association increases monotonically with the capital expenditures of firms, consistent with the notion that firms need to raise more long-term debt when they invest in capital expenditures to start new product lines (Panel C of Figure A.7).

In Table 2.12 , I run the IV analysis using the within-country distance from Mainz as an instrument for historical basic education to explain present-day firm-level outcomes. The effects of regional historical literacy on the present-day innovation, investment, and financing of firms in the IV analysis are similar to the reduced-form estimated effects, which I report in Table A.3 of the Appendix.

In Table A.2 of the Appendix, I report a set of results for alternative measures of firm-level investment and capital structure, as well as for the likelihood that firms declared they faced financial constraints in the two years before participating in the EU-EFIGE/Bruegel survey, run from 2008 to 2010. I document that higher basic education reduces the likelihood that firms declared they had to cut planned investment expenditures in 2010 due to the financial crisis, it increases the ratio between long-term debt and the total liabilities of firms, and it reduces the likelihood that firms declared they faced financial constraints and that they were denied credit after applying. In Figure A.7 of the Appendix, I show that the effect of regional basic education on firm-level innovation is driven by firms in the least technologically-intensive sectors *and* employing low shares of college educated individuals, hence the firms employing more blue-collar workers. Moreover, I show that the effect of regional basic education on firm-level innovation is the same for firms that run any R&D activities inside the firm and those that do not, hence the presence of a R&D department or R&D activities inside the firm is irrelevant to the effect I document in the paper.

Alternative Channels and Explanations

In this section, I discuss a set of channels different from basic education, and alternative explanations of the results in the paper.

I first check if current or historical dimensions that correlate with historical literacy drive the regional associations of historical literacy with present-day regional innovation. In Table 2.13, I consider (i) the ratio of high school-educated inhabitants in the region, (ii) the generalized trust index from the World Value survey at the regional level, (iii) the current dispersion of GDP per capita within regions, which captures income inequality in regions, (iv) the urbanization rate of the region around 1880 to proxy for economic performance at the time when historical literacy is measured, and (v) the index for the quality of past regional institutions proposed by Tabellini (2010). I add each dimension separately as a covariate in regressions of regional patents on regional literacy in 1880 and regional controls. Adding the percentage of inhabitants with a high school degree in column (2) of Table 2.13 decreases the estimated reduced-form effect of historical literacy on regional innovation, which stays marginally significant. Adding each of the other dimensions does not change the effect of historical literacy on innovation significantly. The results suggest that basic education may be a channel that transmits the effect of historical basic education on innovation.

Blue collars or machines? The automation of skilled and unskilled jobs has impacted investment, wages, and income distributions worldwide. Regions and firms with more blue-collar workers may be those with a higher scope for automation of jobs. As workers were substituted by robots, manufacturing firms may have had more investment opportunities. Thus, one may worry about a "reverse basic-education effect": the machines that substituted for the blue-collar workers are what improves the innovation of firms in traditional industries.[38] To assess this interpretation, I test for the effects of historical literacy on regional patents for each year during the period when Europe moved from no automation to its highest levels of automation, that is, from the late 1970s to the mid-1990s (Alesina and Zeira, 2006). Under the "reverse basic-education effect," the effect should increase in size over time: as machines are introduced at a faster pace, new innovation and investment opportunities arise for firms. Instead, under the basic-education interpretation, the effect should, if anything, decrease while machines substitute for blue-collar workers.[39] In Figure A.9 of the Appendix, the pattern of the association between basic education and regional innovation over time is not compatible with an increase in the size of the effect while the machines were introduced in Europe.

Blue collars or managers? Throughout the paper, basic education is interpreted as the education level of blue-collar workers. One may wonder if variation is present in the amount of basic education across managers, although the variation is plausibly larger among blue-collar workers than managers. In the EU-EFIGE/Bruegel database, I do not observe the education level of blue-collar workers and managers separately. I build on Bloom and Van Reenen (2007), who find the managerial practices of family firms that hire professional managers are slightly better than those of the average firm. Instead, the managerial practices of family firms that hire managers within the family are similar to those of the average firm, and they are worse if families use the primogeniture rule. If basic education improves managerial practices unrelated to blue-collar workers' characteristics, the effect should be larger for family firms that hire professional managers. But I find no significant difference in the size of the effects of basic education on innovation if I estimate the specification of column (3) of Table 2.11 separately for family firms with professional managers (0.078, s.e. 0.039) or family managers (0.069, s.e. 0.015). The paper is agnostic on the mechanisms through which blue-collar workers improve the innovation of manufacturing firms. Better-educated blue-collar workers may propose product and process innovations to the management, or they may be better able to implement and experiment with innovative ideas coming from the management. Moreover, different skills of blue-collar workers may be important for innovation. Better scientific knowledge may help them develop innovative technical ideas, but better literary skills also allow them to express their ideas so that managers can understand

[38]This interpretation is not easy to reconcile with the results on the intensive margin of basic education.

[39]In fact, the machines may also enable the blue collars that work with them to contribute to the innovation process.

them clearly.[40]

Financial constraints. One may be concerned that financial constraints alone explain all the results, because research has shown that financial constraints reduce the likelihood that firms innovate.[41] Akcomak and ter Weel (2009) find that early social capital increases current innovation across European regions, and they suggest the financing of innovation is easier in regions where social capital is stronger. The firm-level result that basic education increases the ratio of long-term debt over total debt signals that the channel this paper documents differs from the one studied by Akcomak and ter Weel (2009): if basic education acts only through financing, it should increase the overall debt of firms, whatever the debt's maturity. In Figure A.7 of the Appendix, I offer additional firm-level evidence that financial constraints alone cannot explain the results in this paper. I exploit the fact that the branches of business groups are less likely to be financially constrained than independent firms, because they may obtain funds from internal capital markets. Consistently, the effect of basic education on the likelihood of financial constraints is lower for branches than for independent firms. Instead, the effect of basic education on capital expenditures does not differ across the two groups. This effect should have been lower for branches if financial constraints alone drove it.

First Movers. Basic education may have caused some regions to engage in traditional manufacturing innovation first, and such regions may have perpetuated their primacy in innovation.[42] This interpretation is at odds with a series of evidences in the economic history literature. Sandberg (1982) is among the first to show historical literacy across European countries is not correlated with historical GDP per capita, whereas it is positively correlated with GDP per capita today. Education seems to have been irrelevant to the take up of the Industrial Revolution in Europe (e.g., see Galor (2005) and Allen (2003)). Mokyr (2005a), Mokyr (2005b), and Mokyr and Voth (2009) propose that the upper tail of the knowledge distribution may have determined the early adoption of frontier technologies, and Squicciarini and Voigtlaender (2014) are the first to show that indeed the top of the distribution of education helped technology adoption and income growth in the First Industrial Revolution. Mass education, on the other hand, was less relevant.[43] The industrial organization literature also found results at odds with the first-mover interpretation. For instance, Raymond et al. (2010) find high-tech innovation is path-dependent, but innovation in traditional sectors is spuriously persistent: past innovation by itself does not cause future innovation. In this paper, the placebo distance results show the exogenous variation in the spatial distribution

[40] Unfortunately, to guarantee the selection of schools participating in the program, run by regional authorities, is not biased toward the best regional schools, the region-level results of broad surveys based on standardized tests such as the PISA surveys are not diffused by most countries.

[41] But the evidence on the effects of financial constraints on innovation is mixed. For instance, Gorodnichenko and Schnitzer (2013) find constraints reduce innovation, whereas Almeida et al. (2013) find that they benefit firm-level innovation efficiency.

[42] Literacy should not have also caused the same regions to engage in high-tech innovation first; otherwise, this interpretation is inconsistent with the evidence of no effect of basic education on high-tech innovation.

[43] The evidence on the Second Industrial Revolution is mixed. For instance, Becker et al. (2011) find literacy was important to the industrialization of Prussia.

of past literacy I use is unrelated to the spatial diffusion of development or industrialization at the time the printing press was invented.

2.6 Conclusions

I study a neglected margin of innovation, the innovation of traditional manufacturing firms, and its effects on the investment and financing of firms. I find that a 10% increase in the number of high school graduates in European regions leads to 15% more patents filed in traditional manufacturing industries, and 4% higher investment in capital expenditures. I document these facts with a unique data set on the patented and unpatented innovation of European manufacturing firms. I construct Virtual Regions to show that spatially correlated unobservables that vary within and across regions do not drive the results. To address the issue of reverse causality for the effect of basic education on innovation, I use a shock to the acquisition of basic education by European households before manufacturing started, and I verify the effects of the shock on basic education have persisted for decades.

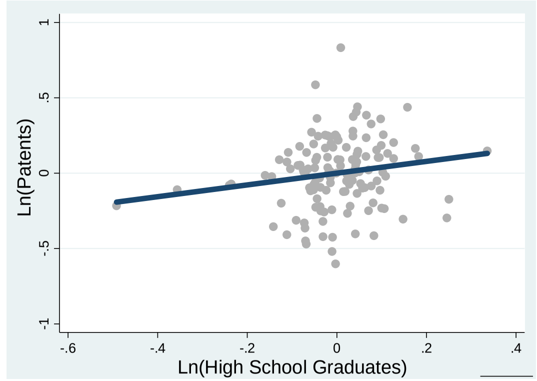
The results suggest that learning-by-doing is not the only type of training that helps firms' innovation in traditional industries: formal basic education is also important. Hence, policies that address high school dropping may be relevant to firms that employ large shares of blue-collar workers. The results have additional policy implications. First, they may help explain why costly place-based transfers are often ineffective in increasing the productivity of firms in depressed areas, which are in most cases traditional manufacturing firms. Policies that address early school leaving may be more effective. An example is direct cash transfers to household heads and children conditional on school attendance (e.g., see Baird et al. (2011) and Bursztyn and Coffman (2012)). The results also have implications for the location decision of firms. U.S. and European firms that move their production overseas to reduce their wage and tax bills should account for the effect of higher basic education in their countries on investment opportunities.

Opening the black box of innovation in traditional industries suggests questions for future research in several fields, such as entrepreneurial finance, strategy, and development economics. For instance, how does the financing of high-tech innovation and of innovation in traditional manufacturing differ? How do managers' and workforce characteristics interact to produce innovation? And to what extent might manufacturing innovation in traditional sectors foster development in countries with no resources to engage in high-tech endeavors. The EU-EFIGE/Bruegel database used in this paper may also help address novel questions in corporate finance and trade, especially if referring to the activities of small and private firms across countries, which allows one to control for national culture and institutions.

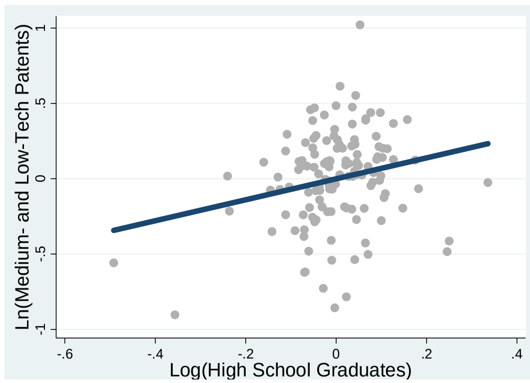
This paper is one of the few in finance to study the long-term effect of historical shocks on financial outcomes. This approach may help financial historians enlarge the scope of their research, and may help finance scholars exploit the time dimension when looking for natural experiments to identify policy effects.

Figure 2.1: Basic Education and Innovation in Regions

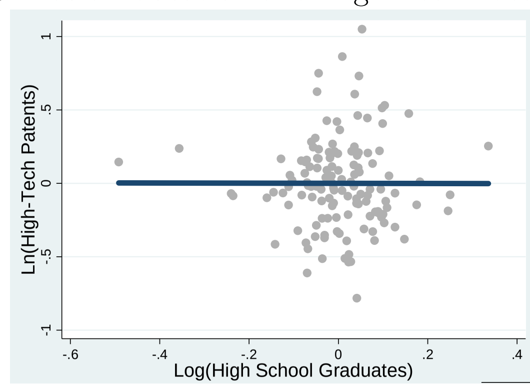
A. Basic Education and Patents in Regions



B. Basic Education and Low-Tech Patents



C. Basic Education and High-Tech Patents



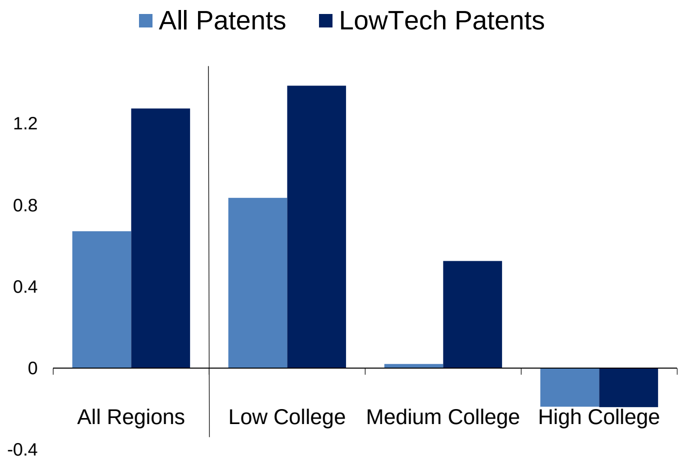
The graphs plot the residuals from the following specification, where the dependent variables are regional patents in Panel A, regional low-tech patents in Panel B, and regional high-tech patents in Panel C:

$$Ln(Patents)_{r,c} = \alpha + \gamma Ln(College)_{r,c} + \theta Ln(NoDegree)_{r,c} + Ln(X)'_{r,c}\delta + \eta_c + \epsilon_{r,c},$$

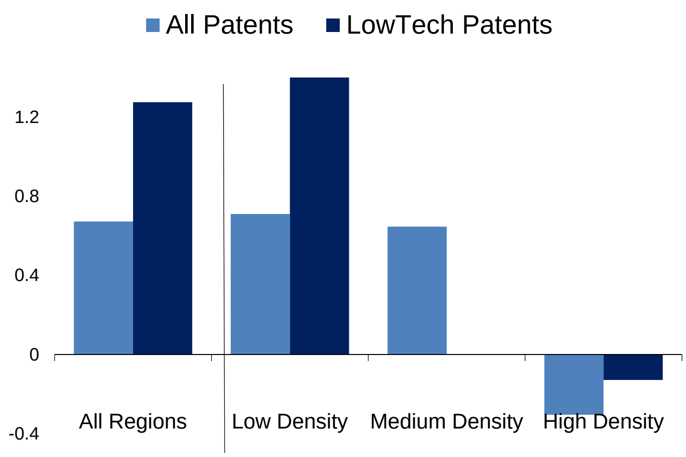
against the residuals from an analogous regression whose dependent variable is the log of high-school graduates ($Ln(HighSchool)_{r,c}$) in regions. X is a set of regional controls that include the log of inhabitants without degrees in 2005, log of latitude, log of area, log of population density in 2005, and the log of the index of the quality of cultivable land from Ramankutty et al. (2002). The split of patents count by technological intensity follows the sectorial taxonomy by the International Patent Office, adopted by Eurostat. The following sectors are categorized as high-tech: Aviation, Computer, Communication Technology, Lasers, Micro-organism and Genetic Engineering, Semi-conductors.

Figure 2.2: Interaction Effects - Regions

A. Effect of high school graduates on patents by share college educated



B. Effect of high school graduates on patents by population density

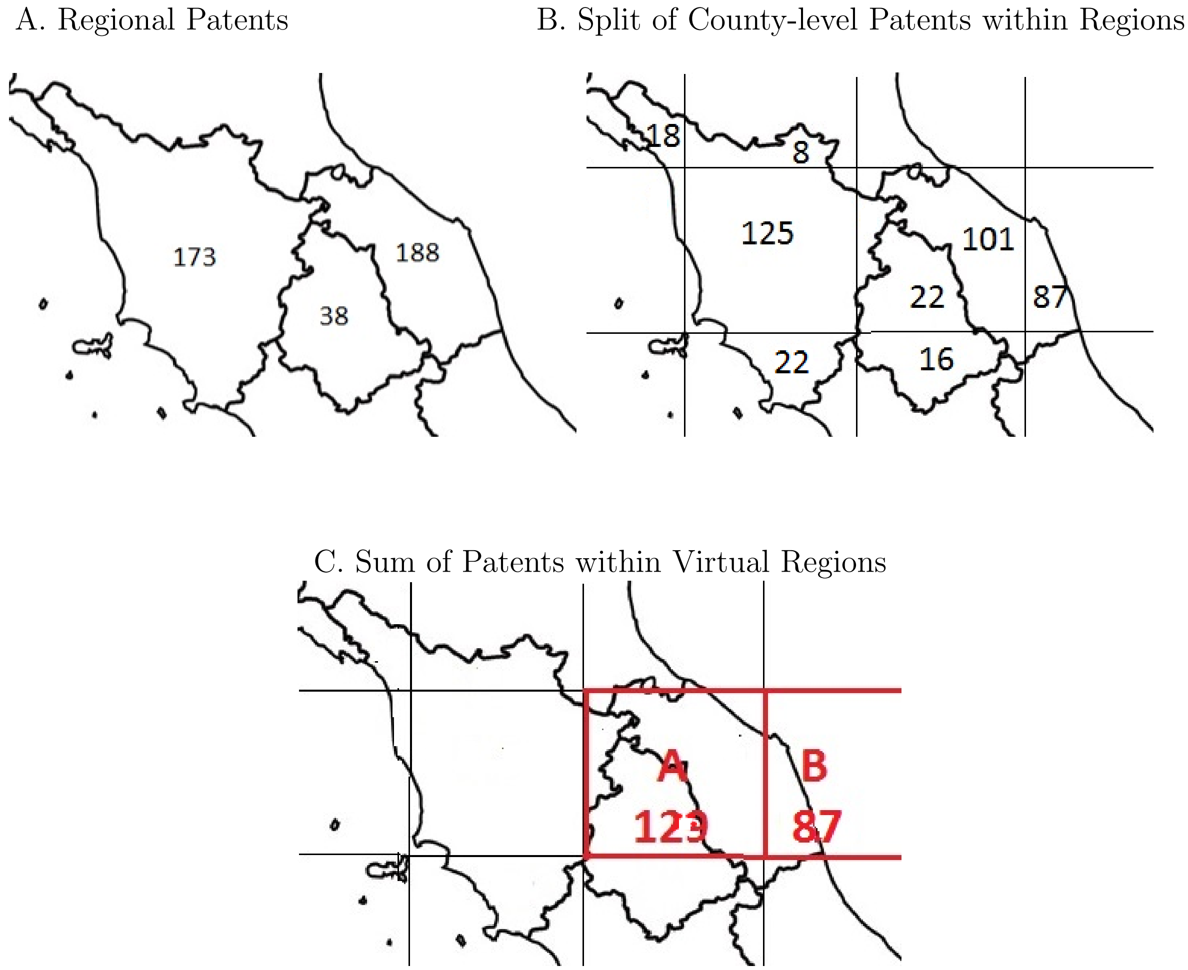


Panels A and B plot the coefficients from the following OLS specification:

$$Ln(Patents)_{r,c} = \alpha + \beta Ln(HighSchool)_{r,c} + \gamma Ln(College)_{r,c} + Ln(X)'_{r,c}\delta + \eta_c + \epsilon_{r,c},$$

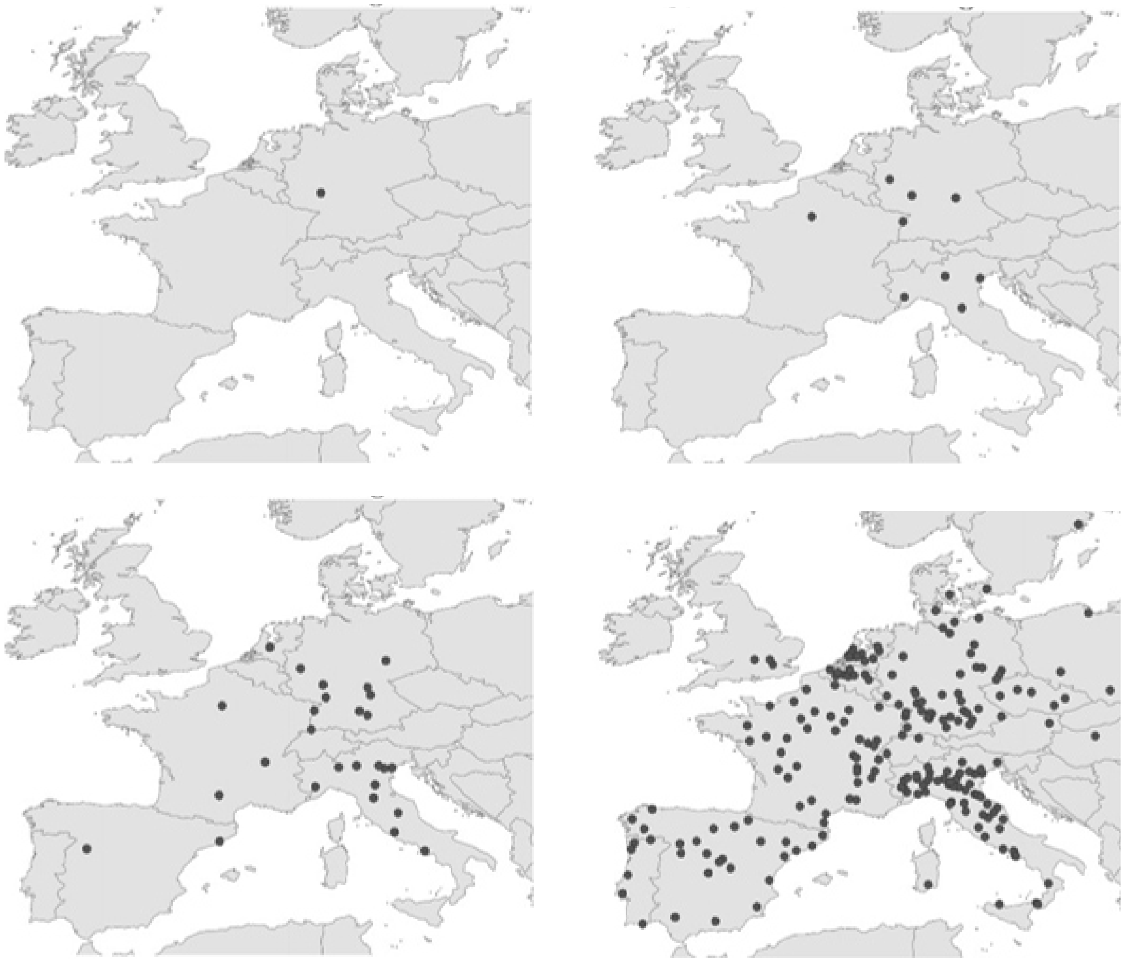
where $Ln(Patents)_{r,c}$ is the log of all patents filed in region r of country c in 2005 (light blue), or the log of mid/low-tech patents (dark blue). $Ln(HighSchool)_{r,c}$ is the log of inhabitants of region r of country c holding a high school degree as of 2005. X is a set of regional controls that include the log of inhabitants without degrees in 2005, log of latitude, log of area, log of population density in 2005, the log of the index of the quality of cultivable land from Ramankutty et al. (2002), and a dummy variable equal to 1 if the region was subject to a communist regime after the Second World War. η_c are a set of country fixed effects. In Panel A, regions are sorted in three equal-size groups based on the share of college-educated inhabitants in the region, and the specification is estimated within each group. In Panel B, regions are sorted in three equal-size groups based on the population density in the region, and the specification is estimated within each group.

Figure 2.3: Virtual Regions



These graphs describe the construction of the Virtual-Region dataset for the variable Patents across three Italian central regions. A grid of size 100 Km by 100 Km is imposed to the map of Europe projected with an equal-area cylindrical projection. Each non-sea portion of a square is a Virtual region. Panel A reports the number of Patents at the regional level for the three regions in the example. Panel B shows how the county-level Patents of each region are assigned to the portions of the regions that enter different Virtual Regions: the county-level values are added across all the counties that enter the same Virtual Region. If a county is split between Virtual Regions, I assign the county value to the Virtual Region that covers its largest part. Panel C shows how the regional portions of Patents are aggregated at the level of Virtual Regions, by summing up the Patents for each regional partition in the Virtual Region.

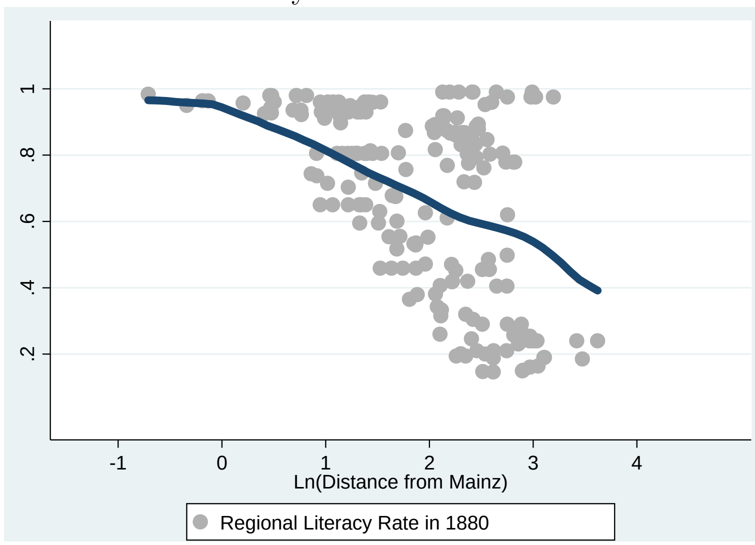
Figure 2.4: Diffusion of the Printing Press after 1450



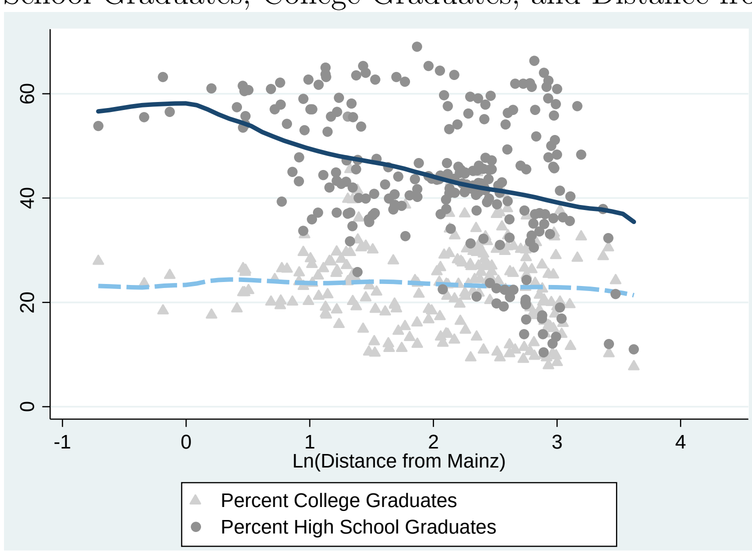
This figure is a one-to-one replication of Figure III of Dittmar (2011). Each point on the maps represent a town where a printing press existed after its invention in 1450 in Mainz (current Germany), which is the only dot in the top-left graph. Each map describes the spatial diffusion of the printing press in the decades after it was invented.

Figure 2.5: Distance from Mainz, Literacy, and Current Education

A. Historical Literacy Rates and Distance from Mainz



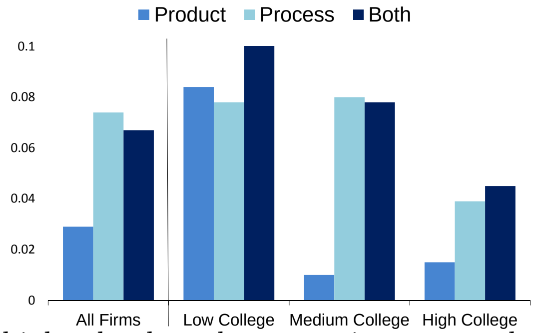
B. High School Graduates, College Graduates, and Distance from Mainz



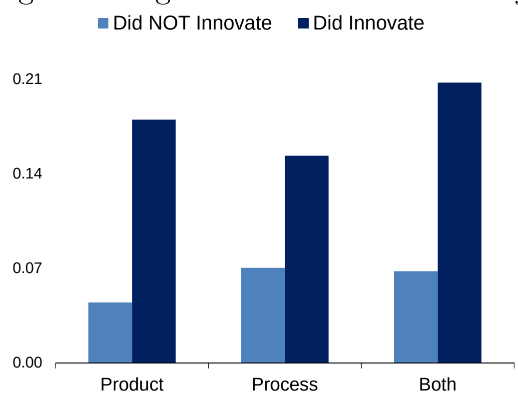
Panel A plots the unconditional correlation of the literacy rate in a European region as of 1880 against the log of the distance of the centroid of the region from the city of Mainz, in current Germany. Each point is a European region for which the historical literacy data is available. The dark blue line is a local linear polynomial fit for the relationship with a bandwidth of 1. Panel B plots the unconditional correlation of the regional percentage of high school graduates (dark points), and of college graduates (light triangles) against the log of the distance of the centroid of the region from the city of Mainz, in current Germany. The solid dark blue line is a local linear polynomial fit with a bandwidth of 1 for the relationship of the percentage of high school graduates and the distance from Mainz. The dashed light blue line is a local linear polynomial fit with a bandwidth of 1 for the relationship of the percentage of college graduates and the distance from Mainz.

Figure 2.6: Interaction Effects - Firms

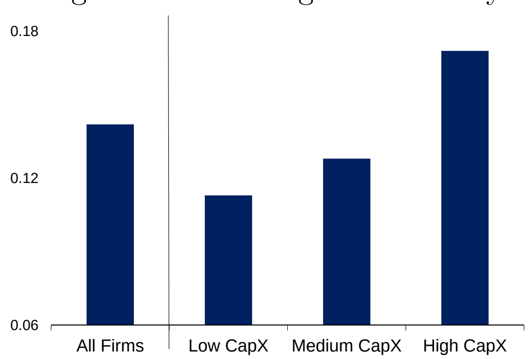
A. Effect of high school graduates on innovation by college employees



B. Effect of high school graduates on investment by innovation



C. Effect of high school graduates on long-term debt by capital expenditures



Panel A plots the estimated coefficient $\hat{\beta}$ from the following probit specification run across three equal-size groups of firms in the EU/EFIGE-Bruegel data set sorted by the share of college-educated employees in the firm:

$$Pr(Innovation = 1)_{frc} = \Phi(\alpha + \beta Ln(HighSchool)_{rc} + X'_{rc}\gamma + F'_{frc}\delta + \eta_c + \eta_a + \eta_s + \eta_l), \qquad (2.12)$$

where $Innovation$ is a dummy equal to 1 if the firm declares it engages in product, process, or both types of innovation. X_{rc} and F_{frc} are region- and firm-level covariates, and η_a, η_s, and η_l are firm age group-, size-, and sector-fixed effects. Panel B plots the estimated coefficient on $Ln(HighSchool)$, the log of regional inhabitants with a high school degree as of 2005, in a OLS regression whose outcome is the capital expenditures of the firms in the EU/EFIGE-Bruegel data set normalized by previous end-of-year assets, and the RHS is the same as in 2.12. The coefficient is estimated separately for firms that did not innovate in the two years before the survey was run (light blue), and firms that did innovate (dark blue). Panel C plots the estimated coefficient on $Ln(HighSchool)$, the log of regional inhabitants with a high school degree as of 2005, in a OLS regression whose outcome is the share of long-term debt over total debt of the firms in the EU/EFIGE-Bruegel data set, and the RHS is the same as in 2.12. The coefficient is estimated separately across three equal-size groups of firms sorted by the capital expenditures of the firm normalized by previous end-of-year assets.

Table 2.1: Summary statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	All Available information			Non-missing information		
Panel A. Regions	Obs	Mean	St. dev	Obs	Mean	St. dev
Current Characteristics						
Latitude	289	48.16	8.370	228	48.73	5.917
Population in 2005 (thousands)	289	1768	1465	228	1893	1565
GDP in 2005 (million €)	288	12824	40261	228	15591	44833
Patents in 2005	289	172.5	350.7	228	211.6	383.5
% High School Graduates in 2005	274	0.479	0.150	227	0.465	0.151
% College Graduates in 2005	274	0.225	0.081	227	0.230	0.078
Avg. Years Schooling if Basic Education	225	11.24	1.488	184	11.14	1.386
Avg. Years Schooling if College Education	225	16.78	4.121	184	16.81	4.123
Land Quality Index (Ramankutty et al, 2002)	280	0.592	0.255	225	0.584	0.238
Competitiveness Index (Annoni and Kozovska, 2010)	267	0.205	0.932	217	0.378	0.782
Trust from the World Value Survey (Wave 9)	234	0.323	0.151	185	0.325	0.137
Communist regime after Second World War	289	0.170	0.376	228	0.145	0.353
Historical Characteristics						
Literacy Rate in 1880	228	0.698	0.263	228	0.698	0.263
Urbanization Rate in the period1860-1880	156	0.132	0.139	155	0.133	0.139
Historical Institutions index (Tabellini, 2010)	156	0.458	2.177	155	0.468	2.180
Panel B. Firms	Obs	Mean	St. dev	Obs	Mean	St. dev
Financials/Ownership						
Total Assets (million €)	12538	14.90	134.1	12458	15.06	135.0
Sales (million €)	10620	22.41	205.3	10594	22.40	205.1
Cash flows (million €)	9449	1.076	17.10	9418	1.074	17.23
Property, Plant, Equipment (million €)	10612	1.919	25.30	10586	1.919	25.31
Tangibility	12434	0.255	0.201	12359	0.255	0.200
Number of Employees	9342	116.6	4029	9313	116.6	4035
Family firm	14760	0.702	0.457	14185	0.709	0.454
Family CEO	14760	0.622	0.485	14185	0.628	0.483
Part of business group	14760	0.185	0.388	14185	0.185	0.388
Exports goods	14760	0.580	0.494	14185	0.579	0.494
Cut investments in 2010	12513	0.429	0.495	12056	0.428	0.494
Innovation						
Product Innovation	14760	0.491	0.499	14185	0.493	0.495
Process Innovation	14760	0.318	0.466	14185	0.322	0.467
Financial constraints						
Declares financial constraints	14760	0.174	0.379	14185	0.176	0.381
Rejected credit application	14760	0.040	0.196	14185	0.040	0.197
Asked for personal guarantees	14760	0.148	0.355	14185	0.148	0.356
Short term/Total debt	11287	0.406	0.394	11224	0.406	0.393

This Table reports summary statistics for the variables in the analysis. In both Panels, columns (1)-(3) provide statistics for all available observations, and columns (4)-(6) for regions and firms for which the historical literacy rates are observed. Panel A reports statistics for variables computed at the level of European regions (NUTS 2). Geographic and demographic variables are from the Eurostat Regional Database, and refer to 2005, the year employed in the cross-sectional analysis of Sections 4, 5, and 6. The regional fractional count of patents filed by applicants in priority year 2005 is from the PATSTAT-Crios database. The average amount of years of schooling are computed from the respondent to the World Value Survey (Wave 9). This measure is not available for the regions of Austria, Belgium, Denmark, Greece, and Portugal. I measure the index of the quality of cultivable land at the regional level averaging the underlying 1 by 1 degree raster data of Ramankutty et al. (2002). I collect the first-wave results of the *EU Regional Competitiveness Index - RCI*, described in Annoni and Kozovska (2010); I use the components on the quality of institutions and the quality of infrastructures. I average the individual level responses from the World Value Survey (Wave 9) at the regional level. Historical Literacy rates in regions as of 1880 are collected from Tabellini (2010), which I integrate with primary and secondary sources: for Greece, I use hand-collected raw population data on literate inhabitants of Greek regions from the 1881 national census; for Germany and the regions of Poland that were part of Prussia in the 19^{th} century, I use the raw data on the literate population by historical Prussian counties from the Ifo-iPEHD project. For Belgium, I compute the historical literacy rates from the province-level data on analphabetism in Ruwet and Wellemans (1978). The historical urbanization rate and the index of the quality of historical institutions are from Tabellini (2010). Panel B reports the statistics for the firms in the EU-EFIGE/Bruegel database, based on a survey run between 2008 and 2010 across seven European countries (Austria, Italy, Spain, UK, Germany, France, Hungary). I create two dummies that equal 1 if a firm declares it engaged in any product or process innovations in the year prior to the interview.

Table 2.2: Basic Education and Regional Innovation

Panel A. Extensive Margin of Basic Education	(1)	(2)	(3)	(4)	(5)	(6)
	All Patents		HighTech Patents		Mid/LowTech Patents	
Log Pop. High School	1.388***	1.008**	0.418	-0.331	2.377***	1.667***
	(0.508)	*(0.433)*	*(0.535)*	*(0.480)*	*(0.794)*	*(0.848)*
Log Pop. College	2.409***	1.382***	2.857***	1.925***	1.776***	0.639
	(0.392)	*(0.356)*	*(0.335)*	*(0.391)*	*(0.521)*	*(0.542)*
Regional controls	X	X	X	X	X	X
Country f.e.	X	X	X	X	X	X
Competitiveness Index, GDP p.c.		X		X		X
Observations	247	247	247	247	247	247
N. of clusters	88	88	88	88	88	88
Adjusted R^2	0.903	0.923	0.856	0.873	0.868	0.888

Panel B. Intensive Margin of Basic Education	(1)	(2)	(3)	(4)	(5)	(6)
	All Patents		HighTech Patents		Mid/LowTech Patents	
Log Avg.Years Basic Edu	1.733***	1.443**	1.274*	1.271	1.732**	1.219**
	(0.643)	*(0.577)*	*(0.645)*	*(0.819)*	*(0.799)*	*(0.590)*
Log Pop. Basic Edu	1.911	3.167	-3.596	-0.529	7.489***	8.710***
	(2.340)	*(2.438)*	*(2.206)*	*(2.783)*	*(3.044)*	*(2.997)*
Log Avg. Years Top Edu	-0.159	0.046	0.023	0.253	-0.032	0.145
	(0.370)	*(0.334)*	*(0.396)*	*(0.403)*	*(0.481)*	*(0.452)*
Log Pop. Top Edu	1.952***	2.304***	1.250	1.620**	3.024***	3.363***
	(0.654)	*(0.724)*	*(0.750)*	*(0.787)*	*(0.847)*	*(0.843)*
Regional controls	X	X	X	X	X	X
Country f.e.	X	X	X	X	X	X
Competitiveness Index, GDP p.c.		X		X		X
Observations	168	168	168	168	168	168
N. of clusters	62	62	62	62	62	62
Adjusted R^2	0.878	0.890	0.816	0.828	0.823	0.832

Panel A reports the estimated coefficients for the following OLS specification (extensive margin of basic education):

$$Ln(Patents)_{r,c} = \alpha + \beta Ln(HighSchool)_{r,c} + \gamma Ln(College)_{r,c} + Ln(X)'_{r,c}\delta + \eta_c + \epsilon_{r,c}, \tag{2.13}$$

where $Ln(Patents)$ are all regional patents in columns (1)-(2), high-tech regional patents in columns (3)-(4), and mid/low-tech regional patents in columns (5)-(6) filed in 2005. $Ln(HighSchool)_{k,r}$ is the log of inhabitants with a high school degree, and $Ln(College)_{k,r}$ is the log of inhabitants with a college degree that reside in region r of country c in 2005. Panel B reports the estimated coefficients for the following OLS specification (intensive margin of basic education):

$$\begin{aligned} Ln(Patents)_{r,c} = \alpha + \beta\theta Ln(s_{B,r,c}) + \beta Ln(BasicEducated)_{r,c} + \gamma\theta Ln(s_{C,r,c}) \\ + \gamma Ln(HigherEducated)_{r,c} + Ln(X)'_{r,c}\delta + \eta_c + \epsilon_{r,c}, \end{aligned} \tag{2.14}$$

where $s_{B,r,c}$ and $s_{C,r,c}$ are the average years of schooling of inhabitants with basic education and higher education, $Ln(BasicEducated)_{r,c}$ is the log of regional inhabitants with a high school degree or lower level of education, and $Ln(HigherEducated)_{r,c}$ is the log of regional inhabitants with more than high school education. The average years of schooling by education levels of regional inhabitants are computed from the respondents to the World Value Survey, which is not available for the full set of European regions. In each Panel, the sample size is restricted to the NUTS 2 regions for which all the regional controls enlisted below are available. In both Panels, Regional controls include the log of inhabitants without degrees, log of latitude, log of area, log of population density, the log of the index of the quality of cultivable land from Ramankutty et al. (2002), and a dummy variable equal to 1 if the region was subject to a communist regime after the Second World War. Odd columns also include the *EU Regional Competitiveness Index - RCI* for the institutional quality and infrastructures components (Annoni and Kozovska, 2010), and the log of GDP per capita. In both Panels, standard errors are clustered at the level of groups of regions (NUTS 1). Statistical significance is shown as follows: ***1%, **5%, *10%.

Table 2.3: Robustness - Basic Education and Regional Innovation

.	(1)	(2)	(3)	(4)
	Extensive Margin of Basic Education		Intensive Margin of Basic Education	
	All Patents	Mid/Low Tech	All Patents	Mid/Low Tech
A. Excluding Top Patenting Regions	0.988	2.087	1.663	1.630
	*0.428***	*0.787****	*0.526****	*0.729***
Country f.e.	X	X	X	X
Other regional controls	X	X	X	X
Observations	233	233	155	155
N. of clusters	86	86	60	60
Adjusted R^2	0.913	0.868	0.872	0.802
B. Excluding Highest GDP Regions	1.712	3.201	1.315	1.313
	*0.499****	*0.736****	*0.591***	*0.775**
Country f.e.	X	X	X	X
Other regional controls	X	X	X	X
Observations	191	191	139	139
N. of clusters	80	80	58	58
Adjusted R^2	0.901	0.855	0.881	0.813
C. Excluding Lowest GDP Regions	1.153	1.876	2.776	2.379
	*0.532***	*0.951***	*1.052***	*1.159***
Country f.e.	X	X	X	X
Other regional controls	X	X		X
Observations	197	197	143	143
N. of clusters	84	84	59	59
Adjusted R^2	0.914	0.876	0.840	0.781
D. Negative Binomial Specifications	0.751	0.712	0.174	0.108
	*0.140****	*0.126****	*0.101**	*0.158*
Country f.e.	X	X	X	X
Other regional controls	X	X	X	X
Observations	247	247	169	169
N. of clusters	88	88	63	63
Unit covariate	1,000 ppl	1,000 ppl	1 year	1 year
E. Patents per capita and Ratios Graduates	0.318	0.380		
	*0.154***	*0.138****		
Country f.e.	X	X	–	–
Other regional controls	X	X		
Observations	247	247		
N. of clusters	88	88		
Adjusted R^2	0.588	0.514		

Panels A-C report the results for estimating log-log regressions of regional patents on the extensive margin (columns (1)-(2)) and the intensive margin (columns (3)-(4)) of basic education and regional controls across alternative subsamples. The outcome variables are all the patents filed in a region in 2005 (columns (1) and (3)), and the mid/low-tech patents filed in 2005 (columns (2) and (4)). In columns (1)-(2), the coefficients are those attached to the log of high school graduates in a region (extensive margin of basic education). In columns (3)-(4), the coefficients are those attached to the log of the average number of years of schooling for individuals with high school degrees or lower levels of education (intensive margin of basic education). The regional controls include the log of inhabitants without degrees, log of latitude, log of area, log of population density, the log of the index of the quality of cultivable land from Ramankutty et al. (2002), a dummy variable equal to 1 if the region was subject to a communist regime after the Second World War, the *EU Regional Competitiveness Index - RCI* for the institutional quality and infrastructures components (Annoni and Kozovska, 2010), and the log of GDP per capita. Panels D-E estimate the relationship using specifications alternative to OLS. In Panel D, I estimate negative binomial regressions to account for the over-dispersed count nature of the patent data. In Panel E, I regress the patents per capita in a region on the ratio of inhabitants with a high school degree, the ratio of those with college degrees, and the other regional covariates. In all Panels, standard errors are clustered at the level of groups of regions (NUTS 1). Statistical significance is shown as follows: ***1%, **5%, *10%.

Table 2.4: Regional unobservables: County-level Analysis

	(1)	(2)	(3)	(4)
Log Pop. High School	1.522***	1.088**	1.124***	1.237***
	(0.421)	*(0.418)*	*(0.395)*	*(0.403)*
Log Pop. College	2.788***	1.956***	1.617***	1.603***
	(0.371)	*(0.380)*	*(0.401)*	*(0.406)*
County-level controls	X	X	X	X
Region f.e.		X	X	X
Excluding cities with >500,000 inhabitants			X	X
Excluding cities with >200,000 inhabitants				X
Observations	1,254	1,254	1,186	677
N. of clusters	92	92	87	65
Adjusted- R^2	0.48	0.86	0.93	0.96

This Table reports the estimated coefficient $\hat{\beta}_k$ in the following OLS specification:

$$Ln(Patents)_{k,r} = \alpha + \beta_k Ln(HighSchool)_{k,r} + \gamma_k Ln(College)_{k,r}$$
$$+Ln(X)'_{k,r}\delta + \eta_r + \epsilon_{k,r},$$

where $Ln(Patents)_{k,r}$ is the log of patents filed in county k of region r in 2005, $Ln(HighSchool)_{k,r}$ is the log of inhabitants with a high school degree in the county in 2005, $Ln(College)_{k,r}$ is the log of inhabitants with a college degree in the count in 2005, and η_r are regional fixed effects. County-level controls include the log of inhabitants without degrees, log of latitude, log of area, log of population density in 2005, the log of the index of the quality of cultivable land from Ramankutty et al. (2002), a dummy variable equal to 1 if the county was subject to a communist regime after the Second World War, and the log of GDP per capita in 2005, all measured at the county level. The *EU Regional Competitiveness Index - RCI* for the institutional quality and infrastructures components is not available at the county level. Standard errors are clustered at the level of groups of regions (NUTS 1). Statistical significance is shown as follows: ***1%, **5%, *10%.

Table 2.5: Unobservables within and across regions: Virtual Regions

	(1)	(2)	(3)	(4)
Log Pop. High School	1.157***	1.087**	1.213***	1.185***
	(0.348)	*(0.412)*	*(0.391)*	*(0.383)*
Log. Pop. College		2.063***		1.976***
		(0.438)		*(0.364)*
Country f.e.	X	X		
Region f.e.			X	X
Virtual-Region controls		X		X
Observations	560	560	560	560
N. of clusters	82	82	82	82
Adjusted- R^2	0.19	0.81	0.94	0.95

This Table reports the estimated $\hat{\beta}_v$ from the following specification:

$$Ln(Patents)_{v,r*} = \alpha + \beta_v Ln(HighSchool)_{v,r*} + \gamma_v Ln(College)_{v,r*}$$
$$+Ln(X')_{v,r*}\delta_v + \eta_{r*} + \epsilon_{v,r*},$$

where $Ln(Patents)_{v,r*}$ are the patents in Virtual Region v, indexed to the region r* that cover its largest part; η_{r*} is a set of regional fixed effects for the underlying $r*$ regions. Virtual Regions are obtained by imposing an arbitrary grid of 100 km by 100 km on the map of Europe. Virtual-Region variables are obtained by aggregating the values of the corresponding variables at the county level, for the counties that enter the Virtual Region. If a county is split across two or more Virtual Regions, I assign the county-level values to the Virtual Region that covers the largest part of the county. Virtual-Region controls include the log of inhabitants without degrees, log of latitude, log of area, log of population density, and the log of the index of the quality of cultivable land from Ramankutty et al. (2002). The *EU Regional Competitiveness Index - RCI* for the institutional quality and infrastructures components is not available at the county level, hence it does not enter the Virtual-Region analysis. The number of regional fixed effects is 234, and the number of groups of regions is 82, because not all the regions of Europe are assigned to one or more Virtual Regions. Standard errors are clustered at the level of groups of regions (NUTS 1). Statistical significance is shown as follows: ***1%, **5%, *10%.

Table 2.6: Virtual Regions: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
	No Top Patents		No smallest Regions		No High Population Density	
Log Pop. High School	1.162***	1.204***	1.209***	1.177***	1.344***	1.469***
	(0.388)	*(0.369)*	*(0.405)*	*(0.391)*	*(0.547)*	*(0.575)*
No Top 5%	X		X		X	
No Top 10%		X		X		X
Virtual-Region controls	X	X	X	X	X	X
Region f.e.	X	X	X	X	X	X
Observations	532	501	532	501	532	501
N. of clusters	79	76	80	78	79	75
Adjusted- R^2	0.90	0.91	0.93	0.94	0.94	0.94

This Table reports the estimated $\hat{\beta}_v$ from the following specification run on a set of subsamples of the full sample of Virtual Regions:

$$Ln(Patents)_{v,r*} = \alpha + \beta_v Ln(HighSchool)_{v,r*} + \gamma_v Ln(College)_{v,r*} \\ + Ln(X')_{v,r*}\delta_v + \eta_{r*} + \epsilon_{v,r*},$$

where $Ln(Patents)_{v,r*}$ are the patents in Virtual Region v, indexed to the region r* that cover its largest part; η_{r*} is a set of regional fixed effects for the underlying $r*$ regions. Virtual Regions are obtained by imposing an arbitrary grid of side 100 km on the map of Europe. All Virtual-Region variables are obtained by aggregating the values of the corresponding variables at the county level, for the counties that enter the Virtual Region. If a county is split across two or more Virtual Regions, I assign the county-level values to the Virtual Region that covers the largest part of the county. Virtual-Region controls include the log of inhabitants without degrees, log of latitude, log of area, log of population density, and the log of the index of the quality of cultivable land from Ramankutty et al. (2002). Standard errors are clustered at the level of groups of regions (NUTS 1). The number of regional fixed effects is 234, and the number of groups of regions is 82, because not all the regions of Europe are assigned to one or more Virtual Regions. Statistical significance is shown as follows: ***1%, **5%, *10%.

Table 2.7: Virtual Regions: Alternative Shapes and Positions of the Grid

Configurations: 4 grid shapes, 27 positions	
Number of configurations	108
Significant coefficients at the 5% level of significance	106 (98.1%)
Median coefficients	1.170
Mean coefficients	1.197
Standard deviation coefficients	0.249
Ratio Mean/Standard deviation	4.79
Mean standard errors	0.351
Standard deviation standard errors	0.042

Simulations: Summary coefficients by shape	
Squares	Triangles
Median: 1.173	Median: 1.175
Mean: 1.171	Mean: 1.167
Diagonal Squares	Parallelograms
Median: 1.162	Median: 1.251
Mean: 1.214	Mean: 1.236

This Table reports the summary of results for estimating the coefficient β_v in the following Virtual-region specification across 108 alternative configurations of the Virtual Regions:

$$Ln(Patents)_{v,r*} = \alpha + \beta_v Ln(HighSchool)_{v,r*} + \gamma_v Ln(College)_{v,r*} \\ + Ln(X')_{v,r*}\delta_v + \eta_{r*} + \epsilon_{v,r*},$$

The configurations use four shapes of the grid, each translated in 27 alternative positions. The shapes include: (i) 100 km by 100 km squares, like those in the analysis so far; (ii) isosceles triangles of side 100 km, obtained by dividing each square in half along its northwest to southeast diagonal; (iii) parallelograms of side 200 km and height 100 km, obtained by merging two triangles across different squares; and (iv) squares whose sides are the diagonals of the original squares. By construction, the four shapes produce Virtual Regions that generally differ from each other in their county composition. As for the positions of the grid, I translate each grid in 27 alternative ways to explore alternative positions with non-overlapping grids: (i) 9 times east, in increments of 10 km; (ii) 9 times north, in increments of 10 km; and (iii) 9 times northeast, in increments of 10 km * $\sqrt{2}$. Translated grids for each shape modify the composition of Virtual Regions as long as counties mainly covered by one square move to another square after the translation. Moreover, the grids produce alternative rules of regional indexing; hence, different groups of Virtual Regions enter the same fixed effect.

Table 2.8: Persistence of Basic Education: Historical Literacy and Current Education

Panel A. Regions	(1)	(2)	(3)	(4)	(5)	(6)
Inhabitants by degree	Log High School		Log College		Log Immigrants	
Literacy 1880	0.049**	0.055**	0.005	-0.007	0.002	-0.007
	(0.023)	*(0.021)*	*(0.065)*	*(0.063)*	*(0.523)*	*(0.503)*
Regional controls	X	X	X	X	X	X
Country f.e.	X	X	X	X	X	X
Exclude no regional variation Literacy 1880		X		X		X
Observations	227	204	227	204	217	194
N. of clusters	78	70	78	70	78	70
(Adj., Pseudo) R^2	0.972	0.988	0.973	0.975	0.199	0.193

Panel B. Individuals	(1)	(2)	(3)	(4)	(5)	(6)
Years of schooling	All inhabitants		High School or lower		Some college	
Literacy 1880	0.921***	0.138	0.665***	0.196	0.045	0.077
	(0.140)	*(0.135)*	*(0.174)*	*(0.142)*	*(0.349)*	*(0.374)*
Individual controls	X	X	X	X	X	X
Regional controls	X	X	X	X	X	X
Country f.e.	X	X	X	X	X	X
Town size f.e.	X	X	X	X	X	X
Degree f.e.		X		X		X
Observations	5383	5383	4327	4327	1037	1037
N. of clusters	31	31	31	31	30	30
(Adj., Pseudo) R^2	0.204	0.580	0.228	0.374	0.083	0.083

Panel A reports results for estimating the following OLS specification at the level of regions (NUTS 2):

$$Ln(HighSchool)_{rc} = \alpha + \beta Literacy1880_{rc} + X'_{rc}\gamma + \eta_c + \epsilon_{rc},$$

where $Ln(HighSchool)_{rc}$ is the number of inhabitants with a high school but not higher degree in region r in country c; $Literacy1880_{rc}$ is the literacy rate in the region in 1880, whose sources are Tabellini (2010) and the national censuses of Greece and Spain for the regions not covered by Tabellini (2010); X_{rc} is a vector of regional controls that include the latitude, area, population density, GDP per capita, and the log of the index of the quality of cultivable land from Ramankutty et al. (2002); η_c is a set of fixed effects at the country level. Panel B reports results for estimating the following OLS specification at the level of household heads which take part for the first time in the World Value Survey from 1981 to 2010:

$$YearsSchooling_{irc} = \alpha + \beta Literacy1880_{rc} + X'_{rc}\gamma + D'_{irc}\delta + \eta_c + \epsilon_{irc},$$

where $YearsSchooling_{irc}$ is the number of years of schooling for household head i in region r and country c measured from the World Value Survey. X_{rc} is a vector of regional controls that include the latitude, area, population density, GDP per capita, and the log of the index of the quality of cultivable land from Ramankutty et al. (2002); D_{irc} is a set of individual-level characteristics that include the gender, age (second polynomial), income (second polynomial); Town size f.e. are a set of 5 fixed effects for the size of the town-city where the individual resides. In both Panels, standard errors are clustered at the level of groups of regions (NUTS 1). Statistical significance is shown as follows: ***1%, **5%, *10%.

Table 2.9: Distance from Mainz: Exclusion Restriction

A. Distance from Mainz and Regional Observables

Region-level observables	(1) Plain Distance	(2) Residual Distance	Firm-level observables	(3) Plain Distance	(4) Residual Distance
Ln(Pop. Density)	-0.090	0.288	Family Firm	-0.031	-0.047
	0.098	*0.215*		*0.031*	*0.025**
Ln(Self Employed)	0.047	0.068	Part of business group	0.020	0.022
	0.070	*0.109*		*0.026*	*0.019*
Ln(College Pop.)	-0.003	-0.090	Did Export	0.001	-0.003
	0.107	*0.124*		*0.028*	*0.024*
Regional Area	0.174	-0.756	Leverage	-0.148	-0.094
	0.161	*0.505*		*0.049***	*0.065*
Past Institution Quality Index	-0.150	-0.083	Tangibility	0.023	0.011
	0.192	*0.138*		*0.012*	*0.006*
Current Institution Quality Index	-0.147	0.011	Ln(Sales)	-0.080	0.005
	0.115	*0.113*		*0.048*	*0.023*
Country f.e.	X	X	Country f.e.	X	X
			Sector f.e.	X	X
			Age f.e.	X	X
			Size f.e.	X	X

B. Distance from Mainz and Historical Literacy in Reduced Form

Dep. Variable: Ln(Patents)			
Ln(Literacy 1880)	0.232		0.302
	*0.073****		*0.064****
Ln(Distance Mainz)		-0.170	-0.122
		*0.080***	*0.074*
Regional controls	X	X	X
Country f.e.	X	X	X
Observations	227	227	227
N. of clusters	78	78	78
Adjusted R^2	0.888	0.888	0.898

C. Placebo First Stages with other relevant Distances

	K-P F stat	A-P chi-sq	Correlation Dist. Mainz
Ln(Distance from Prague)	2.45	2.77	0.567
Ln(Distance from Amsterdam)	4.36	4.92	0.760
Ln(Distance from Madrid)	0.06	0.07	0.488
Ln(Distance from London)	3.90	4.40	0.390
Ln(Distance from Florence)	7.84	8.85	0.606
Ln(Distance from Aix-la-Chapelle)	7.71	8.70	0.881
Ln(Distance from Mainz)	10.51	11.86	1

Table 2.10: 2SLS: Instrumented Historical Literacy and Regional Innovation

Panel A. Second Stage	(1)	(2)	(3)	(4)	(5)
	All Regions	No Communist	No South	No Germany	No all three
Ln(Literacy1880)	0.574***	0.524***	0.616**	0.618***	0.640*
	(0.126)	*(0.123)*	*(0.273)*	*(0.131)*	*(0.320)*
Regional controls	X	X	X	X	X
Country f.e.	X	X	X	X	X
Observations	227	194	173	188	111
N. of clusters	78	66	59	63	37
Adj- R^2	0.877	0.869	0.868	0.847	0.714
Panel B. First Stage					
CD F-Statistic	48.66	42.26	16.38	79.13	35.91
KP F-Statistic	10.51	9.43	7.10	21.96	17.62
AP Chi-sq.	11.86	10.68	8.12	25.13	20.97

Panel A reports the estimated second-stage coefficients from two-stage least squares regressions whose outcome variable is the log of patents filed in a region in 2005, and the log of the literacy rate in the region in 1880 is instrumented with the log of the minimal Euclidean distance of the centroid of a region from the city of Mainz, in Germany. Regional controls include the log of the latitude, area, population density, GDP per capita, and of the index of the quality of cultivable land from Ramankutty et al. (2002). Each column refers to an alternative subsample or regions. Panel B reports the first-stage statistics for the two-stage least square analyses: (i) the Cragg-Donald F-statistic, which is based on i.i.d. standard errors, but is used to compute the critical values reported in Table 5.2. of Stock and Yogo (2005); (ii) the Kleibergen-Paap F-statistic, which is computed for correcting the standard errors for correlation of unknown form at the level of groups of regions; and (iii) the Angrist-Pischke chi-square statistic, which can be used for a rank test of the matrix of the reduced-form equation coefficients and the excluded instruments. Standard errors are clustered at the level of groups of regions (NUTS 1). Statistical significance is shown as follows: ***1%, **5%, *10%.

Table 2.11: Basic Education and Firm-Level Innovation, Investment, and Capital Structure

	(1)	(2)	(3)	(4)	(5)
		Innovation		Investment	Capital Structure
	Product	Process	Both	CapX	LT Debt/ Total Debt
Log Pop. High School	0.029**	0.074***	0.067***	0.135***	0.158***
	(0.013)	*(0.015)*	*(0.013)*	*(0.039)*	*(0.022)*
Log Pop. College	0.016*	0.012	0.011	-0.032	0.019
	(0.009)	*(0.012)*	*(0.009)*	*(0.028)*	*(0.015)*
Tangibility	0.001	0.031***	0.018***	0.048***	0.087***
	(0.006)	*(0.006)*	*(0.005)*	*(0.013)*	*(0.006)*
Leverage	0.015**	0.023**	0.022**	0.014*	0.001
	(0.007)	*(0.010)*	*(0.010)*	*(0.009)*	*(0.002)*
Log Sales	0.016**	0.038***	0.027***	0.232***	-0.022***
	(0.008)	*(0.007)*	*(0.007)*	*(0.022)*	*(0.007)*
Family firm	0.056**	0.038	0.042*	-0.008	-0.017
	(0.022)	*(0.024)*	*(0.021)*	*(0.049)*	*(0.017)*
Family CEO	-0.009	-0.006	-0.005	0.055	0.029*
	(0.020)	*(0.020)*	*(0.018)*	*(0.049)*	*(0.018)*
Part business group	-0.013	-0.011	-0.007	-0.122***	-0.062***
	(0.016)	*(0.015)*	*(0.015)*	*(0.032)*	*(0.016)*
Exports products	0.179***	0.050***	0.065***	-0.108***	-0.009
	(0.012)	*(0.010)*	*(0.010)*	*(0.026)*	*(0.009)*
Sector f.e.	X	X	X	X	X
Age f.e.	X	X	X	X	X
Size f.e.	X	X	X	X	X
Regional controls	X	X	X	X	X
Country f.e.	X	X	X	X	X
Observations	10140	10140	10140	10132	8083
N. of clusters	141	141	141	141	140
(Adj/Pseudo) R^2	0.062	0.020	0.028	0.101	0.174

Columns (1)-(3) report the results for estimating the following probit specification:

$$Pr(Innovation = 1)_{frc} = \Phi(\alpha + \beta Ln(HighSchool)_{rc} + X'_{rc}\gamma + F'_{frc}\delta + \eta_c + \eta_a + \eta_s + \eta_l) \qquad (2.15)$$

where *Innovation* is a dummy that equals 1 if the firms engaged in a product, process, or both types of innovations in the two years prior to participating in the EU/EFIGE-Bruegel-Unicredit survey, which was run from 2008 to 2010. The definitions of product and process innovations are from the *Oslo Manual* (OECD, 2005): a product innovation is "the introduction of a good or service that is new or significantly improved with respect to its characteristics or intended uses." A process innovation is "the implementation of a new or significantly improved production or delivery method"; X is a set of Regional controls which include the log of inhabitants without degrees, log of latitude, log of area, log of population density, and the log of the index of the quality of cultivable land from Ramankutty et al. (2002); F is a set of firm-level controls which include Tangibility (the ratio of tangible assets to total assets), Leverage (the ratio of total debt to shareholders' equity), the log of sales, and a set of dummy variables that equal one if the firm is owned by a family or entrepreneur, if the firm's CEO is a member of the owning family, if the firm is part of a business group, and if the firm engages in exporting of its products outside the country where it operated. Columns (4)-(5) report the results from estimating OLS regressions whose outcomes are the capital expenditures of the firm (property, plant, and equipment, normalized by previous end-of-year assets), and the ratio between long-term debt and total debt of the firm. Firm-level capital expenditures are available in the Amadeus dataset for 10,132 firms of the EU/EFIGE-Bruegel-Unicredit database. The break down of debt into short term and long term is available for 8,083 firms in the EU/EFIGE-Bruegel-Unicredit database. Standard errors are clustered at the level of regions (NUTS 2). Statistical significance is shown as follows: ***1%, **5%, *10%.

Table 2.12: 2SLS: Instrumented Historical Literacy and Firm-level Outcomes

Panel A. Second Stage	(1)	(2)	(3)	(4)	(5)
		Innovation		Investment	Capital Structure
	Product	Process	Both	CapX	LT Debt/ Total Debt
Ln(Literacy 1880)	0.014* *(0.007)*	0.020** *(0.010)*	0.027*** *(0.008)*	0.113*** *(0.026)*	0.085*** *(0.012)*
Regional controls	X	X	X	X	X
Firm-level controls	X	X	X	X	X
Country f.e.	X	X	X	X	X
Sector f.e.	X	X	X	X	X
Size f.e.	X	X	X	X	X
Age group f.e.	X	X	X	X	X
Observations	10,137	10,137	10,137	10,129	8080
N. of clusters	140	140	140	140	139
Adjusted R^2	0.088	0.029	0.029	0.116	0.241
Panel B. First Stage					
KP F-Statistic	68.02	68.02	68.02	68.03	78.31
AP Chi-sq.	68.72	68.72	68.72	68.73	79.18

Panel A reports the estimated second-stage coefficients from two-stage least squares regressions where the log of the literacy rate in a region is instrumented with the log of the minimal Euclidean distance of the centroid of a region from the city of Mainz, in Germany. In columns (1)-(3) *Innovation* is a dummy that equals 1 if the firms engaged in a product, process, or both types of innovations in the two years prior to participating in the Bruegel/EFIGE-Unicredit survey, which was run from 2008 to 2010; X is a set of Regional controls which include the log of inhabitants without degrees, log of latitude, log of area, log of population density, and the log of the index of the quality of cultivable land from Ramankutty et al. (2002); F is a set of firm-level controls which include Tangibility (the ratio of tangible assets to total assets), Leverage (the ratio of total debt to shareholders' equity), the log of sales, and a set of dummy variables that equal one if the firm is owned by a family or entrepreneur, if the firm's CEO is a member of the owning family, if the firm is part of a business group, and if the firm engages in exporting of its products outside the country where it operated. In columns (4)-(5), the outcomes are the capital expenditures of the firm (property, plant, and equipment, normalized by previous end-of-year assets), and the ratio between long-term debt and total debt of the firm. The set of regional and firm-level controls are the same as those described above.

Panel B reports the first-stage statistics for the two-stage least square analyses: (i) the Kleibergen-Paap F-statistic, which is computed for correcting the standard errors for correlation of unknown form at the level of groups of regions; and (ii) the Angrist-Pischke chi-square statistic, which can be used for a rank test of the matrix of the reduced-form equation coefficients and the excluded instruments. Standard errors are clustered at the level of regions (NUTS 2). Statistical significance is shown as follows: ***1%, **5%, *10%.

Table 2.13: Alternative Channels

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: Ln(Patents)					
Ln(Literacy 1880)	0.349 *0.094****	0.162 *0.090**	0.392 *0.090****	0.380 *0.097****	0.346 *0.150***	0.284 *0.125**
Ln(% High School)		0.272 *0.070****				
Ln(% Generalized Trust)			0.042 *0.046*			
Ln(Dispersion GDP within region)				-0.064 *0.044*		
Ln(Urbanization rate 1860-1880)					-0.126 *0.024****	
Ln(Quality Historical Institutions)						-0.263 *0.084***
Regional controls	X	X	X	X	X	X
Observations	224	224	184	216	152	161
R^2	0.846	0.864	0.850	0.850	0.730	0.873

Each column refers to a different OLS specification of the following form:

$$Ln(Patents)_{r,c} = \alpha + \beta Literacy1880_{r,c} + \gamma RegCovar_{r,c} + X'_{r,c}\delta + \epsilon_{r,c},$$

where $RegCovar_{rc}$ is the covariate enlisted at the beginning of each line. X is a set of regional controls that include the log of inhabitants without degrees, log of latitude, log of area, log of population density, and the log of the index of the quality of cultivable land from Ramankutty et al. (2002). Standard errors are clustered at the level of groups of regions (NUTS 1). Statistical significance is shown as follows: ***1%, **5%, *10%.

Chapter 3

Jewish Persecution and Households' Investments

3.1 Introduction

For centuries, Jews have been associated with financial services, largely because Christians and Muslims were banned from lending money at interest.[1] Indeed, when local credit unions and government-owned banks had just started to diffuse across Germany in 1882, 3% of German workers were Jewish, but 23% of the financial sector workforce was Jewish. At the same time, Europeans had persecuted Jewish communities since the Middle Ages. They attacked them in times of distress, ranging from the Black Death of 1349 to the economic crisis of the 1930s.

We consider the geographic heterogeneity of historical Jewish persecution across German counties (*Kreise*) to proxy for the distrust in financial markets of Germans.[2] The monopoly Jewish people had on money-lending had been broken by the 19^{th} century due to the diffusion of credit unions and government-owned banks. But Jews were accused through the Second World War of running financial markets. In the delirious words of Hitler, "In times of distress a wave of public anger has usually arisen against the Jew; [...] the Jew gained an increasing

[1]For instance, Pope Leo IX banned Christians from lending money at interest as far back as 1049, and Gratianus formalized the ban in the *Corpus Iuris Canonici* in 1150: "Usura est, ubi amplius requiritur quam datur: [...] Ecce evidenter ostenditur, quod quicquid ultra sortem exigitur usura est ("Usury is when one requires more than what he gives: [...] It is therefore obvious that whatever is required in addition shall be usury.") The human capital Jews had accumulated since the second century may have facilitated their sorting into trade and finance well before 1049 (Botticini and Eckstein (2005) and Botticini and Eckstein (2012)).

[2]Distrust in finance is a channel for limited stock market participation complementary to generalized trust (e.g., Gennaioli, Shleifer, and Vishny (2014a) and Gennaioli, Shleifer, and Vishny (2014b)).

influence in all economic undertakings by means of his predominance in the stock exchange."[3]

We find that households in counties with higher historical persecution against Jewish communities on average invest less in stocks. One-standard-deviation higher Nazi persecution correlates with a 1.2-percentage-point lower likelihood that households invest in stocks, or 7.5% of the average participation. Pogroms during the Black Death correlate with a 2-percentage-point lower participation, or 12% of the average participation.

These results are robust if we only exploit the variation in Jewish persecution across geographically close counties: counties in the same states (*Bundesländer*), in the same occupation zones after the Second World War, or in the same virtual states. Virtual states are arbitrary partitions of Germany that account for the endogeneity of state borders, and delimit likely homogeneous areas along several dimensions, such as geography, the quality of institutions, and cultural heritage. The results are also similar if we look only at counties in which a Jewish community was documented as far back as the Middle Ages.[4] Hence, the mere presence of Jewish communities that did not trigger persecution cannot explain the results.

Incentives that do not involve ideological hatred or distrust toward the Jewish population may have driven historical persecution. For instance, Christians often took part in pogroms against Jewish communities to seize their wealth. To check whether ideological hatred indeed relates to our findings, we look at the electoral support for the Nazi party. We compare counties in which the economic crisis of the 1930s was less severe - indicating that Nazi votes were likely driven by ideological rather than economic factors (de Bromhead, Eichengreen, and O'Rourke, 2013) - with counties in which the economic crisis was more severe. This test is based on the fact that antisemitism was a pillar of Nazi ideology. Nazi votes are negatively associated to present-day stockholdings only in counties where Nazi votes were likely to signal primarily ideological, rather than primarily economic, support. Nazi votes have no significant association to economic indicators other than stockholdings, and votes for other parties are completely unrelated to present-day stockholdings.

To dig deeper into the scope for a localized cultural norm of distrust in stock markets, we investigate the channels that transmitted the long-term association between Jewish persecution and present-day stockholdings. We run a survey on a sample of 1,000 Germans, and find that respondents in counties with higher Jewish persecution trust the stock market significantly less than others.[5] We consider three demand-side channels. First, Jewish persecution may correlate with other deep-rooted beliefs that we label "backwardness." Backwardness includes cultural cues that make suspicious anything that is not familiar, such as xenophobia and generalized distrust. Second, those German households which are antisemitic may still associate the stock market with Jews, and thus counties where historical and present-day

[3] See *Mein Kampf (James Murphy Translation)*, Vol. I, chapter XI: Race and People.

[4] We thank Nico Voigtländer and Ruben Enikolopov for inspiring these tests.

[5] The survey is run through *ClickWorker*, a German online platform similar to the US-based *mTurk*. The survey is inspired by the Kellogg-Booth Financial Trust Index, described by Sapienza and Zingales (2012). We thank Stefano Della Vigna and Noam Yuchtman for inspiring this test.

antisemitism are higher may display less investment in stocks. Third, deep-rooted distrust in finance, proxied by Jewish persecution, may have transmitted across generations independently from antisemitism: even if households do not associate Jews with finance anymore, they may still distrust finance more in counties where Jewish persecution was higher. All three channels predict lower stock market participation in counties where Jewish persecution was higher. However, the channels have different predictions across education levels: more educated individuals are less prone to backwardness, and are less antisemitic today (e.g., Glaeser (2005), Voigtlaender and Voth (2013b)). We replicate these results with novel data from the *Friedrich Ebert Stiftung*. Education curricula affect pupils' ideology and shape cultural norms (Cantoni et al. (2014)). German curricula do address the issue of antisemitism and the historical persecution of the Jewish minority, decreasing the antisemitism of students. Yet the curricula are less likely to directly address the benefits of stock market participation, because economics or finance are not subjects of study in primary school and in most secondary schools. Therefore, prejudice against financial institutions is unlikely to be decreased with schooling alongside antisemitism. We find that Jewish persecution predicts distrust in finance and stock market participation similarly across *all* education levels. Thus, generic backwardness or present-day antisemitism cannot drive our findings. Instead, the results are consistent with a cultural norm of distrust in finance that has transmitted across generations independently from antisemitism.

We also test for two supply-side channels that may explain our findings through a lower supply or awareness of investment opportunities in the past: the diffusion process of credit unions and local saving banks, and the depletion of human capital due to Jewish persecution.[6] These channels do not explain our findings.

To better understand the properties of the baseline effect, we analyze the heterogeneity of the effect over time and across cohorts. Localized antisemitic sentiment has persisted for at least six centuries (Voigtlaender and Voth, 2012). The effect of Jewish persecution on investments is also stable throughout our sample (1984 to 2010). Moreover, we find that the size of the effect is similar for all cohorts, except for those born before 1945. Everyone in the pre-1945 cohort invests little in stocks, and the variation in Jewish persecution cannot predict the small variation in their stockholdings. The massive antisemitic propaganda of the Nazi period may have eliminated the variation of antisemitic beliefs across German counties.[7] But the experience of a major economic crisis may also drive the result: cohorts that experienced severe stock market downturns invest in stocks less than others (Malmendier and Nagel (2011), Giuliano and Spilimbergo (ming)). We interact cohort dummies with the unemployment rates in 1933, and find results consistent with the experience argument.

In the last part of the paper, we investigate the extent to which the effect of Jewish persecution on stockholdings may be causal. Reverse causality issues do not arise in our setting because of the timing of the outcomes we consider. But unobservables that determined Jewish persecution may affect stockholdings through channels other than the histor-

[6] We thank Joachim Voth and Ulrike Malmendier for suggesting these tests inspired by Pascali (ming).

[7] This explanation is at odds with the results of Voigtlaender and Voth (2012).

ical anti-Jewish sentiment. The results for geographically close counties, and for counties where Jewish communities existed in the past, and the evidence on the channels that have transmitted the effect, reduce the concerns that unobservables drive our results.

The ideal analysis would exploit a quasi-exogenous shock to the occurrence of historical Jewish persecution unrelated to present-day stockholdings. To get close to such a shock, we exploit the forced migrations of the first Jewish communities across the German lands in the Middle Ages because of the Crusades. As far back as the ninth century, Ashkenazi Jews settled along the Rhine and Mosel Valleys in an area that is still known in Hebrew as Ashkenaz. The Crusades of the 11th and 12th centuries forced these communities to leave their hometowns. Over time, Ashkenazi communities have spread widely from the Rhine Valley throughout the German lands. The diffusion has taken centuries because of transportation costs and possibly the properties of migration networks (e.g., see Spitzer (2014)).

We consider a three-stage strategy based on the fact that acts of violence against Jewish communities in the past may be observed if two conditions are fulfilled: (i) a Jewish community existed in the county, and (ii) anti-Jewish sentiment had developed in the county. In the first stage, the distance of a county from the Rhine Valley predicts the probability that a Jewish community existed before the Black Death of 1349. In the second stage, the probability of existence of a Jewish community predicts the extent of Jewish persecution. The rationale is that in counties with no Jewish communities, the ex-ante probability of observing acts of violence against Jews is zero. By contrast, the mere presence of a Jewish community before 1349 makes the probability of Jewish persecution positive. In the third stage, we regress present-day stockholdings on the predicted extent of Jewish persecution. Three-stage least squares estimates based on this strategy confirm the negative effect of Jewish persecution on stockholdings.

A concern with this strategy is that no discontinuity exists in the treatment intensity, captured by the distance from the Rhine Valley. We cannot definitively rule out unobservables correlated with the distance, with Jewish persecution, and with present-day stockholdings. This concern is alleviated by the fact that our results are robust to using only the variation in Jewish persecution across counties close in space, and hence at similar distances from the Rhine Valley. Moreover, we measure the distance from the Rhine Valley for cities at the opposite ends of the Rhine river, and verify that the results are similar when we use the spatial variation across the alternative gradients north-west/south-east and south-west/north-east. We use the distance to capture the variation in Jewish persecution across counties that hosted Jewish communities or not. The distance does not explain the difference between counties that hosted persecuted communities and counties that hosted non-persecuted communities. Nevertheless, the baseline results are robust to limiting the variation in persecution only across counties that hosted a Jewish community in the past. Thus, the mere presence of Jewish communities in a county does not fully explain the results.

At the same time, the distance from the Rhine Valley has several desirable properties that rule out the most plausible sources of heterogeneity that threaten the identification of a causal effect. First, the Rhine Valley was the cradle of Germany's industrialization and is still highly economically developed. If being closer to the wealthy Rhine Valley increased stock market participation, we should observe a negative correlation between distance and participation, which is the opposite of our effect. Second, as in Buggle (2013), we find that distance from the Rhine Valley has a strong negative association with generalized trust, which positively affects stock market participation (Guiso, Sapienza, and Zingales (2009)).[8] Hence, we can disentangle the effects of generalized trust from those of trust in finance, which are likely to be positively correlated in settings different from ours.[9] Third, distance from the Rhine does not predict present-day outcomes that affect stock market participation, and has no autonomous association with stockholdings when added next to Jewish persecution in the baseline specifications. Fourth, we use French households, who live to the west of the Rhine, for a placebo test: Ashkenazi Jews migrated eastwards, and not to the west where crusaders were already persecuting the local communities. If the distance captured spatial variation that is unrelated to the forced migrations, it should affect French stockholdings as well as German stockholdings, but it does not.[10]

This paper makes four contributions. It is one of the few that document a negative long-term financial effect of persecuting a minority. Because of the historically high equity premium, avoiding stock investments reduces households' accumulation of financial wealth. Second, we introduce three novel data sources: (i) we collect historical data on the diffusion of German credit unions in the 19th century, a process worthy of interest in itself; (ii) we use unique data from the *Friedrich Ebert Stiftung* documenting the association between Jews and financial markets next to other dimensions of backwardness for Germans; and (iii) we run our own survey to measure the distrust in finance in Germany. Third, we use variation in the motives to support the Nazi Party to construct a test for the long-term effects of the Nazi ideology. Fourth, we propose an argument for the emergence of Jewish persecution, which may help explain the diffusion of antisemitism in Voigtlaender and Voth (2012) and Voigtlaender and Voth (2013b).

This paper contributes to the research on the economic effects of culture initiated by Banfield (1958) and Putnam (1993). Guiso, Sapienza, and Zingales (2004) show that social capital positively affects stock market participation, illiquid investments, and the use of checks. In Algan and Cahuc (2006), the inherited component of trust by immigrants affects economic growth in the United States. D'Acunto (2015b) finds that cultural stereotypes associated with genders affect the risk preferences and investment decisions of men and women. We use a historical natural experiment to explain present-day economic outcomes (Spolaore and Wacziarg (2013), Nunn (2014)). Historical arguments and the persistence of socio-economic phenomena have recently been used to understand current financial outcomes, as discussed by D'Acunto (2015a): Pascali (ming) shows that Italian provinces where

[8] Buggle (2013) identifies this effect as a result of the Napoleonic institutions in the early 19th century.

[9] Persecution is unrelated to generalized trust as measured from the World Value Survey for geographically close counties.

[10] We thank Sonya Lim for suggesting this test.

Jewish-managed banking developed in the past, and where Christian competing institutions flourished, have a more developed banking system today. Pierce and Snyder (2014) demonstrate that firms in African countries with a larger extraction of slaves during the slave trade period face larger financial constraints and lower access to formal and informal credit today. D'Acunto (2015c) establishes that the quality of basic education across European regions persists for decades, and helps understand the innovation, investment, and capital structures of current traditional manufacturing firms.

The paper relates to the limited stock market participation puzzle. Explanations include background risk (Paxson (1990), Guiso, Jappelli, and Terlizzese (1996)), social interactions (Hong et al., 2004), awareness (Guiso and Jappelli, 2005), generalized trust (Guiso, Sapienza, and Zingales, 2009), financial literacy (van Rooij, Lusardi, and Alessie, 2011), macroeconomic experiences (Malmendier and Nagel, 2011), labor income risk (Betermier, Jansson, Parlour, and Walden (2012)), and corporate scandals (Giannetti and Wang, 2014). Jewish persecution is used to proxy for distrust in finance, but it need not be the *only* source of distrust in finance.

We also build on recent literature on the effects of Jewish persecution, especially on Voigtlaender and Voth (2012). We use their historical data and compute their measures at the county level. Our work differs because (i) we investigate the effects of Jewish persecution on economic outcomes and the channels that drive the effect, and (ii) we propose a strategy to identify the spatial diffusion of historical Jewish communities. Grosfeld, Rodnyansky, and Zhuravskaya (2013) find an effect of the Pale of Settlement on the support for socialist/communist parties after the fall of the Soviet Union, on generalized trust, and on the ratio of self-employed workers, but not on economic outcomes. They interpret the results as evidence of anti-market beliefs generated by ethnic hatred against Jews. In our paper, Jewish persecution is unrelated to generalized trust or political preferences, but is related to economic outcomes. Additionally, the presence of Jews does not explain our results by itself. Recent papers document supply-side effects of Jewish persecution (e.g., Acemoglu, Hassan, and Robinson (2011), Waldinger (2010), Pascali (ming), and Akbulut-Yuksel and Yuksel (ming)). Supply-side channels do not drive our findings.

3.2 Measuring Historical Jewish Persecution

Building on Voigtlaender and Voth (2012) and Voigtlaender and Voth (2013b), we propose four measures of Jewish persecution at the level of German counties.

Intensive and extensive margins of persecution. The first and second measures relate to the persecution of Jews during the Nazi period. They capture an extensive margin and an intensive margin of Jewish persecution. The *extensive* margin, namely, the overall extent of Jewish persecution, is the number of Jews deported from each German county to a concentration or extermination camp after 1933. This measure increases with the size of Jewish communities within a county; hence, it is higher for counties with large cities. The *intensive* margin, namely, the relative extent of Jewish persecution, is the ratio of deported Jews over the total Jewish population in a county as of 1933, when the Third Reich began. German Jews started to hide and live in captivity after the *Kristallnacht* pogrom of 1938. To locate hiding Jews, the Nazis relied on delation by locals. Households in areas where antisemitism was higher were presumably more likely to delate.

Black Death pogroms. The third measure, *Pogrom 1349*, is a dummy that equals 1 if a county experienced at least one anti-Jewish pogrom during the years of the Black Death, around 1349. The Black Death was arguably the worst pandemic in human history: it lasted several years, and up to one third of Europeans may have died. A deep economic and social crisis followed, leading historians to refer to this period as one of the darkest eras in western history. Unsubstantiated theories on the origins of the pandemic diffused all over Europe. Accusations against Jews, the most influential minority group in 14^{th}-century Europe, were common and often led to mass murders and persecution, especially in the German lands. Voigtlaender and Voth (2012) find that pogroms during the Black Death predict the extent of Jewish persecution during the Nazi period at the town level. In contrast to Voigtlaender and Voth (2012), we do not exploit the number of pogroms in a county. The number of pogroms is certainly meaningful in a city-level analysis, because cities that experienced more pogroms presumably developed a stronger anti-Jewish sentiment in the past. But in our county-level analysis dictated by the level of resolution of the stockholding data, the mere number of cities in the county for which data on pogroms are available, as opposed to the concentration of actual pogroms in a city, would increase this margin of persecution. The number of pogroms in a county for which we observe many cities is high even if each city has only experienced one pogrom.

Principal component of measures. We also compute the first principal component of the three measures described above. The principal component isolates the common variation across the other three alternative margins of persecution. For this reason, the principal component is our preferred measure of Jewish persecution, but we report results using all measures throughout the paper.

Data Sources. We collect data from four sources. The characteristics of German households are from the *Socio-Economic Panel* (SOEP) run by the *Deutches Institut für Wirtschaftsforschung Berlin* (DIW). Economics research has already used the SOEP data (e.g., see Fuchs-Schuendeln and Schuendeln (2005) and Burchardi and Hassan (2013)). The SOEP has conducted interviews on a yearly basis since 1984. For each wave, the SOEP includes households that have been interviewed in previous waves as well as new households. Because we are interested in the cross-sectional association of anti-Jewish sentiment and stockholdings, we only include non-repeated observations when running the main analysis. A drawback of the SOEP data is that they do not provide the complete financial portfolios of households; hence, we cannot document how Jewish persecution affects every component of households' financial portfolios, the fraction of wealth invested in stocks, or whether lower stockholdings translate into higher deposits and other safe assets.

We use the data collected by Voigtlaender and Voth (2012) on Jewish persecution in the Middle Ages and during the Nazi period. We aggregate all city-level data at the level of present-day counties, which is the finest geographical level at which we can observe households' characteristics. We obtain county-level historical characteristics from the *Ifo Prussian Economic History Database*, described in detail by Becker et al. (2014). Finally, we collect current county-level controls: socio-demographics from *DeStatis*, the index of land quality of Ramankutty et al. (2002), and the coordinates of the centroid from *Eurostat*.

Table 3.1 reports summary statistics at the household level.[11] In the left Panel, we use all available observations. In the right Panel, we only look at households with non-missing information. The far-right column reports p-values for testing the null hypothesis that the means across the samples are equal. The full sample of non-repeated households in the SOEP county-level data set includes 29,680 observations. The county of residence is not available for 2,655 households. We are also missing county-level historical information for 9,207 households.[12] The other missing observations are due to blanks in demographics, such as the age, sex, and marital status of the household head.

Measures of Jewish persecution do not differ between the whole sample and the subsample without missing information. On average, 57% of households live in counties where at least one pogrom happened during the Black Death. The average ratio of Jews deported during the Nazi period is 34%. This ratio may be higher than 100%, because Jews may have migrated across counties between 1933 and 1945.[13] To compute this measure, we exclude counties where the Jewish population in 1933 was reported to be zero.

The ratio of stock-owning households across counties and over time is about 16% in both samples. The average age of the person who makes financial decisions is 49 years. The two samples do not differ in the geographic characteristics of the counties. They are similar with respect to the average population size and the average number of Jews in 1933. However, households in the running sample are more likely to be homeowners (42% vs. 39%), and they have lower self-reported income (30,500 euros vs. 31,500 euros). The SOEP survey does not ask households for an estimate of their overall wealth. We use income and homeownership to proxy for wealth. Households in the running sample are less likely to be headed by a woman. Their heads are more likely to hold a college degree and less likely to hold only a high school degree.[14] Households in the running sample live in counties with lower unemployment and fewer blue-collar workers in 1933, and a higher proportion of Catholics in 1925, which allows to control for the religious composition of the local population. They are less likely to live in Eastern Germany, and they live in counties with a lower ratio of college graduates in 2005.

Data Properties and Baseline Correlations. Figure 3.2 depicts the properties of Jewish persecution and present-day stock market participation at the county level. Panel

[11] For county-level variables, we attribute the same values to all households living in a given county.

[12] All results are similar if we attribute values of historical variables from neighboring counties to those for which we are missing information.

[13] The ratio exceeds 1 only in one county (Viersen, in North Rhein-Westphalen). Excluding Viersen from the analysis keeps all results unchanged.

[14] The average ratio of household heads who claim to hold a college degree across counties from 1984 to 2010 is extremely low. For robustness, we replicate all the results using a dummy for holding a high school degree or higher levels of education, and all results are similar.

(a) and (b) of Figure 3.2 show the spatial distribution of the ratio of Jews deported during the Nazi Period and of the average ratio of households who own stocks from 1984 to 2010. In both maps, the darker a county is, the higher the value of the variable. The data are not available for blank counties. Jewish persecution was higher in Western counties. Stock market participation is higher in the south and in the north. As expected, participation is much lower in Eastern Germany, which suggests that we average out the differences between east and west in the analysis. Panels (c) and (d) of Figure 3.2 plot the densities of the ratio of Jews deported and present-day stock market participation.[15] The distributions spike around their mean values of 34% and 16% (see Table 3.1). Panel (e) plots the correlation between the ratio of deported Jews and the average ratio of households who own stocks from 1984 to 2010 at the county level in the raw data. The two dimensions are negatively correlated (-0.13, p-value=0.03). Panel (f) of Figure 3.2 shows the average participation across counties with and without pogroms during the Black Death. Although participation is on average higher in counties with no pogroms, a t-test for the difference between the two means does not reject the null that the means are equal. In Figure B.1 of the Appendix, we plot the densities of Jewish persecution for counties within German states (*Bundesländer*). On average, the persecution was higher in states closer to the Rhine Valley than those in Southern, Northern, and Eastern Germany. In Figure B.2 of the Appendix, we plot the correlations between stockholdings and all the other measures of Jewish persecution.

3.3 Jewish Persecution and Current Stockholdings

In the baseline analysis, we estimate the association between the historical Jewish persecution and stock market participation by German households from 1984 until 2010. The following is our most general specification,

$$Pr(HoldsStocks_{ik}|X_{ik}, K_{ik}) = \Phi(\alpha+\beta\times Persecution_k+X'_{ik}\times\gamma+K'_k\times\delta+Income\ deciles+\eta_t), \tag{3.1}$$

where $HoldsStocks_{ik}$ is a dummy that equals 1 if household i in county k holds any stocks, $Persecution_k$ is one of our four measures of Jewish persecution. X_{ik} includes the following individual-level controls: gender, single/marital status; income (2nd degree polynomial); age (2nd degree polynomial); education level; homeownership; and investment in life insurance products. K_k includes the following county-level current and historical controls: latitude; income per capita; share of college-educated population; index of quality of cultivable land; log of population in 1933; log of Jewish population in 1933; unemployment ratio in 1933; share of workers in manufacturing, in retail, and self-employed in 1933; and ratio of Catholic population in 1925. $Income\_deciles$ are dummies indicating the decile of the income distri-

[15] We have excluded Viersen, where the ratio of deported Jews exceeds 100%.

bution to which the household belongs, and Φ is the standard normal cdf.[16] η_t are a set of survey-wave group fixed effects, each capturing a group of four adjacent years.[17] We allow for correlation of unknown form across residuals at the county level, because attributing county-level measures to each household induces a mechanical correlation of residuals across households in a same county.

Table 3.2 reports the average marginal effects for our baseline specification. In columns (1)-(3), $Persecution_k$ is the first principal component of the three measures of Jewish persecution. In column (1), we include only historical socio-demographic controls at the county level, and hence variables that are measured contemporaneously to the Jewish persecution. In column (2), we add individual-level controls, a dummy that equals 1 for households in Eastern Germany, current socio-demographics at the county level, and survey-wave group fixed effects. A one-standard-deviation increase in persecution (1.016) is associated with 1.22-percentage-point lower stock market participation. The effect is statistically significant and economically large. It amounts to 7.5% of the average participation in our sample. The effect is about one fifth the size of the effect of holding a college degree on the likelihood of investing in stocks. Higher education is one of the largest determinants of stock market participation. Households with higher income and singles are more likely to invest in the stock market. Households in Eastern Germany invest less in stocks. We do not find a statistically significant association of gender of the household head on stockholdings after we control for other socio-demographics. In column (3), we estimate the marginal effect in a linear probability model. The parametric assumptions do not drive our results.

Nazi Persecution. In columns (4)-(7) of Table 3.2, we look separately at the extensive and intensive margins of persecution. A one-standard-deviation increase in the logarithm of the Jews deported (2.31) is associated with a 1.16-percentage-point lower likelihood that a household invests in stocks (columns (4)-(5)). In columns (6)-(7), a one-standard-deviation higher ratio of deported Jews over the total Jewish population in 1933 (0.192) is associated with a 0.63-percentage-point drop in the likelihood that a household invests in stocks. Statistical significance is lower for the intensive margin, which excludes counties without Jewish population in 1933.

Black Death Pogroms. In columns (8)-(9) of Table 3.2, we measure the Jewish persecution at the county level as the occurrence of one or more pogroms during the Black Death of 1349. In both the probit and the linear specifications, households that live in counties with at least one pogrom are on average two percentage points less likely to hold stocks. The size and statistical significance of the effect of individual- and county-level observables on present-day stockholdings are comparable across specifications.

Homogeneous Counties. German counties are likely to differ along several dimensions, such as geography, history, and the quality of current and historical institutions. Replicating

[16] All the results are virtually identical if we include polynomial terms of income instead of deciles. As discussed in the previous paragraph, we do not observe a measure of the overall wealth of households in the SOEP sample. Income and homeownership serve as proxies for the overall wealth.

[17] Results do not change if we make the survey-wave fixed effects coarser or finer. Because we exclude repeated observations across survey waves, no time variation exists at the household level.

our baseline results when comparing only geographically close counties, which have arguably similar geographic characteristics, and are exposed to similar institutions and cultural heritages, is important. In columns (1)-(4) of Table 3.3, we report the coefficient for the principal component of the measures of Jewish persecution when augmenting 3.1 with geographic fixed effects, which allows exploiting only the variation in Jewish persecution across counties close in space. In column (1), we only exploit the variation in Jewish persecution across counties that belong to the same state (*Bundesland*). This specification is quite demanding, because each state includes on average only 15 counties. The estimated magnitudes of the effect are similar to our baseline result. Counties in current German states are exposed to the same local institutions and policies, but are not necessarily exposed to the same historical institutions, because the borders of states do not always coincide with the borders of historical administrative regions. Moreover, the size of states is unbalanced: some states comprise few counties, and hence have little variation in Jewish persecution. To address these issues, in columns (2)-(3) of Table 3.3, we divide Germany into 9 and 16 squares of similar size by longitude and latitude, and we only exploit the variation in Jewish persecution across counties that belong to the same squares, which are arbitrary partitions of the country.[18] This method also overcomes the endogeneity of state borders. The baseline results are replicated in both specifications. In column (4) of Table 3.3, we only exploit the variation in Jewish persecution within occupation zones after the Second World War, because Voigtlaender and Voth (2013a) show that the processes of denazification promoted by occupants had different effects on the persistence of antisemitic beliefs across occupation zones. These zones do not perfectly overlap with the state boundaries. This specification also replicates the baseline results.

Counties with documented Jewish communities. Unobservables that favored the settlement of Jewish communities in the distant past could be driving our results thus far, and such determinants may affect stockholdings. To tackle this issue, in column (5) of Table 3.3, we restrict the sample to households in counties with a documented Jewish community before the Black Death. The restriction reduces the sample size by 27%, but unobservables that explain the settlement of Jewish communities in the past cannot be driving the variation in persecution across these counties. This specification replicates the baseline results.

Robustness. In columns (6)-(10) of Table 3.3, we assess the robustness of the baseline results by conducting a set of subsample analyses. In column (6), we estimate the baseline specification at the county level. This test addresses the concern that individual observations in a county may be spatially correlated in a way not properly accounted for by the clustering of the standard errors at the county level. The assumption that correlated realizations are independent could be driving our results. The magnitude of the coefficient on the principal component of persecution is - if anything - larger than in the individual-level specifications, and the precision of the estimates is similar. In column (7), we exclude households in counties in the bottom third of the distribution of present-day income per capita, and in column (8),

[18]Figure B.4 of the Appendix shows how counties are grouped based on these arbitrary partitions of Germany.

we exclude households in counties in the bottom third of the distribution based on the share of college-educated inhabitants. The estimated coefficients are slightly higher than in our baseline analysis, and the precision of the estimates is similar. In columns (9)-(10), we run robustness checks excluding households in counties that were wealthy in the distant past, and specifically, counties that hosted any Hanseatic League cities, or any free imperial cities in the Middle Ages. The estimated coefficients are not different from the baseline results.

Anti-Jewish Ideology, Nazi Votes, and Stockholdings

Incentives unrelated to ideology may have driven Jewish persecution. Individuals may have hoped to obtain rewards based on the number of Jews they helped the Nazis locate. In the Middle Ages, individuals and political leaders may have hoped to seize Jewish property if they took part or promoted the attacks against Jews.

Unlike persecution, voting is unobservable and not verifiable. Voting choices were unlikely to raise the expectation of rewards by the Nazis before the start of the Third Reich. The Nazis rose to power during a long and deep economic crisis: hyperinflation was a major concern, and high unemployment rates plagued several areas. The democratic institutions of the Weimar Republic were perceived as ineffective. Many voters have supported the Nazi party (which had not been in power before 1933) in the hope of improving their economic conditions. de Bromhead et al. (2013) look at the European voting patterns for extreme-right parties in the 1920s and 1930s. They find that persistently depressed economic conditions are a strong predictor of the electoral support of right-wing, anti-system parties. Thus, voting for the Nazi party in 1933 should be a valid proxy for antisemitic sentiment in areas where unemployment in 1933 was low. The economic crisis was less severe in those areas, and votes for the Nazis are more likely to capture the local support for their ideological and political platform. By contrast, votes for the Nazis should be a noisier proxy for antisemitic sentiment in counties where unemployment was high, that is, where voters were likely driven by economic motives when voting for the Nazis. This argument does not imply the Nazis had higher support, on average, in counties with higher unemployment,[19] but it exploits the different motives for supporting the Nazis *within* the group of Nazi voters in 1933.

We expect a negative association between Nazi vote shares in 1933 and present-day stockholdings in counties where unemployment was low in 1933, and a less negative effect in other counties. We find exactly this pattern. In Figure 3.3, the left vertical axis reports the average marginal effect of the Nazi county-level vote share in 1933 elections on stockholdings, and is associated with the histograms. The horizontal axis indicates the percentile of the distribution of counties by the unemployment rate in 1933. We estimate our baseline specification for cumulative percentiles of this distribution. For instance, the histogram labeled "20" reports the average marginal effect for households that live in counties in which the unemployment rate in 1933 was below the 20th percentile. The right vertical axis reports

[19] In fact, King et al. (2008) find voters hit by the economic crisis but without a high risk of unemployment supported the Nazi party in 1933.

standard errors attached to each marginal effect, which are clustered at the county level. They are associated with the black line. Dark brown histograms are marginal effects that are significant at the 1% level or lower; orange histograms are significant at the 5% level; white histograms are not significant at any conventional level. The average marginal effect of Nazi vote shares in 1933 is significantly negative in counties where unemployment in 1933 was low, up to the 45th percentile of the distribution. The magnitude of the effect and its statistical significance decrease up to households in the 80th percentile of the unemployment distribution in 1933. The estimated effect becomes economically small and statistically insignificant once we add households in counties above the 85th percentile of the distribution of unemployment in 1933.

Placeboes. The media affect hatred-based voting behavior. Adena et al. (2013) show that based on who controlled the aired content, the radio had opposite effects on the vote shares of the Nazis. When the content was anti-Nazi (1929-1932), the radio reduced the vote share of the Nazis. Once the Nazis obtained control of the radio, it had a large positive effect on their vote share (March 1933). In Figure 3.4 Panel (a), we find a virtually identical pattern as in Figure 3.3 for the September 1930 elections. For these elections, we have detailed data at the county level. The elections were held after the beginning of the economic crisis but in a time when the radio had a negative effect on the Nazi vote share.

We also propose a placebo analysis to corroborate our interpretation of the evidence. Panels (b) and (c) of Figure 3.4 show that the pattern for Nazi votes is not replicated when we look at votes for the Social-Democrats and the Communists in 1933. Whereas the economic motives to vote for these parties might be similar as for the Nazis, these parties should have not attracted antisemitists more than the Nazis. In Panels (d), (e), and (f) of Figure 3.4, we show that the Nazi votes are not significantly associated with present-day outcomes other than stockholdings conditional on unemployment in 1933. They are uncorrelated with the likelihood that households invest in life insurance products, that the household head is a woman, or with the households' income. In Figure B.5 of the Appendix, we find no significant associations if using the age, the education level, or the homeownership status of the household head as alternative outcomes.

3.4 Channels Mediating the Effect of Jewish Persecution on Stockholdings

In this section, we discuss the channels through which Jewish persecution may affect the likelihood that present-day inhabitants invest in stocks. We consider demand- and supply-side channels. Demand-side channels may mediate the negative effect of Jewish persecution on stock market participation through the present-day demand for stocks. Supply-side channels may mediate the negative effect through a historically lower supply of financial investment opportunities in counties with higher anti-Jewish sentiment.

Validation of the proxies for distrust in finance. We first need to validate our proxies and test if deep-rooted antisemitic sentiment indeed predicts distrust in financial markets. We do not observe distrust in finance for the households in the SOEP sample. For direct validation, we run our own survey on a sample of 1,000 German households, asking them how comfortable they feel with having part of their wealth invested in stocks. Survey respondents scroll a slider that runs from 0 to 100 at increments of one unit. The number they pick is their measure of trust in financial markets.[20] The survey is administered by the company *ClickWorker* on a stratified sample of the German population that sign up to the platform to perform tasks and surveys for pay. The respondents only know they are part of a survey, and they ignore the identity or scopes of the researchers. This procedure is crucial to avoid demand effects invalidating the procedure. We also ask for some demographics including the zip code, which we map into counties. In Figure B.6 of the Appendix, we report the average levels of trust across respondents' gender, age, and education level of respondents. In the analysis below, we restrict the sample to individuals in counties for which we have data on historical anti-Jewish sentiment, and counties for which we have at least five respondents.[21] The average trust in the full sample of respondents (40.91), and in the restricted sample (40.99) are almost identical. The two-sided p-value for a t-test of the difference of the averages across samples is 0.88.

In Table 3.4, we estimate tobit specifications at the individual-respondent level, because trust in finance is measured as an integer between 0 and 100. In column (1), the unconditional correlation of trust in financial markets and the principal component of the measures of Jewish persecution is negative, although we detect no statistical significance. In columns (2)-(4), we add individual respondents' characteristics and current and historical county characteristics, and we limit the variation within state. In all three specifications, a one-standard-deviation increase in Jewish persecution reduces trust in finance by about three percentage points, which is 7.3% of the average trust. The magnitude of the effect is in line with the effect of the principal component of Jewish persecution on stockholdings, which amounts to 7.5% of the average likelihood of holding stocks in the SOEP sample. The similarity of the magnitudes of the effects is important to establish that distrust in finance can drive the effects not only qualitatively, but also quantitatively. In Figure 3.5, we plot the negative unconditional correlation of Jewish persecution and trust in financial markets when averaging our measure at the county level.

Demand-side Channels and Education. We consider three demand-side channels. First, Jewish persecution may capture households' backwardness, which includes cultural cues and beliefs that promote distrust toward the unfamiliar. Second, current households

[20] The screenshot of the survey is in Figure B.6 of the Appendix. The text of the main question reads "Finden Sie die Idee einen Teil Ihres Vermoegens in Aktien zu investieren behaglich?" (Are you comfortable with the idea that part of your wealth is invested in stocks?).

[21] Results are similar if we set the threshold above or below five. Estimated standard errors increase with lower thresholds, consistent with the fact that we add noisier information. Our main specifications use a threshold because we do not find the data representative of the level of trust in a county for counties where we have too few respondents.

which are antisemitic may still associate the stock market with Jews, and thus invest less in stocks. Third, distrust in finance may be a by-product of historical anti-Jewish sentiment. It may have transmitted across generations independently from antisemitism. Even if most households do not associate the financial sector with Jews anymore, they distrust stock market investments more in counties where deep-rooted antisemitic sentiment was higher. We cannot disentangle these three channels based on the main effect: all three channels predict lower stock market participation in counties with higher deep-rooted anti-Jewish sentiment. We exploit the heterogeneity of households by education, a dimension for which the channels have alternative predictions. More educated households are less prone to backwardness, and are less antisemitic today (Glaeser (2005), Voigtlaender and Voth (2013b)). In Panels B and C of Figure 3.6, we replicate the drop in backwardness and antisemitism for German households, using novel data from the *Friedrich Ebert Stiftung*, whose characteristics and properties are described in Decker, Kiess, and Braehler (2012). We use these unique data because they allow us for the first time to observe whether households think Jews have too much influence in financial markets (left graph of Panel B of Figure 3.6). The association of Jews with the stock market decreases with education similarly to other measures of antisemitism and backwardness. But the distrust in finance does not decline monotonically with education, as we show in the left graph of Panel A of Figure 3.6.[22]

Backwardness or present-day antisemitism could explain the effect only if the size of the effect decreased with the education level of households. But the main effect and the distrust in finance are strikingly similar across education levels. In Table 3.5, we test for the interaction of our baseline effect with the level of education of household heads. Panel A of Table 3.5 shows that the baseline effect of Jewish persecution on stockholdings is not attenuated for college-educated household heads. The same holds true if we test for differential effects of college and high school education.[23] In Panel B of Table 3.5, we run the interaction analysis for the ratio of college graduates in a county. We interact our baseline measures of persecution with dummies capturing counties in the top and the bottom third by the ratio of college graduates. The effect of persecution on stockholdings does not vary systematically with the average education level. This path is not consistent with generic backwardness or present-day antisemitism, but it is consistent with a cultural norm of distrust in finance that has transmitted independently from antisemitism across generations. Education curricula affect pupils' ideology and shape cultural norms (Cantoni et al. (2014)). German curricula address the issue of antisemitism, decreasing the antisemitism of students. But education curricula seldom address the benefits of stock market participation, meaning that prejudice against financial institutions is not decreased alongside antisemitism.

Supply-side Channels. We also study two supply-side channels. Starting in 1843, credit unions (*Volks-* and *Raiffeisenbanken*) have diffused across Germany. They were and still are specialized in financing local businesses and collecting households' savings. If credit

[22]To make patterns easily comparable, we have defined the measure of distrust as (100-Trust)/20, where Trust is the survey-based measure of trust in finance from the previous paragraph.

[23]This result also helps disentangle distrust in finance from generalized trust, the effect of which is driven by the least educated (Guiso et al., 2004).

unions diffused early into more antisemitic areas, current households in those areas might be less aware of stock investment. We collect novel data on the foundation dates of credit unions across German counties from the proprietary registry of the *Hoppenstedt Firmendatenbank*.[24] Panel A of Figure 3.7 shows the diffusion path of credit unions. Before 1860, few credit unions existed in the current southern state of Baden-Württemberg and in the eastern states of Brandenburg and Sachsen-Anhalt. Credit unions diffused in the south and center of the country, and a few banks were founded in the west. As of 1900, credit unions had diffused throughout the whole west and the whole south, as well as the north and the center of Germany. The bottom-right picture of Panel A shows that credit unions can be found in every German county today. In Panel B of Figure 3.7, we plot the year when the first credit union is documented in a county against the ratio of deported Jews over the 1933 Jewish population at the county level. The two dimensions are not negatively correlated.

Another supply-side channel concerns the skill structure of counties. Persecution may have reduced local financial services because a large share of finance workers were Jewish. If the depletion of human capital needed to run financial institutions drove our baseline effect, the effect should be larger in counties with a higher ratio of Jewish workers in finance. In Panel A of Figure 3.8, we find no association of the ratio of Jews in finance as of 1882 on present-day stockholdings. In Panel B, we proxy the ratio of Jews in finance as of 1933 with the ratio of Jews in the total population in 1933, and we find no association between the main effect and the ratio of Jews in a county.

3.5 Properties of the Effect of Jewish Persecution on Stockholdings

Effect over time. First, we investigate whether the magnitude of the effect varies over time. The SOEP data allow us to observe stock market participation from 1984 until 2010, that is, for more than 25 years. In these years, stock market participation varied substantially. The German stock market has also witnessed periods of booms and busts: the crash of 1987, the bull markets of the "dot-com bubble," and the financial crisis of 2008.

In Table 3.6, we estimate 3.1 adding a set of interactions between each measure of Jewish persecution and wave-group dummies capturing periods of four years.[25] All other controls are the same as in Table 3.2. The omitted category is the most recent group of waves, from 2008 to 2010. If the effect of antisemitic violence on financial decision-making faded away over time, we would find that the coefficients on the interaction terms are significantly negative. Their magnitude should decrease monotonically over time. If the effect were stable over time, the coefficients attached to the interaction terms would be indistinguishable from

[24]The registry reconstructs the chains of mergers and acquisitions over the decades for currently-existing German banks. They collect the foundation date, the type, and other characteristics of any entity involved in these chains as far back as any information is retraceable.

[25]Results are the same if we change the number of years covered by each dummy.

zero. In all the columns of Panel A of Table 3.6, the effect of Jewish persecution for the period 2008 to 2010, as captured by the baseline coefficient, is similar to the baseline effect in Table 3.2 in terms of magnitude and statistical significance. None of the coefficients attached to the interaction terms differ significantly from zero.

Effect across cohorts. Different cohorts have been exposed to vastly different macroeconomic conditions throughout their lives. Individuals born before the end of the Second World War have also experienced the massive antisemitic Nazi propaganda.

In Panel B of Table 3.6, we test how the impact of Jewish persecution on financial decision-making varies across cohorts. We divide the households into three cohorts based on the year the household head was born: the *Nazi-period cohort* includes households whose head was born before 1945;[26] the *Postwar cohort* includes those born between 1945 and 1965; the *Post 1965 cohort* includes those born after 1965. Averaging out time effects is crucial because the Nazi-period cohort observations are concentrated in the first half of the survey waves, whereas the Post 1965 observations are concentrated in the second half. The average stock market participation differs substantially across the two periods. The effects we estimate for the Postwar cohort are similar to the baseline estimates in Table 3.2 in terms of magnitude and statistical significance. The interaction effect disappears for households in the Nazi-period cohort, because everyone in this cohort invests little in stocks, and the variation in Jewish persecution has no predictive power. Two explanations seem plausible. The massive antisemitic propaganda of the Nazis may have eliminated the cross-sectional variation of anti-Jewish sentiment across counties, making it impossible for the variation in Jewish persecution to explain stockholdings.[27] Alternatively, the economic crises and social disruptions witnessed by the Nazi-period cohort may be responsible for the generalized lower investment in stocks. In a parallel historical event, Malmendier and Nagel (2011) show that U.S. cohorts that experienced the Great Depression are less likely to invest in stocks. The effect of persecution on financial decision-making could be of second order compared to personal experiences of large business-cycle fluctuations. To disentangle the two hypotheses, we exploit the cross-sectional variation of the depth of the economic crisis in the 1930s at the county level. We look at the stock market participation of Nazi-period cohort households sorted by the county-level unemployment rate in 1933. We split counties into three groups with low, medium, or high levels of unemployment in 1933. In untabulated results, we estimate triple interactions of persecution with the households' cohort and with the dummies for the severity of unemployment. Counties where the level of unemployment was high drive the reversal of the effect of persecution for the Nazi-period cohort. The estimated coefficient on the interaction between the Nazi-period cohort and persecution is significantly positive (0.021, s.e. 0.008), but the coefficient on the triple interaction of the Nazi-period cohort, persecution, and low unemployment in 1933 is significantly negative (-0.029, s.e. 0.008). Because the baseline effect is negative (-0.017, s.e. 0.006), Nazi-period cohort households that

[26] If we define the cohorts using different time cuts, results are similar.

[27] This explanation would be inconsistent with the 600-year-long persistence in the heterogeneity of anti-Jewish sentiment documented by Voigtlaender and Voth (2012).

lived in areas with a less dramatic economic crisis during the 1930s are negatively affected by historical anti-Jewish violence, similar to households in other cohorts. The coefficients on the triple interactions are noisier for other measures of persecution. Overall, the evidence suggests the effect of Jewish persecution on financial decision making is stable over time and across cohorts, and it may be of second order to the effect of dramatic economic lifetime experiences.

3.6 Forced Migrations, Jewish Persecution, and Stockholdings

Unobservable characteristics of German counties may have jointly determined Jewish persecution and stockholdings. Comparing geographically close counties helps alleviate this concern, but this test may not be enough to establish a causal effect of Jewish persecution on present-day financial decisions. Ideally, we would exploit a shock to the extent of Jewish persecution unrelated to stockholdings. Because localized anti-Jewish sentiment has persisted since the Middle Ages, the shock should have happened before the Black Death of 1349.

To get close to such an ideal shock, we look at the forced migrations of Ashkenazi Jews out of the Rhine Valley after the 11^{th} century. In Panel (a) of Figure 3.9, the darker a county is, the older the first Jewish community documented in the county. Blank counties are those with missing data. The earliest Jewish presence in the German lands was found in the cities of Trier, along the Mosel, and Cologne, along the Rhine. Archaeologists date this presence to the ninth century. Research has found evidence of Jewish communities in the 10^{th} century along the entire Rhine Valley.[28] The Jewish population in other areas of current Germany was sparse before the 11^{th} century (Engelman (1944)). At the onset of the Crusades, Jewish communities were persecuted by Christian knights traveling to the Holy Land. Several towns on the Rhine expelled Jews, causing a massive Jewish migration towards Eastern, Northern, and Southern Germany. Evidence of sizable Jewish communities dates back to the late 13^{th} and 14^{th} centuries in Munich (south) and Berlin (east) (Toch (2012)). Migrations accelerated with the Black Death.[29] In Panels (b) and (c) of Figure 3.9, we show the location of the cities of Trier, on the Mosel, and Emmerich, on the northern end of the German Rhine. The age of the first documented Jewish community in a county increases as one moves towards each of these cities.

We argue that the distance from the Rhine Valley determined the existence of Jewish communities at the time of medieval persecutions. In a first step, we use the distance of a county from the Rhine Valley to predict the probability that a Jewish community existed in the county before the Black Death. In a second step, we use the predicted probability

[28] We refer to Toch (2012) as a comprehensive economic history of European Jews in the Middle Ages.

[29] Only in the 15^{th} century did Ashkenazi Jews merge with the communities of Khazar origin who had moved from the Black Sea to current Poland. See van Straten (2004) for archaeological evidence and Elhaik (2012) for genetic-based evidence.

of existence to predict the extent of Jewish persecution. The rationale is as follows: in counties with no Jewish communities before the Black Death, violence against Jews cannot have emerged, because no targets for such violence existed. In counties where early Jewish communities existed, the probability of a historical pogrom against the local Jews is strictly positive because of the mere presence of Jews. Of course, we will not necessarily observe a positive realization, that is, a pogrom in all of these counties ex post. In a third step, we use the predicted probability of persecution at the county level to predict the likelihood of stock investments at the household and county levels.

Distance from the Rhine Valley. We consider three measures of the distance of a county from the Rhine Valley. They are the residuals ($\hat{\epsilon_k}$) from a regression of the geographic distance between the centroid of a county and a point in the Rhine Valley on observable proxies of economic activity at the county level in the Middle Ages,

$$DistanceRhine_k = \alpha + K_k' \times \delta + \epsilon_k, \tag{3.2}$$

where K_k is a set of dummy variables for whether the county hosted a monthly market, a bishop siege, free imperial cities, Hanseatic League cities, a city incorporated before 1349, and geographic characteristics (latitude, the land quality index of Ramankutty et al. (2002), and whether a navigable river existed in the county, whether the county lies in eastern Germany). We compute the minimal Euclidean distance of a county from the cities of Trier, on the Mosel, and Emmerich, on the northern end of the German Rhine (Figure 3.9). The lowest of these two distances is our third measure. The shortest distance is about 2 km, whereas the greatest distance is 1100 km. The alternative measures aim to capture two gradients of the distance from the Rhine Valley, that is, the south-west/north-east gradient and the north-west/south-east gradient. As Panels (b) and (c) of Figure 3.9 show, counties lying at the same distance from each of the two cities, that is, on the same *isodistance curves*, are different; hence, our measures indeed capture two alternative gradients of the distance from the Rhine Valley. Across both gradients, the likelihood that a Jewish community existed in the Middle Ages increases towards the Rhine Valley. To make our specifications directly comparable to the baseline results, we include the latitude of counties when computing the residual distances, but all the results are virtually identical if we exclude the latitude from the vector K_k in 3.2.

Identification concerns. The shock helps us identify a causal effect of Jewish persecution on present-day stockholdings only if the residual distance from the Rhine Valley does not affect stockholdings through channels different from the persecution of Jewish communities. We propose a set of results to assess the plausibility of this exclusion restriction. In Table 3.7, we regress several county- and individual-level observables that correlate with stock market participation on our distance measures. Different from customary reporting, each entry refers to a different regression. We enlist the outcome variables in the left column and report the main covariate at the top of each column. In odd columns, we use the residuals for the three distance measures regressed on the medieval characteristics of counties as in 3.2. In even columns, we also condition on observables at the county level in 1933, before the start of the Third Reich. Most distance residuals are not significantly associated with current county- and individual-level observables. In most cases, the size of the estimated coefficients is small, which suggests the imprecision of our estimates does not drive the non-results. Two exceptions stand out: (i) the ratio of college graduates at the county level is positively associated with the residuals on the distance from Emmerich, but the estimated coefficient is negative when we use the distance from Trier; (ii) the coefficients when age is the outcome variable have a non-negligible size, but we detect no statistical significance.

In Panel A of Table 3.8, we estimate the reduced-form effect of the distances on the ratio of households that own stocks when the distances enter as regressors instead of antisemitic violence in the past, and when both the distances and one of the violence measures enter separately. All the coefficients refer to OLS regressions. In columns (1), (4), and (7), all three distances are positively associated with the likelihood that households hold stocks. Once the measures of Jewish persecution enter the reduced-form specifications, the estimated autonomous associations of the distances with stockholdings drop in magnitude, whereas the estimated standard errors for each coefficient stay virtually identical. Hence, the insignificant effects are not due to the imprecision of our estimates.

In Panel B of Table 3.8, we propose a placebo test for an association of the distance from the Rhine Valley and stockholdings through channels other than Jewish migrations. We look at the effect of the distance on the likelihood that French households own stocks. Data for French households are from the *Enquete Patrimoine* run by the Banque of France in 2004. If the distance from the Rhine captures anything peculiar to the spatial diffusion of development, we should observe an effect of the distance on the stockholdings by French households who live to the west of the Rhine. Across all our measures, we find no economically or statistically significant association between the distance from the Rhine and the stockholdings of French households. In the Appendix, we derive the conditions for identifying our three-stage OLS system presented below and detect no correlation between the instrument and the residuals from the structural equations.

3SLS results. Table 3.9 reports results for estimating the stages described above in a three-stage OLS framework (see Becker and Woessmann (2009)) at the county level,[30]

$$Community1096_k = \alpha + \beta \times LogDistanceRhine_k + K_k' \times \delta + \epsilon_k \tag{3.3}$$

$$Persecution_k = \alpha + \beta \times Commu\hat{n}ity1096_k + K_k' \times \delta + \epsilon_k \tag{3.4}$$

$$RatioStockhold_k = \alpha + \beta \times Perse\hat{c}ution_k + K_k' \times \delta + \epsilon_k, \tag{3.5}$$

where $Commu\hat{n}ity1096_k$ and $Perse\hat{c}ution_k$ are the predicted values for county k when estimating the system of three simultaneous equations. Columns (1)-(3) of Table 3.9 report the coefficient on $LogDistanceRhine_k$ when estimating the system using the three alternative measures of distance from the Rhine Valley. In column (1), a one-standard-deviation increase in the residual distance from Trier (0.69) reduces by nine percentage points the likelihood

[30]Results are similar if we estimate the system at the individual level. Because the first and second stages involve outcomes that vary only at the county level, we find results for variables aggregated at the county level in all stages to be more conservative.

that an early Jewish community existed in a county. This figure is 27% of the average likelihood that a county hosted a Jewish community in 1096 (33%), the date before the Black Death for which we have data on most counties. Estimated magnitudes when using the distance from Emmerich and the minimal distance are -13 and -11 percentage points.[31] Hence, the farther a German county is from the Rhine, the less likely a Jewish community existed there in 1096. In columns (4)-(6) of Table 3.9, we report results for the second stage. A one-standard-deviation increase in the likelihood of an early Jewish community increases the principal component of the persecution measures by 0.53, which is about one half of a standard deviation. The magnitudes are similar when we predict the likelihood of a Jewish community with the other measures of distance. Columns (7)-(9) of Table 3.9 document the third stage. Consistent with the reduced-form results, an increase in the instrumented persecution of Jews significantly reduces stock market participation at the county level.

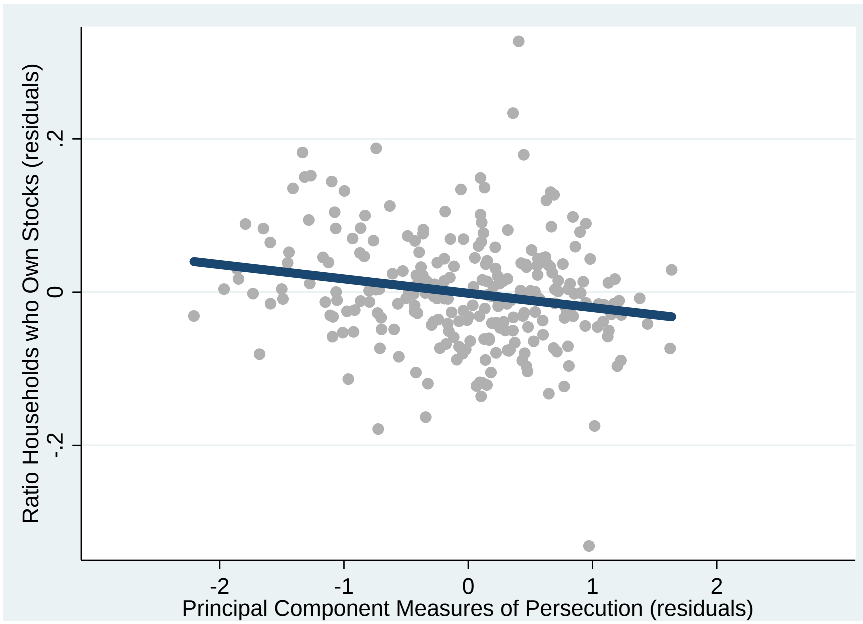
3.7 Conclusions

We use the geographic variation in Jewish persecution across German counties to proxy for localized distrust in finance. A one-standard-deviation increase in persecution reduces by 1.2 percentage points the likelihood that a German household invests in stocks, which is 7.5% of the average participation rate. Households in counties that experienced at least one pogrom in 1349 are two percentage points less likely to invest in stocks. The results are robust to using only the variation in persecution across geographically close counties. The magnitude of the effect is stable from 1984 until 2010 and across cohorts. Plausible supply-side channels do not mediate the effect of Jewish persecution on stockholdings. The results are not driven by present-day antisemitism, or by the backwardness of households. The results are consistent with the notion that distrust in finance has transmitted across generations independently from antisemitism, because the effect of Jewish persecution on stockholdings does not vary across education levels.

The results contribute to the interdisciplinary debate on hatred beliefs and their long-term consequences on societies. They suggest that policies designed to eliminate race- and religious-based persecution should not only be justified by the promotion of human rights, but also by the wealth of the broader population. The results may suggest an economic paradigm to justify policies against hatred beliefs. While several non-Western cultures do not share the ethical notion of the primacy of individual interests over the collective ones, promoting the wealth of societies is universally accepted. Cultures that do not promote human rights may be persuaded by evidence that the persecution of minorities reduces not only the long-term wealth of the persecuted, but of the persecutors as well.

[31] The standard deviations of the distance from Emmerich and the minimal distance are 0.62, and 0.77.

Figure 3.1: Jewish Persecution and Present-Day Stock Market Participation across German Counties

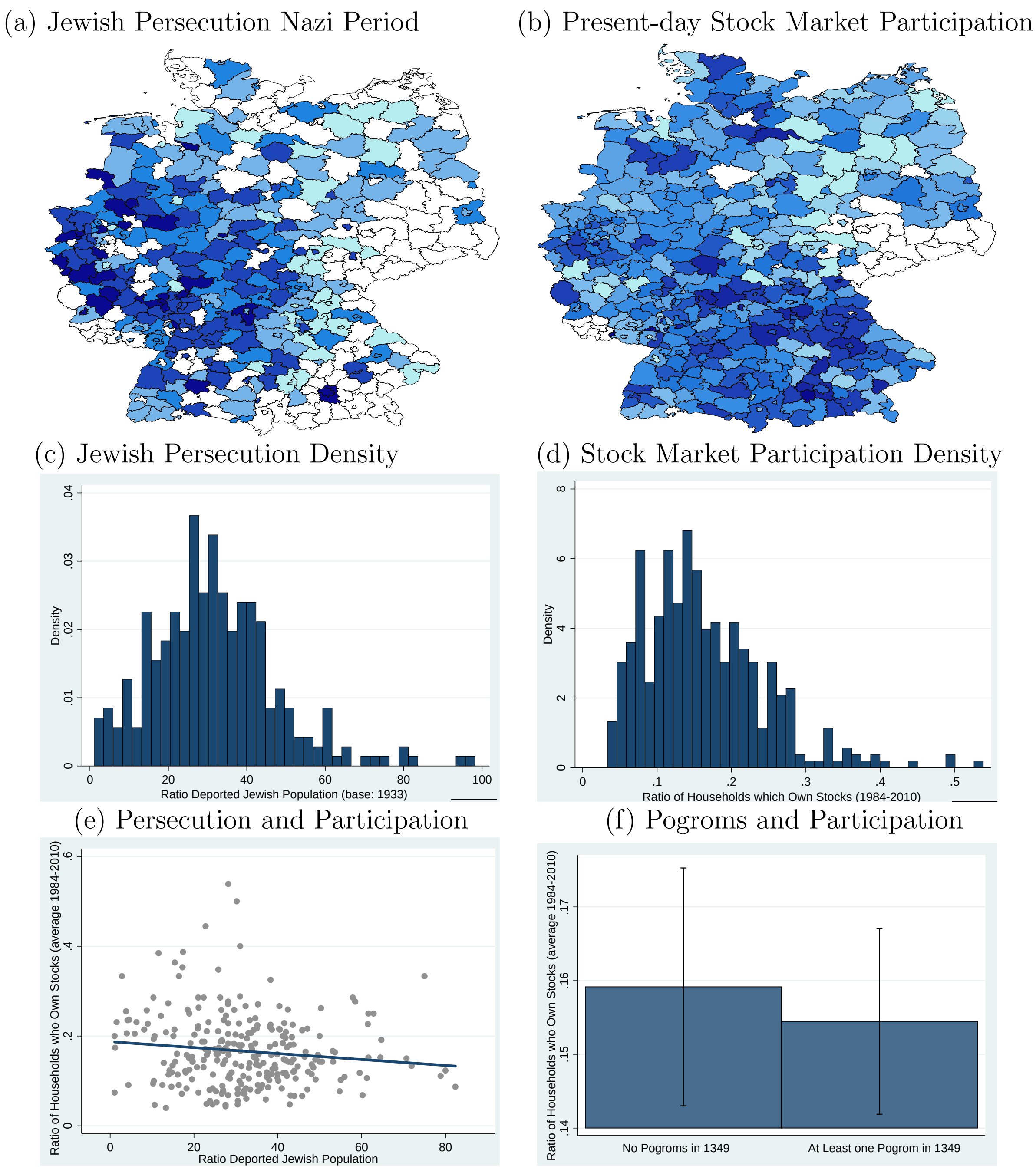


Each point is a German county. The vertical axis plots the residuals from estimating the following equation,

$$RatioStockhold_k = \alpha + K_k' \times \delta + \epsilon_k,$$

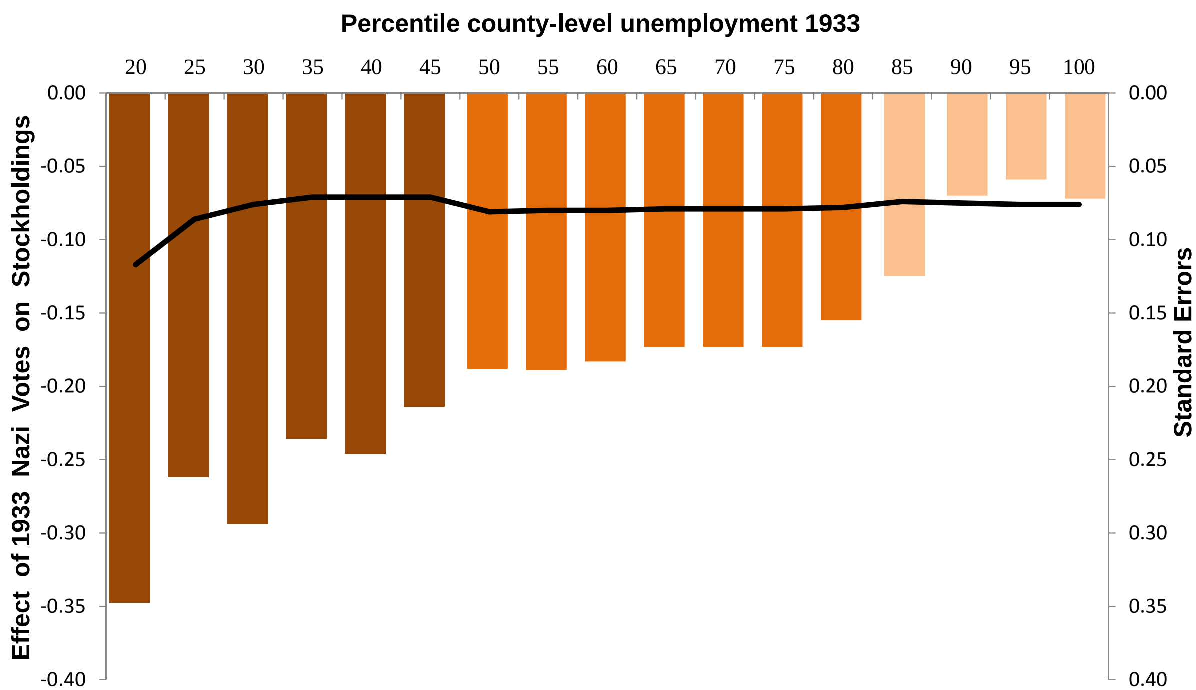
where K_k is the set of county-level observables of Equation 1 in the paper, against the residuals of a regression of the principal component of the measures of Jewish persecution on the same set of covariates, K_k.

Figure 3.2: Data Properties: Jewish Persecution and Stock Market Participation



In Panels (a) and (b), the darker a county is, the higher the value of the depicted variable. Blank counties are those for which the data are not available. Panel (a) plots the ratio of Jews deported during the Nazi period over the total Jewish population in 1933. Panel (b) plots the average yearly ratio of households who invest in stocks from 1984 to 2010. Panels (c) and (d) plot the distributions of the same measures as above. Panel (e) depicts the unconditional correlation between stock market participation and the ratio of Jews deported across German counties. Panel (f) shows the mean stock market participation in counties that experienced and did not experience a pogrom in 1349.

Figure 3.3: Nazi Votes, Economic Crisis, and Stock Market Participation



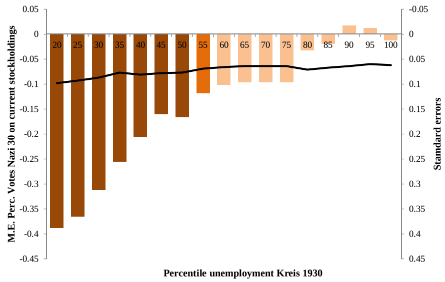
This graph plots the average marginal effects of the vote share for the Nazi Party in 1933 when estimating the following probit specification across subsamples of households sorted by the unemployment rate in their county as of 1933:

$$Pr(HoldsStocks_{ik}|X_i, K_{ik}) = \Phi(\alpha + \beta \times VoteShareNazi1933_k + X_i' \times \gamma + K_{ik}' \times \delta + \eta_t + \epsilon_{ik}).$$

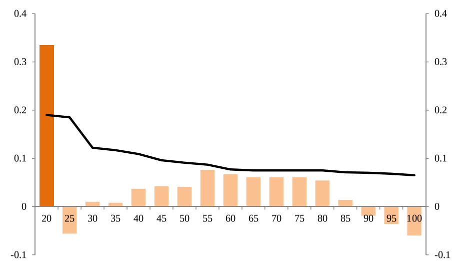
Each observation is a German household interviewed for the first time by the SOEP between 1984 and 2010. The left vertical axis reports the average marginal effect of $VoteShareNazi1933$, and it is associated with the histograms. The horizontal axis indicates the percentile of the distribution of counties by the unemployment rate in 1933 below which the estimation is performed. For instance, the histogram labeled "20" reports the average marginal effect for estimating the probit model for households in counties where the unemployment rate in 1933 was below the 20th percentile; the histogram labeled "30" reports the average marginal effect for estimating the probit model for households in counties where the unemployment rate in 1933 was below the 30th percentile. The right vertical axis reports standard errors attached to each marginal effect and clustered at the county level. They are associated with the black line. Dark brown histograms are marginal effects that are significant at the 1% level or lower; orange histograms are significant at the 5% level; white histograms are not significant at any conventional level.

Figure 3.4: Economic Crisis and Stock Market Participation: Other Parties and Placebo Outcomes

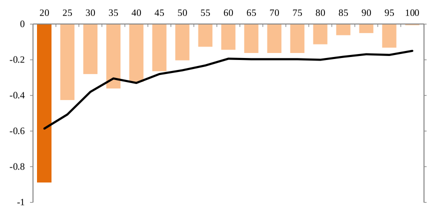
(a) Nazi Votes 1930



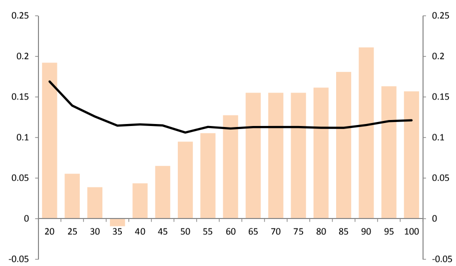
(b) Social-Democratic Votes 1933



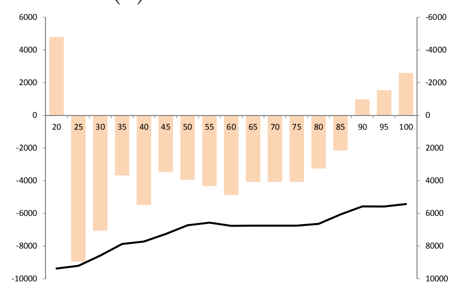
(c) Communist Votes 1933



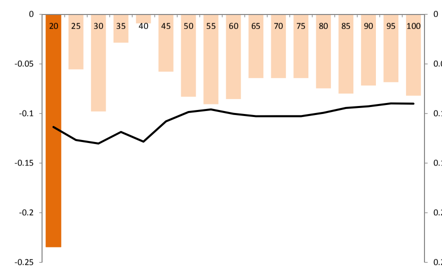
(d) Investments in Life Insurance



(e) Household Income



(f) Female Household Head

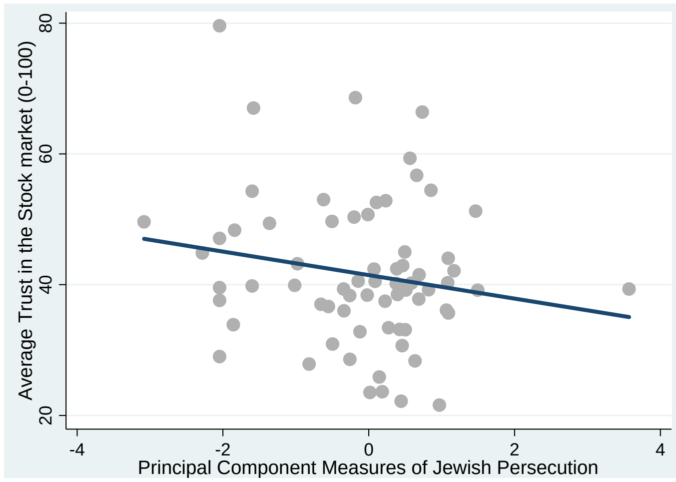


These graphs plot the average marginal effects of the vote share for the Nazi Party in 1933 on several outcomes computed from the following OLS specifications across subsamples of households sorted by the unemployment rate in their county of residence as of 1933:

$$DepVar_{ik} = \alpha + \beta \times VoteShareNazi1933_k + X_i' \times \gamma + K_{ik}' \times \delta + \eta_t + \epsilon_{ik}$$

Each observation is a German household interviewed by the SOEP any time between 1984 and 2010. $DepVar$ is indicated on top of each graph. In each graph, the left vertical axis reports the OLS coefficient on $VoteShareNazi1933$, and it is associated with the histograms. The horizontal axis indicates the percentile of the distribution of counties by the unemployment rate in 1933 below which the estimation is performed. For instance, the histogram labeled "20" reports the OLS coefficient for estimating the OLS specification only for households which live in counties where the unemployment rate in 1933 was below the 20th percentile; the histogram labeled "30" reports the average marginal effect for estimating the OLS specification only for households who live in counties where the unemployment rate in 1933 was below the 30th percentile. The right vertical axis reports standard errors attached to each coefficient and clustered at the county level. They are associated with the black line. Dark brown histograms are marginal effects that are significant at the 1% level or lower; orange histograms are significant at the 5% level; white histograms are not significant at any conventional level.

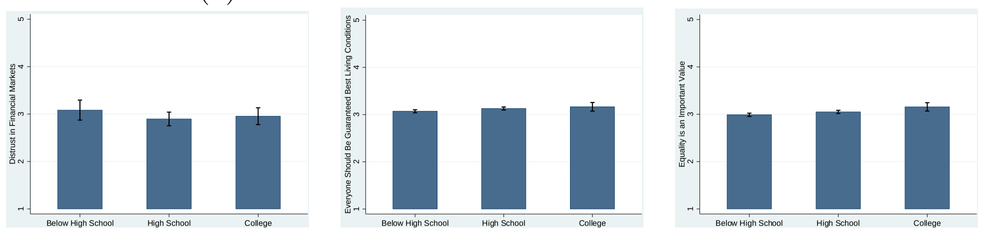
Figure 3.5: Jewish Persecution and Trust in Finance: County-level Averages



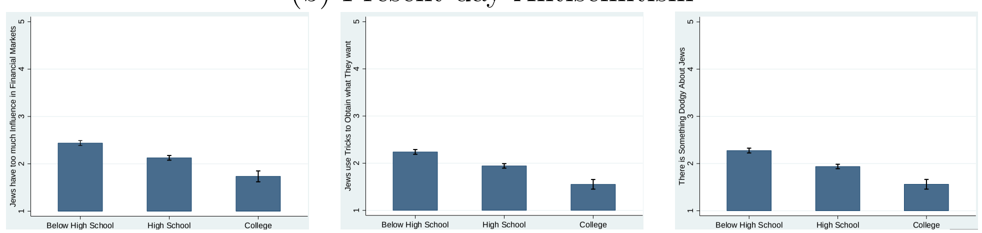
In this graph, the vertical axis reports the county-level average trust in financial markets self-reported on an integer scale between 0 and 100 by a pool of German respondents surveyed by the authors through the company *Clickworker*. The horizontal axis reports the principal component of the measures of Jewish persecution at the county level.

Figure 3.6: Distrust in Finance, Antisemitism, and Backwardness by Education Levels

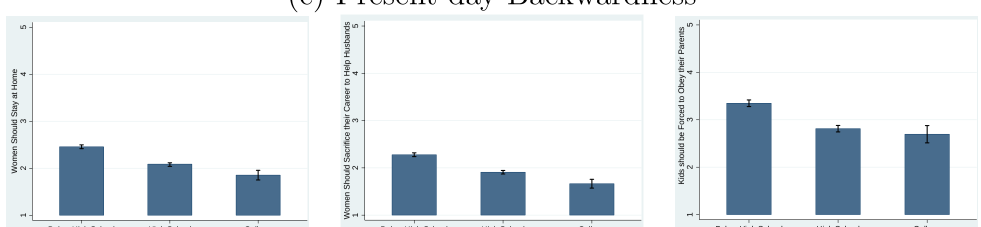
(a) Distrust in Finance and Neutral Values



(b) Present-day Antisemitism



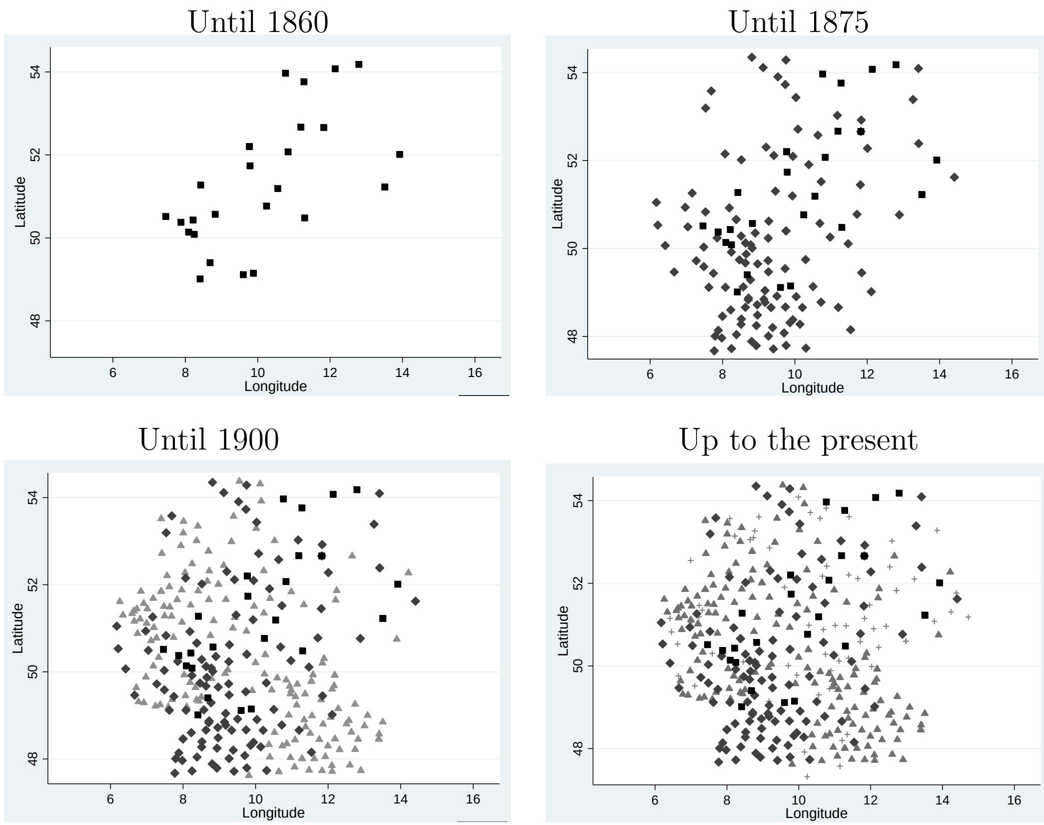
(c) Present-day Backwardness



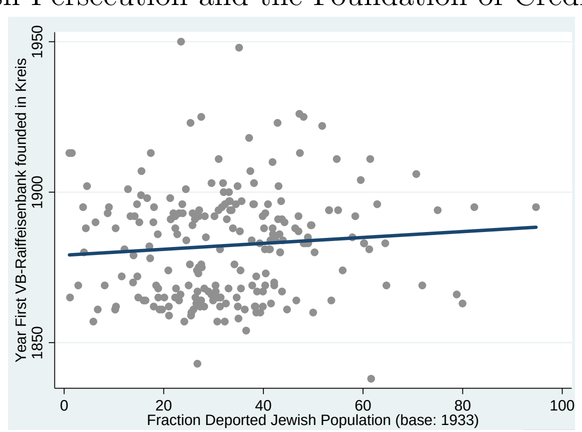
Each graph reports the sample mean of the outcomes described on the vertical axes split in three groups sorted by the respondent's education level: *Below High School* includes respondents who have a *Realschule* degree or lower level of education; *High School* includes respondents who have an *Abitur* or correspondent level of education in the German system. *College* includes respondents with a *Vordiplom* or higher level of education.

Figure 3.7: Supply-side Channel 1: Jewish Persecution and the Diffusion of Credit Unions

(a) Diffusion Path of Credit Unions across Space and Over Time



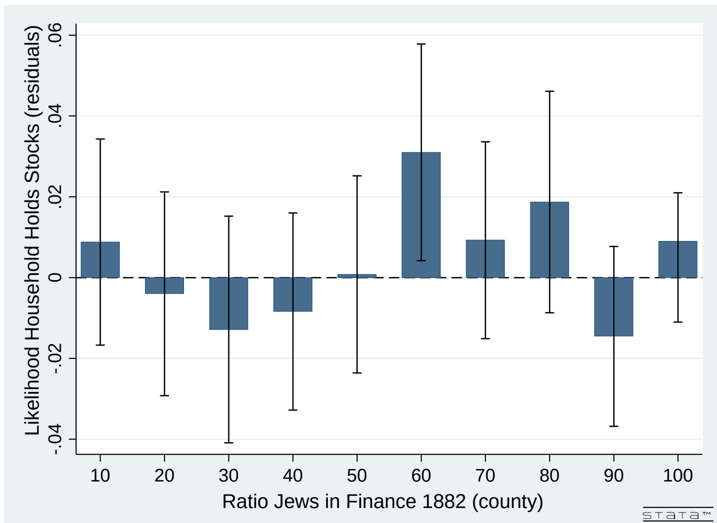
(b) Jewish Persecution and the Foundation of Credit Unions



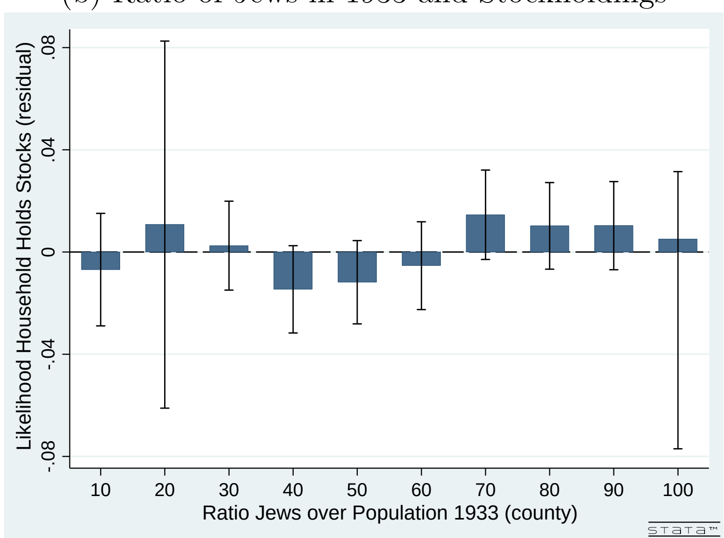
Panel (a) plots the Germans counties where credit unions were first introduced at different points in time. Panel (b) plots the correlation between the year when the first credit union of a county was founded and the ratio of Jews deported during the Nazi period over the total Jewish population in 1933. Each point is a German county.

Figure 3.8: Supply-side Channel 2: Jews in Finance and Stockholdings

(a) Ratio of Jews in Finance in 1882 and Stockholdings



(b) Ratio of Jews in 1933 and Stockholdings



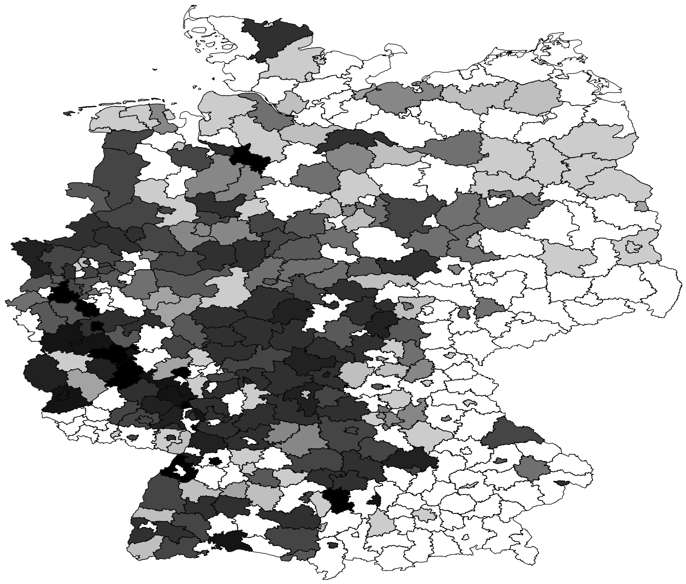
Panel (a) and Panel (b) plot the average residuals from the following OLS equation,

$$HoldsStocks_i = \alpha + X_i' \times \delta + K_{ik}' \times \delta + \eta_t + \epsilon_i,$$

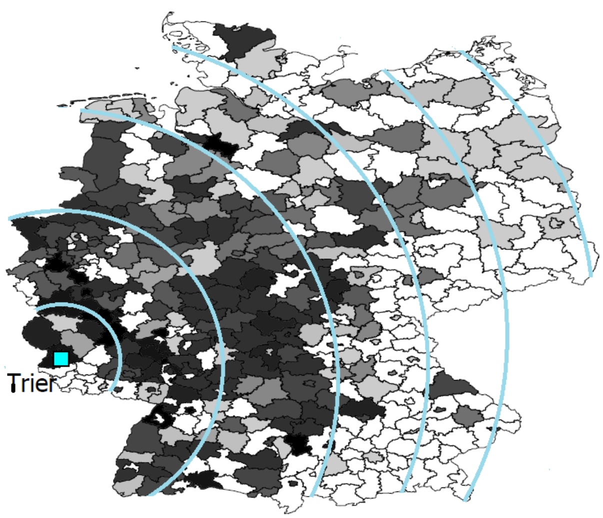
across the deciles of the distribution of the share of financial employees of the Jewish religion in a county as of 1882 and of the ratio of Jews over the total German population in a county as of 1933. Intervals represent 95% confidence intervals for the estimated averages.

Figure 3.9: Distance from the Rhine and Jewish Communities in the Middle Ages

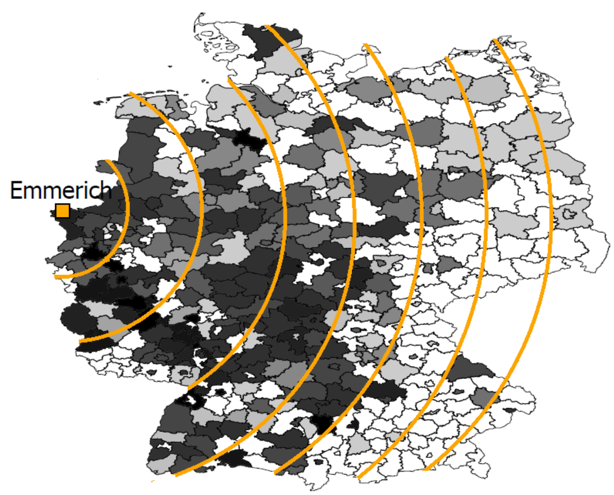
(a) Year when the first Jewish community was documented



(b) Distance from Trier



(c) Distance from Emmerich



In the maps, the darker a county is, the earlier a Jewish community was documented in that county. Blank counties are those for which the data are not available. The bottom maps show the location of the cities of Trier, on the Mosel, and Emmerich, on the northern end of the German Rhine. The isodistance curves out of the two cities emphasize which counties are at the same distance from Trier or from Emmerich.

Table 3.1: Summary Statistics

	All available observations					Non-missing information					p-value
	Obs.	Mean	St. dev.	Min.	Max.	Obs.	Mean	St. dev.	Min.	Max.	
Persecution of Jews											
Pogrom 1349	17818	0.576	0.494	0	1	13870	0.570	0.495	0	1	0.358
Log deported Jews	17818	4.956	2.311	0	10.93	13870	4.950	2.295	0	10.93	0.881
Ratio deported Jews	17477	34.06	19.61	0	163.5	13599	34.34	19.81	0	163.5	0.158
Principal component persecution	17477	-0.223	1.106	-3.082	3.571	13599	-0.222	1.112	-3.082	3.571	0.980
Household characteristics											
Holds Stocks	26761	0.163	0.369	0	1	13870	0.158	0.364	0	1	0.092
Homeowner	27064	0.391	0.488	0	1	13870	0.417	0.493	0	1	0.009
Has life insurance	26761	0.469	0.499	0	1	13870	0.462	0.499	0	1	0.198
Income	26761	31522	26614	-36	986400	13870	30514	27061	-36	986400	0.000
Age	21981	48.62	17.64	17	97	13870	48.49	17.58	17	96	0.371
Female	21982	0.490	0.500	0	1	13870	0.473	0.499	0	1	0.000
Single	27064	0.178	0.382	0	1	13870	0.218	0.413	0	1	0.000
High School or higher	27079	0.766	0.423	0	1	13870	0.718	0.450	0	1	0.000
College	27075	0.016	0.127	0	1	13867	0.022	0.146	0	1	0.000
County characteristics											
Log Population 1933	17818	11.57	1.371	7.94	14.29	13870	11.56	1.368	7.94	14.29	0.638
Log Jews 1933	17818	6.271	2.047	0	11.99	13870	6.261	2.041	0	11.99	0.708
Percentage unemployed 1933	17818	18.21	8.768	2.618	40.52	13870	18.01	8.838	2.618	40.52	0.028
Percentage blue collars 1933	17818	44.85	10.56	16.49	72.40	13870	44.67	10.68	16.49	72.40	0.023
Percentage self employed 1933	17818	20.01	4.211	9.096	32.74	13870	20.08	4.265	9.096	32.74	0.064
Percentage Catholics 1925	17818	40.11	32.71	0.502	98.77	13870	41.27	32.53	0.517	98.77	0.006
Eastern Germany	27079	0.129	0.335	0	1	13870	0.071	0.258	0	1	0.000
Latitude	24959	50.69	1.727	47.95	54.03	13870	50.65	1.642	47.95	54.03	0.643
Land quality index	24959	0.562	0.149	0.306	0.870	13870	0.566	0.149	0.306	0.870	0.628
Income p.c. 2005	22748	18033	2301	12846	27253	13870	18099	2220	13115	25027	0.452
Ratio College graduates 2005	24766	24.40	4.929	17.60	34.60	13870	23.55	4.463	17.60	34.60	0.010

This Table reports summary statistics for the historical persecution of Jews in the German lands, and for the characteristics of households and counties where households live. Each observation is a German household interviewed by the SOEP for the first time between 1984 and 2010. For each variable, the table reports the number of observations for which the variable is observed, its mean, standard deviation, minimal, and maximal values. The left Panel reports statistics for households interviewed by the SOEP whose county of residence is known. The right Panel reports statistics for the subsample of households for which we are not missing any household or county-level characteristics. The column on the far right reports p-values for tests of the difference in the means across the two samples.

Table 3.2: Jewish Persecution and Stock Market Participation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Principal Component Measures of Persecution			Log Deported Jews (*extensive margin*)		Ratio Deported Jews (*intensive margin*)		Pogrom 1349	
	Probit	Probit	OLS	Probit	OLS	Probit	OLS	Probit	OLS
Persecution of Jews	-0.011	-0.012	-0.013	-0.005	-0.006	-0.032	-0.033	-0.021	-0.021
(column heading)	*0.005***	*0.004****	*0.004****	*0.002***	*0.002****	*0.017**	*0.017**	*0.009***	*0.009***
Log Jews 1933	0.045	-0.054	-0.017	0.022	-0.014	-0.020	-0.032	-0.025	-0.031
	0.047	*0.052*	*0.052*	*0.038*	*0.038*	*0.056*	*0.052*	*0.037*	*0.037*
% Catholics 1925	0.029	0.006	0.004	0.003	-0.002	0.004	0.002	0.001	-0.001
	*0.013***	*0.011*	*0.012*	*0.011*	*0.011*	*0.011*	*0.011*	*0.011*	*0.011*
Age		-0.002	-0.002	-0.003	-0.003	-0.002	-0.002	-0.003	-0.003
		*0.001**	*0.001**	*0.001***	*0.001***	*0.001**	*0.001***	*0.001***	*0.001***
Age^2 /100		0.004	0.004	0.004	0.005	0.004	0.004	0.005	0.004
		*0.001****	*0.001****	*0.001****	*0.001****	*0.001****	*0.001****	*0.001****	*0.001****
Female		-0.036	-0.010	-0.014	-0.020	0.003	-0.010	-0.015	-0.020
		0.060	*0.061*	*0.060*	*0.060*	*0.006*	*0.061*	*0.060*	*0.060*
Single		0.058	0.052	0.057	0.052	0.058	0.052	0.058	0.052
		*0.010****	*0.009****	*0.010****	*0.009****	*0.010****	*0.010****	*0.010****	*0.009****
College		0.058	0.091	0.058	0.091	0.058	0.091	0.059	0.093
		*0.021****	*0.032****	*0.021****	*0.031****	*0.021****	*0.032****	*0.020****	*0.031****
Eastern Germany		-0.028	-0.024	-0.035	-0.031	-0.029	-0.026	-0.026	-0.021
		0.021	*0.019*	*0.021*	*0.019*	*0.022*	*0.020*	*0.020*	*0.018*
Income p.c. 2005		0.033	0.037	0.035	0.039	0.037	0.040	0.028	0.031
		*0.016***	*0.017***	*0.016***	*0.017***	*0.017***	*0.017***	*0.017**	*0.017**
% College graduates 2005		0.128	0.110	0.162	0.145	0.129	0.113	0.090	0.065
		0.142	*0.142*	*0.141*	*0.140*	*0.145*	*0.146*	*0.128*	*0.126*
Income deciles	X	X	X	X	X	X	X	X	X
Other historical controls	X	X	X	X	X	X	X	X	X
Wave groups f.e.		X	X	X	X	X	X	X	X
Regional controls		X	X	X	X	X	X	X	X
Observations	17,360	13,596	13,596	13,867	13,867	13,596	13,596	13,867	13,867
N. of clusters	287	261	261	270	270	261	261	270	270
(Pseudo-) R^2	0.01	0.11	0.09	0.11	0.09	0.11	0.09	0.11	0.09

This Table reports the average marginal effects computed after estimating the following probit specification:

$$Pr(HoldsStocks_{ik}|X_{ik}, K_{ik}) = \Phi(\alpha + \beta \times Persecution_k + X'_{ik} \times \gamma + K'_{ik} \times \delta + Incomedeciles_{ik} + \eta_t + \epsilon_{ik}).$$

Each observation is a German household interviewed by the SOEP for the first time between 1984 and 2010. In all columns, the dependent variable is a dummy that equals 1 if the household holds stocks. The main covariate of interest, $Persecution$, is the measure of Jewish persecution reported at the top of each column. S.e. are clustered at the county level. Statistical significance is reported as follows: *10%, **5%, ***1%.

Table 3.3: Homogeneous Counties and Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Homogeneous Counties						Robustness			
	Within States	Within 9 Virtual States	Within 16 Virtual States	Within II World War Zone	Only if community	County level	No poorest counties	No least educated counties	No Hanseatic League	No Free Imperial Cities
Principal Component	-0.009	-0.010	-0.009	-0.011	-0.010	-0.017	-0.018	-0.020	-0.010	-0.012
Persecution of Jews	*0.004***	*0.004****	*0.004***	*0.004***	*0.004****	*0.005****	*0.005****	*0.006****	*0.004***	*0.005****
Individual contr.	X	X	X	X	X		X	X	X	X
Historical contr.	X	X	X	X	X	X	X	X	X	X
Regional controls	X	X	X	X	X	X	X	X	X	X
Wave group f.e.	X	X	X	X	X		X	X	X	X
Observations	13,599	13,599	13,599	13,599	9,881	261	9,269	8,228	11,289	10,817
N. of clusters	261	261	261	261	190		153	164	235	216
(Pseudo-) R^2	0.10	0.08	0.11	0.09	0.09	0.23	0.10	0.10	0.09	0.09

This Table reports the average marginal effects computed after estimating the following probit specification:

$$Pr(HoldsStocks_{ik}|X_{ik}, K_{ik}) = \Phi(\alpha + \beta \times Persecution_k + X'_{ik} \times \gamma + K'_{ik} \times \delta + Incomedeciles_{ik} + \eta_t + \epsilon_{ik}),$$

adding the geographic fixed effects described on the column headings (columns (1)-(5), or across subsamples defined as described by the column headings (columns (6)-(10)). Each observation is a German household interviewed by the SOEP for the first time between 1984 and 2010. In all columns, the dependent variable is a dummy that equals 1 if the household holds stocks. The main covariate of interest, $Persecution$, is the principal component of the three measures of Jewish persecution described in the main text. S.e. are clustered at the county level. Statistical significance is reported as follows: *10%, **5%, ***1%.

Table 3.4: Jewish Persecution and Present-day Trust in Finance: Individuals

	(1)	(2)	(3)	(4)
P.C. Persecution Jews	-1.812	-2.151	-3.205	-3.000
	(1.031)	(1.083)**	(1.412)**	(1.551)*
Male		9.182	8.767	8.885
		(2.436)***	(2.437)***	(2.573)***
Log Jews 1925			1.175	1.218
			(0.950)	(1.044)
% unemployed 1933			0.030	0.031
			(0.189)	(0.202)
Age group f.e.		X	X	X
Education group f.e.		X	X	X
Income p.c. quintiles		X	X	X
Land f.e.				X
Observations	700	618	611	611
N. of clusters	69	64	63	63
(Pseudo-) R^2	0.01	0.01	0.01	0.01

This Table reports the coefficients for estimating four tobit specifications whose outcome variable is an integer measure between 0 and 100 of trust in financial markets reported by 1,000 Germans surveyed by the authors through the company *Clickworker*. The sample is restricted to respondents in counties for which the data on Jewish persecution in the Middle Ages are available, and in counties for which we had at least five respondents. Standard errors are clustered at the county level. S.e. are clustered at the county level (Kreis). Statistical significance is reported as follows: *10%, **5%, ***1%.

Table 3.5: Non-sensitivity of the effect to education levels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	P.C. Measures of Persecution		Log deported Jews (extensive margin)		Ratio deported Jews (intensive margin)		Pogrom 1349	
Panel A. Effect by individual level of education								
Persecution	-0.011	-0.010	-0.005	-0.005	-0.030	-0.015	-0.013	-0.024
	*0.005***	*0.007*	*0.002***	*0.003*	*0.018**	*0.037*	*0.009*	*0.014**
Persecution*College	-0.029	-0.029	-0.009	-0.009	-0.017	-0.016	-0.045	-0.041
	0.026	*0.0027*	*0.011*	*0.012*	*0.014*	*0.014*	*0.059*	*0.061*
Persecution*High School		-0.002		-0.001		-0.002		0.006
		0.006		*0.003*		*0.019*		*0.013*
College	0.105	0.106	0.157	0.159	0.163	0.162	0.138	0.137
	*0.031****	*0.0032****	*0.059****	*0.063***	*0.053****	*0.055****	*0.046****	*0.047****
High School		0.013		0.015		0.019		0.009
		0.008		*0.018*		*0.015*		*0.012*
Panel B. Effect by ratio college graduates in county								
Persecution	-0.020	-0.020	-0.008	-0.008	-0.093	-0.087	-0.029	-0.030
	*0.006****	*0.006****	*0.003****	*0.003****	*0.035****	*0.035***	*0.013***	*0.014***
Persecution*High College	0.007	0.005	0.007	0.006	0.129	0.107	-0.015	-0.014
	0.009	*0.009*	*0.004**	*0.004*	*0.056***	*0.056**	*0.018*	*0.020*
Persecution*Low College	0.013	0.012	0.004	0.003	0.069	0.063	0.024	0.024
	*0.007***	*0.007**	*0.004*	*0.004*	*0.036**	*0.036**	*0.017*	*0.017*
Individual, historical controls	X	X	X	X	X	X	X	X
Wave groups f.e.	X	X	X	X	X	X	X	X
Current regional controls	X	X	X	X	X	X	X	X

This Table reports average marginal effects computed after estimating the following probit spefication:

$$Pr(HoldsStocks_{ik}|X_i, K_{ik}) = \Phi(\alpha + \beta \times Persecution_k + X_i' \times \gamma + K_{ik}' \times \delta + \eta_t + \epsilon_{ik}).$$

Each observation is a German household interviewed by the SOEP for the first time between 1984 and 2010. In all columns, the dependent variable is a dummy equal to 1 if the household holds stocks. The main covariate of interest, $Persecution$, is the measure of Jewish persecution described at the top of each column. In Panel A, $Persecution$ is interacted with dummies that equal 1 if the household head holds a college degree or a high school degree. In Panel B, $Persecution$ is interacted with dummy variables for two groups of households: those in the top third and those in the bottom third of the distribution based on the county-level share of college graduates. S.e. are clustered at the county level (Kreis). Statistical significance is reported as follows: *10%, **5%, ***1%.

Table 3.6: Properties of the Effect: Over Time and Across Cohorts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Effect over time	P.C. Measures of Persecution		Log deported Jews *(extensive margin)*		Ratio deported Jews *(intensive margin)*		Pogrom 1349	
	Probit	OLS	Probit	OLS	Probit	OLS	Probit	OLS
Persecution	-0.016	-0.015	-0.005	-0.006	-0.047	-0.044	-0.025	-0.024
	*0.006****	*0.006****	*0.003**	*0.003***	*0.025**	*0.022***	*0.013**	*0.013**
Persecution*1984-1987	0.017	0.014	0.005	0.004	0.102	0.101	0.003	0.001
	0.013	*0.011*	*0.007*	*0.006*	*0.057**	*0.053**	*0.025*	*0.021*
Persecution*1988-1991	0.001	-0.001	-0.003	-0.004	0.028	0.024	0.002	0.001
	0.015	*0.017*	*0.008*	*0.008*	*0.082*	*0.089*	*0.034*	*0.038*
Persecution*1992-1995	0.004	0.003	-0.002	-0.002	0.040	0.042	0.006	0.004
	0.013	*0.014*	*0.006*	*0.007*	*0.059*	*0.064*	*0.029*	*0.033*
Persecution*1996-1999	-0.004	-0.008	0.001	0.001	-0.034	-0.059	0.010	0.019
	0.014	*0.017*	*0.008*	*0.009*	*0.088*	*0.098*	*0.029*	*0.033*
Persecution*2000-2003	0.006	0.004	-0.003	-0.004	0.018	0.001	0.023	0.029
	0.007	*0.009*	*0.003*	*0.004*	*0.042*	*0.049*	*0.019*	*0.022*
Persecution*2004-2007	0.008	0.007	0.001	0.001	0.009	0.010	0.013	0.009
	0.008	*0.007*	*0.004*	*0.003*	*0.047*	*0.040*	*0.018*	*0.017*
Panel B. Effect across cohorts								
Persecution	-0.017	-0.018	-0.006	-0.007	-0.047	-0.043	-0.026	-0.027
	*0.006****	*0.006****	*0.002****	*0.002****	*0.022***	*0.022***	*0.010***	*0.011***
Persecution*Nazi Cohort	0.011	0.012	0.007	0.007	0.085	0.084	0.031	0.033
	*0.007**	*0.007**	*0.002****	*0.002****	*0.026****	*0.027****	*0.012***	*0.013***
Persecution*Post 1965	0.005	0.006	-0.004	-0.004	-0.068	-0.053	-0.010	-0.008
Cohort	*0.008*	*0.007*	*0.002**	*0.002***	*0.027***	*0.022***	*0.013*	*0.012*
Individual, historical controls	X	X	X	X	X	X	X	X
Wave groups f.e.	X	X	X	X	X	X	X	X
Current regional controls	X	X	X	X	X	X	X	X
Observations	13,559	13,559	13,870	13,870	13,559	13,559	13,870	13,870
N. of clusters	261	261	270	270	261	261	270	270
(Pseudo-) R^2	0.10	0.08	0.10	0.08	0.10	0.08	0.10	0.08

This Table reports average marginal effects computed after estimating the following probit specification:

$$Pr(HoldsStocks_{ik}|X_i, K_{ik}) = \Phi(\alpha + \beta \times Persecution_k + X_i' \times \gamma + K_{ik}' \times \delta + \eta_t + \epsilon_{ik}).$$

Each observation is a German household interviewed by the SOEP for the first time between 1984 and 2010. In all columns, the dependent variable is a dummy equal to 1 if the household holds stocks. The main covariate of interest, $Persecution$, is the measure of persecution of Jews described at the top of each column. In Panel A, $Persecution$ is interacted with a set of dummy variables each including four waves of SOEP surveys. In Panel B, it is interacted with dummy variables for three cohorts of households: those who experienced the Nazi period, those born after the Second World War, and those born after 1965. S.e. are clustered at the county level (Kreis). Statistical significance is reported as follows: *10%, **5%, ***1%.

Table 3.7: Distance from the Rhine Valley and Observables

Dependent Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Distance Trier		Distance Emmerich		Minimal distance	
Log County Income p.c.	0.009	0.011	-0.007	-0.008	0.006	0.009
	0.008	*0.009*	*0.011*	*0.013*	*0.007*	*0.008*
% College graduates	-0.096	-0.224	0.718	0.795	0.230	0.147
	0.293	*0.303*	*0.341***	*0.439**	*0.250*	*0.276*
Land Quality Index	-0.003	-0.003	-0.005	-0.005	-0.003	-0.003
	0.017	*0.020*	*0.014*	*0.019*	*0.014*	*0.016*
Female	0.001	-0.006	0.006	-0.003	0.004	-0.002
	0.012	*0.012*	*0.011*	*0.014*	*0.010*	*0.010*
Single	0.009	0.000	0.015	0.008	0.009	0.002
	0.012	*0.014*	*0.010*	*0.014*	*0.010*	*0.011*
Home owner	-0.006	-0.012	0.016	-0.000	-0.002	-0.013
	0.020	*0.023*	*0.020*	*0.025*	*0.017*	*0.019*
Log Income	0.006	0.019	-0.000	0.010	0.007	0.019
	0.023	*0.025*	*0.019*	*0.024*	*0.018*	*0.021*
Age	-0.401	-0.124	-0.630	-0.228	-0.434	-0.179
	0.634	*0.787*	*0.463*	*0.717*	*0.501*	*0.661*
High School or higher	-0.008	-0.005	-0.004	0.000	-0.007	-0.005
	0.009	*0.009*	*0.009*	*0.011*	*0.007*	*0.008*
Eastern Germany f.e.	X	X	X	X	X	X
Res. medieval	X		X		X	
Res. medieval and 1930s		X		X		X

This Table documents the association between the measures of distance from the Rhine Valley and current county- and individual-level observables. The measures of distance are the residuals ($\hat{\epsilon_k}$) from the following OLS specification,

$$DistanceRhine_k = \alpha + K_k' \times \delta + \epsilon_k, \tag{3.6}$$

where *DistanceRhine* is one of three measures indicated above each column, and K_k is a set of observables capturing the economic characteristics of counties in the past. In odd columns, K_k includes all observable proxies for economic growth in the Middle Ages (whether the county hosted a monthly market, a bishop siege, any free imperial city, any Hanseatic League city, any city incorporated before 1349) and geographic characteristics (latitude, the land quality index of Ramankutty et al. (2002), whether a navigable river existed in the county, whether the county is in eastern Germany). In even columns, K_k additionally includes the log of Jewish inhabitants as of 1933 and a set of socio-demographics of counties in 1933 (unemployment rate, ratio of workers in blue-collar jobs, in self-employment, ratio of Roman Catholics). Each line reports the estimated β in the following OLS specification,

$$Depvar_{ik} = \alpha + \beta \times \hat{\epsilon_k} + C_{ik}' \times \delta + \epsilon_{ik},$$

where *Depvar* is a county- or individual-level observable enlisted in the left column for each line, $\hat{\epsilon_k}$ are the residuals from 3.6, and C are current individual and county-level observables. S.e. are clustered at the county level (equivalent to Huber-White heteroskedasticity-robust s.e. in specifications at the county level). Statistical significance is reported as follows: *10%, **5%, ***1%.

Table 3.8: Placebo Tests with the Distance from the Rhine Valley

Panel A. Reduced Form	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Distance Trier			Distance Emmerich			Minimal distance		
Log Distance	0.014	0.005	0.002	0.017	0.009	0.009	0.017	0.009	0.008
	*0.006*****	*0.006*	*0.006*	*0.007*****	*0.007*	*0.007*	*0.005*******	*0.06*	*0.006*
P.C. Persecution		-0.011			-0.011			-0.011	
		*0.004*****			*0.004*****			*0.004*****	
Pogrom 1349			-0.019			-0.018			-0.019
			*0.009*****			*0.009*****			*0.009*****
Historical contr.	X	X	X	X	X	X	X	X	X
County controls	X	X	X	X	X	X	X	X	X
Individual controls	X	X	X	X	X	X	X	X	X
Observations	13870	13426	13679	13870	13426	13679	13870	13426	13679
Adjusted R^2	0.08	0.09	0.09	0.08	0.09	0.09	0.08	0.09	0.09
Panel B. Placebo Distance	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	French Households: West of the Rhine Valley, no Jewish migrations								
	Distance Trier			Distance Emmerich			Minimal distance		
Log Distance	-0.003	-0.005	-0.004	0.005	-0.003	-0.003	-0.003	-0.005	-0.004
	0.013	*0.011*	*0.011*	*0.028*	*0.024*	*0.024*	*0.013*	*0.011*	*0.011*
Regional controls	X	X	X	X	X	X	X	X	X
Individual controls		X	X		X	X		X	X
Income quintiles		X	X		X	X		X	X
Education quintiles		X	X		X	X		X	X
Size of town			X			X			X
Observations	9692	9383	9383	9692	9383	9383	9692	9383	9383
Adjusted R^2	0.01	0.13	0.13	0.01	0.13	0.13	0.01	0.13	0.13

Panel A reports the average marginal effects computed after estimating the following probit specification,

$$Pr(HoldsStocks_i|X_i, K_k) = \Phi(\alpha + \beta_1 \times DistanceRhine_k + \beta_2 \times PersecutionJews_k + X_i' \times \delta + K_k' \times \delta + \eta_t + \epsilon_{ik}), \qquad (3.7)$$

where $DistanceRhine$ is the measured distance from the Rhine Valley indicated above each column, and $PersecutionJews$ is the principal component of the margins of persecution against Jews in columns (2), (5), and (8), or an indicator for counties with at least a pogrom against the local Jewish community in 1349 in columns (3), (6), and (9). Panel B reports the coefficient on the distance measures for estimating 3.7 for French households interviewed by the *Enquete Patrimoine* run by the Bank of France in 2004. S.e. are clustered at the county level. Statistical significance is reported as follows: *10%, **5%, ***1%.

Table 3.9: Instrumenting Jewish Persecution: Three-stage OLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	First stage: Existence Jewish community in 1096			Second stage: Persecution of Jews			Third stage: Holds stocks		
	Dist. Trier	Dist. Emmerich	Minimal dist.	Dist. Trier	Dist. Emmerich	Minimal dist.	Dist. Trier	Dist. Emmerich	Minimal dist.
Residual Distance Rhine	-0.136	-0.213	-0.148						
	*0.044****	*0.043****	*0.037****						
Existence Jewish community in 1096				1.063	1.508	1.274			
				*0.504***	*0.430****	*0.490****			
Persecution of Jews							-0.063	-0.050	-0.057
							*0.027***	*0.021***	*0.024***
Historical controls	X	X	X	X	X	X	X	X	X
Regional controls	X	X	X	X	X	X	X	X	X
Observations	268	268	268	268	268	268	268	268	268
Adjusted R^2	0.09	0.08	0.06	0.33	0.35	0.25	0.16	0.16	0.16

This Table reports OLS coefficients for the three-stage county-level instrumental variable procedure described in Section 4. In the first stage (columns (1)-(3)), the probability that a Jewish community existed in 1096, before the Black Death, in each German county is predicted by the residuals of the distance of the county from the Rhine Valley regressed on observables for the medieval economic conditions of counties. In the second stage (columns (4)-(6)), the principal component of the measures of Jewish persecution is predicted by the predicted probability that a Jewish community existed in a county in 1096. In the third stage (columns (7)-(9)), the ratio of households who own stocks in each county is predicted with the predicted extent of Jewish persecution. In all stages, observations are German counties and coefficients are estimated with OLS. Hubert-White s.e. are reported below each coefficient. Statistical significance is reported as follows: *10%, **5%, ***1%.

Bibliography

Acemoglu, D. and D. Autor (2012). What Does Human Capital Do? *Journal of Economic Literature 50(2)*, 426–463.

Acemoglu, D., F. Gallego, and J. Robinson (2014). Institutions, Human Capital, and Development. *Annual Review of Economics*.

Acemoglu, D., T. Hassan, and J. Robinson (2011). Social Structure and Development: A Legacy of the Holocaust in Russia. *Quarterly Journal of Economics 126(2)*.

Acemoglu, D., S. Johnson, and J. Robinson (2001). The Colonial Origins of Comparative Development: an Empirical Investigation. *American Economic Review 91*, 1369–1401.

Acharya, V., R. Baghai, and K. Subramanian (2013). Labor Laws and Innovation. *Journal of Law and Economics 56(4)*, 997–1037.

Adena, M., R. Enikolopov, M. Petrova, V. Santarosa, and E. Zhuravskaya (2013). Radio and the Rise of the Nazis in Prewar Germany. *PSE Working Paper 2013-32*.

Aghion, P., C. Meghir, and J. Vandenbussche (2006). Growth, Distance to Frontier and the Composition of Human Capital. *Journal of Economic Growth 11*, 97–127.

Agrawal, A., A. Galasso, and A. Oettl (2014). Roads and Innovation. *Working Paper*.

Akbulut-Yuksel, M. and M. Yuksel (forthcoming). The Long-Term Direct and External Effects of Jewish Expulsions in Nazi Germany. *American Economic Journal: Economic Policy*.

Akcomak, I. and B. ter Weel (2009). Social capital, innovation and growth: Evidence from Europe. *European Econonomic Review 53*, 544–567.

Alesina, A. and J. Zeira (2006). Technology and Labor Regulations. *NBER Working Paper no. 12581*.

Algan, Y. and P. Cahuc (2006). Inherited Trust and Growth. *Journal of Political Economy 114(2)*, 280–316.

Allen, R. (2003). Progress and Poverty in Early Modern Europe. *Econonomic History Review 56(3)*, 406–443.

Almeida, H., P. Hsu, and D. Li (2013). Less is More: Financial Constraints and Innovation Efficiency. *Working Paper*.

Altomonte, C. and T. Aquilante (2012). The EU-Efige/Bruegel Dataset. *Bruegel Working Paper Series*.

Ambrose, B., P. Eichholtz, and T. Lindenthal (2013). House Prices and Fundamentals: 355 Years of Evidence. *Journal of Money, Credit, and Banking 45(2-3)*, 477–491.

Annoni, P. and K. Kozovska (2010). EU Regional Competitiveness Index 2010. *JRC Scientific and Technical Reports*.

Aristotle (320B.C.). Oikonomika.

Baird, S., C. McIntosh, and B. Ozler (2011). Cash or Condition? Evidence from a Cash Transfer Experiment. *Quarterly Journal of Economics 126(4)*, 1709–1753.

Banfield, E. (1958). The Moral Basis of a Backward Society. *The Free Press*.

Barbier, F. (2006). L'Europe de Gutenberg: Le Livre et l'Invention de la Modernite' Occidentale. *Paris: Belin Editions*.

Barro, R. (1991). Economic Growth in a Cross-section of Countries. *Quarterly Journal of Economics 106*, 407–443.

BEA, B. o. E. A. (2014). Gross-Domestic-Product-by-Industry data, 1997-2013.

Beck, T., A. Demirguc-Kunt, and R. Levine (2003). Law, Endowments, and Finance. *Journal of Financial Economics 70*, 137–181.

Becker, G. (1962). Investment in Human Capital: a Theoretical Analysis. *Journal of Political Economy 70(5)*, 9–49.

Becker, S., F. Cinnirella, E. Hornung, and L. Woessmann (2014). iPEHD - The Prussian Ifo Economic History Database. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 47(2)*, 57–66.

Becker, S., E. Hornung, and L. Woessmann (2011). Education and Catch-up in the Industrial Revolution. *American Economic Journal: Macroeconomics 3(3)*, 92–126.

Becker, S. and L. Woessmann (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics 124(2)*, 531–596.

Betermier, S., T. Jansson, C. Parlour, and J. Walden (2012). Hedging Labor Income Risk. *Journal of Financial Economics 105(3)*, 622–639.

Bloom, N. and J. Van Reenen (2007). Measuring and Explaining Management Practices Across Firms. *QuarterlyJournal of Economics 122(4)*, 1351–1408.

Botticini, M. and Z. Eckstein (2005). Jewish Occupational Selection: Education, Restriction or Minorities? *Journal of Economic History 65(4)*, 922–948.

Botticini, M. and Z. Eckstein (2012). The Chosen Few. *Princeton University Press*.

Buggle, J. (2013). Law and Social Capital: Evidence from the Code Napoleon in Germany. *Working Paper*.

Bunn, O. and R. Shiller (2014). Changing Times, Changing Values: A Historical Analysis of Sectors within the US Stock Market 1872-2013. *NBER Workin Paper no. 20370*.

Burchardi, K. and T. Hassan (2013). The Economic Impact of Social Ties: Evidence from German Reunification. *Quarterly Journal of Economics 128(3)*, 1219–1271.

Bursztyn, L. and L. Coffman (2012). The Schooling Decision: Family Preferences, Intergenerational Conflict, and Moral Hazard in the Brazilian Favelas. *Journal of Political Economy 120(3)*, 359–396.

Calomiris, C. and L. Schweikart (1991). The Panic of 1857: Origins, Transmission, and Containment. *Journal of Economic History 51(4)*, 807–834.

Cameron, C. and D. Miller (2011). Robust Inference with Clustered Data. *Handbook of Empirical Economics and Finance*, 1–28.

Campbell, J. and R. Shiller (1998). Valuation Ratios and the Long-Run Stock Market Outlook. *Journal of Portfolio Management*, 11–26.

Campbell, J. and R. Shiller (2001). Valuation Ratios and the Long-Run Stock Market Outlook: an Update. *NBER Working Paper no. 8221*.

Cantoni, D., Y. Chen, D. Yang, N. Yuchtman, and J. Zhang (2014). Curriculum and Ideology. *Working paper*.

Capen, E., R. Clapp, and W. Campbell (1971). Competitive Bidding in High Risk Situations. *Journal of Petroleum Technology 23*, 641–653.

Carlino, G. and W. Kerr (2014). Agglomeration and Innovation. *NBER Working Paper 20367*.

Chatterji, A., E. Glaeser, and W. Kerr (2013). Clusters of Entrepreneurship and Innovation. *NBER Working Paper 19013*.

Chitu, L., B. Eichengreen, and A. Mehl (2014). History, Gravity and International Finance. *Journal of international Money and Finance*.

Ciccone, A. and E. Papaioannou (2009). Human Capital, the Structure of Production, and Growth. *Review of Economics and Statistics 91(1)*, 66–82.

Cochrane, J. (2001). Asset Pricing. *Princeton University Press*.

Coffano, M. and G. Tarasconi (2014). CRIOS-Patstat Database: Sources, Contents, and Access Rules. *CRIOS Working Paper n.1*.

Conley, T. (1999). "GMM Estimation with Cross-Sectional Dependence". *Journal of Econometrics 105(1)*, 59–83.

D'Acunto, F. (2015a). History & Finance: A Nascent Field. *Working Paper*.

D'Acunto, F. (2015b). Identity, Overconfidence, and Investment Decisions. *Working Paper*.

D'Acunto, F. (2015c). Innovating to Invest: The Role of Basic Education. *Working Paper*.

D'Acunto, F., M. Prokopczuk, and M. Weber (2015). Distrust in Finance Lingers: Jewish Persecution and Households' Investments. *Working Paper*.

Davis, L., R. Hughes, and S. Reiter (1960). Aspects of Quantitative Research in Economic History. *Journal of Economic History 20(4)*, 539–547.

de Bromhead, A., B. Eichengreen, and K. O'Rourke (2013). Political Extremism in the 1920s and 1930s: Do the German Lessons Generalize? *Journal of Economic History 73(2)*, 371–406.

Decker, O., J. Kiess, and E. Braehler (2012). Die Mitte im Umbruch: Rechtextreme Einstellungen in Deutschland 2012. *Bonn: Verlag J.H.W. Dietz Nachf. GmbH*.

Dell, M. (2009). GIS Analysis for Applied Economists. *Manuscript*.

Dell, M. (2010). The Persistent Effect of Peru's Mining Mita. *Econometrica 78(8)*, 1863–1903.

Desmet, K. and E. Rossi-Hansberg (2012). Innovation in Space. *American Economic Review P. and P. 102(3)*, 447–452.

Dittmar, J. (2011). Information Technology and Economic Change: The Impact of the Printing Press. *Quarterly Journal of the Economics 126(3)*, 1133–1172.

Easterlin, R. (1981). Why isn't the Whole World Developed? *Journal of Economic History 41(1)*, 1–17.

Easterly, W. and R. Levine (2003). Tropics, Germs, and Crops: How Endowments Influence Economic Development. *Journal of Monetary Economics 50(1)*, 3–40.

Eichengreen, B. (2014a). Financial Development and Economic Growth: the Long View. *in Financial Development and Cooperation in Asia and the Pacific*.

Eichengreen, B. (2014b). Why Economics Needs History. *INET Interviews: 30 Ways to Be an Economist*.

Elhaik, E. (2012). The Missing Link of Jewish European Ancestry: Contrasting the Rhineland and the Khazarian Hypotheses. *Genome Biology and Evolution 5(1)*, 61–74.

Engelman, U. Z. (1944). The Rise of the Jew in the Western World. *New York: Behrman's Jewish Book House*.

EuropeanCommission (2013). A Recovery on the Horizon? Annual Report on Eropean SMEs.

Fogel, R. (1964). Railroads and American Economic Growth: Essays in Econometric History. *Baltimore: Johns Hopkins Press*.

Fogel, R. and S. Engerman (1974). Time on the Cross: The Economics of American Negro Slavery. *New York: W.W. Norton and Co.*.

Frehen, R., W. Goetzmann, and K. Rouwenhorst (2013). New Evidence on the First Financial Bubble. *Journal of Financial Economics*.

Fuchs-Schuendeln, N. and M. Schuendeln (2005). Precautionary Savings and Self-Selection: Evidence from the German Reunification Experiment. *Quarterly Journal of Economics*, 1085–1120.

Galetovic, A., S. Haber, and R. Levine (2014). Patent Holdup: do Patent Holders Holdup Innovation. *Working Paper*.

Galor, O. (2005). From Stagnation to Growth: Unified Growth Theory. *Handbook of Economic Growth 1*, 171–293.

Gennaioli, N., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2013). Human Capital and Regional Development. *Quarterly Journal of Economics 128(1)*, 105–164.

Gennaioli, N., A. Shleifer, and R. Vishny (2014a). Finance and the Preservation of Wealth. *Quarterly Journal of Economics, forthcoming*.

Gennaioli, N., A. Shleifer, and R. Vishny (2014b). Money Doctors. *Journal of Finance, forthcoming*.

Giannetti, M. and T. Wang (2014). Corporate Scandals and Household Stock Market Participation. *Working Paper*.

Giuliano, P. and A. Spilimbergo (forthcoming). Growing Up in a Recession. *Review of Economic Studies*.

Giusti, G., C. Noussair, and H. Voth (2013). Recreating the South Sea Bubble: Lessons from an Experiment in Financial History. *CEPR Discussion Paper no.9652*.

Glaeser, E. (2005). The Political Economy of Hatred. *Quarterly Journal of Economics 120(1)*, 45–86.

Glaeser, E., R. LaPorta, F. Lopez-de Silanes, and A. Shleifer (2004). Do Institutions Cause Growth? *Jurnal of Economic Growth 9(3)*, 271–303.

Glaeser, E., S. Pekkala Kerr, and W. Kerr (2014). Entrepreneurship and Urban Growth: an Empirical Assessment with Historical Mines. *Review of Economics and Statistics, forthcoming*.

Goetzmann, W., C. Labio, K. Rouwenhorst, and T. Young (2013). The Great Mirror of Folly: Finance, Culture, and the Crash of 1720. *New Haven: Yale Unversity Press*.

Goldin, C. (1995). Cliometrics and the Nobel. *Journal of Economic Perspectives 9(2)*, 191–208.

Goldin, C. and L. Katz (2008). The Race Between Education and Technology. *Harvard University Press*.

Gompers, P. (1995). Optimal Investmen, Monitoring, and the Staging of Venture Capital. *Journal of Finance 50*, 1461–1489.

Gompers, P. and J. Lerner (2001). The Venture Capital Revolution. *Journal of Economic Perspectives 15(2)*, 145–168.

Gorodnichenko, Y. and M. Schnitzer (2013). Financial Constraints and Innovation: Why Poor Countries don't Catch Up. *Journal of the European Economic Association 11*, 1115–1152.

Greif, A. (1997). Cliometrics After 40 Years: Microtheory and Economic History. *American Economic Review P. and P.*.

Grosfeld, I., A. Rodnyansky, and E. Zhuravskaya (2013). Persistent Antimarket Culture: A Legacy of the Pale of Settlement after the Holocaust. *American Economic Journal: Economic Policy 5(3)*, 189–226.

Grosjean, P. (2011). The Institutional Legacy of the Ottoman Empire: Islamic Rule and Financial Development in South Eastern Europe. *Journal of Comparative Economics 39*, 1–16.

Guiso, L. and T. Jappelli (2005). Awareness and Stock Market Participation. *Review of Finance 9(4)*, 537–567.

Guiso, L., T. Jappelli, and D. Terlizzese (1996). Income Risk, Borrowing Constraints and Portfolio Choice. *American Economic Review 86*, 158–172.

Guiso, L., P. Sapienza, and L. Zingales (2004). The Role of Social Capital in Financial Development. *American Economic Review 94(3)*.

Guiso, L., P. Sapienza, and L. Zingales (2009). Trusting the Stock Market. *Journal of Finance 63(6)*, 2557–2600.

Hanushek, E. and L. Woessmann (2012). Do Better Schools Lead to More Growth? *Journal of Economic Growth 17*, 267–321.

Hong, H., J. Kubic, and J. Stein (2004). Social Interaction and Stock Market Participation. *Journal of Finance*.

Justinian (533). Institutiones. *Constitutio "Imperatoriam maiestatem"*.

Kaplan, S. and P. Stromberg (2003). Financial Contract theory meets the Real World: an Empirical Analysis of Venture Capital Contracts. *Review of Economic Studies 70*, 281–315.

Kerr, W., J. Lerner, and A. Schoar (2014). The Consequences of Entrepreneurial Finance: Evidence from Angel Financings. *Review of Financial Studies 27(1)*, 20–55.

King, G., O. Osen, M. Tanner, and A. Wagner (2008). Ordinary Economic Voting Behavior in the Extraordinary Election of Adolf Hitler. *Journal of Economic History 68(4)*, 951–996.

Kogan, L., D. Papanikolau, A. Seru, and N. Stoffman (2013). Technological Innovation, Resource Allocation and Growth. *Working Paper*.

Kortum, S. and J. Lerner (2000). Assessing the Contribution of Venture Capital to Innovation. *RAND Journal of Economics 31*, 674–692.

Koudijs, P. (forthcoming). The Boats that did not Sail: Asset Price Volatility and Market Efficiency in a Natural Experiment. *Journal of Finance*.

Koudijs, P. and B. Golez (2014). Dividends and Discount Rates in the Very Long Run. *Working Paper*.

La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. Vishny (1998). Law and Finance. *Journal of Political Economy 1113-1155*, 106(6).

Lamoreaux, N. and K. Sokoloff (2007). Financing Innovation in the United States, 1870 to Present. *Cambridge: The MIT Press*.

Lerner, J. and U. Malmendier (2013). With a Little Help from my Random Friends: Success and Failure in Post-Business School Entrepreneurship. *Review of Financial Studies 26(10)*, 2411–2452.

Lerner, J., M. Sorensen, and P. Stromberg (2011). Private Equity and Long-Run Investment: The Case of Innovation. *Journal of Finance 66(2)*, 445–477.

Levine, R. (1998). The Legal Environment, Banks, and Long-Run Economic Growth. *Journal of Money, Credit, and Banking 30(3)*, 596–613.

Levitt, S., J. List, and C. Syverson (2013). Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant. *Journal of Political Economy 121(4)*, 643–681.

Lucas, R. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics 22*, 3–42.

Malmendier, U. (2002). Societas Publicanorum: staatliche Wirtschaftsaktivitaeten in den Haenden privater Unternehmer. *Koeln-Weimar-Wien: Verlag Boehlau*.

Malmendier, U., E. Moretti, and F. Peters (2012). Winning by Losing: Evidence on Overbidding in Mergers. *Working Paper*.

Malmendier, U. and S. Nagel (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking? *Quarterly Journal of Economics 126(1)*, 373–416.

Manso, G. (2011). Motivating Innovation. *Journal of Finance 66*, 1823–1869.

McCloskey, D. (1976). Does the Past Have Useful Economics? *Journal of Economic Literature 14(2)*, 434–461.

Michalopoulos, S. (2012). The Origins of Ethnolinguistic Diversity. *American Economic Review 102(4)*, 1508–1539.

Michalopoulos, S. and E. Papaioannou (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica 81(1)*, 113–152.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy 66(4)*, 281–302.

Mokyr, J. (2005a). Long-Term Economic Growth and the History of Technology. *Handbook of Economic Growth 1*, 1113–1180.

Mokyr, J. (2005b). The Intellectual Origin of Modern Economic Growth. *Journal of Economic History 65(2)*, 285–351.

Mokyr, J. and J. Voth (2009). Understanding Growth in Early Modern Europe. *Cambridge Economic History of Europe Cambridge:Cambridge University Press.*

Moretti, E. (2012). The New Geography of Jobs. *New York: Houghton Mifflin Harcourt*.

Moretti, E. and D. Wilson (forthcoming). State Incentives for Innovation, Star Scientists and Jobs: Evidence from Biotech. *Journal of Urban Economics*.

Moser, P. (2005). How do Patent Laws Influence Innovation? Evidence from the Nineteenth-Century World Fairs. *American Economic Review 95(4)*, 1214–1236.

Nelson, R. and E. Phelps (1966). Investment in Humans, Technological Diffusion and Economic Growth. *American Economic Review 56(1)*, 69–75.

Nunn, N. (2014). Historical Development. *Handbook of Economic Growth*.

OECD (2013). Education at a Glance. *OECD Publishing*.

OECD, O. f. E. C. D. (2005). Oslo Manual - Guidelines for Collecting and Interpreting Innovation Data.

Opler, T. and S. Titman (1994). Financial Distress and Corporate Performance. *Journal of Finance 49(3)*, 1015–1040.

Pascali, L. (forthcoming). Banks and Development: Jewish Communities in the Italian Renaissance and Current Economic Performance. *Review of Economics and Statistics, forthcoming*.

Paxson, C. (1990). Borrowing Constraints and Portfolio Choice. *Quarterly Journal of Economics 105(2)*, 535–543.

Pierce, L. and J. Snyder (2014). The Historical Roots of Firm Access to Finance: Evidence from the African Slave Trade. *Working Paper*.

Putnam, R. (1993). Making Democracy Work. *Princeton University Press.*

Putterman, L. and D. Weil (2010). Post-1500 Population Flows and the Long-Run Determinants of Economic Growth. *Quarterly Journal of Economics 125(4)*, 1627–1682.

Ramankutty, N. et al. (2002). The Global Distribution of Cultivable Lands. *Global Ecology and Biogeography 11*, 377–392.

Raymond, W., P. Mohnen, F. Palm, and S. Schim van der Loeff (2010). Persistence of Innovation in Dutch Manufacturing: is it Spurious? *Review of Economics and Statistics 92(3)*, 495–504.

Ruwet, J. and Y. Wellemans (1978). L'Analphabetisme en Belgique (XVIIIeme-XIXeme Siecles. *Bibliothque de l'Universite de Louvain.*

Sandberg, L. (1982). Ignorance, poverty and economic backwardness in the early stages of European industrialization: Variations on Alexandre Gerschenkron's grand theme. *Journal of European Economic History 11(3)*, 675–697.

Sapienza, P. and L. Zingales (2012). A Trust Crisis. *International Review of Finance 12(2)*, 123–131.

Saxenian, A. (1994). Regional Advantage: Culture and Competition in Silicon Valley and Route 128. *Harvard University Press.*

Schoellman, T. (2012). Education Quality and Development Accounting. *Review of Economic Studies 79(1)*, 388–417.

Schumpeter, J. (1911). The Theory of Economic Development. *Harvard University Press*.

Smith, A. (1776). An Inquiry into the Nature and Causes of the Wealth of Nations.

Spitzer, Y. (2014). Pogroms, Networks, and Migration. *Working Paper*.

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

Squicciarini, M. and N. Voigtlaender (2014). Human Capital and Industrialization: Evidence from the Age of Enlightenment. *Working Paper*.

Stock, J. and M. Yogo (2005). Testing for Weak Instruments in Linear IV Regression. *Identification and Inference for Econometric Models. New York: Cambridge University Press*.

Tabellini, G. (2010). Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association 8(4)*, 677–716.

Toch, M. (2012). The Economic History of European Jews: Late Antiquity and Early Middle Ages. *Leiden: Brill Academic Publishers*.

USPTO, U. S. P. T. O. (2012). U.S. Patenting Trends by NAICS Industry Categories, 1963-2012.

van Rooij, M., A. Lusardi, and R. Alessie (2011). Financial Literacy and Stock Market Participation. *Journal of Financial Economics 101(2)*, 449–472.

van Straten, J. (2004). Jewish Migrations from Germany to Poland: the Rheinland Hypothesis Revisited. *The Mankind Quarterly*, 367–384.

Veblen, J. (1904). The Theory of Business Enterprise. *New York: Charles Scribner's Sons*.

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

Voigtlaender, N. and H. Voth (2013a). Exorcizing Hitler: Denazification and Anti-Semitism in Germany, 1933-2006. *CEPR Discussion Paper 8935*.

Voigtlaender, N. and H. Voth (2013b). Married to Intolerance: Attitudes towards Intermarriage in Germany, 1900-2006. *American Economic Review Papers and Proceedings 103(3)*, 79–85.

Waldinger, F. (2010). Quality Matters: The Expulsion of Professors and the Consequences for PhD Students Outcomes in Nazi Germany. *Journal of Political Economy 118(4)*, 787–831.

WIPO, W. I. P. O. (2006). International Patent Classification - Colre Level, Volume 5.

Xenophon (360B.C.). Oikonomikos.

Appendix A

Appendix - Innovating to Invest

Figure A.1: Distribution of Regional Variables Within Countries

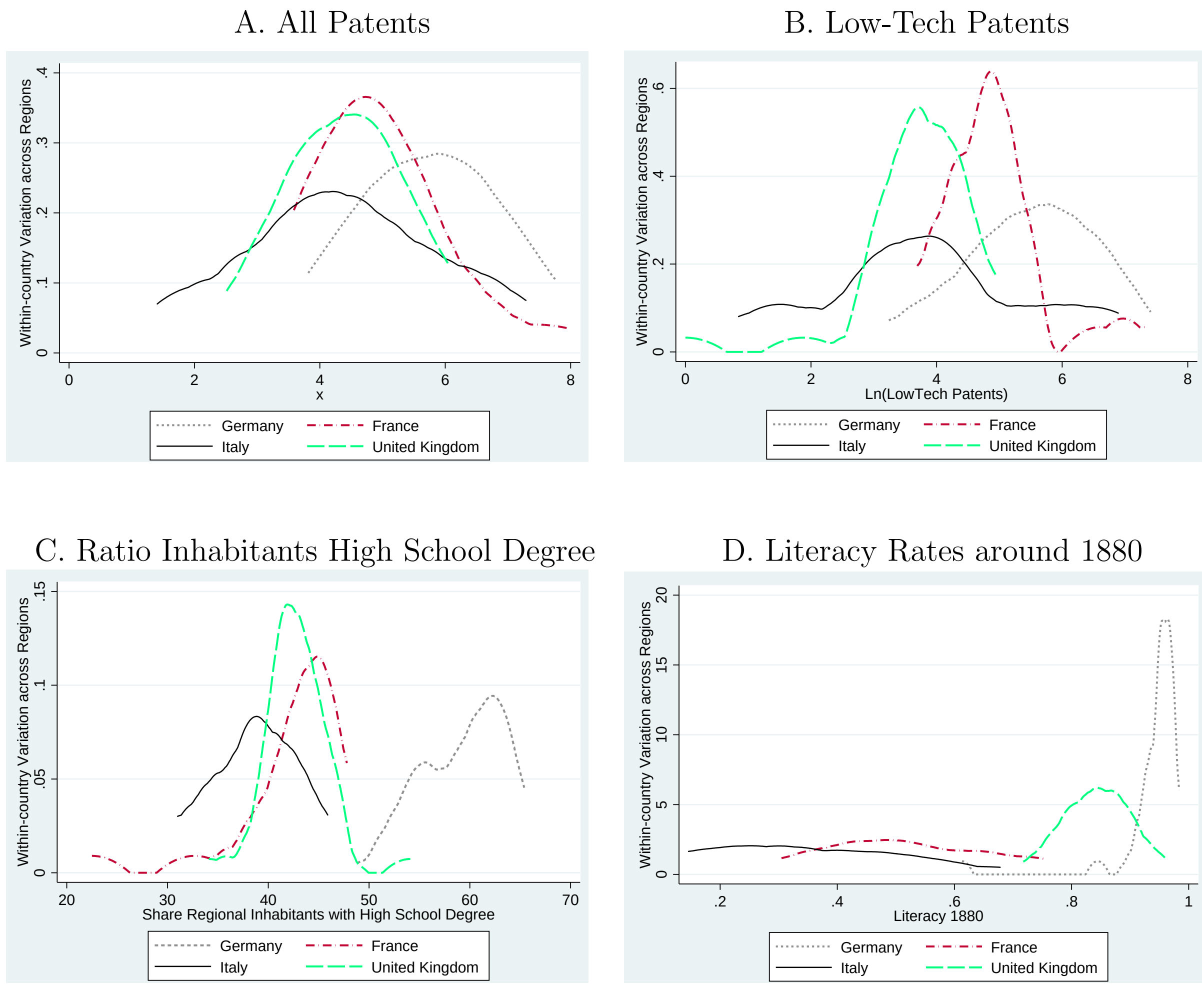
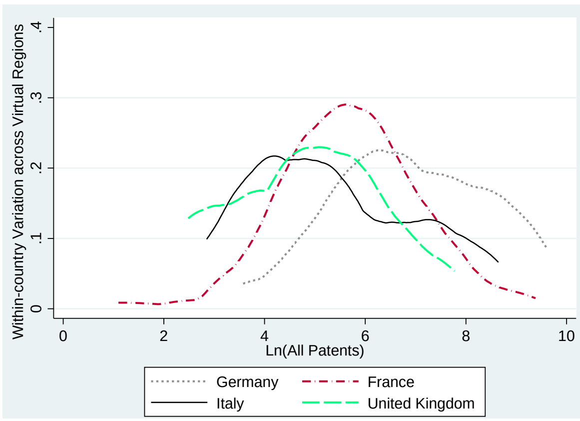


Figure A.2: Distribution of Virtual-Region Variables Within Countries

A. All Patents



B. Ratio Inhabitants High School Degree

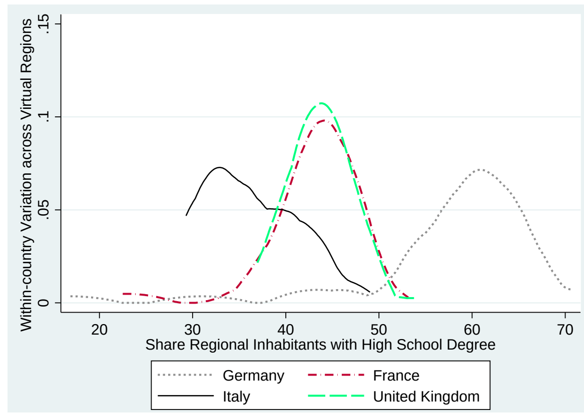


Figure A.3: Distribution of Firm-level Variables Within Countries

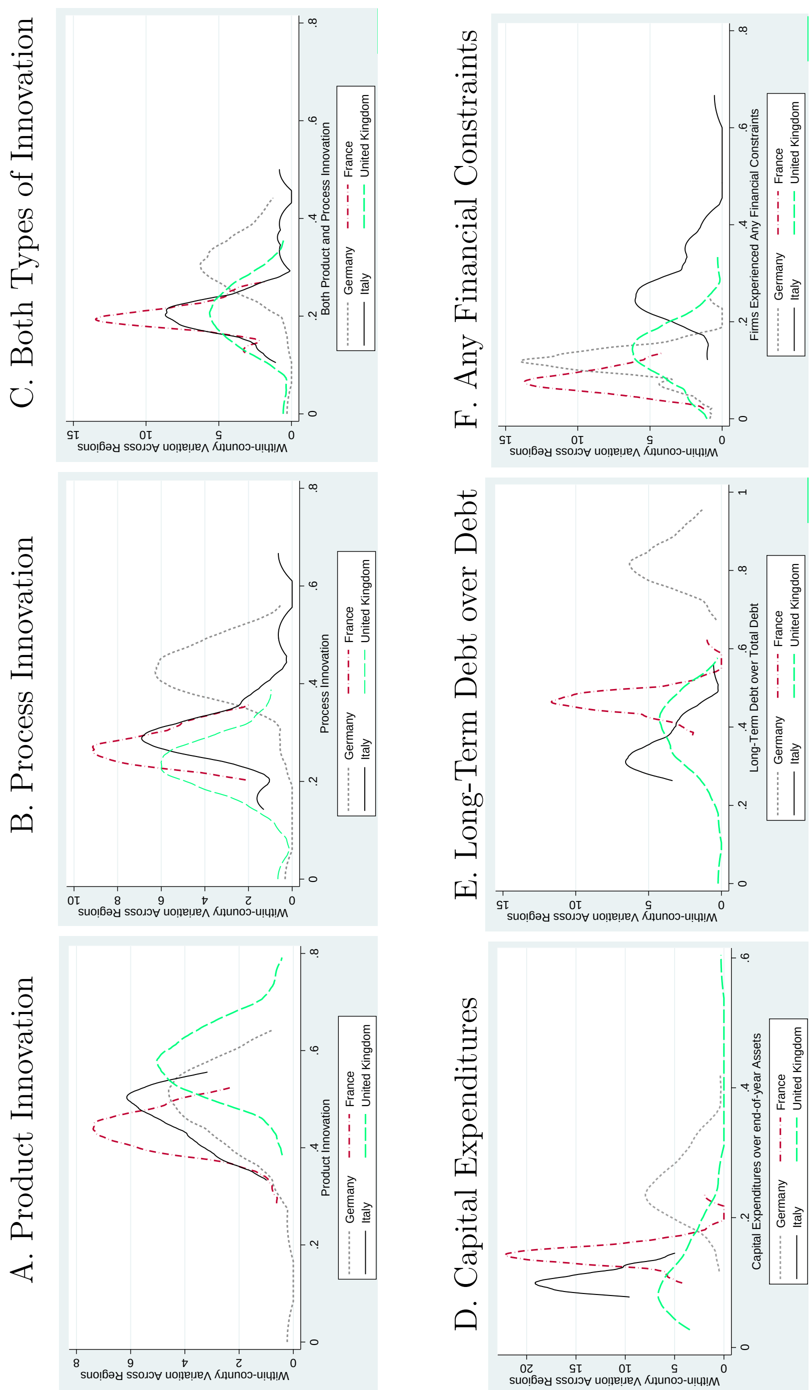
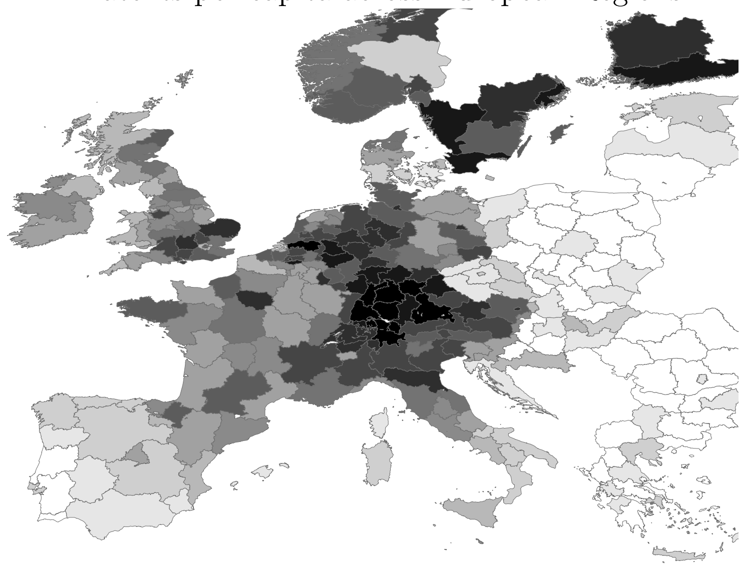
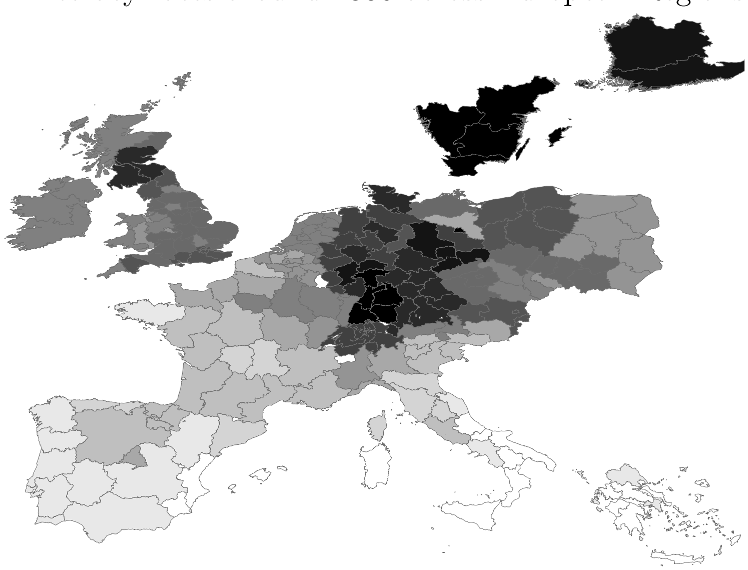


Figure A.4: Spatial Distribution of Patents per Capita and Historical Literacy

A. Patents per capita across European Regions

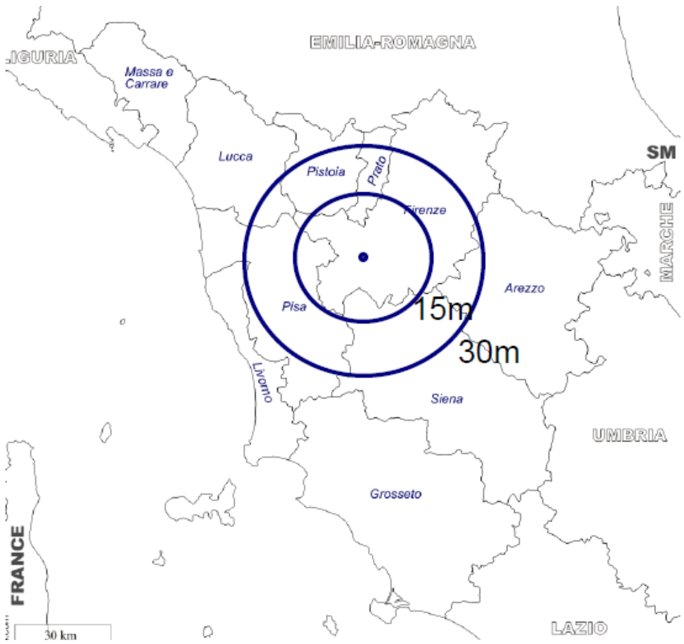


B. Literacy rates around 1880 across European Regions



Panel A plots the spatial distribution of the ratio of patents per capita for European Regions (NUTS 2) in 2005. The fractional regional patent count by applicant in priority year 2005 is from the PATSTAT-Crios database, based on the raw data of the European Patent Office database. Panel B plots the spatial distribution of literacy rates in 1880. Literacy rates are collected from Tabellini (2010), which I integrate with primary sources for the regions not covered by Tabellini (2010): for Greece, I compute the literacy rates based on the hand-collected raw population data on literate and illiterate inhabitants of Greek regions from the 1881 national census; for Germany and the regions of Poland that were part of Prussia in the 19^{th} century, I use the raw data on the literate population by historical Prussian counties from the Ifo-iPEHD project. For Belgium, I compute the historical literacy rates from the province-level data on analphabetism in Ruwet and Wellemans (1978). In both panels the darker a region is, the higher the value of the plotted variable.

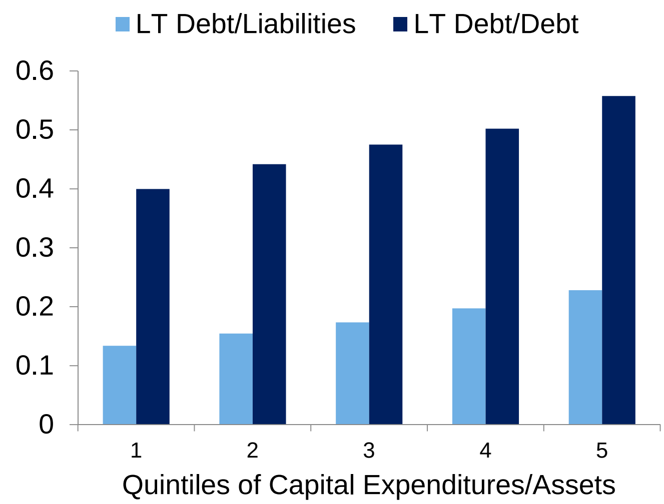
Figure A.5: Commuting across Counties within Regions



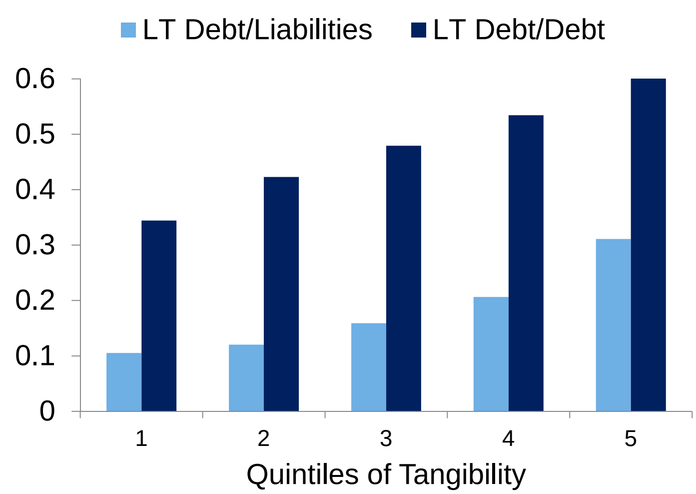
This picture illustrates the commuting times across counties of one of the 10% largest European regions, Tuscany. The blue point in the picture is the Empoli valley, in the Firenze county, an area that hosts several traditional manufacturing Small and Medium Enterprises (SMEs). Commutes as short as 15 minutes allow workers from 4 other counties to access the area: Pisa, Siena, Prato, and Pistoia. Commutes of about 30 minutes allow workers from 2 additional counties to access the area: Arezzo and Lucca.

Figure A.6: Long-Term Debt by Capital Expenditures and Tangibility

A. Long-Term Debt by quintiles of Capital Expenditures



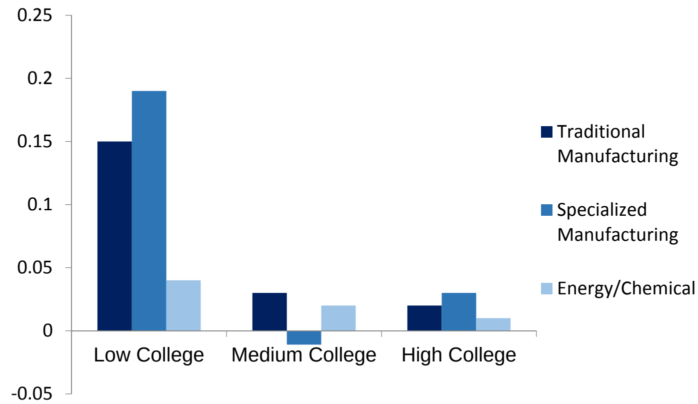
B. Long-Term Debt by quintiles of Tangibility



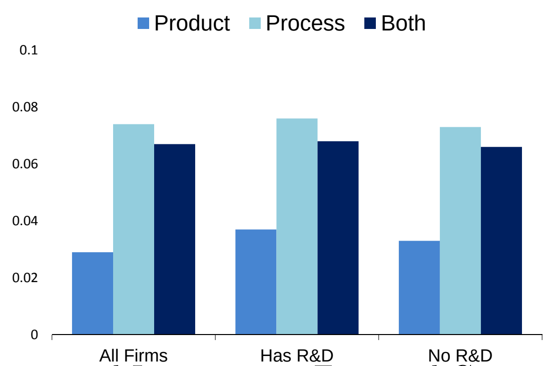
Panel A plots the average ratio of long-term debt over liabilities (light blue), and of long-term debt over total debt (dark blue) for the firms in the EU-EFIGE/Bruegel database across five equal-size groups of firms, sorted by their capital expenditures normalized by previous end-of-year total assets. Panel B plots the average ratio of long-term debt over liabilities (light blue), and of long-term debt over total debt (dark blue) for the firms in the EU-EFIGE/Bruegel database across five equal-size groups of firms, sorted by their tangibility, measured as the ratio between tangible assets and total assets.

Figure A.7: Additional Interaction Results

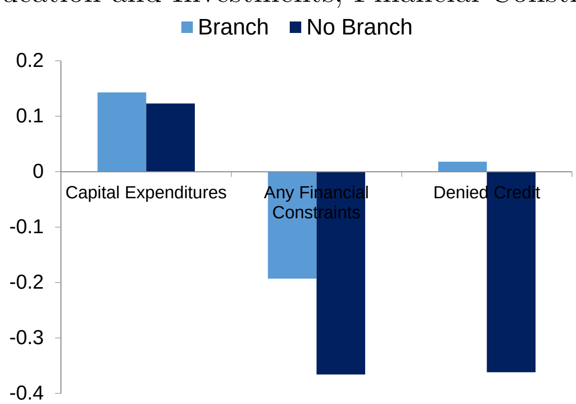
A. Basic Education and Product Innovation: Sectors and College Employees



B. Basic Education and Firm-level Innovation by R&D activities inside the firm

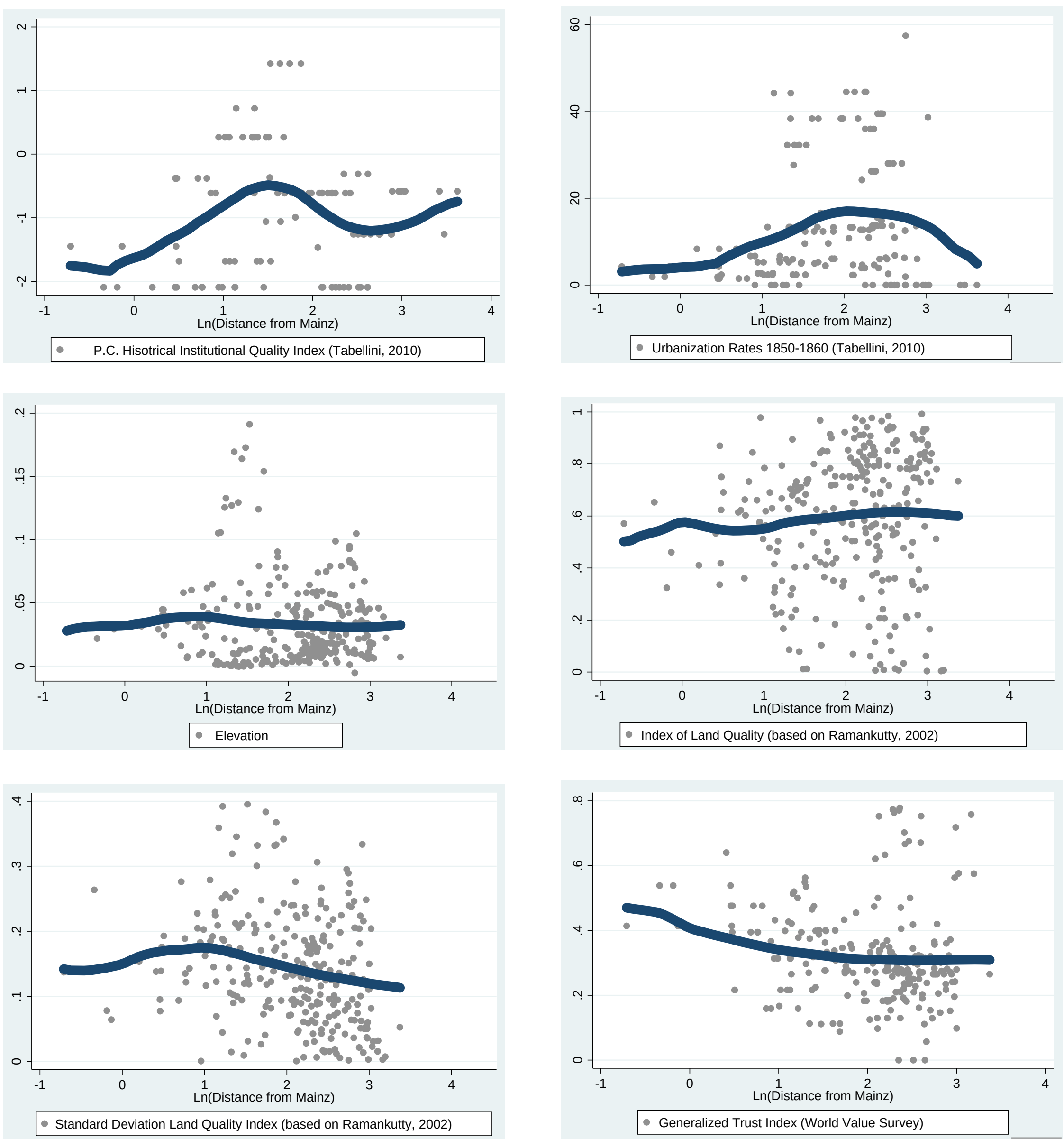


C. Basic Education and Investments, Financial Constraints by Branch Status



Panel A plots the estimated effect of regional high-school graduates on the likelihood that a firm in the EU/EFIGE-Bruegel data set engages in product innovation for firms sorted twice: across three equal-size groups based on the share of college-graduate employees in the firm, and across three Pavitt sectorial taxonomy groups based on the technological intensity of the sectors. Traditional manufacturing (dark blue) is the least technologically-intensive sector, whereas Energy/Chemical is the most technologically intensive sector. Panel B plots the effect of regional high-school graduates on the likelihood that a firm in the EU/EFIGE-Bruegel data set engages in product, process, or both types of innovation for firms sorted by whether the firm engages in R&D activities or not. Panel C plots the effect of regional high-school graduates on the capital expenditures, the likelihood of financial constraints, and the likelihood of being denied credit, based on the survey responses of firms in the EU/EFIGE-Bruegel database. The effects are estimated separately for firms that are part of a business group, or branches (light blue), and firms that are autonomous (dark blue).

Figure A.8: Distance from Mainz and other Historical, Current Outcomes



The graphs above plot the association between the dimension described in each graph and the log of the distance from Mainz across the European regions for which each dimension is observed. Dimensions include the Index of the quality of historical regional institutions from Tabellini (2010), the urbanization rate in a region in the period 1860-1880 from Tabellini (2010), the average elevation of the region, the index of quality of cultivable lands based on Ramankutty et al. (2002), the standard deviation of the indices of land quality for the cells that are included in each region by Ramankutty et al. (2002), and the index of generalized trust in a region computed as the regional average of the individual responses to the World Value Survey (Wave 9).

Figure A.9: Blue-collar Workers or Machines?

This graph reports the coefficient of the effect of historical literacy in a region as of 1880 on yearly regional patent counts from 1978 to 1996 estimated with negative binomial regressions of regional patents on the literacy rate in 1880 and a set of geographic and historical regional dimensions.

Table A.1: High School Degrees by European Country

EU Country	**ISCED 3 Terminal certification ("High School")**
Germany	Beruflicher Abschluss, Lehrabschluss, Hochschulreife, Abitur
France	CAP, BEP, BP, BAC professionnell, BAC general et technologique
Italy	Diploma istruzione professionale, Diploma istruzione secondaria superiore
Spain	Bachiller, Tecnico, Certification, Titulo Profesional
United Kingdom	GSCE, General National Vocational Qualification, GCE Advanced Level
Austria	Reifepruefung, Diplompruefung (excl last 2 grades)
Netherlands	Vocational Qualification Level 3-4, HAVO/VWO qualification
Finland	Ammattitutkinto, Lukiokoulutus
Denmark	HF-enkeltfagsprove, HHX-eksamen, HTX-eksamen, Studentereksamen
Greece	Geniko Lykio, Epagelmatiko Lykeio (EPAL), Epagelmatiki Sxoli (EPAS)
Ireland	VTOS (Core), Leaving Certificate Vocational, NFQ Levels 4-5
Bulgaria	Profesionalni programi 8 klas, Sredno obshto obrazovanieprofilirano
Czech Republic	Vysvědčení o maturitní zkoušce, výuční list
Estonia	Üldkeskharidus, Kutsekeskharidus põhihariduse basil, Kutsekeskharidusõpe
Hungary	Párhuzamos oktatás szakközépiskolában, Gimnázium 9-12
Latvia	Vispārējā izglītība, Arodizglītība, Profesionālā vidējā izglītība
Lithuania	Vidurinio ugdymo programos, Profesinio mokymo programos
Poland	Technikum, liceum, szkoła równorzędna
Romania	Invatamant liceal, Invatamant profesional
Slovakia	Výučný list, Vysvedčenie o maturitnej skúške
Slovenia	Program srednjega poklicnega izobraževanja-dualni system, Maturitetni tečaj

This Table reports the terminal degrees which are classified as level 3 (high school) by the *International Standard Classification of Education* (ISCED) of Unesco. The number of individuals who hold such degrees and certifications in a region or county as of 2005 are classified as inhabitants with a high school degree in the empirical analysis.

Table A.2: Basic Education and Additional Firm-level Results

	(1)	(2)	(3)	(4)
	Investment	Capital Structure	Financial Constraints	
	Cut Investment in 2010	LT Debt/ Liabilities	Any constraint	Denied credit
Log Pop. High School	-0.152***	0.142	-0.090	-0.030
	*0.021****	*0.016****	*0.016****	*0.008****
Log Pop. College	0.195	0.020	-0.021	-0.009
	*0.027****	*0.009***	*0.014*	*0.007*
Tangibility	0.039	0.069	0.018	0.001
	*0.014****	*0.006****	*0.004****	*0.002*
Leverage	0.001	-0.001	0.002	0.001
	0.007	*0.001*	*0.002*	*0.001*
Log Sales	-0.005	-0.017	-0.014	-0.016
	0.008	*0.005****	*0.007**	*0.005****
Family firm	0.039	0.012	0.011	0.004
	*0.019***	*0.011*	*0.022*	*0.008*
Family CEO	-0.039	0.002	-0.022	-0.008
	*0.020**	*0.011*	*0.022*	*0.010*
Part business group	0.049	-0.032	-0.052	-0.003
	*0.018****	*0.006****	*0.014****	*0.008*
Exports products	0.039	0.006	0.038	0.016
	*0.014****	*0.005*	*0.008****	*0.006****
Sector f.e.	X	X	X	X
Age f.e.	X	X	X	X
Size f.e.	X	X	X	X
Regional char.	X	X	X	X
Country f.e.	X	X	X	X
Observations	8538	10045	10140	10140
N. of clusters	141	141	141	141
(Adj, Pseudo) R^2	0.036	0.234	0.046	0.059

This Table refers to the firms participating in the EU/EFIGE-Bruegel-Unicredit survey, which was run from 2008 to 2010. The table reports results from estimating OLS regressions whose outcomes are a dummy for whether firms declared they had to cut their planned investment expenditures in 2010 due to the financial crisis (column (1)), the ratio between the long-term debt and the liabilities of firms (column (2)), and dummies equal to one if firms declared they faced financial constraints in the two years prior to the interview (column (3)), and if they received less credit than required and needed (column (4)). In all regression, the set of Regional controls include the log of inhabitants without degrees, log of latitude, log of area, log of population density, and the log of the index of the quality of cultivable land from Ramankutty et al. (2002); the firm-level controls include Tangibility (the ratio of tangible assets to total assets), Leverage (the ratio of total debt to shareholders' equity), the log of sales, and a set of dummy variables that equal one if the firm is owned by a family or entrepreneur, if the firm's CEO is a member of the owning family, if the firm is part of a business group, and if the firm engages in exporting of its products outside the country where it operated. Standard errors are clustered at the level of regions (NUTS 2). Statistical significance is shown as follows: ***1%, **5%, *10%.

Table A.3: Historical Literacy and Current Regional and Firm-level Outcomes

A. Historical Literacy and Current Regional Innovation

	(1)	(2)	(3)	(4)
	All Patents		HighTech	Mid/Low Tech
Ln(Literacy 1880)	0.397***	0.232***	0.093	0.251***
	(0.080)	*(0.073)*	*(0.075)*	*(0.077)*
Country f.e.		X	X	X
Regional controls		X	X	X
Observations	228	227	222	222
N. of clusters	79	78	78	78
(Adjusted) R^2	0.184	0.888	0.832	0.864

B. Historical Literacy and Current Firm-level Outcomes

	(1)	(2)	(3)	(4)	(5)
		Innovation		Investment	Capital Structure
	Product	Process	Both	CapX	LT Debt/ Total Debt
Ln(Literacy 1880)	0.004	0.014*	0.017***	0.068***	0.056***
	(0.007)	*(0.007)*	*(0.006)*	*(0.017)*	*(0.008)*
Sector f.e.	X	X	X	X	X
Age f.e.	X	X	X	X	X
Size f.e.	X	X	X	X	X
Regional controls	X	X	X	X	X
Country f.e.	X	X	X	X	X
Observations	10137	10137	10137	10129	8080
N. of clusters	140	140	140	140	139
(Adj, Pseudo) R^2	0.066	0.023	0.028	0.118	0.246

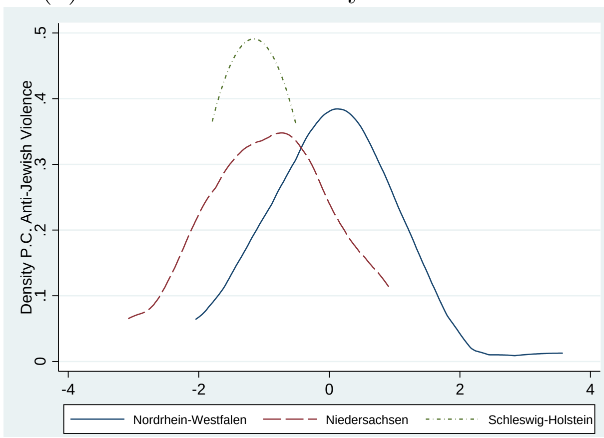
Panel A reports results for estimating OLS regressions of the log of regional patents on the log of regional literacy rates around 1880. Regional controls include the log of inhabitants with college degrees, the log of inhabitants without degrees, log of latitude, log of area, log of population density in 2005, the log of the index of the quality of cultivable land from Ramankutty et al. (2002), and a dummy variable equal to 1 if the region was subject to a communist regime after the Second World War. In all columns of both Panels, the variation within countries. Standard errors are clustered at the level of groups of regions (NUTS 1). Panel B reports results for estimating firm-level specifications, where the log of regional high-school graduates is replaced with the log of the regional literacy rate around 1880. Standard errors are clustered at the level of regions (NUTS 2). In both Panels, statistical significance is shown as follows: ***1%, **5%, *10%.

Appendix B

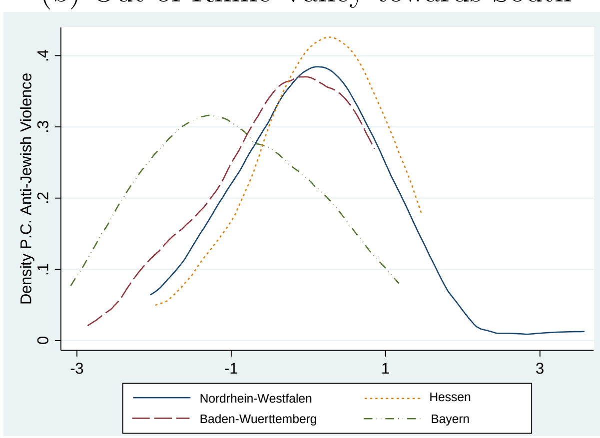
Appendix - Jewish Persecution

Figure B.1: Densities of County-level Persecution across German States

(a) Out of Rhine Valley towards North



(b) Out of Rhine Valley towards South



(c) Farthest Distance from Rhine Valley in Both Directions

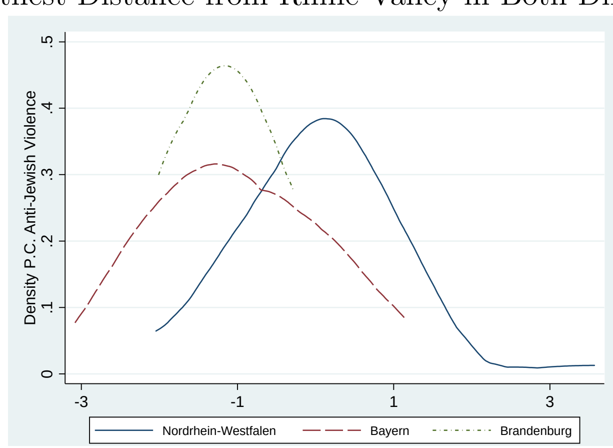


Figure B.2: Raw Data - Correlations

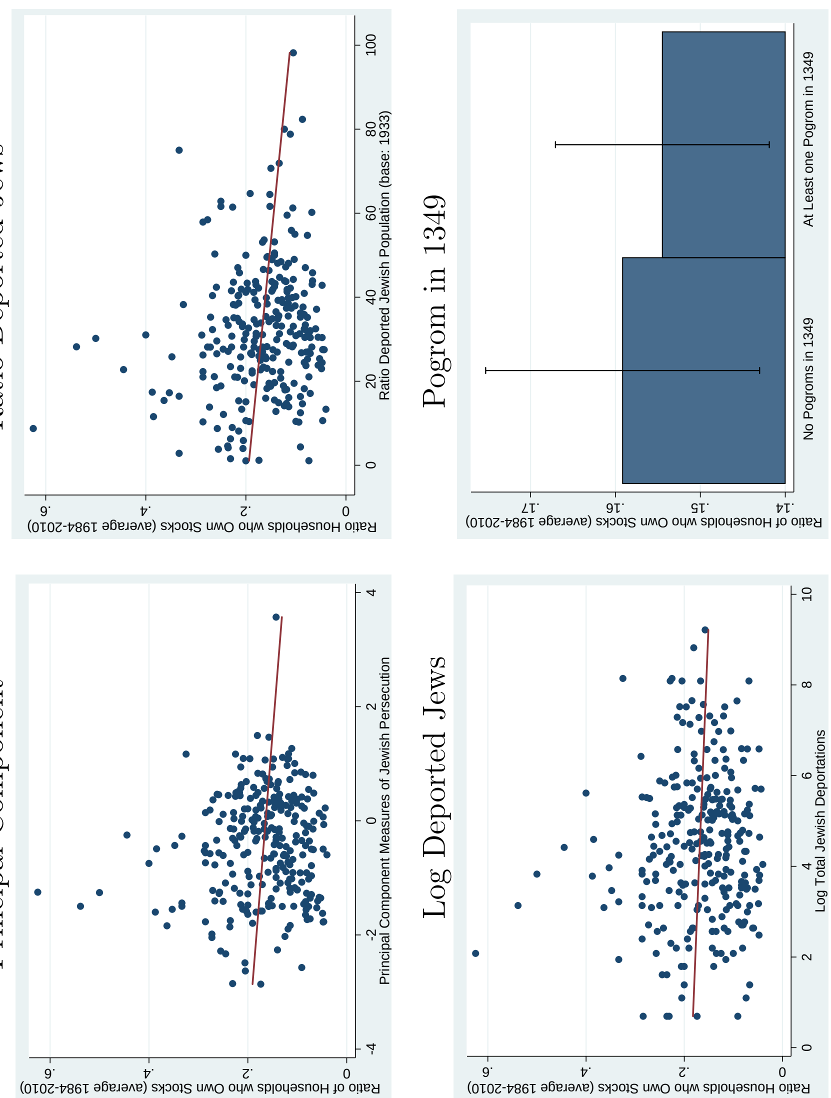


Figure B.3: Raw Data - Stockholdings, Principal Component by Pogrom in 1349

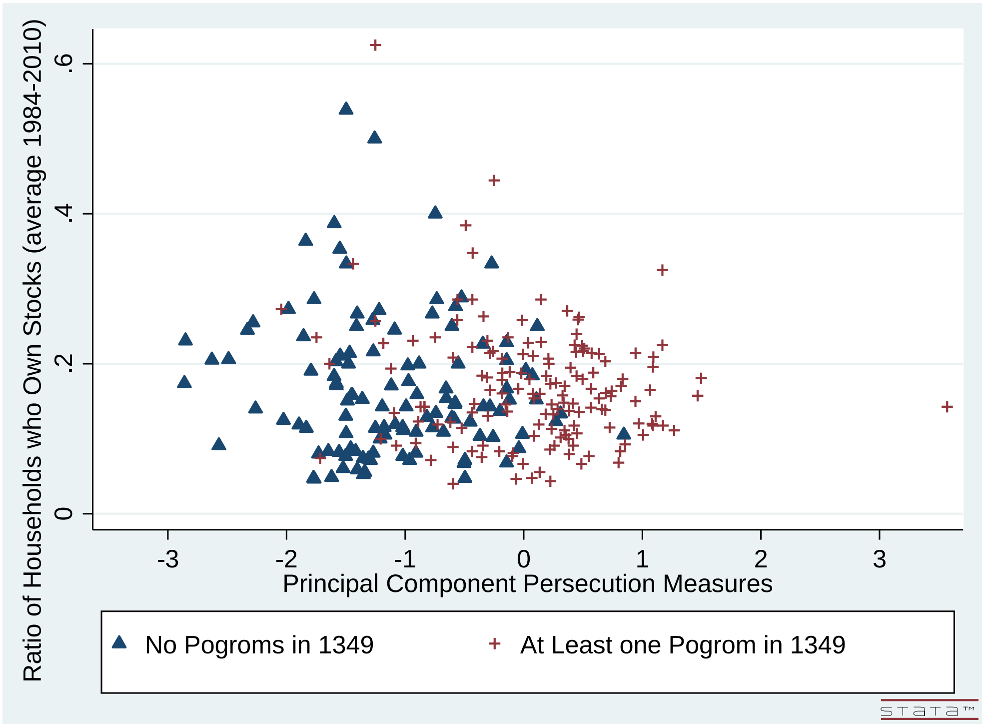
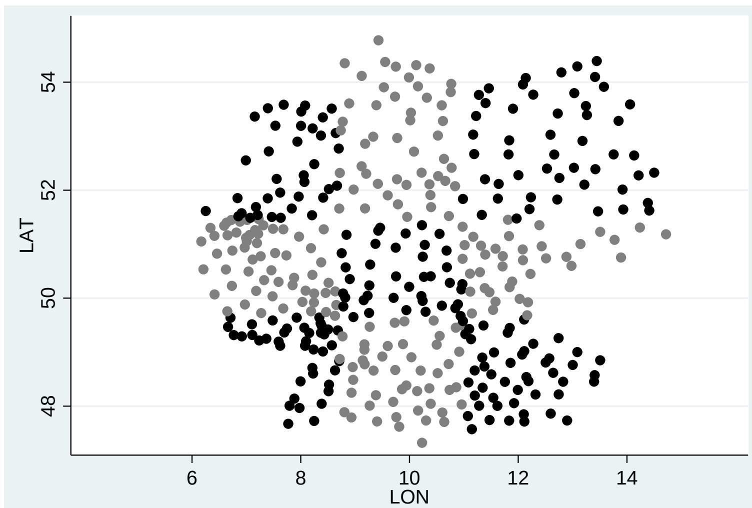


Figure B.4: Virtual States

Virtual States 1: 9 arbitrary partitions



Virtual States 2: 16 arbitrary partitions

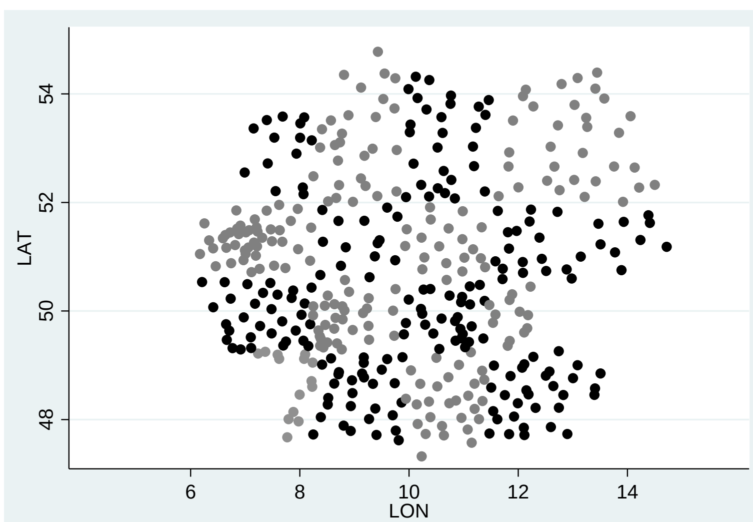
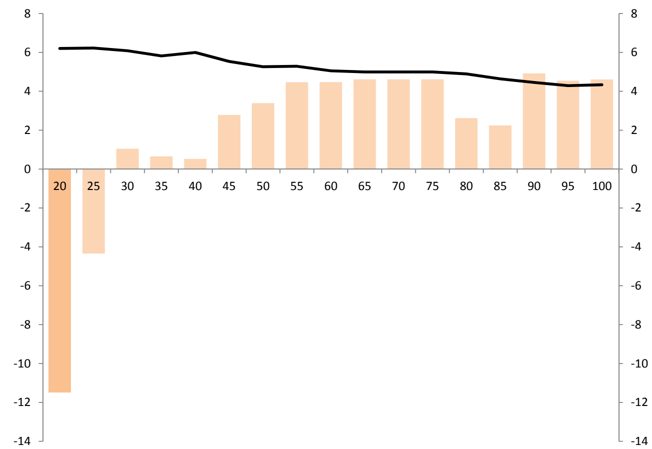
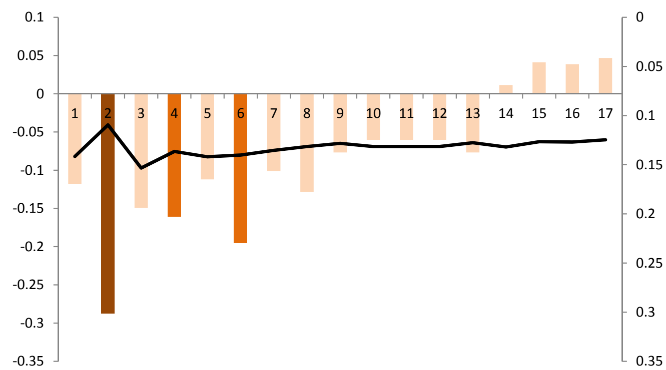


Figure B.5: Economic Crisis and Alternative Outcomes

(a) Households' Head Age



(b) Home Owner



(c) High School Degree or Higher Education

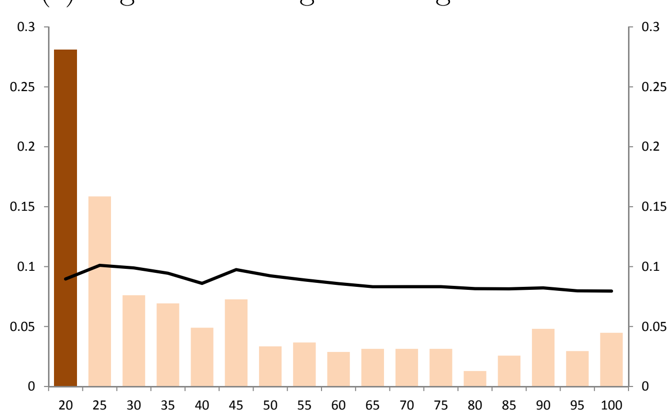
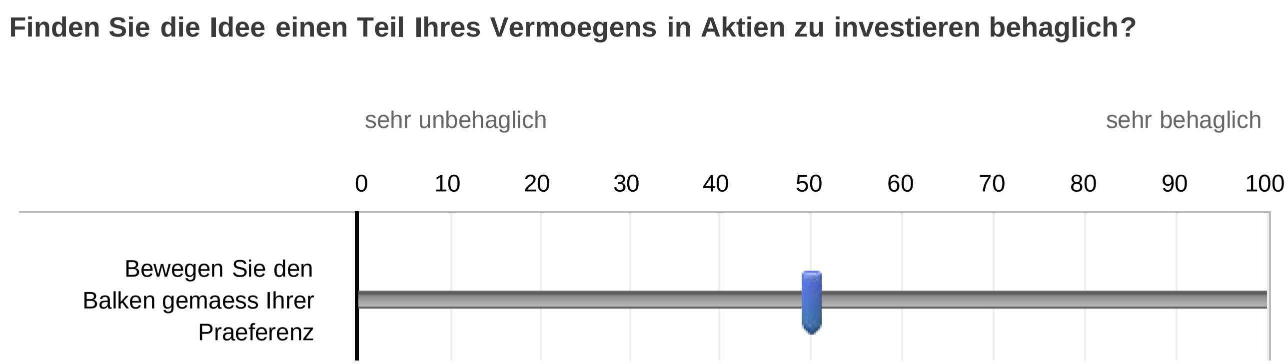
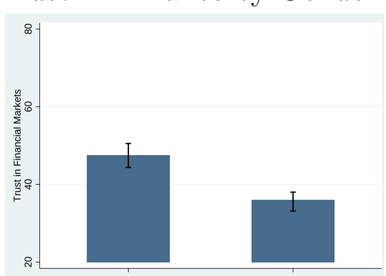


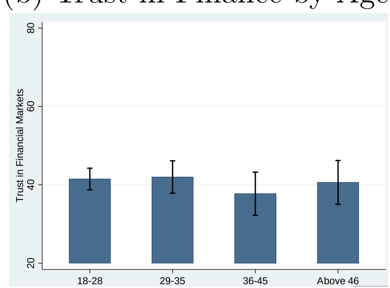
Figure B.6: Trust in Finance: Survey and Demographics



(a) Trust in Finance by Gender



(b) Trust in Finance by Age



(c) Trust in Finance by Education

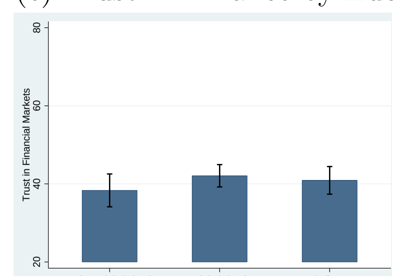
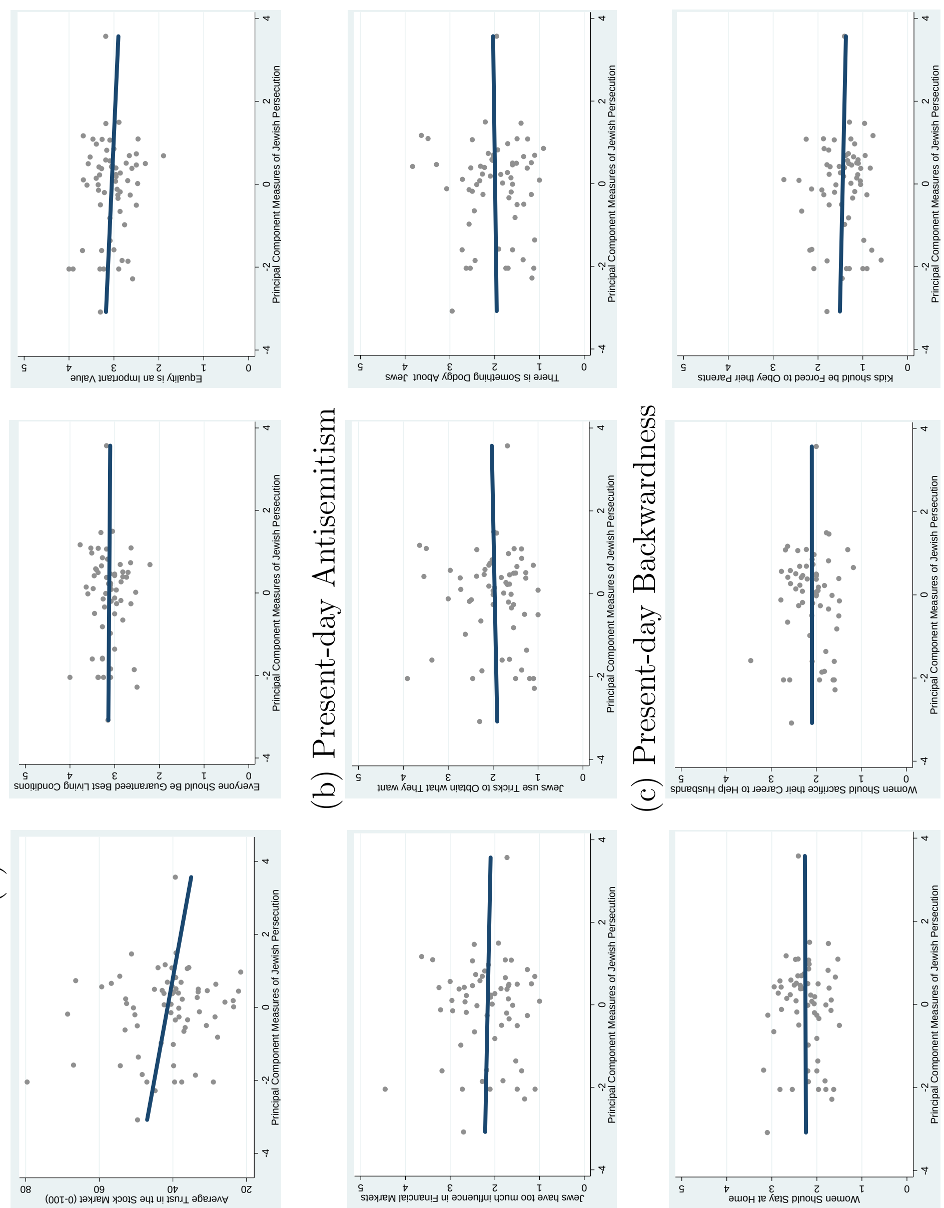


Figure B.7: Distrust in Finance, Antisemitism, Backwardness, and Jewish Persecution



(a) Distrust in Finance and Neutral Values

(b) Present-day Antisemitism

(c) Present-day Backwardness

Table B.1: Residuals from Structural Equations and Purged Distance

	(1)	(2)	(3)	(4)	(5)	(6)
	Second stage residuals: Persecution of Jews on Existence 1096			Third stage residuals: % Hold stocks on Persecution of Jews		
	Dist. Trier	Dist. Emmerich	Minimal dist.	Dist. Trier	Dist. Emmerich	Minimal dist.
P.C. Persecution Jews	-0.053	-0.179	-0.092	-0.001	0.006	0.001
	0.115	*0.103**	*0.099*	*0.009*	*0.007*	*0.007*
Extensive Margin	-0.007	-0.067	-0.046	-0.002	0.009	0.001
	0.115	*0.106*	*0.093*	*0.009*	*0.007*	*0.007*
Intensive Margin	-1.946	-3.620	-2.463	-0.002	0.006	0.001
	2.204	*2.759*	*2.140*	*0.009*	*0.007*	*0.007*
Pogrom 1349	0.011	-0.060	-0.005	-0.001	0.010	0.002
	0.047	*0.045*	*0.041*	*0.009*	*0.007*	*0.007*
Historical contr.	X	X	X	X	X	X
Regional controls	X	X	X	X	X	X

This Table documents the association between the measures of distance from the Rhine Valley and the residuals from the second- and third-stage structural equations of the 3SLS system. The measures of distance are the estimated residuals from the following OLS specification,

$$DistanceRhine_k = \alpha + K_k' \times \delta + \epsilon_k,$$

where $DistanceRhine$ is one of three measures indicated above each column, and K is a set of observables capturing the economic characteristics of counties in the past. Columns (1)-(3) regress the residuals from the following structural equation on the distance measures,

$$PersecutionJews_k = \alpha_2 + \beta_2 \times Exist1096_k + K_k' \times \delta_2 + \epsilon_{2k},$$

where $PersecutionJews$ is the measure of persecution enlisted in the left column for each line, and $Exist1096$ is an indicator for whether a Jewish community existed in a county as of 1096. Columns (4)-(6) regress the residuals from the following structural equation on the distance measures,

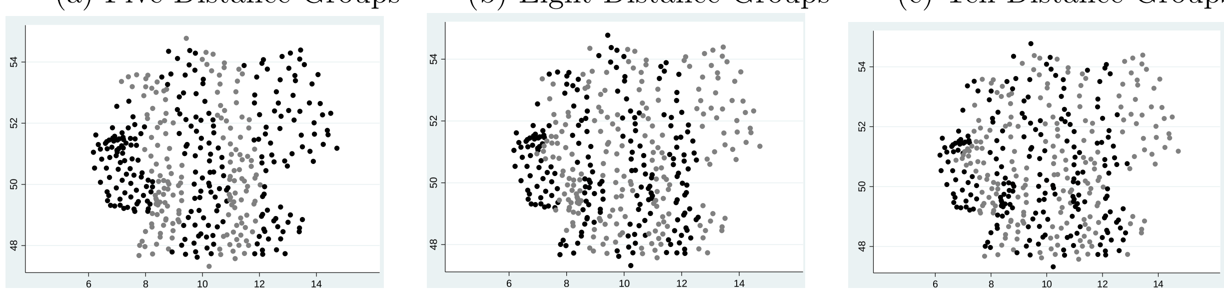
$$RatioHoldStocks_k = \alpha_3 + \beta_3 \times PersecutionJews_k + K_k' \times \delta_3 + \epsilon_{3k},$$

Hubert-White s.e. are reported below each coefficient. Significance is reported as follows: *10%, **5%, ***1%.

Table B.2: Keeping the Distance from the Rhine Valley Constant: Within-distance-group analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Principal Component Measures of Persecution			Log deported Jews *(extensive margin)*			Ratio deported Jews *(intensive margin)*			Pogroms 1349		
Persecution of Jews	-0.011	-0.010	-0.010	-0.006	-0.006	-0.005	-0.027	-0.025	-0.021	-0.018	-0.015	-0.018
(see columns heading)	*0.004****	*0.004***	*0.004***	*0.002****	*0.002****	*0.002****	*0.017*	*0.017*	*0.016*	*0.009***	*0.009**	*0.009***
Distance group f.e.	X	X	X	X	X	X	X	X	X	X	X	X
Number distance groups	5	8	10	5	8	10	5	8	10	5	8	10
Avg. st.dev.distance within groups (Km)	135	127	120	135	127	120	137	127	120	137	127	120
Average n. counties across groups	52	41	31	54	42	33	52	41	31	54	42	33
Individual, historical controls	X	X	X	X	X	X	X	X	X	X	X	X
Wave groups f.e.	X	X	X	X	X	X	X	X	X	X	X	X
Observations	13,599	13,599	13,599	13,870	13,870	13,870	13,599	13,599	13,599	13,870	13,870	13,870
N. of clusters	261	261	261	270	270	270	261	261	261	270	270	270
(Pseudo-) R^2	0.09	0.09	0.10	0.10	0.09	0.09	0.09	0.09	0.10	0.09	0.09	0.10

(a) Five Distance Groups (b) Eight Distance Groups (c) Ten Distance Groups



This Table reports the average marginal effects computed after estimating the following probit specification,

$$Pr(HoldsStocks_{ik}|X_{ik}, K_{ik}) = \Phi(\alpha + \beta \times Persecution_k + X'_{ik} \times \gamma + K'_{ik} \times \delta + \eta_t + \eta_s + \epsilon_{ik}).$$

Each observation is a German household interviewed by the SOEP for the first time between 1984 and 2010. In all columns, the dependent variable is a dummy that equals 1 if the household holds stocks. *Persecution* is the measure of Jewish persecution described at the top of each column. η_s are a set of 5, 8, or 10 groups of counties at similar distances from the Rhine Valley. S.e. are clustered at the county level. Statistical significance is reported as follows: *10%, **5%, ***1%.