



DEPARTMENT OF ECONOMICS

Working Paper Series

Civic capital and the size distribution of plants: Short-run dynamics and long-run equilibrium

Mathias Burkert
OECD, Paris
G. Alfredo Minerva
University of Bologna

March 28, 2012

Paper # 12-3

We characterize how the size distribution of plants, within narrowly defined industries, changed in Italy over a ten-year time span, and relate this to the stock of civic capital at the provincial level. Data on plant size come from the 1991 and 2001 Italian censuses. Civic capital turns out to have a positive effect on both the average and standard deviation of size. Looking at several precise points of the plant size distribution, we find that it shifts toward the right and becomes more dispersed where civic capital is high. The potential endogeneity of current civic capital is addressed by instrumenting it with historical variables. Our main conclusion is that the geographic variation in the stock of civic capital poses substantial constraints on plants' ability to expand. Understanding this is the key for the implementation of effective industrial policies.

Department of Economics
One Shields Avenue
Davis, CA 95616
(530)752-0741

http://www.econ.ucdavis.edu/working_search.cfm

Civic capital and the size distribution of plants: Short-run dynamics and long-run equilibrium *

Matthias Bürker[†] and G. Alfredo Minerva[‡]

March 21, 2012

Abstract

We characterize how the size distribution of plants, within narrowly defined industries, changed in Italy over a ten-year time span, and relate this to the stock of civic capital at the provincial level. Data on plant size come from the 1991 and 2001 Italian censuses. Civic capital turns out to have a positive effect on both the average and standard deviation of size. Looking at several precise points of the plant size distribution, we find that it shifts toward the right and becomes more dispersed where civic capital is high. The potential endogeneity of current civic capital is addressed by instrumenting it with historical variables. Our main conclusion is that the geographic variation in the stock of civic capital poses substantial constraints on plants' ability to expand. Understanding this is the key for the implementation of effective industrial policies.

JEL Classification: A13, D23, L20, R12.

Keywords: Civic capital; Cooperation; Opportunism; Plant size distribution; Trust.

*This paper is a substantially revised version of working paper n. 755/2011, "Explaining the size distribution of plants: An approach based on civic capital", Department of Economics, University of Bologna. We are grateful to Paolo Buonanno, Robert Putnam, and Paolo Vanin for having provided us some of the data. We thank for their comments Erich Battistin, Guido de Blasio, Stefano Federico, Roberto Golinelli, Saul Lach, Sauro Mocetti, Giovanni Peri, Fabiano Schivardi, and participants to the 3rd GLUNLAB workshop in Rimini and to seminars at the University of Bologna and University of Paris 1 Panthéon-Sorbonne. Part of this research was carried out while Minerva was visiting the Department of Economics of the University of California, Davis, whose hospitality is gratefully acknowledged.

[†]Centre for Entrepreneurship, SMEs and Local Development, OECD, Paris. E-mail: matthias.buerker@oecd.org

[‡]Corresponding author. Department of Economics, University of Bologna. E-mail: ga.minerva@unibo.it

1 Introduction

There is by now consensus among economists that civic capital is conducive to economic development. Its underlying values and beliefs reduce transaction costs, facilitate the anonymous exchange among private agents on the marketplace and increase returns to investment. Moreover, countries endowed with high stock of civic capital better protect property rights, enforce contracts more efficiently and display less corrupt governments. Recent studies (Algan and Cahuc, 2010; Tabellini, 2010) have shown that civic capital has a causal impact on both the level and the growth rate of GDP.¹

Despite this well-established positive relationship between civic capital and growth at the aggregate level, empirical evidence on the specific channels through which civic capital operates is still limited. The aim of this paper is to identify a channel through which civic capital affects economic development. We achieve this by analyzing the relationship between the stock of civic capital and the dynamics of the size distribution of plants. Specifically, employing Italian census data reveals that civic capital spurs larger plant size at a very detailed level of spatial and sectoral disaggregation. As emphasized by Bloom et al. (2012), it is well known that, in order to drive inefficient firms out of the market, productive firms have to increase in size and expand their market share. To the extent that the stock of civic capital facilitates or hinders this process, the dynamics of aggregate productivity turns out to be boosted or impaired. Since industry growth is a genuine source of economic growth, our analysis uncovers a specific mechanism through which civic capital enhances aggregate economic development.

A novel issue that we address in this paper concerns the increase in the dispersion of the size distribution in areas where civic capital is high. To the best of our knowledge we are the first to investigate this aspect. One of the implications of a recent literature on misallocation and productivity (see, for example, Guner et al., 2008; Hsieh and Klenow, 2009) is that decreasing distortions in the use of factors of production shall lead to an increase in the dispersion of the distribution of plant size.² Civic capital tends to reduce transaction costs and distortions affecting labor employment at the plant-level, so our results confirm these findings. We will discuss in detail the sources of labor inefficiencies related to the lack of civic capital below.

All in all, the broad motivation for our paper is that learning the determinants of the size distribution is important since it can inform us on the constraints faced by plants in particular environments, and, implicitly, about the determinants of aggregate productivity.

We study the effect of civic capital on plant size in the Italian context because it bears several advantages. First, thanks to several studies that analyze civic capital in Italy, we can fall back on a range of well established proxies available for each of the 103 Italian provinces we work with.³ Specifically, we proxy the stock of provincial civic capital by referenda turnout, blood donation and volunteering, with the latter two being standardized by population.⁴ In section 2, we motivate these proxies and outline their relationship with generalized trust. Second, analyzing plant size growth within one country allows to exclude confounding factors such as differences in tax regimes, subsidies or trade policies, as they are fixed at the national level.

¹Throughout the paper, we avoid using the loaded term social GDP. Rather, we follow Guiso et al. (2011) and use the notion of civic capital. See section 2 for a definition and its relationship with trust.

²See Table 6 in Guner et al. (2008) and Figure III in Hsieh and Klenow (2009).

³See for example Banfield (1958).

⁴See Putnam (1993) for a conceptual underpinning of these proxies. Other empirical studies which use these proxies include Buonanno et al. (2009), de Blasio and Nuzzo (2010), Guiso et al. (2004), Guiso et al. (2008a).

Third, we can use data from the 7th and 8th Censuses of economic activity (carried out in 1991 and 2001, respectively) which provide information on the 3.6 millions plants operating in 1991 and on the 4.4 millions plants operating in 2001. Hence, our analysis produces results which are valid for all sectors of economic activity and that do not suffer from sample selection issues, given that we work with the universe of Italian establishments.

In order to give to our coefficients a causal interpretation, we have to address several empirical challenges. The first concerns measurement error in the stock of civic capital. Similarly to Tabellini (2010) we tackle this issue by extracting the first principal component of the three above mentioned proxies. Second, to ensure that our results are not driven by the pronounced geographic inequalities in socio-economic development existing in Italy, we include a dummy for each of the twenty NUTS-2 regions. Similarly, a set of 4-digit NACE fixed effects guarantees that we are looking at changes in the size distribution within narrowly defined industries, thus avoiding the risk of attributing to civic capital what is indeed the effect of specialization of certain provinces in certain industries. Moreover, a set of well-considered provincial controls hold the effect of several confounding factors constant. These controls include financial development, a measure of formal contract enforcement, the stock of human capital, final demand and GDP as well as the urbanization rate in 1861. Finally, we have to exclude reversed causation. The so-called modernization theory emphasizes that economic development impacts on the formation of cultural and societal norms and values (Inglehart and Baker, 2000). Accordingly, high stocks of civic capital might be the result of plant growth, or, more generally, economic development. In order to invalidate this claim, we provide evidence from two-stage least squares regressions. We use historical information to identify our instruments. The first set of instruments dates back to the middle of the 19th and early 20th century, and are basically deeply lagged values of civic capital. The second relies on Putnam (1993) and Guiso et al. (2008a): we go further back in time, and employ the number of free city states per province that existed in 1300.

Our results provide a detailed picture on how civic capital shifts the plant size distribution over time. We employ a dynamic specification (an error-correction model) under the assumption that growth (the short-run dynamics) in the dependent variables is the outcome of disequilibrium from the long-run relationship and of other shocks. The econometric model allows to retrieve the underlying parameters of the long-run relationship, that are confronted with the predictions of our conceptual framework. We find that civic capital increases the growth rate of average plant size, implying that the stock of civic capital plays a crucial role for industry growth. We then estimate the long-run parameters, according to which civic capital exerts a sizeable positive impact on both the first and the second moment of the plant distribution. Checking the effect at various, precise points of the distribution we find that civic capital is important for growth of both small and large plants. Comparing the magnitude of the effect, however, reveals that growth is more pronounced for large firms, a result which is consistent with the increase in dispersion that we document in high civic capital areas. We attribute our findings to the fact that civic capital is associated with a reduction in opportunistic behavior of agents in the local area, which in turn leads to more cooperation in intra-firm relationships, and hampers the incidence of principal-agent problems. We also explain why the effect should be more pronounced in large plants.

The rest of the paper is organized as follows. Below, we state exactly the contribution of our paper with respect to the existing literature. In section 2, we provide a definition of civic capital and outline how it

relates to opportunistic behavior and trust. In section 3, we present the data and describe the construction of the variables. Section 4 presents the baseline OLS analysis. Sections 6 and 7 provide robustness analysis and two-stage least square results, respectively. Finally, section 8 concludes.

1.1 Contribution to the literature

Several studies have demonstrated that trust is associated with improved economic performance at the aggregate level (Knack and Keefer, 1997; Zak and Knack, 2001; Dearmon and Grier, 2009). Recently, Algan and Cahuc (2010) and Tabellini (2010) proved that the link between civic capital and both the level and the growth rate of GDP is causal. Based mostly on cross-country data, these studies have established civic capital as an important determinant of economic development. However, they remain silent on the specific mechanisms at the micro-level by which civic capital enhances economic development.

Another strand of the literature has emphasized the role of trust for the size of firms. Fukuyama (1995) was among the first to stress the role of trust in sustaining large scale organizations. Employing survey and accounting data on manufacturing businesses for a cross-section of countries in 2006, Bloom et al. (2012) find a positive effect of trust on two aspects of business organization: the plant-level degree of decentralization, and average firm size. As far as the relationship between trust and size is concerned, we think that there are important differences with our analysis that are worth highlighting. First, provided that they work with databases such as Bureau van Dijk's Amadeus, they have a sample which is representative of firms with at least 50 employees. Looking at 2001 Italian Census data of firms for the whole spectrum of economic activity (we do not limit our analysis to manufacturing), we find that, out of more than 15 millions employees in private organizations, 32.7% of them were employed in firms with at least 50 employees, while 67.3 % were employed in firms with less than 50 employees. Then, in a country like Italy, 2/3 of employment is concentrated in organizations below the 50 employees threshold, and it is crucial to assess the relationship between trust and size for the entire span of the size distribution. A second issue that deserves attention is that their dependent variable in the trust-size regression is the average size of manufacturing firms at the regional level. In this manner, their analysis is informative of the fact that manufacturing firms in high-trust regions in 2006 were on average larger than manufacturing firms in low-trust regions. There are many possible sources of endogeneity affecting this relationship. One occurs in the case high-trust regions specialize (for reasons not related to trust) in those manufacturing industries where average firm size is large. We believe it is much more transparent to look at the variation in the size distribution within (possibly narrowly-defined) industries, instead of manufacturing as a whole. As a third point, we want to mention that our identification strategy does not rely on cross-sectional comparisons, but instead is based on a dynamic equation of the size distribution over the 10-year period covered by two consecutive censuses (1991 and 2001). We think this is an important contribution since cross-sectional analyses may suffer from the possibility that findings are the result of a particular shock in a particular year, rather than being the outcome of a pattern consistent over decades. We have used the last two available censuses because in earlier ones data are not available at the 4-digit NACE level. Finally, the last difference consists in our analysis of the dispersion of the size distribution, something which is not examined in their paper.

Another contribution of our paper stems from the fact that the positive relationship between civic

capital and the average size of organizations that we document is not granted a priori. The transaction costs literature (Williamson, 1979, 1985) stresses that, when faced with the make-or-buy decision, firms choose the form of exchange that minimizes transaction costs. If civic capital decreases opportunistic behavior, contract incompleteness should decrease as well. Consequently, as market transactions become more attractive relative to in-house production, some activities are outsourced to external suppliers, and the average size of plants could decrease.⁵ Moreover, the literature on institutional determinants of vertical integration (a recent paper is Acemoglu et al., 2009) points out the fact that, for various reasons, the theoretical relationship between contractual institutions and vertical integration is a rather subtle one. Again, to the extent that the size of organizations is increasing in the degree of vertical integration, this literature delivers somewhat ambiguous predictions also for size. Our estimates that unequivocally reveal an overall positive effect of civic capital on the average and the dispersion of the size distribution, and our focus on civic capital and trust instead of formal contractual institutions complement this literature.

2 Conceptual framework

2.1 The concept of civic capital

Guiso et al. (2011) define civic capital as “those persistent and shared beliefs and values that help a group overcome the free rider problem in the pursuit of socially valuable activities”.⁶ Two immediate and inter-related predictions can be derived from this definition. First, civic capital is associated with a reduction in opportunism. This is an immediate implication from the definition, as free-riding in collective endeavors is a genuine form of opportunistic behavior. Second, areas with more civic capital are expected to have higher levels of trust. Opportunistic behavior, defined as “self-interest seeking which guile” includes “calculated efforts to mislead, distort, disguise, obfuscate, or otherwise confuse” (Williamson, 1985, p. 47). It is therefore reasonable to assume that in an area in which opportunistic behavior is widespread, individuals form adverse beliefs about the trustworthiness of others, resulting in a distrusting environment. The literature has stressed that the decision to trust another group or person depends on the belief in the trustworthiness of others and on individual preferences (Fehr, 2009; Sapienza et al., 2007). To the extent that civic capital is associated with widespread pro-social preferences, this is another channel conducive to high trust (Fehr, 2009). Given the interrelatedness between civic capital and trust, in what follows we will use the two terms interchangeably.⁷

⁵In a study (Bürker and Minerva, 2010) related to this paper it is shown that civic capital favors outsourcing of business services. Specifically, relative to overall sales, firms acquire more services on the market in areas where civic capital is high.

⁶This deviation from strict self-interest can be explained by pro-social attitudes such as altruism (Andreoni and Miller, 2002), inequity aversion (Fehr and Schmidt, 1999), preferences for reciprocal fairness (Falk and Fischbacher, 2006) or more generally by strongly internalized values which constitute a moral obligation not to defect (Portes, 1998). An alternative mechanism stresses the social component of civic capital. Specifically, individuals cooperate in order not to be socially ostracized by their fellows (Portes, 1998).

⁷It is important to stress that in the agency literature trust is usually modelled as an endogenous outcome of firm strategies. The enhancement of trust can be obtained through motivational schemes (Casadesus-Masanell, 2004) or by facilitating social interactions outside the workplace during sport activities, holidays, etc. (Spagnolo, 1999). We abstain from all these aspects and consider trust as exogenously determined by civic capital. Empirical evidence corroborates the validity of this assumption. Employing data on a large Italian bank, Ichino and Maggi (2000) demonstrate that individual background is a chief determinant

2.2 Civic capital, trust and the size distribution of organizations

The reasoning above suggests that civic capital improves organizational efficiency of plants by facilitating cooperation and reducing the scope of agency problems and dilemmas of collective actions, which are centered around opportunistic behavior. Arrow (1968) has made this point plainly clear by stating that “one of the characteristics of a successful economic system is that the relations of trust and confidence between principal and agent are sufficiently strong so that the agent will not cheat even though it may be ‘rational economic behavior’ to do so”.⁸ We now outline some specific mechanisms of how civic capital operates. First, we look at the decision of the principal to delegate decision rights to the agent. Next, we consider team production, while the third mechanism by which civic capital impacts on plant size relies on the fact that, as organizations get larger, cooperation in one-shot transactions becomes increasingly more important.

2.2.1 Trust and decentralization

The first mechanism of how civic capital affects establishment size works through the delegation of decision authority by CEOs to subordinate managers. In theoretical models of allocation of decision rights, trust makes the objective functions of the principal and the agent more similar. The rationale behind is that a reduction in opportunistic behavior implies that the agent is less likely to take actions which merely serve his private goals rather than the benefit of the organization as a whole. In Aghion and Tirole (1997) these congruence effects increase the agent’s willingness to pay for authority relative to the principal which leads to more decentralization. Alternatively, in Dessein (2002), higher similarity in objective functions reduces the incurred loss of control by the principal and improves communication. Under standard assumptions on the uncertainty of the environment, the decrease in the loss of control outweighs the communication effect which leads to more decentralization. Delegation of tasks and decision rights, in turn, is crucial for the growth of the size of plants. Penrose (1959) stressed that CEOs face resource constraints in terms of time and cognitive abilities if they have a wide span of control. This acts as a stumbling block on establishment growth, as the CEO has to invest costly resources such as time and effort to manage and decide upon complex business operations. Delegation then reduces the overload of CEOs and frees resources necessary for expansion.⁹

2.2.2 Trust and shirking in team work

When individual contributions in team production are unobservable and the team members’ remuneration is linked to overall team output, individuals have an incentive to free-ride on their colleagues (Alchian and Demsetz, 1972; Holmstrom, 1982). As suggested by its definition, civic capital plays an important role in reducing individual shirking in collective endeavors. Ichino and Maggi (2000) show in the Italian context that shirking, defined as absenteeism and misconduct, is higher in the low civic capital regions in the South.

of cooperative behavior at the workplace. Along these lines, we assume that intra-plant cooperation is exclusively determined by the stock of civic capital in the province where the plant is located, as embodied by the workers.

⁸See Arrow (1968) on p. 538.

⁹Chandler (1962) provides historical evidence that the growth of several North-American corporations was spurred by the creation of separated divisions within firms and the corresponding delegation of decision authority to subordinates.

Becker and Murphy (1992) illustrate how uncooperative behavior in team work reduces the size of plants. In their model firm output is the cooperative outcome of a team of workers. In order to produce a unit of final output, a given number of complementary tasks has to be performed. As both workers and tasks are assumed to be identical, each worker performs an equally large set of tasks. The working time of each member consists in acquiring task-specific skills and then performing the production tasks. As the size of the team increases, each worker performs a smaller set of tasks which increases its proficiency in production (for example, due to learning by doing). On the other hand, as specialization increases, so does the probability that the production chain collapses as the demand for coordination increases in the number of workers. Optimal team size is then given by the trade-off between gains from specialization and coordination costs. In a framework like Becker and Murphy (1992), civic capital can be thought to reduce shirking and coordination costs, and to allow the formation of larger teams.

2.2.3 Trust and cooperation in one-shot transactions

Relational contracts play an important role in governing intra-firm relationships (Baker et al., 2002). If employees interact frequently, the threat to terminate the relationship in the case of defection acts as an effective enforcement device. In particular, as long as the discounted gains outweigh short-run benefits from defection, agents have an incentive to stick to informal agreements. However, the frequency of interactions between two given employees depends on the size of the plant. In large establishments in which production is fragmented into numerous divisions, the probability that agents interfere with “strangers” is much larger than in smaller ones. Hence, the role of reputation and implicit contracts decreases because interactions are less likely to be repeated. Again, if transactions take place in trusting environments, cooperative outcomes can be sustained even if the game is not repeated (Fehr and Fischbacher, 2002).¹⁰

2.2.4 Testable implications

We have identified above the mechanisms that improve the efficiency of workers’ interactions at the plant-level. These mechanisms are likely to sustain, as a first order effect, larger plant size in areas where the stock of civic capital is high. We are well aware that, as pointed out in section 1.1, through a reduction in opportunism between independent contractors and a decrease in vertical integration, civic capital may actually exert a negative effect on average plant size. Acknowledging this arouses even more the interest in the empirical analysis, because we can verify the claim that the forces that pull toward an increase in size are stronger than those that pull toward a negative or ambiguous effect.

A related topic that can be addressed concerns whether civic capital affects uniformly plants along the size distribution. Our working hypothesis is that the cooperation-enhancing effect of civic capital (working through a facilitation in the delegation of authority, a reduction in shirking, and a higher propensity to cooperate in one-shot transactions) is more important for the functioning of large plants. The reason is that in small establishments cooperation can be sustained by other mechanisms such as direct monitoring, kinship bonds, relational contracts. For these reasons, we predict a stronger impact of civic capital in larger organizations.

¹⁰When explaining differences in the size of firms across countries, Fukuyama (1995) stresses this important property of trust.

Finally, the fact that, in a recent literature that has looked at the productivity effects of misallocation of factors of production, more efficient allocations are associated to a more dispersed distribution of plant size (see the simulation results in Guner et al., 2008, when distortions in the use of labor are removed) supports the expectation that plant size distribution should be more dispersed where civic capital is high. The reason is that civic capital (and the high level of trust and cooperation it generates) translates into a reduction of distortions at the plant-level.¹¹

3 Data and variables

In this section, we describe the different data sources used. We then discuss in detail the procedure that we follow to build the variables. First of all, let us describe the industrial and geographic scales we work with. Industries are classified according to the Statistical Classification of Economic Activities in the European Community (NACE), rev. 1. We work with a disaggregation at the 4-digit level, which classifies the entire spectrum of economic activity into 503 different industries. We keep in the sample only establishments which are owned by private entities, and we exclude establishments owned by non-profit organizations and public institutions.¹² Concerning the geographic disaggregation, we work with the 103 Italian provinces existing in 2001.¹³ We work with provinces, and not with smaller spatial units, because data on civic capital are not available at a smaller level of geography. Working with provinces is particularly convenient provided that it allows us to have a sufficiently large number of observations for each 4-digit industry.¹⁴

3.1 Plant size distribution

The data on plant size come from the 7th and 8th Censuses of economic activity carried out by the Italian national statistics authority (Istat) in 1991 and 2001, covering the universe of Italian establishments. In the censuses, plants are classified into 12 different size bins, according to the size of the plant in terms of workers.¹⁵ For each province and each industry, the data set then provides information on the total number of plants and the total number of workers in a given size bin.¹⁶

3.1.1 Average size

Let us now turn to the description of the dependent variables that we employ in the analysis. Following Kumar et al. (1999) and Laeven and Woodruff (2007), we define weighted average plant size (APS) in

¹¹A word of caution seems appropriate here. Exact predictions on the dispersion of the size distribution of plants, in the absence of a formal model, are hard to derive for us, and for this reason we rely on the literature on misallocation and productivity to qualify our results. However, our paper cannot be considered a full-fledged effort to verify these models.

¹²This choice implies that are excluded from the analysis industries where plants are exclusively managed by non-profit or government institutions, like health, social services and public administration.

¹³Italian provinces correspond to the NUTS-3 partitioning.

¹⁴The average number of observations per year for each industry is 70, which amounts to say that, on average, a 4-digit industry can be found in 70 provinces out of a total of 103.

¹⁵The bin categories are the following: bin number 1 (0 workers), 2 (1 worker), 3 (2 workers), 4 (3-5 workers), 5 (6-9), 6 (10-19), 7 (20-49 workers), 8 (50-99), 9 (100-199), 10 (200-499), 11 (500-999), 12 (1000 and more).

¹⁶The information about plant size in the Istat Censuses is similar to the one provided in the County Business Patterns for the U.S.

industry i , province j , and time t as:

$$APS_{i,j,t} = \sum_{b=1}^n \left(\frac{N_{b,i,j,t}^{emp}}{N_{b,i,j,t}^{estab}} \right) \times \left(\frac{N_{b,i,j,t}^{emp}}{N_{i,j,t}^{emp}} \right), \quad (1)$$

where $N_{b,i,j,t}^{emp}$ is the total number of workers in bin b , industry i , province j , and time t , $N_{b,i,j,t}^{estab}$ is the total number of plants in a given bin-province-industry-time, while $N_{i,j,t}^{emp}$ is the total number of workers in industry i , province j , and time t for all size bins. Hence, in this formula we weight the average size per bin, given by the first fraction in equation (1), by the share of workers working in that bin over the total number of workers in the province-industry, P-I henceforth. The literature has emphasized that the rationale for this weighting scheme is to put more weight on those plants which carry out the bulk of economic activity in a given P-I, and to weaken the impact on average plant size of a large number of small plants. The growth rate of APS can be then easily obtained as the difference between time t and $t-1$ of the logarithm of APS,

$$\Delta \ln APS_{i,j,t} = \ln APS_{i,j,t} - \ln APS_{i,j,t-1}.$$

In the robustness checks, we also consider a simplified version of (1), in which each size bin is weighted by the share of firms in that bin:

$$\overline{APS}_{i,j,t} = \sum_{b=1}^n \left(\frac{N_{b,i,j,t}^{emp}}{N_{b,i,j,t}^{estab}} \right) \times \left(\frac{N_{b,i,j,t}^{estab}}{N_{i,j,t}^{estab}} \right) = \frac{N_{i,j,t}^{emp}}{N_{i,j,t}^{estab}}. \quad (2)$$

This amounts to a simple division of the total number of workers by the total number of plants. We refer to $\overline{APS}_{i,j,t}$ as unweighted average plant size. Moreover, we will also look at the effect of civic capital on total employment in the P-I, $N_{i,j,t}^{emp}$, and on the total number of establishments, $N_{i,j,t}^{estab}$.

3.1.2 Standard deviation of size

Another feature of the size distribution of plants in a given P-I that we want to characterize is the degree of dispersion around the mean. For this reason, we consider the standard deviation of weighted plant size in industry i , province j , and time t . It is defined as:

$$SDPS_{i,j,t} = \sqrt{\sum_{b=1}^n \left[\left(\frac{N_{b,i,j,t}^{emp}}{N_{b,i,j,t}^{estab}} \right) - APS_{i,j,t} \right]^2 \times \left(\frac{N_{b,i,j,t}^{emp}}{N_{i,j,t}^{emp}} \right)}. \quad (3)$$

This variable puts more weight in the computation of dispersion on size bins that carry out the bulk of economic activity. Again, the growth rate of SDPS can be easily computed as the difference between the logarithm of SDPS at time t and $t-1$. The counterpart of (2) which will be employed as a test of robustness is:

$$\overline{SDPS}_{i,j,t} = \sqrt{\sum_{b=1}^n \left[\left(\frac{N_{b,i,j,t}^{emp}}{N_{b,i,j,t}^{estab}} \right) - \overline{APS}_{i,j,t} \right]^2 \times \left(\frac{N_{b,i,j,t}^{estab}}{N_{i,j,t}^{estab}} \right)}. \quad (4)$$

3.1.3 Specific points of the size distribution

In addition to using the growth of the first and second moments of the size distribution at the level of each P-I, we characterize its evolution providing a comprehensive analysis of the cumulative distribution function (cdf hereafter). We follow two distinct routes. First, as in Campbell and Hopenhayn (2005), we evaluate cdf at the following upper boundaries of plant-level employment: 9, 19, and 49 employees. We then look

at how $F(9)$, $F(19)$ and $F(49)$ have changed over time by computing their first differences over time. For example, for $F(9)$ we get

$$\Delta F(9)_{i,j,t} = F(9)_{i,j,t} - F(9)_{i,j,t-1}.$$

Since the range of the cdf is the interval $[0,1]$ here we do not take logs, otherwise we would miss observations whenever the cdf equals zero.¹⁷

Under the second approach, similarly to what Syverson (2004) has done in the context of productivity analysis, we compute the logarithm of size at the 25th, 50th, and 75th percentiles of the cdf in each year, as well as a standard measure of dispersion, the interquartile range (the difference between size at the 75th and 25th percentiles). The growth rate of these variables between 1991 and 2001 is then easily obtained by taking logs and then the difference.

3.1.4 Graphs and maps

In the baseline analysis a given industry in a given province is included in the final sample only if at least three plants are located there.¹⁸ Figure 1 provides a graphical representation of the empirical size distribution of plants. The graphs show the size distribution for the Manufacture of Tools (NACE rev. 1 code 2862) in two provinces of Tuscany, Pistoia and Siena, in 1991 and 2001.

[Insert Figure 1 about here]

Size bins are plotted on the horizontal axis. On the vertical axis we measure the value of the weight given to the bins. For each bin, the two bars correspond to the two different weights. The dark gray bars provide the weights based on employees, as in equations (1) and (3). The light gray bars provide the weights based on the number of establishments, as in equations (2) and (4). The latter also represent the empirical probability distribution of the different size bins. By putting more weight on large plants, the employee-based weighting scheme increases both the average and standard deviation of plant size. For example, for the 2001 data, in Siena APS is 18.04, while \overline{APS} is halved in magnitude, taking the value of 9.3. As for the standard deviation, $SDPS$ is 10.42, and \overline{SDPS} is 9.02.

The maps displayed in Figure 2 and 3 show the growth of APS and $SDPS$ between 1991 and 2001 across Italian provinces. The panels (a) show for each province the average growth rate across 4-digit industries. In panels (b) the averages are taken over residuals of a regression of the growth rate on industry fixed effects and initial 1991 level.

[Insert Figures 2, 3 about here]

Panel (a) in Figure 2 reveals that growth in average size has been somewhat stronger in the Center-North part of the country. This pattern is not replicated in the panel (a) of Figure 3: in the case of the growth of standard deviation there is not any difference between the North and the South of the country.

¹⁷Imagine that in a certain P-I there are only three plants with an equal size of 8 employees in 1991, and 10 employees 2001. In this case, taking logs would induce a missing value for $F(9)$ in 2001, and we would get a missing observation also for the difference between 1991 and 2001.

¹⁸The rationale for this choice is to exclude observations where the dependent variable is computed from very few plants. In the robustness and sensitivity analysis we also include observations based on one or two plants only.

As far as panels (b) are concerned, it is interesting to see that, once industry-fixed effects and initial 1991 values are taken into account, the evolution over time of the first two moments of the size distribution shows remarkable similarities all over the country. This is suggestive of a common mechanism behind the dynamics of the two variables. Some provinces of Emilia-Romagna, Lombardy, and Veneto can be singled out for the stronger growth of the two variables.

3.2 Measurement of civic capital

We employ several proxies to measure the stock of civic capital in a given province. These are blood donations and the number of volunteers in non-profit organizations, both standardized by population, as well as electoral participation in referenda between 1946 and 1987.¹⁹ Although quite different, all three activities share common properties which make them suitable proxies for the stock of civic capital in a province. First, individuals who donate blood, participate in volunteering or vote incur a non-negligible cost. These costs often exceed the mere opportunity cost of time devoted to each of these activities.²⁰ Second, and most important, none of these activities provide financial or legal incentives, so individuals do not obtain any economic pay-off. Rather, individuals who donate blood, engage in volunteering, or vote in referenda express a concern for some common good, triggered either by social preferences or by social pressure.²¹ As outlined in section 2, the diffusion of these traits among the local population accounts for the stock of civic capital in a province.

Figures 4, 5, 6 show the geographic distribution of blood donation, volunteering and electoral turnout, respectively. All three maps reveal that civic capital is highest in the Central and Northern part of the country and lowest in the South and Islands.

[Insert Figures 4, 5, 6 about here]

Table 1 provides evidence that there exists a pronounced positive correlation between each of the three proxies. The relationships are roughly equally strong in magnitude but nevertheless far from being perfect. Despite broad common patterns, the maps reveal some differences. This indicates that none of the three proxies can be taken as a precise measure of the stock of local civic capital. Rather, each variable is blurred by idiosyncratic factors which induce a certain geographic participation pattern, although they are orthogonal to the prevalence of civic capital.²²

[Insert Table 1 about here]

¹⁹In Appendix 9.1 we provide a detailed description of the variables and their sources.

²⁰For example, donating blood imposes a substantial physical limitation in the short-run, voting requires information gathering and personal evaluation of alternatives.

²¹See Putnam (1993) p. 93 for reasons why turnout referenda is more suitable than participation in “normal” political elections. For a general motivation of the choice of electoral turnout to proxy civic capital see again Putnam (1993).

²²For example, let us think to the 1987 referenda, dealing with some laws regulating the installation of nuclear plants in Italy, which assumed in the political debate the status of a vote against the use of nuclear energy. It is well known that the local community is strongly against the presence of nuclear plants in his own territory (this is sometimes referred to as the ‘nimby’ syndrome). Then, participation rates in 1987 referenda across Italian provinces could be driven by the presence of nuclear plants in the territory, in addition to the stock of civic capital.

In order to remove the noise, we look at the common component of the three proxies. The last row in Table 1 shows the correlations of the proxies with the first principal component.²³ The strong correlations suggest that all three proxies have a pronounced common dimension. Put differently, the first principal component identifies the behavioral attitude of the local population that simultaneously underlies the choice of whether to donate blood, participate on a voluntary basis in non-profit organizations, or vote in referenda; that is, it identifies the values and beliefs that account for the stock of civic capital.²⁴

Figure 7 shows the geographic distribution across Italian provinces of the measure of civic capital based on the first principal component. As before, we find that civic capital is the highest in the Center-North, and the lowest in the Southern mainland and Sicily. It is important to stress that our identification strategy relies on the variation of civic capital at the provincial level within Italian regions, because we will introduce in the empirical analysis dummy variables at the regional level.²⁵ Figure 8 illustrates this variation within regions. Specifically, the map shows the residuals from a regression of the principal component of civic capital on a set of regional dummies. The fact that provinces with very high and very low residuals, as evidenced by white and black-colored areas, are dispersed all over the country indicates that the variation of civic capital that we exploit is equally pronounced in all parts of Italy.

[Insert Figures 7, 8 about here]

In the analysis we also use historical variables of civic capital as instruments. The first historical measure is the number of mutual aid societies in 1873, standardized by population. These predominantly urban associations served craftsmen and artisans as a form of insurance against economic and social calamities. The second measure is average electoral turnout in elections during the period 1919-1921. Both variables are available at the regional level.²⁶ The last historical instrument is related to the type of early political institutions that were prevalent in the territory of a given province in 1300. More specifically, we employ the number of cities in each province that were free city-states in the year 1300. This variable is based on data from Guiso et al. (2008a).²⁷ In section 7 we justify the choice of these variables as instruments for current civic capital.

3.3 Other explanatory variables

The size distribution of establishments may be affected by a wide range of factors. For this reason, we control for as many determinants as possible.²⁸ The first control variable that we consider is the size of the local market. Melitz and Ottaviano (2008) show in a monopolistically competitive model with firm heterogeneity

²³In Appendix 9.2 we briefly review how to derive the first principal component.

²⁴Another paper where the first principal component is used to summarize common cultural traits at the regional level is Tabellini (2010).

²⁵An exception is the instrumental variables analysis, where, due to data constraints, we introduce macro-regional dummies instead of regional dummies. The macro-regions are five: North-West, North-East, Center, South, Islands. Given the geographic pattern of civic capital across Italy, macro-regional dummies are still able to neutralize the North-South divide.

²⁶For a detailed description consult the Appendix 9.1 and Putnam (1993).

²⁷The authors, in order to reduce the cost of collecting historical data at the town level, analyze the history of only the 400 biggest cities in terms of 1871 population in the Center-North. They confine data collection to this area because free-city states were a peculiarity of this part of the Italian peninsula.

²⁸A detailed description of the data and the corresponding sources is provided in Appendix 9.1

that average firm size (both in terms of output and sales) is larger in bigger markets. This theoretical result is in accordance with the empirical findings for the U.S. (Syverson, 2004; Campbell and Hopenhayn, 2005). In addition, Melitz and Ottaviano (2008) show that the dispersion of the plant size distribution should increase with market size. We use two variables for the dimension of the local market. The first is given by the provincial population, weighted by the relevance of final demand for that particular industry, as derived from the Italian Input-Output Use tables. The second measure is provincial gross domestic product.

Another possible determinant of the size distribution is the local stock of human capital. We see human capital as an outcome of education. For this reason our regressions control for the share of university graduates in the population of a given province.

Next, we take the efficiency of the legal system into account, understood as the quality of contract enforcement. We proxy the quality of the legal system in a given year computing the average number of days it took to reach a verdict in first-degree trials in labor-related affairs that were completed in that year in each of the 165 Italian labor courts.²⁹ Laeven and Woodruff (2007) have estimated that firm size is increasing with the quality of the legal system in Mexico. However, the effect of contract enforcement on firm's internal organization may actually go in the opposite direction. In particular, following the Transaction Cost Economics literature, a simple argument could run as follows. Provided that it is difficult to sign binding agreements where the quality of contract enforcement is weak, firms will tend to be more vertically integrated there, and their size will also be larger. Employing cross-country data, Acemoglu et al. (2009) find a positive and statistically significant correlation between contracting costs and vertical integration, but the significance disappears once industry dummies are plugged in.

Apart from the direct link through contract enforcement, formal institutions influence the size of economic organizations through the development of financial markets (Beck and Levine, 2003; La Porta et al., 1997): well-functioning financial markets allow firms to grow and to increase in terms of size.³⁰ As in Benfratello et al. (2008) we use the number of bank branches per province, normalized by population, as a proxy for the degree of financial development.

Finally, the degree of urbanization could be correlated with the plant size distribution: even after controlling for the dimension of the local market, plant size could still be the outcome of the overall degree of urbanization.³¹ Correlation between the plant size distribution and urbanization could also go through the different geographic characteristics of the provinces, when they translate into nature-given starting or stumbling blocks for economic activity, and might confound our results. Urbanization in 1861 is supposed to capture such time-constant provincial characteristics.³²

In Table 2 we provide a full set of descriptive statistics for the main variables in our database.

[Insert Table 2 about here]

²⁹Since the number of courts is larger than the number of provinces we developed a procedure to retrieve the average duration of trials at the provincial level. See Appendix 9.1 for details.

³⁰However Rajan and Zingales (1998) point out that developed financial markets not only allow firms to grow faster, but also increase the birth rate of new firms, which are generally quite small. Hence, the overall effect of financial development on firm size is a priori ambiguous.

³¹A classical reference linking plant growth and urbanization is Jacobs (1969).

³²It will be explained below that picking historical urbanization is also particularly convenient for the first stage of our instrumental variables approach, where we regress civic capital on historical instruments and other regressors.

4 The effect of civic capital on the plant size distribution: OLS results

4.1 Growth in average plant size

We base the regression analysis on a dynamic specification, namely an error-correction model.³³ The equation employed is the so-called Bårdsen transformation:

$$\Delta \ln APS_{i,j,t} = \alpha_0 + \alpha_1 \ln APS_{i,j,t-1} + \ln X'_{j,t-1} \bar{\beta}_1 + \beta_2 \ln CC_j + \Delta \ln X'_{j,t} \bar{\alpha}_2 + \kappa_r + \kappa_i + \epsilon_{i,j,t} \quad (5)$$

where $\ln APS_{i,j,t}$ is the logarithm of employment-weighted average plant size in province j and industry i at time t , $\ln CC_j$ is the log of the measure of civic capital in province j , $\ln X_{j,t}$ is the log of a vector of provincial controls at time t , κ_r denotes the region's dummy, κ_i is a 4-digit unobserved industry effect, and $\epsilon_{i,j}$ is the error term at time t . The plant growth regression takes scale dependency of the growth rates into account through the term $\ln APS_{i,j,t-1}$.³⁴

To better understand the usefulness of our approach, let us rewrite our model in the following manner:

$$\Delta \ln APS_{i,j,t} = \alpha_0 + \alpha_1 (\ln APS_{i,j,t-1} - \ln X'_{j,t-1} \bar{k}_1 - k_2 \ln CC_j - \gamma_r - \gamma_i) + \Delta \ln X'_{j,t} \bar{\alpha}_2 + \epsilon_{i,j,t} \quad (6)$$

Working with a specification like (6) is dictated by the following reasoning.³⁵ We assume average plant size in 2001 adjusts from its value in 1991 in response to the disequilibrium from the long-run relationship in 1991, represented by the term in brackets in (6), and the change in covariates, $\Delta \ln X'_{j,t}$. Two are the prominent features. The first is that we are positing the existence of a long-run equilibrium relationship between average plant size, a set of covariates, and civic capital. The long-run relationship implied by our model, which obtains when there are no stochastic shocks and all the variables are constant, is:

$$\ln APS_{i,j} = k_0 + \ln X'_{j,t} \bar{k}_1 + k_2 \ln CC_j + \gamma_r + \gamma_i \quad (7)$$

The second important feature is that we see the realization of average size in 2001 as a dynamic process which is the outcome, among other things, of the disequilibrium that occurred in 1991 between actual size and long-run equilibrium size. In our specification, the growth pattern of plant size is intertwined with the long-run equilibrium relationship. Investigating the impact of civic capital on the 10-year dynamic adjustment process, we are also able to retrieve the long-run static relationship between civic capital and size.

We are interested in the estimates of the long-run parameters of equation (7), with k_2 being our primary target. After having estimated through OLS the Bårdsen transformation (5), k_2 is simply

$$k_2 = \frac{\beta_2}{-\alpha_1} \quad (8)$$

We apply to equation (6) a fixed effect analysis, in the sense that we interpret our data set as a panel, where observations pertaining to a given industry i are collected repeatedly over different provinces. During

³³Useful references on this topic are Hendry et al. (1984) and Banerjee et al. (1993). See also Bårdsen (1989).

³⁴On the importance of scale dependency see Rossi-Hansberg and Wright (2007) and references therein.

³⁵When we measure civic capital by the principal component our equation is the following:

$$\Delta \ln APS_{i,j,t} = \alpha_0 + \alpha_1 (\ln APS_{i,j,t-1} - \ln X'_{j,t-1} \bar{k}_1 - k_2 PC_j + \gamma_r + \gamma_i) + \Delta \ln X'_{j,t} \bar{\alpha}_2 + \epsilon_{i,j,t}.$$

Here PC_j is the principal component of the log of the three proxy variables, rather than the log of some variable.

the statistical inference process, we take into account the potential correlation among the regression error terms using standard errors that are clustered both at the provincial level and at the industry level.³⁶ It is useful at this point to spend some words about the source of identification of the effect of civic capital in (5), and to differentiate it from cross-sectional estimates. In our dynamic specification, the identification of the effect of civic capital comes from the change in the dependent variable that occurred, for a given P-I, between 1991 and 2001. In terms of Figure 1, we are comparing the manufacture of tools industry in Pistoia in 1991 with the same industry in the same province in 2001. Growth is assumed to be a consequence of the disequilibrium of the dependent variable from its long-run value and of other shocks. On the contrary, in a cross-sectional analysis, what is done is to compare the level of the dependent variable in a province for a given industry with its level in another province for the same industry. In terms of Figure 1, the manufacture of tools industry in Pistoia in 2001 would be compared with the same industry in Siena in 2001. Working with a dynamic specification allows to control for some factors that could induce spurious correlations. To see this, suppose that plants with some unobservable characteristics (say plants of a particular age) are established in provinces with a high stock of civic capital. In a cross-section, we would estimate a spurious coefficient of civic capital on average size, even if civic capital had no real impact, if this unobservable characteristic were related to the level of size.³⁷

In the tables that follow we report the estimates of equation (5). At the bottom of each table we also detail the values of the long-run parameter for civic capital. Since these parameters are ratios of coefficients estimated from (5), the inference is based on the “delta method”.³⁸

[Insert Table 3 about here]

Results confirm our hypothesis that the level of civic capital is positively related to average plant size. In columns from (1) to (4) of Table 3 we estimate a restricted model where the vector of coefficients $\bar{\alpha}_2$ is set to zero. We regress the growth in average plant size on four different civic capital variables: blood donations, volunteers, electoral turnout in referenda, and the first principal component of all these three measures. Except for the case of the referenda turnout, the results are always statistically significant. The coefficient is most precisely estimated when the principal component is used. This is what we expected, given that the motivation to extract the first principal component is getting an accurate measure of civic capital out of the three proxy variables. Also final demand turns out to affect positively plant size. The latter result is in accordance with theoretical models such as Melitz and Ottaviano (2008) and empirical analyses such as Syverson (2004), Campbell and Hopenhayn (2005) and Laeven and Woodruff (2007). The fact that α_1 is negative is consistent with mean reversion: small establishments grow faster than large establishments.

³⁶The two-way clustering procedure that we adopt tackles two issues. On the one side, the correlation between error terms within provinces could be the result of disturbances at the local level, which could descend, for instance, from unobservable provincial characteristics. On the other side, the correlation in the error terms within industries could still survive the inclusion of the fixed effects γ_i . Think, for example, to some random event that led some big plants in an industry to locate in some province. This could induce correlation in the error terms for observations in that particular industry. On inference with clustered data see Cameron and Miller (2010).

³⁷There are forms of spurious correlation that are difficult to address even with our dynamic specification. For example, if high civic capital provinces attract younger firms, and younger firms do grow more not just because they are smaller in size, controlling for the initial level of size would not be enough to avoid a spurious correlation.

³⁸On this method see, for example, Greene (1999).

In columns from (5) to (8) of Table 3 we estimate the full model (5). Except for column (7), coefficient estimates are statistically different from zero in all the specifications and confirm that civic capital exerts a positive effect on the average size of plants. In addition to final demand, in this specification there is a positive correlation between the length of trials (an inverse measure of contract enforcement quality) and plant size growth. This evidence corroborates the transaction cost theory of the firm, since lengthy trials can be expected to induce firms to internalize more activities and to operate on a larger scale.

We also provide in the bottom part of Table 3 the magnitude of k_2 , the long run parameter on civic capital, given by the formula (8). According to equation (7), this parameter is actually an elasticity, because both the dependent variable and the regressor are logarithms. Then, in the case of blood donations, an estimated elasticity equal to 0.10 means that a 10% difference in the number of blood donations at the provincial level is associated to a 1% difference in long-run weighted average size of plants at the level of 4-digit industries. The elasticity is equal to 0.14 in the case of volunteers. It is only in the case of the referenda turnover that the coefficient is no longer significant. This fact can be taken as evidence that, once the regional effects are washed out by the dummy variables, there is not enough variability left in provincial turnout to allow a coefficient estimate which is statistically different from zero.³⁹ The following example gives an idea of the economic magnitude of the effect of civic capital when the measure is the first principal component. Milan is a high civic capital province (it ranks 21st in Italy), while Naples is the province with the lowest civic capital in Italy. An estimated long-run coefficient equal to 0.10 implies that average plant size would be 58% larger in the long-run equilibrium if Naples had the same stock of civic capital of Milan.⁴⁰ This is a sizeable impact, and it is very close to the actual difference of 47% in average size that comes out from the 2001 Census data.

The last column of Table 3 provides the estimation results of (5) when the dependent variable is un-weighted plant size. Also in this case civic capital is positively related to average plant size.

4.2 Growth in the standard deviation of plant size

Our empirical analysis proceeds with the estimation of the link between the standard deviation of plant size and civic capital. The idea is simply to replace in equation (5) $\ln APS_{i,j,t}$ with $\ln SDPS_{i,j,t}$, the logarithm of the standard deviation of plant size in industry i and province j at time t , computed according to equation (3). The results, provided in Table 4, in columns from (1) to (4) are obtained with the restricted model where the coefficients on the provincial controls' growth are set to zero. In columns from (5) to (8) we estimate the full model. The following results stand out. As before, the proxies of civic capital are statistically significant, with the exception of electoral turnout in referenda. The value of the long-run

³⁹When the proxy of civic capital is the first principal component of the above mentioned three variables, it is not straightforward to provide an interpretation in terms of elasticity of the coefficient of civic capital. As we outline in Appendix 9.2, the extraction of the principal component is a statistical procedure whose output, starting from the log of our three proxies, is a variable which has no observable counterpart. However, we think that this procedure is particular appropriate in a framework as ours where we want to identify in the most accurate way the common behavioral attitude of the local population toward trusting others, and in view of this a certain artificiality of the measure can be tolerated.

⁴⁰The principal component for Naples is -3.347, while for Milan it is +1.236. The contribution from the difference in civic capital to the difference of average plant size in log terms for the two locations is then equal to 0.458, which corresponds to a difference of 58% in terms of average size measured in levels.

elasticity parameter for standard deviation with respect to civic capital is 0.10 for blood donations and 0.14 for volunteers. The last column in the table presents the estimates when the dependent variable is growth in unweighted standard deviation. Again, the coefficients are statistically significant, both that belonging to the main equation and the long-run parameter. The results suggest that civic capital increases the second moment of the distribution of plant size, in addition to the first. As to the other controls, final demand (a proxy for local market size) increases the standard deviation, in the spirit of Melitz and Ottaviano (2008). Cumbersome contract enforcement, proxied by lengthy trials, is associated to a more dispersed distribution in terms of size (probably due to an increase in the number of big plants in the P-I relatively to smaller ones).

[Insert Table 4 about here]

5 The evolution of the size distribution over time: further analysis

We then look at the plant size distribution employing a wide array of measures to understand how the shape has changed at specific points of the size distribution at the P-I level. We evaluate cdf at the following upper boundaries of plant-level employment: 9, 19, and 49 employees. These values (bounded between zero and one) are then used as dependent variable in our specification (5). Columns (1), (2), and (3) of Table 5 show a negative coefficient on civic capital, and a negative value for the long-run parameters. This means that the cdf of plant size shifts toward the right. So, on average, plants in areas with a higher stock of civic capital are larger. This is perfectly in line with the findings we obtained when our dependent variable was average size in the P-I and is also coherent with our conceptual framework. Regarding the magnitude and the statistical significance of the estimates, we note that the shift is particularly evident in the left part of the distribution (up to 19 employees), while it is reduced in magnitude for larger plants. We attribute this attenuation to the increase in the number of zeros in the dependent variable as the size threshold goes up. Actually, as the threshold goes up, the cdf is more likely to be equal to one in both periods of time, the change in cdf over ten years is more likely to be zero, and the variability is reduced. This pattern is documented quite precisely by the descriptive statistics of Table 2.

[Insert Table 5 about here]

In the columns from (4) to (6) we evaluate plant size at given percentiles of its distribution within each P-I. The idea here is to see whether the impact of civic capital is homogeneous across plants, or whether plants at different percentiles of the distribution are affected differently. The results prove that the strongest and statistically most significant effect occurs for plants at the 75th percentile, the largest ones. Plants at the 25th percentile and plants at the 50th percentile also become larger in areas with a higher stock of civic capital, but with a long-run parameter equal to 0.01 and 0.02 respectively (the coefficient at the 50th percentile is not statistically different from zero).⁴¹ We conclude that civic capital prompts plant

⁴¹The discrete nature of the data may explain the reduced statistical significance of the civic capital coefficients in the regressions based on percentiles. Taking as an example NACE industry 2862 in 2001 in Pistoia, firms at the 25th percentile of

size growth across all the spectrum of the size distribution. However, the effect is stronger and statistically more significant in the case of larger plants. A corollary of this result, provided that the increase in size is more pronounced in the right tail of the distribution, is the increment in dispersion. In column (7) the interquartile range, measured by the difference in the log of size of firms at the 75th and 25th percentiles, goes up with a long-run parameter equal to 0.05.

Let us go back to Figure 1, where we plot the distribution for NACE 2862 in Siena and Pistoia in 1991 and 2001, to provide some more insight into our results. We want to provide a stylized graphical interpretation of how civic capital shifts the distribution of plant size. The two provinces are roughly equal in terms of overall size (measured by total GDP) and belong to the same region (Tuscany). However, they differ markedly in the endowment of civic capital. While Pistoia's endowment ranks 42nd in Italy in terms of the first principal component, Siena has the 3rd highest endowment. The figure confirms that the higher it is the stock of civic capital the higher it is the increment in the number of plants in the upper tail of the size distribution at the end of the 10-year period. The value of $1-F(9)$, which captures the share of plants with at least 10 employees (obtained summing light gray bars in size bins 6 and 7), goes up in Pistoia from 0.13 to 0.20; that is, by 7 percentage points. The same variable soars by 17 percentage points in Siena (from 0.25 to 0.42). The values $1-F(19)$ and $1-F(49)$ do not change in Pistoia, because there are no plants with at least 20 employees in 1991 or 2001. On the contrary, in Siena $1-F(19)$ goes up by 3 percentage points (from 0.05 to 0.08, size bin 7). Also in Siena there are no plants with at least 50 employees at the beginning or at the end of the 10-year period, so $1-F(49)$ does not change. Turning to the interquartile range variable (the difference between size at the 75th and 25th percentiles of the distribution) it is Siena again to show a more pronounced rise over the period. In this province the range increases from 5.5 in 1991 (7.5 is size at the 75th percentile and 2 is size at the 25th) to 12.5 (14.5 at the 75th minus 2 at the 25th). In Pistoia, over the same lapse of time, we observe a much smaller increment, from 5.4 (7.4 minus 2) to 5.5 (7.5 minus 2). Finally, size at the 75th percentile almost doubles in Siena, while it stays almost constant in Pistoia. Size at the 25th percentile stays constant in both provinces.

The empirical evidence based on the estimates for APS and on those for the cdf points to a shift toward the right of the size distribution. The rise in average size is more pronounced for larger plants than for smaller ones. This can be justified on the basis of our theoretical framework to the extent that the mechanisms outlined there are more relevant for larger organizations, and it comes as no surprise that such a stretch of the distribution is associated to an increase in dispersion, measured both in terms of SDPS and interquartile range. The increase in dispersion is also consistent with recent models featuring distortions in the use of factors of production.

To fully characterize the short-run dynamics of the plant distribution, it would be useful to know whether civic capital impacts on the size distribution through factors like plant size at entry, or through the margin associated with growth in existing establishments. Related to this, it would be useful to know whether, in response to the growth in the average size of incumbents, some smaller establishments are forced to exit the size distribution (in bin 3 in this case) are actually between the 13th and 33th percentiles, those at the 50th percentile (bin 4) are actually between the 33th and 67th, those at the 75th (bin 5) are actually between the 67th and the 80th. Obviously, the exact intervals where the 25th, 50th and 75th percentiles fall will in general vary across different P-I cells, with this bringing considerable noise to the measure.

the market. Data limitations prevent us from making definitive statements on issues like these. However, looking at variables such as total employment and the total number of plants may help us in qualifying the results. Column (8) shows that total employment grows in provinces endowed with more civic capital. The coefficient estimate is significant at the 10% level only. According to back-of-the-envelope calculations, the estimated long-run parameter for civic capital implies a total employment advantage for an industry in Milan equal to 44% with respect to Naples.⁴² In column (9), where the dependent variable is growth in the total number of establishments, we get an estimate which is not statistically different from zero (the standard error is large in this case). Overall, we interpret the evidence as pointing to a moderate growth of industry employment in areas with higher civic capital, with the effect being particularly pronounced in larger plants. It is hard to assess with the data at hand the relative importance between two possible channels, which are larger size of plants at entry (with simultaneous exit of smaller establishments) vis-à-vis the increase in the size of incumbents.

6 Robustness and sensitivity analysis

Our robustness and sensitivity checks are presented in Table 6. A simple partition of 4-digit industries can be obtained by grouping them in two different sets of industries, namely manufacturing (NACE section D) and services (NACE sections from E onward). We then ask whether manufacturing is any different from services. For both manufacturing and services, civic capital increases average plant size (columns (1) and (2)) and the standard deviation (columns (6) and (7)). The effect is more pronounced for manufacturing, but we can conclude that our results are not peculiar to a particular sector.

[Insert Table 6 about here]

In columns (3) and (8) we check whether results are affected by the number of variables we employ in the computation of the first principal component. Specifically, one may wonder whether excluding referenda turnout (whose coefficient is not statistically significant when we employ it as a measure of civic capital) from the computation of the first principal component affects the results. We prove that this is not the case. The estimated coefficients are just slightly smaller in magnitude than in the baseline regressions, but still highly significant.

In columns (4) and (10) we add yet another provincial control, namely GDP. The inclusion of GDP reduces the statistical significance of the other controls such as final demand. However, the estimates related to the impact of civic capital are affected in a negligible manner.

In columns (5) and (11) we consider the full set of available observations. We mentioned previously that we exclude from the baseline regressions those observations where the dependent variable is based just on one or two plants. In these columns we check whether adding these observations to the sample changes the results in some manner. The change in the coefficient estimates (both for the short and long-run parameters) with respect to our baseline analysis is tiny.

⁴²From 2001 census data we have computed that total employment at the 4-digit industry level is actually on average 99% larger in Milan than in Naples.

Finally, in columns (6) and (12) we run regressions after excluding from the sample the observations from the provinces of Rome, Milan and Turin. These are provinces where a considerable amount of economic activity is concentrated: Rome is the capital and base of ministries and government agencies, Milan is the economic and financial center, Turin is where FIAT group and automobile industry is headquartered. Our estimates prove that their exclusion from the sample bears no change to our results.

7 The effect of civic capital: evidence from 2SLS estimation

7.1 Motivation and identifying assumptions

In addition to problems of measurement error in our main explanatory variable that we have addressed computing the first principal component of three different proxies of civic capital, there are other reasons which might prevent the correct identification of the impact of civic capital on the plant size distribution. Consider the case of reverse causation. If a large plant locates in a province average plant size increases, and, arguably, the local population is economically better off. As a consequence of this improvement in material well-being, individuals develop civic virtues, i.e. they donate blood, vote in referenda and engage as volunteers. Moreover, although we control for a range of confounding factors, we cannot completely exclude that other omitted variables bias our estimates.

In order to prove that our results are neither driven by omitted variables nor by reversed causality, we provide evidence from two-stage least square estimates. The first set of instruments are lagged proxy variables of civic capital, namely members of mutual aid societies in 1873, and electoral turnout around the 1920s. The second set of instruments goes even further back in time. It relies on Putnam (1993), who argues that the huge differences in the endowment of civic capital across Italian regions can be traced back to different political regimes prevailing at the beginning of the second millennium. In particular, he stresses the role of free-city states that emerged in the Center-North of the Italian peninsula. Testing Putnam's theory, Guiso et al. (2008a) show that the free-city state experience has had a causal effect on the accumulation of civic capital.

We have good reasons to believe that these historical measures are highly correlated with the current stock of civic capital. The literature has stressed the persistence of civic capital over long periods of time, highlighting the crucial role of intergenerational transmission of values and beliefs from parents to their offspring (Tabellini, 2008; Guiso et al., 2008b). Hence, the stock of today's civic capital in a certain local area can be explained to a large extent by historical values of civic capital in that area. The high persistence of civic capital over time translates, as we show below, into a strong first stage relationship.

The other requirement for the validity of our instruments is that they must be uncorrelated with the error term $\epsilon_{i,j}$ in equation (5). This amounts to saying that, conditional on the other regressors, the historical instrumental variables have had no direct effect on the first and second moment of the plant size distribution. This could be a somewhat strong assumption, as our historical instruments might have affected current economic development and hence current plant size through channels other than the current stock of civic capital. In order to address this threat, we include the log of provincial GDP as an additional regressor in all 2SLS specifications. The rationale is that this variable captures all the possible effects of the

historical instrumental variables on the plant size distribution that work through economic development. This makes us more confident that the exclusion restriction is not violated.

7.2 Results

7.2.1 Civic capital in the 19th and early 20th century

The first set of instruments dates back to the middle of the 19th and early 20th century. The 19th century experienced an unusual movement in popular sociability, not only in Italy but in whole Western Europe. As a response to the new economic and social calamities brought about by the process of industrialization, traditional associations like guilds and religious societies were replaced by more civic, charitable and educational organizations. One prominent manifestation of this new collectivism was the creation of mutual aid societies. Their members enjoyed benefits such as medical care, insurance against work accidents and basic instruction. Importantly, the functioning of mutual aid societies relied solely on the principle of reciprocity: absent any kind of formal enforcement and coordination, their members joined individual forces for mutual benefit. Putnam (1993) describes them as “a locally organized, underfunded, self-help version of what the twentieth century would call the welfare state”. Therefore, the membership rate in mutual aid societies in an area is a good proxy for the stock of local civic capital. The year to which this information refers to is 1873.

The second measure of civic capital is electoral turnout in the 1920s. Specifically, we average turnout over four political elections: national elections from 1919 and 1921 and provincial and communal elections from 1920. The choice of these elections is dictated by the fact that they were the only elections with universal suffrage before the advent of fascism.

Again, using lagged proxy variables of civic capital we have to exclude reverse causation. In other words, the pattern of civic capital in the past should not be the outcome of economic well-being at that time. We address this issue by analyzing the relationship between urbanization rate in 1861, which relates to economic development at that time, and the two instruments.⁴³ The correlation coefficient between past urbanization and members of mutual aid societies is -0.05. This implies that for the year 1873 we have, if anything, a *negative* relationship between civic capital and material well-being. Consequently, we are confident that this measure of civic capital is exogenous from the point of view of our analysis. As for turnout in the 1920s, the correlation coefficient is -0.24.⁴⁴

Our potentially endogenous variable (the current level of civic capital, as given by the principal component) is expressed in the following way in the first stage regression for APS:

$$PC_j = \omega_0 + \omega_1 \ln APS_{i,j,t-1} + \omega_2 \ln CC_r^{Aid} + \omega_3 \ln CC_r^{Turn} + \ln X'_{j,t-1} \bar{\omega}_4 + \Delta \ln X'_{j,t} \bar{\omega}_5 + \gamma_{mr} + \epsilon_j, \quad (9)$$

where $\ln CC_r^{Aid}$ is the stock of civic capital in region r proxied by the log of the membership in mutual aid societies in 1873 (normalized by population), while $\ln CC_r^{Turn}$ is the log of average regional turnout of

⁴³It is Tabellini (2010) who argues that past urbanization is a good proxy for past economic development. In any case, we signal that some caution has to be paid when dealing with historical urbanization in the context of Southern Italy. As explained by Malanima (2005), in 1861 many big cities in Southern Italy were actually *agrotowns*, since a large share of inhabitants was employed in agriculture, and not in manufacturing or services.

⁴⁴We acknowledge that this negative relationship is less reliable for our purposes, as we are comparing urbanization rate in 1861 with the measure of civic capital 60 years later.

elections in the 1920s. These two terms are meant to capture the persistent component of civic capital. Since these measures are only available at the regional level we cannot include regional dummy variables. Instead, we replace them with macro-region dummies, γ_{mr} . The vectors $\ln X'_{j,t-1}$ and $\Delta \ln X'_{j,t}$ contain all provincial covariates, including provincial GDP.

The bottom part of Table 7 (Panel B) contains the results from the first stage. Columns (2) and (3) refer to APS, while (7) and (8) refer to SDPS. Columns (2) and (7) report the estimates when the only instrument employed is membership in mutual aid societies. In columns (3) and (8) we have the results when also turnout in the 1920s elections is employed. The coefficient estimates of both instruments have the expected positive sign and are statistically different from zero. Moreover, the instruments explain a very large share of the variation in current civic capital, as evidenced by the high first-stage R^2 and F statistics. Given the high persistence of civic capital across time, this comes as no surprise. As for the other covariates, it is important to stress the role of urbanization in 1861 in equation (9).⁴⁵ As mentioned by Putnam (1993), mutual aid societies predominantly served craftsmen and artisans in cities as a form of insurance against economic and social calamities. As such, they were mostly an urban phenomenon. Hence, including historical urbanization in 1861 among the regressors improves the reliability of the instruments as it controls for the degree of urbanization at that period, something which might confound the results.

[Insert Table 7 about here]

Results from the second stage are depicted in Panel A of Table 7. In order to provide comparable OLS results, we have re-estimated equation (5) with macro-regional, instead of regional, dummies. The corresponding results are shown in columns (1) and (6). As for the 2SLS estimates, they are qualitatively similar to the OLS results. The coefficient of civic capital is positive and statistically different from zero at the 1% level. In columns (3) and (8), given that we have two instruments, we are allowed to perform a test of overidentification to assess whether one of the two is not exogenous. The p-value of the test statistic is equal to 0.55 for APS and 0.16 for SDPS and it is sufficiently high to reject the alternative hypothesis, according to which at least one of the two instruments is not exogenous. Hence, the data support the assumption that the exclusion restriction is not violated.⁴⁶

7.2.2 Free-city state experiences during the Middle Age

The second strategy to instrument the current level of civic capital relies on the number of free-city states experiences that were present in the territory of each province in 1300. Putnam (1993) argues that the huge differences in the endowment of civic capital across Italian regions are due to the emergence of free-city states in the Center-North of the country in the Middle Age. Lacking any centralized form of government, their citizens had to collaborate to provide solutions to problems of common interest, first and foremost to defend their cities against foreign invaders. As a consequence, individuals developed a sense of cooperation and concern for common issues resulting in a high stock of civic capital.

⁴⁵We have not reported its estimate in Table 7, and those of other provincial controls, due to space constraints.

⁴⁶The test statistics have to be interpreted with caution, though. While testing the exogeneity of, say, the number of mutual aid societies in 1873, the overidentification test explicitly *assumes* that the other instrument, turnout in the 1920s, is truly exogenous. Therefore, the test cannot detect whether both instruments are invalid.

We exploit information on free-city states in 1300 provided by Guiso et al. (2008a). However, we perform a slightly different first-stage regression. Since we work at the province level instead of the city level, for each province we count the number of cities that were a communal republic in the year 1300, which gives us a minimum of zero and a maximum of three free-city states per province.⁴⁷ Figure 9 shows the number of free-city states per province across Italy. The information is only available for the Center-North. The reason is that free-city states were a peculiarity of this part of the Italian peninsula, while Southern Italy and Sicily were controlled by kings of Norman descent. Thanks to the inclusion of macro-regional dummies, our instrumental variables strategy looks within each of the three macro-regions in Center-North of Italy (specifically, North-East, North-West, and Center) to exploit the fact that in the territory of some provinces belonging to the same macro-region there were more free-city states than in other provinces, and so the prediction is that civic capital today should be higher where the number of free-city states was higher in 1300. In a sense, we are allowing for the possibility of local spillover effects from political independence of single cities at the level of the whole provincial territory.

In the first stage regression we plug three different dummy variables capturing the number of free-city states (be it 1, 2 or 3). We exclude from the regression the dummy variable for the case when there are no free-city states, which becomes our reference group. The specification in the first stage in the case of APS is the following:

$$PC_j = \omega_0 + \omega_1 \ln APS_{i,j,t-1} + \omega_2 City_j^1 + \omega_3 City_j^2 + \omega_4 City_j^3 + \ln X'_{j,t-1} \bar{\omega}_5 + \Delta \ln X'_{j,t} \bar{\omega}_6 + \gamma_{mr} + \epsilon_j \quad (10)$$

where $City_j^1$, $City_j^2$, and $City_j^3$ are the dummy variables for 1, 2, or 3 free-city states in the province, respectively. Macro-regional dummies and provincial covariates are also included. Columns (5) and (10) in Panel B show that these instruments are good predictors for the current level of civic capital. As expected, ceteris paribus, provinces with two free-city state in 1300 display a higher level of civic capital than provinces with one or no free-city states. Performing an F -test on the free-city instruments reveals that they fit civic capital data well, although not as well as in the case of aid societies and turnout in the 1920s (see the bottom part of Panel B). Also the first-stage R^2 is large, but somewhat reduced with respect to the other first stage regressions of columns (2), (3), (7) and (8).

The second stage results in Panel A show a positive and statistically significant effect of civic capital on both average plant size and the standard deviation of size. In order to provide comparable OLS results, in columns (4) and (9) we have re-estimated equation (5) on the sub-sample of provinces in the Center-North of Italy. Quantitatively, the 2SLS estimates of civic capital are larger than the OLS results. Then, OLS are downward biased. This downward bias is also what Guiso et al. (2008a) get in their regressions where the dependent variable is the level of per capita income. These results may seem puzzling at a first glance. For this reason we looked carefully at the literature about development and institutions and interestingly we found notable cases where instrumental variables estimates are larger than the OLS ones. We mention here the papers by Acemoglu et al. (2001) and Gorodnichenko and Roland (2011). The explanation provided by the authors is to plead for the presence of an attenuation bias due to measurement error. The same

⁴⁷Obviously, one needs to control for the size of the province since the number of free-city states could be a function of the dimension of the province. This is achieved through the inclusion in the first stage of variables such as GDP, final demand, urbanization in 1861. Notice that there are just three provinces with 3 free-city states (Alessandria, Cuneo and Turin) and they are all located in Piedmont.

argument applies to our case. In other terms, our 2SLS may help to further reduce the measurement error in our variable of civic capital.

7.2.3 Evidence from specific points of the size distribution

In Table 8 we present the 2SLS results where the dependent variables are exact points from the size distribution. Again, Panel B shows that our instruments do a fairly good job in explaining the level of civic capital at the provincial level. The estimates from the second stage in Panel A confirm the OLS results of Table 5: civic capital shifts toward the right the size distribution, as documented by the columns from (1) to (6). Similarly to the OLS estimates, there is some attenuation as we move up in the size distribution. Looking at plant size at the 25th and 75th percentiles, we see an increase in both of them, but the increase for establishments at the 75th percentile is more pronounced. It is then consequential that we witness an increase in the dispersion of the size distribution, as documented by the rise in interquartile range. Overidentification tests carried out when we instrument for membership in aid societies and turnout in the 1920s reject the null that at least one of the two variables is not exogenous.

8 Conclusion

Employing two censuses on the universe of establishments in Italy, we have shown how trust, captured by the geography of civic capital, shapes the size distribution of plants. More specifically, our results reveal that civic capital increases the average and the dispersion of the size distribution in 4-digit industries across provinces. The effect is very robust across different specifications addressing several potential threats. First, a set of appropriate controls excludes that our results are driven by other determinants of plant size such as market demand, financial development, human capital, judicial inefficiency and past urbanization. Moreover, fixed effects at a highly disaggregated industry and spatial level are included. Third, we explicitly address potential sources of endogeneity of civic capital. On the one hand, we extract the first principal component of three well-established measures thus minimizing the problem of measurement error. On the other hand, we employ the variation of civic capital in the middle of the 19th and the beginning of the 20th century, and free-city state experiences in 1300 to instrument for current trust. We get the conclusion that it is unlikely that our results are driven by reversed causality or omitted variables bias.

Even if we do not present a formal model, we have carefully assessed in the paper the mechanisms that can explain our empirical findings. The hallmark of civic capital is to increase cooperation in intra-plant transactions, even when economic incentives to do so are weak (for example in large organizations). This is expected to boost average plant size. In addition to this, the decrease in the distortions at the plant-level induced by a more trusting environment can be expected, according to a recent literature on the consequences of misallocation, to increase the dispersion of the size distribution.

Our study brings the following policy implication. We have proved that, even within the territory of the same region, the geographic variation in the stock of civic capital poses substantial constraints on plants' ability to expand. This is an important starting point for the implementation of effective industrial policies. Imagine that a regional government disposes of a limited budget in a program designed to subsidize the expansion of small or medium-sized plants. Our study suggests that regional authorities should concentrate

resources in the dark areas of Figure 8, or at least limit those going in the white areas of the map.

9 Appendix

9.1 Detailed description of the data

9.1.1 Dependent variables

The construction of the dependent variables is described in detail in the main text. The information is taken from the 7th and 8th Istat Censuses of economic activity, corresponding to the years 1991 and 2001.

9.1.2 Measures or instruments for civic capital

Blood donations: The number of blood donations per 1000 inhabitants, disaggregated by province. The data are collected from the health authorities of Italian regions. In each region, regional health authorities collect data on blood donations and subsequently send this information to the High Institute for Health (*Istituto Superiore di Sanità*) which, in turn, maintains a National and Regional Registry of Blood and Plasma. Provincial data on blood donations are not available for Apulia and Lazio. For the provinces of these two regions we take the total regional value. Data refer to the year 2002 and the source is Cartocci (2007).

Volunteers: It is the number of volunteers in non-profit organizations. Data refer to the year 2000 and the source is de Blasio and Nuzzo (2010).

Referenda turnout: It is the average provincial electoral turnout for the referenda on the choice between republic and monarchy (1946), divorce (1974), public financing of political parties (1978), public security and anti-terrorism measures (1981), abortion (1981), wage escalator regulations (1985) and nuclear power and hunting regulations (1987). The following eight provinces were created after 1995: Biella, Lecco, Lodi, Rimini, Prato, Crotone, Vibo Valentia, Verbano-Cusio-Ossola. The provinces to which they belonged before 1995 and whose value has been assigned to them appear in parenthesis: Biella (Vercelli), Lecco (simple average of Bergamo and Como), Lodi (Milan), Rimini (Forli-Cesena), Prato (Firenze), Crotone (Catanzaro), Vibo Valentia (Catanzaro), Verbano-Cusio-Ossola (Novara). The source of data for referendum turnout is the Ministry of the Interior.

Mutual aid societies in 1873: It is the number of the members in mutual aid societies in 1873 at the regional level, standardized by 100,000 inhabitants. Data for Valle d'Aosta, Trentino-Alto Adige and Friuli-Venezia Giulia are missing. We adopt the values of Piedmont for Valle d'Aosta, the region from where it was split off. For the latter two regions, we adopt the values of Veneto, which is socio-geographically the closest one. Additionally, there is no data for Molise, for which we take the value of Abruzzo, the region from where it was split off. The source is Putnam (1993).

Turnout in 1920s: It is the average electoral turnout at the regional level in the national elections of 1919 and 1921, provincial elections in 1920 and communal elections in 1920. There is no data for the regions of Valle d'Aosta, Trentino-Alto Adige and Friuli-Venezia Giulia. We adopt the values of Piedmont for Valle d'Aosta, the region from which it was split off. For the latter two regions, we adopt the values of Veneto, which is socio-geographically the closest one. The source of these data is Putnam (1993).

Number of free-city states in 1300: It is the number of free-city state experiences in the territory of each province in 1300. Data are from Guiso et al. (2008a). In order to reduce the cost of collecting historical data at the town level, the authors analyze the history of only the 400 biggest cities in terms of 1871 population in the area that was under the Holy Roman Empire at the beginning of the second Millennium (basically, the Center-North of Italy).

9.1.3 Other explanatory variables

Final demand: The variable is constructed as follows. First of all, we assign each 4-digit industry to the corresponding 2-digit sector. For each 2-digit industry, we compute from Input-Output Use tables for Italy the share of output being purchased by households, public institutions and non-profit organizations. This share is then multiplied by total provincial population to get an exact measure of the dimension of the local market in terms of final demand for all the industries belonging to a given 2-digit sector (the variable then varies by 2-digit sector and by province). Both Input-Output Use tables and provincial population are from Istat.

Urbanization in 1861: This variable is the share of total provincial population living in cities with more than 10,000 inhabitants in 1861. Data on city size come from the "Italian Urban Population Database

1300-1861", provided by Paolo Malanima (http://www.paolomalanima.it/default_file/Page646.htm). The number of total provincial population in 1861 is taken from Populstat (<http://www.populstat.info/>).

University graduates: It is the number of university graduates per province, divided by total provincial population. Data are from Istat.

Bank branches: It is the number of bank branches per 1000 inhabitants, disaggregated by province. The source of the data is the Bank of Italy.

Length of labor trials: It is the number of days it takes to complete a first degree trial in labor affairs in each of the 165 Italian labor courts. The data refer to the years from 1995 to 2001 (earlier years are not available) and are provided by Istat in the data base Territorial Information System on Justice (Sistema Informativo Territoriale sulla Giustizia). Since there are more courts than provinces and since in some cases the territory of a court belongs to two different provinces we proceed as follows. First, we assign to each city in the province the value of the court to which the city belongs. This information is then averaged over all the cities belonging to the same province to get a provincial variable.

GDP: It is the provincial nominal gross domestic product, expressed in thousands of euros. The source is Istat.

9.2 Derivation of the first principal component

The intuition of principal component analysis (PCA) in our context is the following: given the three proxies of civic capital, each province corresponds to a point in a three dimensional vector space. The idea of PCA is to find a linear combination of the three variables which re-expresses the original data set in such a way that it captures most of the common variance. This linear combination corresponds to the first principal component.

In general terms, the first principal component can be derived as follows (see Jolliffe, 2002): vector \mathbf{x} denominates the data consisting of p random variables (the three proxies of civic capital in our case) and vector α_1 consists of p constants, $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}$. Consider the linear function $\alpha_1' \mathbf{x}$:

$$\alpha_1' \mathbf{x} = \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1p}x_p = \sum_{j=1}^p \alpha_{1j}x_j \quad (11)$$

Finding the first principal component amounts to determine the elements of α_1 which maximize the variance of $Var[\alpha_1' \mathbf{x}] = \alpha_1' \mathbf{S} \alpha_1$, where \mathbf{S} is the covariance matrix of \mathbf{x} . The vector α_1 is constrained to have unit length, which implies that $\alpha_1' \alpha_1 = 1$. The corresponding Lagrange maximization function takes the following form:

$$\alpha_1' \mathbf{S} \alpha_1 - \lambda(\alpha_1' \alpha_1 - 1). \quad (12)$$

Maximizing (12) with respect to α_1 gives

$$(\mathbf{S} - \lambda I_p) \alpha_1 = 0, \quad (13)$$

in which the Lagrange multiplier λ is the eigenvalue of \mathbf{S} and the corresponding eigenvector is α_1 . I_p is the p -dimensional identity matrix. Because the quantity to be maximized is $\alpha_1' \mathbf{S} \alpha_1 = \alpha_1' \lambda \alpha_1 = \lambda$, the eigenvector with the highest eigenvalue is chosen. The first principal component is then $\alpha_1' \mathbf{x}$. In our data, the highest eigenvalue takes the value of 2.48. The associated eigenvector explains 75% of the total variance.

References

- Acemoglu, D., S. Johnson, and T. Mitton, 2009: Determinants of Vertical Integration: Financial Development and Contracting Costs. *Journal of Finance*, **64**(3), 1251–1290.
- Acemoglu, D., S. Johnson, and J. Robinson, 2001: The colonial origins of comparative development: An empirical investigation. *American Economic Review*, **91**(5), 1369–1401.
- Aghion, P., Y. Algan, P. Cahuc, and A. Shleifer, 2010: Regulation and Distrust. *Quarterly Journal of Economics*, **125**(3), 1015–1049.
- Aghion, P. and J. Tirole, 1997: Formal and real authority in organizations. *Journal of Political Economy*, **105**(1), 1–29.
- Alchian, A. A. and H. Demsetz, 1972: Production, Information Costs, and Economic Organization. *American Economic Review*, **62**(5), 777–795.

- Algan, Y. and P. Cahuc, 2010: Inherited Trust and Growth. *American Economic Review*, **100**(5), 2060–92.
- Andreoni, J. and J. Miller, 2002: Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, **70**(2), 737–753.
- Arrow, K. J., 1968: The Economics of Moral Hazard: Further Comment. *American Economic Review*, **58**(3), pp. 537–539.
- Baker, G., R. Gibbons, and K. Murphy, 2002: Relational contracts and the theory of the firm. *Quarterly Journal of Economics*, **117**(1), 39–84.
- Banerjee, A., J. J. Dolado, J. W. Galbraith, and D. F. Hendry, 1993: *Co-integration, error correction, and the econometric analysis of non-stationary data*. Oxford University Press, Oxford.
- Banfield, E., 1958: *The Moral Basis of a Backward Society*. The Free Press, New York.
- Bårdsen, G., 1989: Estimation of Long Run Coefficients in Error Correction Models. *Oxford Bulletin of Economics and Statistics*, **51**(2), 345–350.
- Beck, T. and R. Levine, 2003: Legal Institutions and Financial Development. NBER Working Papers 10126, National Bureau of Economic Research.
- Becker, G. S. and K. M. Murphy, 1992: The Division of Labor, Coordination Costs, and Knowledge. *Quarterly Journal of Economics*, **107**(4), 1137–1160.
- Benfratello, L., F. Schiantarelli, and A. Sembenelli, 2008: Banks and innovation: Microeconomic evidence on Italian firms. *Journal of Financial Economics*, **90**(2), 197–217.
- Bloom, N., R. Sadun, and J. V. Reenen, 2012: The organization of firms across countries. Revise and resubmit, *Quarterly Journal of Economics*.
- Bürker, M. and G. A. Minerva, 2010: Civic Capital, Firms’ Size and Purchased Service Intensity in Italy. Paper presented at the 51st Annual Congress of the Italian Economic Association, University of Catania, Italy.
- Buonanno, P., D. Montolio, and P. Vanin, 2009: Does Social Capital Reduce Crime? *Journal of Law & Economics*, **52**(1), 145–170.
- Cameron, A. C. and D. L. Miller, 2010: Robust Inference with Clustered Data. Working Papers 107, University of California, Davis, Department of Economics.
- Campbell, J. and H. Hopenhayn, 2005: Market size matters. *Journal of Industrial Economics*, **53**(1), 1–25.
- Cartocci, R., 2007: *Mappe del tesoro: Atlante del capitale sociale in Italia*. Il Mulino, Bologna.
- Casadesus-Masanell, R., 2004: Trust in agency. *Journal of Economics & Management Strategy*, **13**(3), 375–404.
- Chandler, A., 1962: *Strategy and Structure*. MIT Press, Cambridge, MA.
- de Blasio, G. and G. Nuzzo, 2010: Historical traditions of civicness and local economic development. *Journal of Regional Science*, **50**(4), 833–857.
- Dearmon, J. and K. Grier, 2009: Trust and development. *Journal of Economic Behavior and Organization*, **71**(2), 210–220.
- Dessein, W., 2002: Authority and communication in organizations. *Review of Economic Studies*, **69**(4), 811–838.
- Falk, A. and U. Fischbacher, 2006: A theory of reciprocity. *Games and Economic Behavior*, **54**(2), 293–315.
- Fehr, E., 2009: On The Economics and Biology of Trust. *Journal of the European Economic Association*, **7**(2-3), 235–266.
- Fehr, E. and U. Fischbacher, 2002: Why Social Preferences Matter - The impact of Non-Selfish Motives on Competition, Cooperation and Incentives. *Economic Journal*, **112**(478), 1–33.

- Fehr, E. and K. Schmidt, 1999: A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, **114**(3), 817–868.
- Fukuyama, F., 1995: *Trust: The Social Virtues and the Creation of Prosperity*. Free Press, New York.
- Gorodnichenko, Y. and G. Roland, 2011: Culture, Institutions and The Wealth of Nations. Working paper, University of California, Berkeley.
- Greene, W. H., 1999: *Econometric analysis*. Prentice-Hall, Upper Saddle River, 4th edition.
- Guiso, L., P. Sapienza, and L. Zingales, 2004: The Role of Social Capital in Financial Development. *American Economic Review*, **94**(3), 526–556.
- Guiso, L., P. Sapienza, and L. Zingales, 2008a: Long Term Persistence. NBER Working Papers 14278, National Bureau of Economic Research.
- Guiso, L., P. Sapienza, and L. Zingales, 2008b: Social capital as good culture. *Journal of the European Economic Association*, **6**(2-3), 295–320.
- Guiso, L., P. Sapienza, and L. Zingales, 2011: Civic Capital as the Missing Link. In *Handbook of Social Economics*, Benhabib, J., Bisin, A., and Jackson, M. O., editors, volume 1. North-Holland, Princeton NJ, 417–480.
- Guner, N., G. Ventura, and Y. Xu, 2008: Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, **11**(4), 721–744.
- Hendry, D. F., A. R. Pagan, and D. J. Sargan, 1984: Dynamic specification. In *Handbook of Econometrics*, Griliches, Z. and Intriligator, M. D., editors, volume 2. North-Holland, Princeton NJ, 1023–1100.
- Holmes, T. J. and J. J. Stevens, 2002: Geographic Concentration and Establishment Scale. *Review of Economics and Statistics*, **84**(4), 682–690.
- Holmstrom, B., 1982: Moral Hazard in Teams. *Bell Journal of Economics*, **13**(2), 324–340.
- Hsieh, C.-T. and P. J. Klenow, 2009: Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, **74**(4), 1403–1448.
- Ichino, A. and G. Maggi, 2000: Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm. *Quarterly Journal of Economics*, **115**(3), 1057–1090.
- Inglehart, R. and W. Baker, 2000: Modernization, cultural change, and the persistence of traditional values. *American Sociological Review*, **65**(1), 19–51.
- Jacobs, J., 1969: *The economy of cities*. Vintage, New York.
- Jolliffe, I. T., 2002: *Principal Component Analysis*. Springer, New York.
- Knack, S. and P. Keefer, 1997: Does social capital have an economic payoff? A cross-country investigation. *Quarterly Journal of Economics*, **112**(4), 1251–1288.
- Kumar, K. B., R. G. Rajan, and L. Zingales, 1999: What Determines Firm Size? NBER Working Papers 7208, National Bureau of Economic Research.
- La Porta, R., F. L. de Silanes, A. Shleifer, and R. W. Vishny, 1997: Legal Determinants of External Finance. *Journal of Finance*, **52**(3), 1131–50.
- Laeven, L. and C. Woodruff, 2007: The Quality of the Legal System, Firm Ownership, and Firm Size. *Review of Economics and Statistics*, **89**(4), 601–614.
- Malanima, P., 2005: Urbanisation and the Italian economy during the last millenium. *European Review of Economic History*, **9**, 97–122.
- Melitz, M. J. and G. I. P. Ottaviano, 2008: Market size, trade, and productivity. *Review of Economic Studies*, **75**(1), 295–316.
- Penrose, E. T., 1959: *The Theory of the Growth of the Firm*. Oxford University Press, Oxford.
- Pinotti, P., 2012: Trust, Regulation and Market Failures. Technical report.

- Portes, A., 1998: Social Capital: Its origins and applications in modern sociology. *Annual Review of Sociology*, **24**, 1–24.
- Putnam, R., 1993: *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton.
- Rajan, R. G. and L. Zingales, 1998: Financial Dependence and Growth. *American Economic Review*, **88**(3), 559–86.
- Rossi-Hansberg, E. and M. L. J. Wright, 2007: Establishment Size Dynamics in the Aggregate Economy. *American Economic Review*, **97**(5), 1639–1666.
- Sapienza, P., A. Toldra Simats, and L. Zingales, 2007: Understanding Trust. CEPR Discussion Papers 6462, Centre for Economic Policy Research.
- Spagnolo, G., 1999: Social relations and cooperation in organizations. *Journal of Economic Behavior & Organization*, **38**(1), 1–25.
- Syverson, C., 2004: Market structure and productivity: A concrete example. *Journal of Political Economy*, **112**(6), 1181–1222.
- Tabellini, G., 2008: The scope of cooperation: Values and incentives. *Quarterly Journal of Economics*, **123**(3), 905–950.
- Tabellini, G., 2010: Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association*, **8**(4), 677–716.
- Williamson, O. E., 1979: Transaction-Cost Economics: The Governance of Contractual Relations. *Journal of Law and Economics*, **22**(2), 233–261.
- Williamson, O. E., 1985: *The Economic Institutions of Capitalism*. The Free Press, New York.
- Zak, P. J. and S. Knack, 2001: Trust and growth. *Economic Journal*, **111**(470), 295–321.

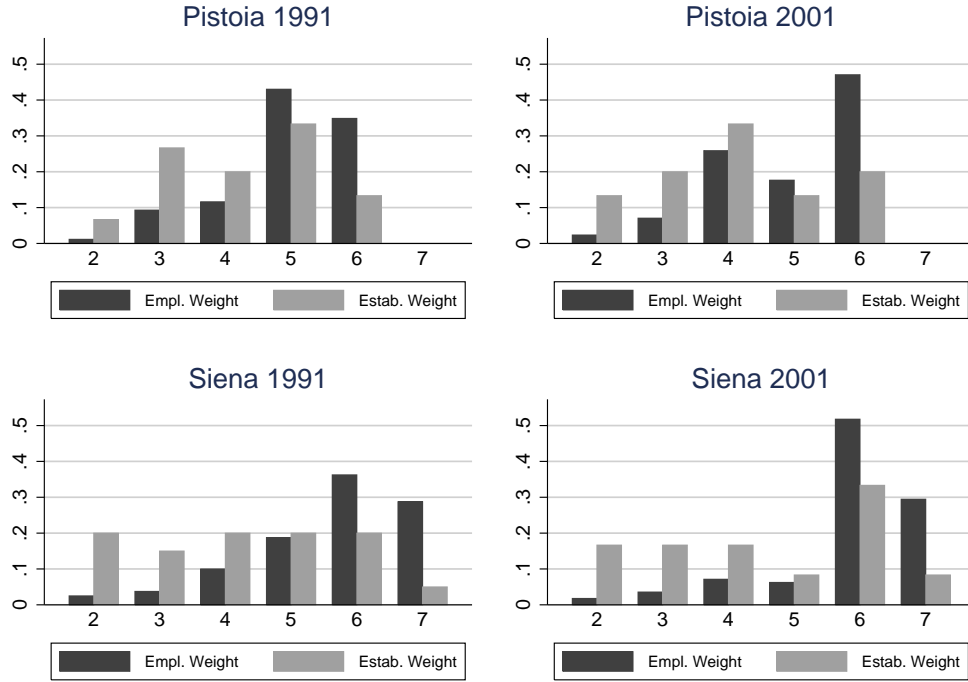


Figure 1: Empirical distribution of plant size in NACE industry 2862 (Manufacture of tools) in two provinces of Tuscany (Siena and Pistoia) in 1991 and 2001.

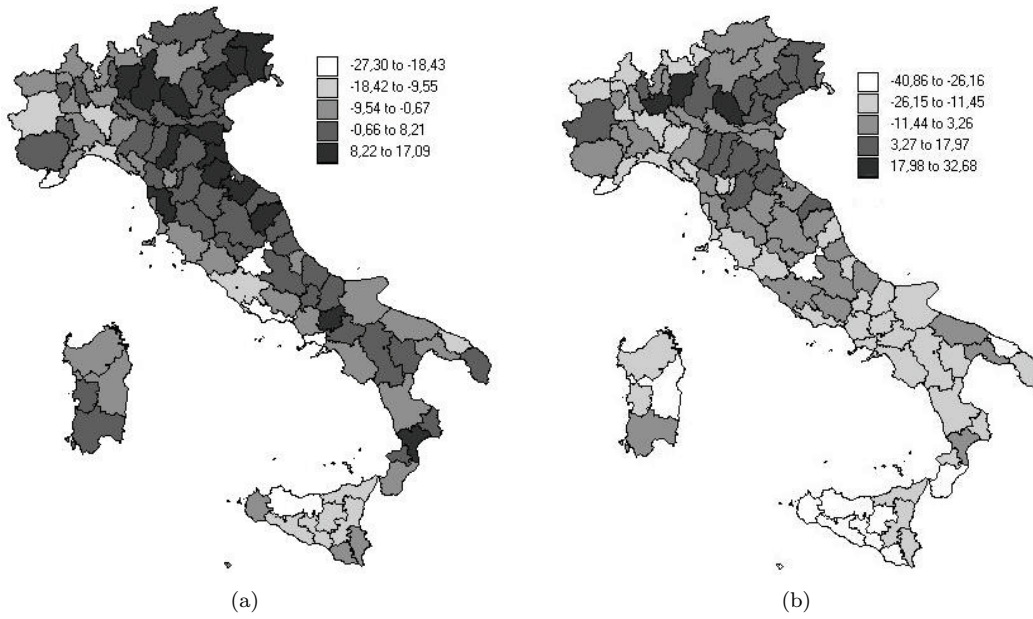


Figure 2: Growth in weighted average plant size in percentage points. Panel (a) are provincial averages over 4-digit industries. Panel (b) are provincial averages controlling for industry fixed effects and average size in 1991.

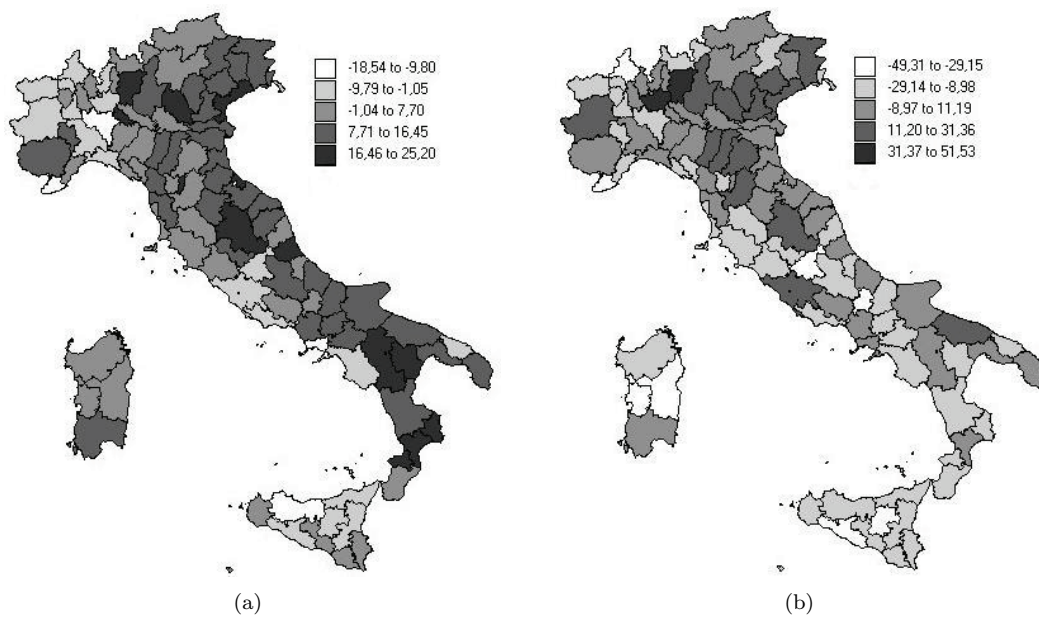


Figure 3: Growth in the standard deviation of plant size in percentage points. Panel (a) are provincial averages over 4-digit industries. Panel (b) are provincial averages controlling for industry fixed effects and standard deviation of size in 1991.

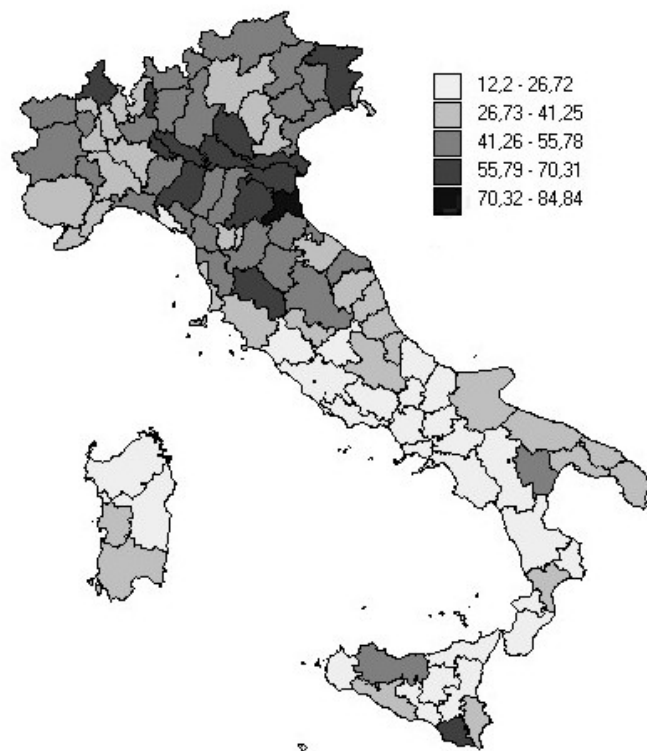


Figure 4: Blood donations per 1000 inhabitants.

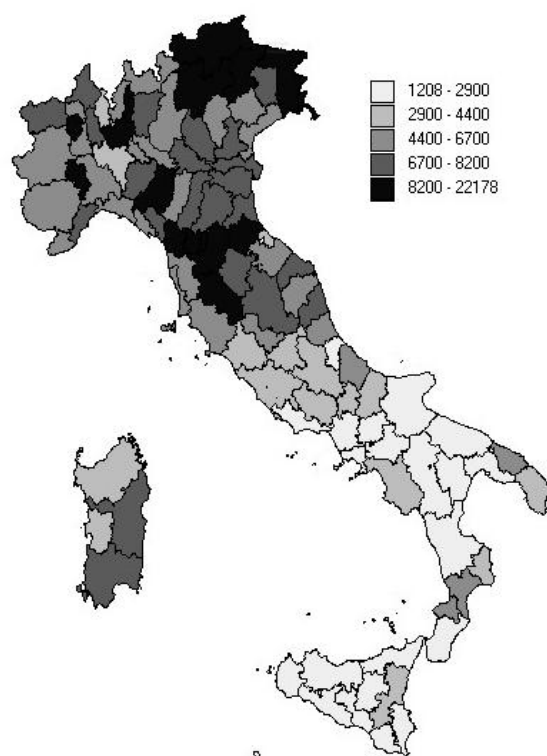


Figure 5: Number of volunteers in non-profit organizations per 100,000 inhabitants.

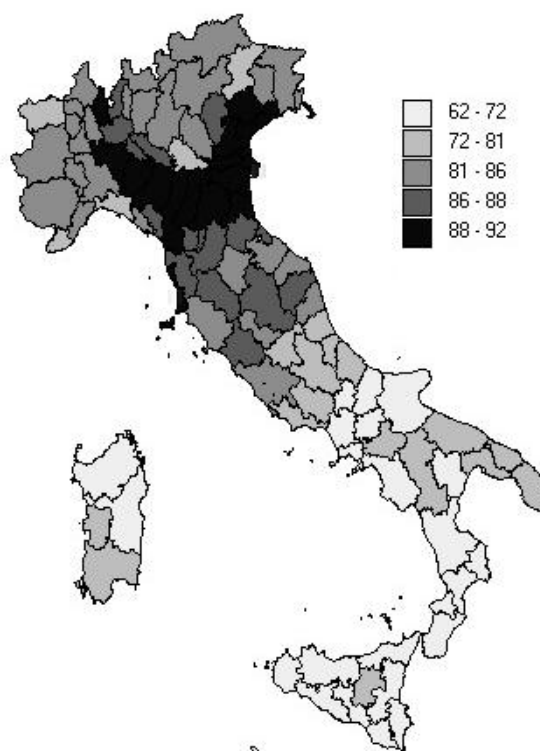


Figure 6: Electoral turnout in referenda, averaged over 7 referenda that took place between 1946 and 1987.

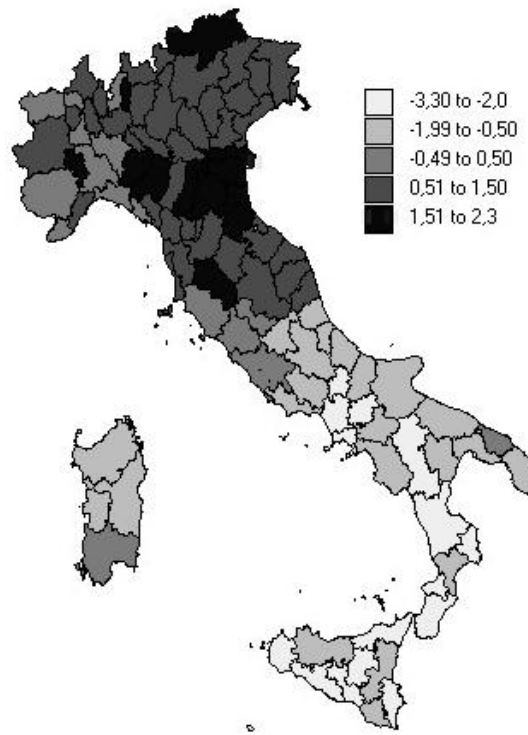


Figure 7: Map of civic capital measured by the first principal component of blood donations, volunteering, and electoral turnout.

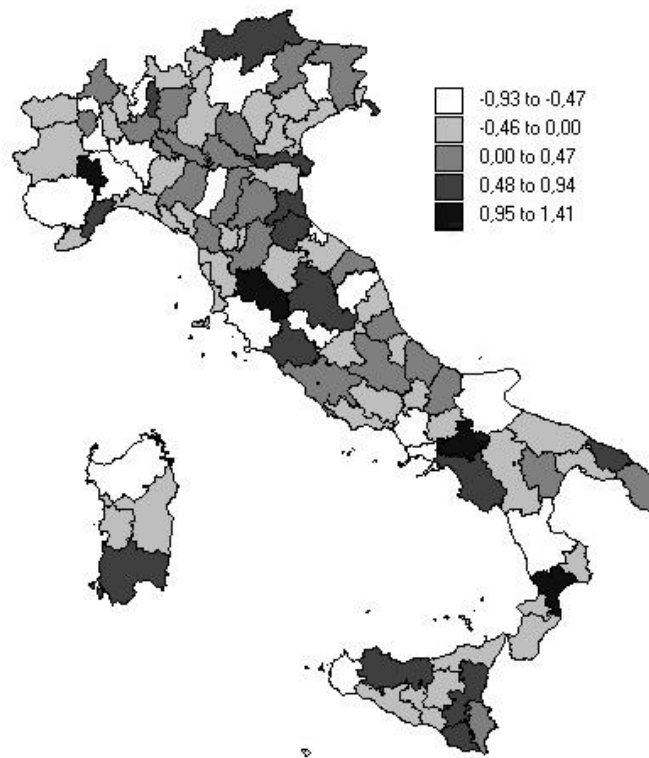


Figure 8: Map of civic capital measured by the variation of the first principal component within regions. The figure plots the residuals of a regression of the provincial principal component on regional dummies.

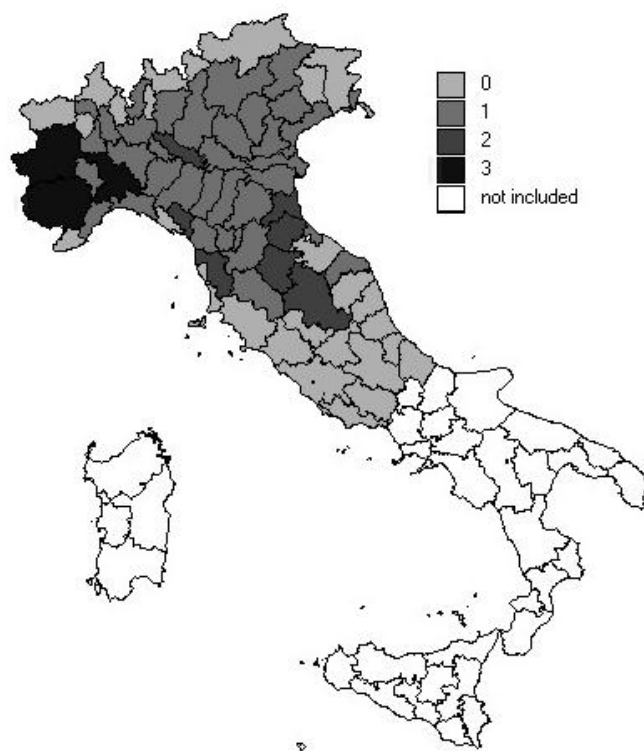


Figure 9: Number of free-city states by province in 1300. South of Italy and Islands are not included.

Table 1: Correlation among the proxy variables of civic capital

	Referenda turnout (log)	Volunteers (log)	Blood donations (log)
Volunteers (log)	0.69	1	
Blood donations (log)	0.61	0.57	1
Principal component	0.89	0.87	0.84

Note: The number of observations is 103. *Blood donations* is the log of the number of blood donations per 100,000 inhabitants in 2002; *Volunteers* is the log of the number of volunteers in non-profit institutions per 100,000 inhabitants in 2000; *Referenda turnout* is the log of the average electoral turnout in referenda between 1946 and 1987; *Principal component* is the first principal component of the above mentioned three proxies of civic capital. All correlations are statistically different from zero at the 1% level.

Table 2: Descriptive statistics

Variable	Obs.	Mean	S.D.	Min	Max	1st Quartile	Median	3rd Quartile
<i>Dependent variables (first differences 1991-2001)</i>								
Average Plant Size (log)	35,846	-0.008	1.018	-6.561	5.923	-0.439	0	0.435
Standard Deviation Plant Size (log)	30,112	0.059	1.074	-5.314	5.342	-0.495	0.049	0.624
F(9)	35,862	-0.008	0.258	-1	1	-0.087	0	0.05
F(19)	35,862	-0.009	0.207	-1	1	-0.026	0	0
F(49)	35,862	-0.007	0.147	-1	1	0	0	0
Size at 25th pct. (log)	35,709	-0.134	0.803	-8.356	7.149	-0.182	0	0
Size at 75th pct. (log)	35,845	-0.109	0.894	-6.315	5.796	-0.609	0	0.137
Interquartile range (log)	24,925	-0.036	0.915	-6.292	5.784	-0.539	0	0.335
Total employment (log)	35,846	0.053	1.026	-6.617	7.386	-0.405	0.025	0.502
Total establishments (log)	35,862	0.126	0.755	-3.932	4.779	-0.288	0.044	0.511
<i>Civic capital variables</i>								
Blood donations (log)	103	3.568	0.415	2.501	4.44	3.285	3.635	3.842
Volunteers (log)	103	8.533	0.627	7.097	10.007	8.146	8.599	8.928
Referenda turnout (log)	103	4.378	0.108	4.129	4.516	4.3	4.419	4.466
Principal component	103	0	1.5	-3.347	2.307	-1.231	0.439	1.13
<i>Historical instrumental variables</i>								
Aid societies in 1873 (log)	20	4.03	0.092	3.824	4.208	3.97	4.028	4.08
Turnout in 1920s (log)	20	6.247	0.976	4.159	7.465	5.408	6.54	7.071
Number of free-city states	67	0.85	0.78	0	3	0	1	1
<i>Other explanatory variables (level in 2001)</i>								
Final demand (log)	103	10.884	2.971	0	15.098	10.735	11.666	12.387
GDP (log)	103	15.793	0.798	14.051	18.565	15.234	15.709	16.245
University graduates (log)	103	1.854	0.175	1.455	2.453	1.733	1.846	1.935
Bank branches (log)	103	1.594	0.444	-0.661	2.327	1.274	1.722	1.914
Length of labor trials (log)	103	6.595	0.449	4.635	7.437	6.348	6.64	6.853
Urbanization in 1861 (log)	103	2.366	0.896	0	4.276	1.919	2.245	2.91

Note: The table provides descriptive statistics for some of the variables used in the regressions. The dependent variables (in first differences) are: *Average Plant Size* is the log of average size, with each size bin weighted by the corresponding share of employees; *Standard Deviation Plant Size* is the log of the standard deviation of plant size, with each size bin weighted by the corresponding share of employees; *F(9)* is the share of plants with at most 9 employees; *F(19)* is the share of plants with at most 19 employees; *F(49)* is the share of plants with at most 49 employees; *Size at 25th pct.* is the log of size at the 25th percentile of the distribution; *Size at 75th pct.* is the log of size at the 75th percentile of the distribution; *Interquartile range* is the log of the difference of size at the 25th and 75th percentiles; *Total employment* is the log of the total number of employees; *Total establishments* is the log of the total number of establishments. We consider the following variables to measure civic capital: *Blood donations* is the log of the number of blood donations per 100,000 inhabitants in 2002; *Volunteers* is the log of the number of volunteers in non-profit institutions per 100,000 inhabitants in 2000; *Referenda turnout* is the log of the average electoral turnout in referenda between 1946 and 1987; *Principal component* is the first principal component of the above mentioned three proxies of civic capital. The historical instruments are: *Aid societies in 1873* is the log of the members of mutual aid societies per 100,000 inhabitants in 1873; *Turnout in 1920s* is the log of turnout in the elections that took place in Italy in the 1920s; *Number of free-city states* is the number of free-city states in the territory of each province in 1300. The provincial controls (we report only their 2001 values to save on space) are: *Final demand* is the log of weighted population, with weights obtained from Input-Output Use tables; *University graduates* is the log of the share of university graduates over total population; *Bank branches* is the log of the number of bank branches per 1000 inhabitants; *Length of labor trials* is the log of the number of days it takes to complete a first-degree trial in labor-related affairs; *Urbanization in 1861* is the log of the share of inhabitants living in cities with at least 10,000 people in 1861.

Table 3: Growth in average plant size: baseline analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Blood Coef./se	Volunteers Coef./se	Referenda Coef./se	Principal comp. Coef./se	Blood Coef./se	Volunteers Coef./se	Referenda Coef./se	Principal comp. Coef./se	Unweighted size Coef./se
Blood donations (log)	0.045* (0.025)				0.044* (0.023)				
Volunteers (log)		0.056*** (0.017)				0.058*** (0.016)			
Referenda turnout (log)			-0.063 (0.194)				-0.029 (0.172)		
Principal component				0.039*** (0.012)				0.041*** (0.011)	0.025*** (0.008)
Weighted size 1991 (log)	-0.428*** (0.015)	-0.428*** (0.015)	-0.428*** (0.015)	-0.428*** (0.015)	-0.428*** (0.015)	-0.429*** (0.015)	-0.428*** (0.015)	-0.429*** (0.015)	
Unweighted size 1991 (log)									-0.423*** (0.016)
Final demand 1991 (log)	0.149*** (0.015)	0.151*** (0.015)	0.148*** (0.015)	0.154*** (0.015)	0.133*** (0.016)	0.135*** (0.015)	0.133*** (0.016)	0.137*** (0.015)	0.022*** (0.009)
Final demand growth					0.950*** (0.202)	1.016*** (0.204)	0.938*** (0.211)	1.021*** (0.190)	0.551*** (0.143)
Bank branches 1991 (log)	-0.022 (0.044)	-0.014 (0.040)	-0.018 (0.045)	-0.017 (0.041)	-0.014 (0.044)	-0.007 (0.038)	-0.007 (0.044)	-0.007 (0.040)	0.011 (0.027)
Bank branches growth					0.009 (0.010)	0.008 (0.011)	0.014 (0.011)	0.005 (0.009)	-0.004 (0.008)
University graduates 1991 (log)	-0.008 (0.038)	-0.009 (0.042)	0.002 (0.036)	-0.024 (0.041)	0.087 (0.057)	0.084 (0.056)	0.089 (0.057)	0.076 (0.055)	-0.025 (0.044)
University graduates growth					0.137 (0.105)	0.123 (0.104)	0.127 (0.108)	0.143 (0.101)	0.008 (0.072)
Length of trials 1995 (log)	0.008 (0.017)	0.015 (0.017)	0.009 (0.018)	0.012 (0.016)	0.074** (0.032)	0.070** (0.031)	0.079** (0.035)	0.067** (0.030)	0.030 (0.023)
Length of trials growth					0.055** (0.027)	0.042 (0.027)	0.058** (0.029)	0.043* (0.026)	0.019 (0.021)
Urbanization in 1861 (log)	-0.001 (0.012)	0.004 (0.012)	0.000 (0.013)	0.002 (0.012)	0.007 (0.011)	0.012 (0.010)	0.008 (0.011)	0.009 (0.010)	-0.004 (0.007)
Long-run parameter k_2					0.104* (0.054)	0.136*** (0.038)	-0.067 (0.402)	0.096*** (0.026)	0.058*** (0.020)
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (within)	0.210	0.210	0.210	0.210	0.211	0.211	0.211	0.211	0.235
Obs.	28,694	28,694	28,694	28,694	28,694	28,694	28,694	28,694	28,766

Note: The table provides the results of the fixed effect estimator where the panel variable are 4-digit industries. 20 regional dummies are also included. Standard errors are clustered at the province and industry level. Standard errors of the long-run parameters are based on the delta method. The dependent variable is the growth of $APSI_{i,t}$, weighted average plant size, except for the last column, where growth in unweighted average size is employed. We use the following variables to measure civic capital: *Blood donations* is the log of the number of blood donations per 1000 inhabitants in 2002; *Volunteers* is the log of the number of volunteers in non-profit institutions per 100,000 inhabitants in 2000; *Referenda turnout* is the log of the average electoral turnout in referenda between 1946 and 1987; *Principal component* is the first principal component of the above mentioned three proxies of civic capital. We also include the following provincial controls: *Final demand* is the log of weighted population, with weights obtained from Input-Output Use tables; *University graduates* is the log of the share of university graduates over total population; *Bank branches* is the log of the number of bank branches per 1000 inhabitants; *Length of trials* is the log of the number of days it takes to complete a first-degree trial in civil affairs; *Urbanization in 1861* is the log of the share of inhabitants living in cities with at least 10,000 people in 1861. ***, **, * denote significance at the 1%, 5%, 10% level, respectively.

Table 4: Growth in the standard deviation of plant size: baseline analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Blood Coef./se	Volunteers Coef./se	Referenda Coef./se	Principal comp. Coef./se	Blood Coef./se	Volunteers Coef./se	Referenda Coef./se	Principal comp. Coef./se	Std. dev. unweight. Coef./se
Blood donations (log)	0.047 (0.029)				0.048* (0.027)				
Volunteers (log)		0.062*** (0.021)				0.066*** (0.018)			
Referenda turnout (log)			0.035 (0.217)				0.083 (0.186)		
Principal component				0.045*** (0.013)				0.049*** (0.012)	0.041*** (0.011)
Standard dev. weight. size 1991 (log)	-0.464*** (0.015)	-0.464*** (0.014)	-0.464*** (0.014)	-0.464*** (0.015)	-0.465*** (0.015)	-0.465*** (0.014)	-0.465*** (0.015)	-0.465*** (0.015)	-0.439*** (0.015)
Standard dev. unweight. size 1991 (log)									
Final demand 1991 (log)	0.271*** (0.017)	0.273*** (0.017)	0.271*** (0.017)	0.276*** (0.017)	0.249*** (0.017)	0.251*** (0.017)	0.250*** (0.017)	0.254*** (0.017)	0.128*** (0.014)
Final demand growth					1.318*** (0.234)	1.393*** (0.236)	1.312*** (0.243)	1.403*** (0.215)	1.016*** (0.181)
Bank branches 1991 (log)	-0.020 (0.047)	-0.011 (0.042)	-0.014 (0.048)	-0.015 (0.043)	-0.013 (0.046)	-0.006 (0.041)	-0.003 (0.047)	-0.012 (0.043)	-0.000 (0.036)
Bank branches growth					0.005 (0.012)	0.003 (0.012)	0.009 (0.013)	-0.001 (0.010)	0.003 (0.009)
University graduates 1991 (log)	-0.039 (0.042)	-0.040 (0.046)	-0.034 (0.042)	-0.058 (0.044)	0.108* (0.059)	0.105* (0.058)	0.107* (0.061)	0.095* (0.056)	0.051 (0.053)
University graduates growth					0.230** (0.113)	0.215* (0.115)	0.223* (0.117)	0.238** (0.110)	0.184* (0.101)
Length of trials 1995 (log)	-0.001 (0.022)	0.007 (0.023)	0.001 (0.023)	0.003 (0.021)	0.078** (0.037)	0.074** (0.037)	0.084** (0.041)	0.070** (0.034)	0.062** (0.029)
Length of trials growth					0.063* (0.032)	0.049 (0.031)	0.067* (0.035)	0.049 (0.030)	0.044* (0.025)
Urbanization in 1861 (log)	0.004 (0.014)	0.009 (0.013)	0.005 (0.014)	0.007 (0.013)	0.014 (0.012)	0.019* (0.011)	0.016 (0.012)	0.017 (0.010)	0.009 (0.009)
Long-run parameter k_2					0.103* (0.057)	0.143*** (0.039)	0.178 (0.401)	0.106*** (0.026)	0.094*** (0.024)
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (within)	0.229	0.229	0.229	0.229	0.230	0.230	0.230	0.230	0.221
Obs.	27,995	27,995	27,995	27,995	27,995	27,995	27,995	27,995	28,047

Note: The table provides the results of the fixed effect estimator where the panel variable are 4-digit industries. 20 regional dummies are also included. Standard errors are clustered at the province and industry level. Standard errors of the long-run parameters are based on the delta method. The dependent variable is the growth of $SDPS_{i,t}$, the weighted standard deviation, except for the last column, where growth in unweighted standard deviation is employed. We use the following variables to measure social capital: *Blood donations* is the log of the number of blood donations per 100,000 inhabitants in 2002; *Volunteers* is the log of the number of volunteers in non-profit institutions per 100,000 inhabitants in 2000; *Referenda turnout* is the log of the average electoral turnout in referenda between 1946 and 1987; *Principal component* is the first principal component of the above mentioned three proxies of civic capital. We also include the following provincial controls: *Final demand* is the log of weighted population, with weights obtained from Input-Output Use tables; *University graduates* is the log of the share of university graduates over total population; *Bank branches* is the log of the number of bank branches per 1000 inhabitants; *Length of trials* is the log of the number of days it takes to complete a first-degree trial in labor-related affairs; *Urbanization in 1861* is the log of the share of inhabitants living in cities with at least 10,000 people in 1861. ***, **, * denote significance at the 1%, 5%, 10% level, respectively.

Table 5: The evolution of the size distribution of plants: further analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	F(9)	F(19)	F(49)	25th pct.	50th pct.	75th pct.	Interquart. range	Total empl.	Total establ.
	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se
Principal component	-0.007*** (0.002)	-0.007*** (0.002)	-0.003** (0.002)	0.009* (0.005)	0.015 (0.010)	0.024** (0.011)	0.028** (0.014)	0.020* (0.010)	-0.005 (0.008)
F(9) - cdf threshold 9 employees 1991	-0.615*** (0.014)								
F(19) - cdf threshold 19 employees 1991		-0.551*** (0.015)							
F(49) - cdf threshold 49 employees 1991			-0.478*** (0.015)						
Size of plants at 25th pct. 1991 (log)				-0.842*** (0.014)					
Size of plants at 50th pct. 1991 (log)					-0.688*** (0.013)				
Size of plants at 75th pct. 1991 (log)						-0.587*** (0.014)			
Interquartile range 1991 (log)							-0.601*** (0.014)		
Total number of workers 1991 (log)								-0.244*** (0.015)	
Total number of plants 1991 (log)									-0.221*** (0.013)
Final demand 1991 (log)	-0.040*** (0.003)	-0.027*** (0.003)	-0.018*** (0.002)	-0.023*** (0.007)	-0.018* (0.011)	0.005 (0.011)	0.004 (0.014)	0.253*** (0.019)	0.226*** (0.015)
Final demand growth	-0.264*** (0.045)	-0.178*** (0.036)	-0.081*** (0.031)	0.198* (0.110)	0.398** (0.200)	0.646*** (0.228)	0.679** (0.295)	1.370*** (0.232)	0.881*** (0.215)
Long-run parameter k_2	-0.012*** (.004)	-0.013*** (0.004)	-0.007** (0.003)	0.011* (0.006)	0.022 (0.015)	0.041** (0.018)	0.046** (0.023)	0.081* (0.042)	-0.024 (0.035)
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other prov. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (within)	0.322	0.274	0.222	0.467	0.355	0.303	0.308	0.131	0.156
Obs.	28,858	28,842	28,842	28,762	28,789	28,753	23,037	28,782	28,748

Note: The table provides the results of the fixed effect estimator where the panel variable are 4-digit industries. 20 regional dummies are also included. Standard errors are clustered at the province and industry level. Standard errors of the long-run parameters are based on the delta method. The dependent variable in column (1) is the share of plants with less than 10 employees, in column (2) with less than 20 employees, in column (3) with less than 50 employees, in column (4) it is the size of plants at the 25th percentile of the size distribution, in column (5) at the 50th percentile, in column (6) at the 75th percentile, in column (7) it is the interquartile range, in column (8) it is the log of total employment, and in column (9) it is the log of the total number of establishments. See previous tables for a description of other variables. ***, **, * denote significance at the 1%, 5%, 10% level, respectively.

Table 6: Robustness and sensitivity analysis

	Growth of average size				Growth of standard deviation of size							
	(1) Manuf. Coef./se	(2) Services Coef./se	(3) Two var. Coef./se	(4) GDP Coef./se	(5) Full sample Coef./se	(6) No TO,MI,RM Coef./se	(7) Manuf. Coef./se	(8) Services Coef./se	(9) Two var. Coef./se	(10) GDP Coef./se	(11) Full sample Coef./se	(12) No TO,MI,RM Coef./se
Princ. comp.	0.066*** (0.017)	0.028** (0.013)		0.037*** (0.010)	0.036*** (0.013)	0.039*** (0.011)	0.070*** (0.019)	0.037** (0.015)		0.044*** (0.011)	0.046*** (0.013)	0.048*** (0.012)
Princ. comp. (two var.)			0.034*** (0.009)						0.038*** (0.010)			
Av. size 1991 (log)	-0.359*** (0.014)	-0.534*** (0.026)	-0.429*** (0.015)	-0.429*** (0.015)	-0.401*** (0.013)	-0.439*** (0.015)						
Std. dev. 1991 (log)							-0.386*** (0.015)	-0.567*** (0.021)	-0.465*** (0.015)	-0.466*** (0.015)	-0.454*** (0.014)	-0.478*** (0.014)
Final demand 1991 (log)	0.095*** (0.022)	0.193*** (0.020)	0.135*** (0.015)	-0.014 (0.053)	0.151*** (0.014)	0.127*** (0.017)	0.179*** (0.027)	0.341*** (0.022)	0.251*** (0.017)	0.057 (0.057)	0.257*** (0.017)	0.248*** (0.020)
Final demand growth	1.557*** (0.324)	0.619** (0.275)	1.003*** (0.192)	0.802*** (0.202)	0.893*** (0.213)	1.144*** (0.186)	1.880*** (0.355)	1.043*** (0.356)	1.378*** (0.222)	1.048*** (0.251)	1.488*** (0.228)	1.524*** (0.213)
GDP 1991 (log)				0.148*** (0.050)						0.192*** (0.054)		
GDP growth				-0.016 (0.147)						0.126 (0.173)		
Long-run parameter k_2	0.183*** (0.049)	0.053** (0.024)	0.079*** (0.021)	0.085*** (0.024)	0.090*** (0.032)	0.090*** (0.024)	0.182*** (0.049)	0.065** (0.026)	0.082*** (0.021)	0.093*** (0.024)	0.101*** (0.028)	0.101*** (0.025)
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other prov. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.194	0.244	0.211	0.212	0.191	0.217	0.212	0.261	0.230	0.231	0.224	0.237
Obs.	10,601	15,393	28,694	28,694	35,482	27,496	10,392	14,979	27,995	27,995	29,803	26,805

Note: The table provides the results of the fixed effect estimator where the panel variable are 4-digit industries. 20 regional dummies are also included. Standard errors are clustered at the province and industry level. Standard errors of the long-run parameters are based on the delta method. The dependent variable is the growth of $APSt_{i,j,t}$ in columns (1)-(5), and the growth rate of $SDPS_{i,j,t}$ in columns (6)-(10). The variable employed to measure civic capital in all the columns except (3) and (8) is the first principal component of the log of the number of blood donations per 100,000 inhabitants, the log of the number of volunteers in non-profit institutions per 100,000 inhabitants in 2000, the log of the average electoral turnout in referenda between 1946 and 1987. In columns (3) and (8) civic capital is the first principal component of blood donations and volunteers in non-profit institutions. In columns (4) and (10) we include as an additional provincial control GDP , the log of provincial gross domestic product. See previous tables for a description of other variables. ***, **, * denote significance at the 1%, 5%, 10% level, respectively.

Table 7: 2SLS regressions of the growth of APS and SDPS

	Panel A: Second stage									
	Growth of average size					Growth of standard deviation of size				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Principal component	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se
Weighted size 1991 (log)	0.056*** (0.011)	0.058** (0.026)	0.068*** (0.021)	0.057*** (0.013)	0.090*** (0.028)	0.065*** (0.012)	0.049* (0.028)	0.070*** (0.014)	0.066*** (0.014)	0.109*** (0.030)
Standard dev. weight. size 1991 (log)	-0.428*** (0.015)	-0.428*** (0.015)	-0.428*** (0.015)	-0.401*** (0.016)	-0.401*** (0.016)	-0.465*** (0.015)	-0.465*** (0.015)	-0.466*** (0.015)	-0.436*** (0.015)	-0.436*** (0.015)
Final demand 1991 (log)	-0.042 (0.050)	-0.040 (0.051)	-0.032 (0.049)	0.006 (0.047)	0.044 (0.053)	0.026 (0.055)	0.013 (0.061)	0.031 (0.055)	0.073 (0.054)	0.123** (0.062)
Final demand growth	0.570*** (0.207)	0.579*** (0.214)	0.607*** (0.205)	0.663*** (0.235)	0.760*** (0.249)	0.802*** (0.254)	0.755*** (0.275)	0.818*** (0.255)	0.863*** (0.298)	0.991*** (0.295)
GDP 1991 (log)	0.174*** (0.048)	0.172*** (0.048)	0.165*** (0.046)	0.146*** (0.044)	0.113*** (0.051)	0.218*** (0.054)	0.229*** (0.059)	0.213*** (0.053)	0.188*** (0.051)	0.145*** (0.058)
GDP growth	0.217 (0.140)	0.212 (0.145)	0.192 (0.146)	0.108 (0.164)	-0.004 (0.187)	0.376** (0.152)	0.408** (0.159)	0.364** (0.158)	0.208 (0.176)	0.058 (0.197)
Long-run parameter k_2	0.130*** (0.026)	0.137** (0.060)	0.159*** (0.049)	0.143*** (0.033)	0.225*** (0.073)	0.140*** (0.026)	0.106* (0.060)	0.151*** (0.051)	0.151*** (0.032)	0.250*** (0.072)
Macro-regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other prov. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (within)	0.210	0.210	0.210	0.193	0.192	0.229	0.229	0.229	0.210	0.210
Obs.	28,694	28,694	28,694	19,737	19,737	27,995	27,995	27,995	19,355	19,355
Panel B: First stage (Dependent var. is Principal component)										
Weighted size 1991 (log)	0.012 (0.008)	0.012 (0.008)	0.012 (0.008)	0.006 (0.005)	0.006 (0.005)	0.015** (0.007)	0.015** (0.007)	0.014** (0.007)	0.015** (0.007)	0.007* (0.004)
Standard dev. weight. size 1991 (log)	-3.079 (2.249)	-3.498* (2.039)	-3.498* (2.039)	-4.808** (2.028)	-4.808** (2.028)	-3.061 (2.247)	-3.061 (2.247)	-3.491* (2.036)	-3.491* (2.036)	-4.802** (2.029)
Final demand 1991 (log)	-0.655 (0.519)	-0.627 (0.449)	-0.627 (0.449)	-0.868** (0.380)	-0.868** (0.380)	-0.661 (0.527)	-0.661 (0.527)	-0.637 (0.456)	-0.637 (0.456)	-0.880** (0.384)
GDP growth	2.423* (1.306)	2.557** (1.236)	2.557** (1.236)	3.506** (1.390)	3.506** (1.390)	2.450* (1.307)	2.450* (1.307)	2.586** (1.234)	2.586** (1.234)	3.541** (1.395)
GDP 1991 (log)	0.518 (0.509)	0.509 (0.437)	0.509 (0.437)	0.804** (0.358)	0.804** (0.358)	0.515 (0.518)	0.515 (0.518)	0.510 (0.443)	0.510 (0.443)	0.813** (0.360)
Aid societies in 1873 (log)	0.876*** (0.192)	0.469** (0.221)	0.469** (0.221)	3.796*** (1.235)	3.796*** (1.235)	0.876*** (0.193)	0.876*** (0.193)	0.467** (0.222)	0.467** (0.222)	0.876*** (0.193)
Turnout in 1920s (log)	0.295* (0.148)	0.295* (0.148)	0.295* (0.148)	0.758*** (0.169)	0.758*** (0.169)	0.295* (0.148)	0.295* (0.148)	0.758*** (0.169)	0.758*** (0.169)	0.293* (0.148)
1 free-city state in 1300	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)
2 free-city states in 1300	0.293* (0.148)	0.293* (0.148)	0.293* (0.148)	0.758*** (0.169)	0.758*** (0.169)	0.293* (0.148)	0.293* (0.148)	0.758*** (0.169)	0.758*** (0.169)	0.293* (0.148)
3 free-city states in 1300	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)	-0.361 (0.277)
Macro-regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other prov. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.836	0.836	0.836	0.601	0.601	0.836	0.836	0.836	0.836	0.602
F-stat. of excluded instruments	20.76	13.29	13.29	8.41	8.41	20.68	20.68	13.19	13.19	8.31
Obs.	28,694	28,694	28,694	19,737	19,737	27,995	27,995	27,995	27,995	19,355

Note: The table provides the results of the fixed effect two-stages least squares estimator where the panel variable are 4-digit industries. 5 macro-regional dummies are also included. Standard errors are clustered at the province and industry level. Panel A shows the second stage, while Panel B shows the first stage. In the first stage we use two sets of instrumental variables. The first set of instrumental variables are: *Aid societies in 1873* is the log of the members of mutual aid societies per 100,000 inhabitants in 1873; *Turnout in 1920s* is the log of turnout in the elections that took place in Italy in the 1920s. The second set of instrumental variables are dummies for the number of free-city states at the provincial level in 1300. ***, **, * denote significance at the 1%, 5%, 10% level, respectively.

Table 8: 2SLS regressions of the growth of exact points of the size distribution

	Panel A: Second stage											
	F(9)			F(19)			F(49)			25th pct.		
	(1) Coef./se	(2) Coef./se	(3) Coef./se	(4) Coef./se	(5) Coef./se	(6) Coef./se	(7) Coef./se	(8) Coef./se	(9) Coef./se	(10) Coef./se	(11) Coef./se	(12) Coef./se
Principal component	-0.017*** (0.005)	-0.023*** (0.007)	-0.008** (0.004)	-0.009* (0.005)	-0.005* (0.003)	-0.004 (0.003)	0.014 (0.010)	0.029* (0.016)	0.052** (0.020)	0.074** (0.031)	0.037 (0.026)	0.074** (0.032)
F(9) - cdf threshold 9 employees 1991	-0.601*** (0.014)	-0.571*** (0.016)										
F(19) - cdf threshold 19 employees 1991			-0.536*** (0.015)	-0.512*** (0.016)								
F(49) - cdf threshold 49 employees 1991					-0.468*** (0.015)	-0.452*** (0.015)						
Size of plants at 25th pct. 1991 (log)							-0.842*** (0.014)	-0.830*** (0.016)				
Size of plants at 75th pct. 1991 (log)									-0.587*** (0.014)	-0.577*** (0.015)		
Interquartile range 1991 (log)											-0.601*** (0.014)	-0.590*** (0.015)
Long-run parameter k_2	-0.028*** (0.008)	-0.040*** (0.013)	-0.015** (0.007)	-0.017* (0.009)	-0.011* (0.006)	-0.008 (0.008)	0.016 (0.012)	0.035* (0.020)	0.088** (0.034)	0.128** (0.053)	0.061 (0.042)	0.125** (0.055)
Macro-regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Provincial controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (within)	0.311	0.282	0.261	0.238	0.210	0.200	0.466	0.436	0.301	0.285	0.307	0.293
Obs.	28,379	19,476	28,555	19,615	28,699	19,732	28,762	19,778	28,753	19,771	23,037	16,450
Panel B: First stage (Dependent var. is Principal component)												
F(9) - cdf threshold 9 employees 1991	-0.077** (0.037)	-0.042* (0.022)										
F(19) - cdf threshold 19 employees 1991			-0.075* (0.041)	-0.039 (0.028)								
F(49) - cdf threshold 49 employees 1991					-0.036 (0.055)	-0.028 (0.043)						
Size of plants at 25th pct. 1991 (log)							-0.001 (0.014)	0.015 (0.010)				
Size of plants at 75th pct. 1991 (log)									0.021* (0.011)	0.006 (0.008)		
Interquartile range 1991 (log)											0.019* (0.010)	0.008 (0.008)
Aid societies in 1873 (log)	0.466** (0.221)		0.468** (0.221)		0.470** (0.221)		0.470** (0.221)		0.469** (0.221)		0.453** (0.223)	
Turnout in 1920s (log)	3.818*** (1.237)		3.806*** (1.236)		3.800*** (1.236)		3.804*** (1.237)		3.800*** (1.235)		3.771*** (1.231)	
1 free-city state in 1300		0.296* (0.148)		0.297** (0.148)		0.295** (0.148)		0.294* (0.148)		0.297** (0.148)		0.290* (0.146)
2 free-city states in 1300		0.758*** (0.170)		0.760*** (0.169)		0.760*** (0.169)		0.760*** (0.169)		0.759*** (0.169)		0.749*** (0.169)
3 free-city states in 1300		-0.356 (0.279)		-0.359 (0.278)		-0.361 (0.278)		-0.360 (0.278)		-0.361 (0.278)		-0.361 (0.276)
Macro-regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Provincial controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.851	0.603	0.851	0.602	0.851	0.602	0.851	0.602	0.851	0.602	0.843	0.595
F-stat. of excluded instruments	13.22	8.34	13.28	8.45	13.30	8.47	13.30	8.46	13.32	8.45	13.33	8.27
Obs.	28,379	19,476	28,555	19,615	28,699	19,732	28,762	19,778	28,753	19,771	23,037	16,450

Note: The table provides the results of the fixed effect two-stages least squares estimator where the panel variable are 4-digit industries. 5 macro-regional dummies are also included. Standard errors are clustered at the province and industry level. Panel A shows the second stage, while Panel B shows the first stage. In the first stage we use two sets of instrumental variables. The first set of instrumental variables are: *Aid societies in 1873* is the log of the members of mutual aid societies per 100,000 inhabitants in 1873; *Turnout in 1920s* is the log of turnout in the elections that took place in Italy in the 1920s. The second set of instrumental variables are dummies for the number of free-city states at the provincial level in 1300. ***, **, * denote significance at the 1%, 5%, 10% level, respectively.