



Alma Mater Studiorum - Università di Bologna  
DEPARTMENT OF ECONOMICS

**Explaining the Size Distribution of Plants:  
An Approach Based on Civic Capital**

Matthias Bürker  
Gaetano Alfredo Minerva

*Quaderni - Working Paper DSE N° 755*



# Explaining the Size Distribution of Plants: An Approach Based on Civic Capital \*

Matthias Bürker<sup>†</sup> and G. Alfredo Minerva<sup>‡</sup>

May 31, 2011

## Abstract

We show that the distribution of plant size within narrowly defined industries is affected by the variation in the stock of civic capital that occurs at the provincial level. Data on plant size come from the 2001 Italian Census of Manufacturing and Services. Civic capital turns out to have a positive effect on both the average and the standard deviation of the plant size distribution. This effect is stronger in labor-intensive industries. The potential endogeneity of current civic capital is addressed by instrumenting it with historical variables. Our interpretation for these results is that civic capital is associated with reduced opportunistic behavior, which improves intra-firm cooperation and hampers the incidence of principal-agent problems, thus allowing plants to operate on a larger scale.

**JEL Classification:** D23, Z13, L20, A13

**Keywords:** Agency problems; Civic capital; Cooperation; Opportunism; Plant size distribution.

---

\*We are grateful to Paolo Buonanno, Robert Putnam, and Paolo Vanin for having provided us some of the data used in this paper. We also thank for their comments Erich Battistin, Jeff Butler, Guido de Blasio, Stefano Federico, 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.

<sup>†</sup>Department of Economics, University of Bologna. E-mail: matthias.burker@unibo.it

<sup>‡</sup>Corresponding author. Department of Economics, University of Bologna. E-mail: ga.minerva@unibo.it

# 1 Introduction

What determines the size distribution of economic organizations in different industries? Oligopoly theory stresses the role of market size: in large markets, competitive pressures are intense and reduce the mark-up of firms. In order to recover the fixed costs of production, each firm has to operate on a larger scale, which increases average size. Similarly, models of monopolistic competition with heterogeneous firms such as Melitz and Ottaviano (2008) predict that market size increases both average size and the standard deviation of size. Empirical evidence has confirmed these theoretical predictions. Employing data on 13 retail industries in the US, Campbell and Hopenhayn (2005) show that market size has a magnifying effect on average establishment size. Moreover, they find that the dispersion in plant size is positively related to the size of the local market in some retail industries, but not in all of them. Similarly, Syverson (2004) shows that, for the concrete industry in the US, market size increases plant size.

In this paper, we empirically identify another fundamental determinant of the size distribution of establishments: the stock of local civic capital.<sup>1</sup> Employing Italian 2001 Census data, we show that civic capital allows economic organizations to operate on a larger scale, leading to higher average plant size. Moreover, our results document an increase in the dispersion of establishment size. We motivate these findings by an important property of civic capital in that it serves as a check against opportunistic behavior and hence increases the likelihood of cooperative outcomes in collective endeavors. Therefore, civic capital plays a crucial role in intra-firm relationships, where a smooth flow of information is required for decision making, where CEOs have to efficiently coordinate a wide variety of tasks, where decisions have to be communicated in a transparent way, and diverging individual interests need to be bridged for the sake of the firm as a whole. In short, the internal efficiency of organizations depends crucially on the extent to which individuals cooperate. In principle, there exist several mechanisms which sustain cooperation, such as relational contracts, external enforcement, or kinship bonds. However, interactions within firms often take place in contexts where (i) external enforcement is hampered by unobservability of actions or problems of performance measurement, (ii) interactions are not repeated frequently, and (iii) agents are not related to their partner by personal bonds. In those situations cooperation among individuals hinges crucially on civic capital and its associated high level of trust. In addition, the importance of civic capital in sustaining cooperation depends on the size of the organization in that the scope for opportunistic defection increases with the number of workers. Standard models of moral hazard predict that the probability of shirking increases with team size (Holmstrom, 1982), as each member gets a lower share of output. Moreover, relational contracting is less feasible because interactions between any two members of an organization become less frequent.

Our choice of focusing on Italian data bears several important advantages. First, the variation in the endowment of civic capital across Italy has received ample interest in social sciences (Banfield, 1958; Putnam, 1993; Guiso et al., 2004, 2008a). This allows us to circumvent the contentious measurement issue, as we can fall back on a range of well-established proxies which are available at a geographically disaggregated level. Specifically, we build on Guiso et al. (2004) and measure civic capital by the number of blood donations and electoral turnout in referenda.<sup>2</sup> The extent to which individuals engage in volunteering in non-profit

---

<sup>1</sup>Although conceptually closely related, we avoid using the loaded term of social capital. See section 2 for a definition of civic capital and its relationship with trust.

<sup>2</sup>Other studies which use these measures include Buonanno et al. (2009), and de Blasio and Nuzzo (2010).

organizations completes the set of measures. Second, working on Italy allows us to use data from the 8th Census on Industry and Services (2001) which provides detailed information on the 4.4 millions plants operating in Italy. Hence, our analysis does not suffer from sample selection issues, provided that we are working with the universe of Italian establishments. Third, by focusing on one country, other factors typically affecting plant size like government subsidies, employment regulation or tax regimes are held constant.

In order to claim a causal impact of civic capital on economic outcomes, it is crucial to overcome several potential sources of endogeneity. The first problem concerns measurement error. None of our three proxies accurately reflects the true stock of civic capital. In order to minimize measurement error and the resulting attenuation bias, we extract the first principle component of the three variables of civic capital.

The second problem is the issue of reversed causality. According to the modernization theory, economic development has a direct impact on the formation of cultural and social values, thereby determining the level of civic capital (Inglehart and Baker, 2000). Similarly, trust in others is likely to be triggered by trust in political and legal institutions (Rothstein, 2000). Accordingly, efficient law enforcement both produces high levels of civic capital and favors economic activity, resulting in large plants. Similarly to Buonanno et al. (2009) and de Blasio and Nuzzo (2010), we use historical data to overcome the problem of reversed causality and omitted variables. In particular, we use information on civic capital at the middle of the 19th and beginning of the 20th century. Moreover, relying on Putnam (1993) and Guiso et al. (2008a), a second set of instruments is based on the number of free-city states in the territory of a province during the Middle Age.

Several controls ensure that our results are not driven by confounding factors. To separate civic capital from confounding unobservables related to socio-economic development, we exploit only its variation at the provincial level within the 20 Italian regions, by including a dummy for each region. In addition, we interpret our data set as a panel, where observations pertaining to a given industry are collected repeatedly over different provinces. Industry-specific heterogeneity is accordingly removed by four-digit Nace industry fixed effects. Moreover, based on Input-Output Use tables, we construct a proxy for local final demand. The quality of contract enforcement, human capital, financial development, urbanization rates in 1861, and GDP, all defined at the province level, complete the set of controls.

Results reveal that the variation of civic capital has a magnifying impact on both the mean and the dispersion of the plant size distribution. Consistently with this, we find that the number of large plants is higher in high-civic capital areas. We attribute this increase 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 discuss at length these mechanisms below. Empirically, the effect of civic capital is heterogeneous across different industries: we find that the impact is stronger in industries where the capital-labor ratio is low. Concerning the economic magnitude of our results, the model predicts that the difference in the endowment of civic capital between the provinces of Naples (low civic capital) and Milan (high civic capital) is associated with a difference in the average size of plants of 25%.

The positive relationship between civic capital and the size of organizations is a priori not straightforward, though. In principle, there are at least two mechanisms why civic capital could actually be associated with

a *decrease* in average plant size. On the one hand, the influential 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, market transactions become more attractive relative to in-house production, and so some activities are outsourced to external suppliers and the average size of plants could decrease.<sup>3</sup> On the other hand, civic capital is positively related to entrepreneurship. Tabellini (2010) shows that the prevailing cultural traits in Southern Italy are such that they discourage individual initiative and effort, which is supposed to hinder the foundation of new firms.<sup>4</sup> Accordingly, de Blasio and Nuzzo (2010) and Percoco (2011) show that in the Italian context civic capital increases entrepreneurship, which could then decrease the average size of plants, provided that newly established firms are smaller.

All in all, the Census data we employ unambiguously reveal that the overall effect of civic capital on the average size of organizations is positive, despite the fact that the mechanisms just mentioned point to a possible negative effect. This is one contribution of our paper.

Another contribution is that we identify a novel determinant that affects the size distribution of plants, beside local market size. Moreover, while papers like Syverson (2004) and Campbell and Hopenhayn (2005) limit their analysis to a narrow set of industries, our study establishes a regularity which holds for the whole spectrum of economic activity.

Next, whereas the literature on the determinants of firm size has stressed the role of formal institutions such as the quality of contract enforcement or the protection of property rights (Kumar et al., 1999; Laeven and Woodruff, 2007; Lu and Tao, 2009), our analysis implies that informal institutions matter directly in this respect.

Finally, the idea that informal institutions such as trust are important for the functioning of large organizations has been formulated by Fukuyama (1995). However, he does not corroborate his claim with some strong empirical evidence. La Porta et al. (1997) provide a cross-country analysis of the relative importance of large organizations and the level of trust. They analyze the sales of the top twenty largest firms relative to national GDP. Differently from them, our analysis focuses on the absolute size of the universe of plants in Italy. Bloom et al. (2009) are closer to our approach, since they show, in a cross-country analysis, that the average of firm size for manufacturing as a whole, measured at the regional level in terms of NUTS 2 spatial units, is larger in regions where trust is higher. Our paper substantially improves their analysis in terms of precision in at least two ways. First, we are able to assess how the size distribution within each single 4-digit Nace industry in the Italian Census changes according to the variation in the level of trust. Second, we do it at a very detailed spatial scale (provinces, which correspond to NUTS 3 spatial units).

The rest of the paper is organized as follows. 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 5 and 6 provide

---

<sup>3</sup>In a related study (Bürker and Minerva, 2010) we show 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.

<sup>4</sup>Similarly, in his illustrative study on social and economic life in a small village in Southern Italy, Banfield (1958) has observed a mentality of the local population characterized by resignation and helplessness.

robustness analysis and two-stages least square results, respectively. Finally, section 7 concludes.

## 2 Conceptual framework

### 2.1 The concept of civic capital

Guiso et al. (2010) 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”.<sup>5</sup> Two immediate and interrelated predictions of civic capital 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. The literature has stressed that the decision to trust another group or person depends on both the belief in the trustworthiness of others and on individual preferences (Fehr, 2009; Sapienza et al., 2007). 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. Similarly, to the extent that civic capital is associated with widespread pro-social preferences, this should also lead to high trust (Fehr, 2009). Given the interrelatedness between civic capital and trust, in what follows we will use the two terms interchangeably.<sup>6</sup>

### 2.2 Civic capital, trust and the size of organizations

The reasoning above suggests that civic capital improves overall organizational efficiency of firms by facilitating cooperation and reducing the scope of agency problems and dilemmas of collective actions, which are centered around opportunism. 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”.<sup>7</sup>

In this section we outline three specific mechanisms of how civic capital improves cooperation in intra-firm relationships such that it increases establishment size. However, our analysis does not allow to make a

---

<sup>5</sup>This 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).

<sup>6</sup>It 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 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.

<sup>7</sup>See p. 538.

qualitative assessment of their relative importance. Rather, the aim is to show that our empirical findings are well-grounded in the predictions of some theoretical models. 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 the delegation of authority**

The first mechanism of how civic capital translates into larger 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 delegation. 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 delegation. 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.<sup>8</sup>

### **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.

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

---

<sup>8</sup>Chandler (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.

coordination costs. In a framework like Becker and Murphy (1992), civic capital can be thought to reduce shirking and coordination costs, and to allow for 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 from future trade 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).<sup>9</sup>

### 2.2.4 Testable implications

These considerations allow us to derive our main working hypotheses. First of all, the discussion above makes clear why we expect that, in areas where the stock of civic capital is high, plants should be larger. As to the standard deviation of average size, the way it may react to civic capital is very much in line with the way market size increases the standard deviation in the theoretical model of Melitz and Ottaviano (2008). If in areas where civic capital is high we have a larger average size of plants, this may also bring an increase in the standard deviation due to the direct magnifying effect on the size distribution. We will describe more concretely this issue in the discussion of the empirical results.

The second implication of our framework is that we expect some heterogeneity in the impact of civic capital on the plant size distribution across different industries. Specifically, we expect industries where there is an intense interaction among workers to be more reactive to the stock of local civic capital. The reason simply lies on the fact that the mechanisms of how civic capital affects relationships at the level of plants are firmly grounded in interpersonal contacts. Conversely, we expect plants where the production process is based on standardized routines and automation to be less sensitive to civic capital.<sup>10</sup>

## 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

---

<sup>9</sup>When explaining differences in the size of firms across countries, Fukuyama (1995) stresses this important property of trust.

<sup>10</sup>Automation is the use of control systems and information technologies to reduce the need for the intervention of human work in the production of goods and services.



which are owned by private entities, and we exclude establishments owned by non-profit organizations and public institutions.<sup>11</sup>

Concerning the geographic disaggregation, we work with the 103 Italian provinces existing in 2001.<sup>12</sup> We work with provinces, and not with smaller spatial units, because data on civic capital are not available at a smaller level of geography. However, working with provinces is particularly convenient provided that it allows us to have a sufficiently large number of observations for each 4-digit industry.<sup>13</sup>

### 3.1 Plant size distribution

The data on plant size come from the 8th Census of Industry and Services carried out by the Italian national statistic authority (Istat) in 2001, covering the universe of Italian establishments. In the Census, plants are classified into 12 different size bins, according to the size of the plant in terms of workers.<sup>14</sup> 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.<sup>15</sup>

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), the first dependent variable that we define is the weighted average of plant size in 4-digit industry  $i$  and province  $j$ . In formal terms we compute average plant size (APS) as:

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

where  $N_{b,i,j}^{emp}$  is the total number of workers in bin  $b$ , industry  $i$ , and province  $j$ ,  $N_{b,i,j}^{estab}$  is the total number of plants in a given bin-province-industry, while  $N_{i,j}^{emp}$  is the total number of workers in industry  $i$  and province  $j$  for all size bins. Hence, in this formula we weight the average size per bin, as 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. 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 province-industry, and to weaken the impact on average plant size of a large number of small plants.

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:

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

This amounts to a simple division of the total number of workers by the total number of plants for each province-industry combination. We refer to  $\overline{APS}_{i,j}$  as the simple average plant size. Moreover, we will also look at the effect of civic capital on total employment in the province-industry,  $N_{i,j}^{emp}$ , on the total

<sup>11</sup>This 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.

<sup>12</sup>Italian provinces correspond to the NUTS 3 partitioning.

<sup>13</sup>The average number of observations 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.

<sup>14</sup>The 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).

<sup>15</sup>The information about plant size in the Istat Census is similar to the one provided in the County Business Patterns for the U.S. County Business Patterns data are used, for example, in Holmes and Stevens (2002).

number of establishments,  $N_{i,j}^{estab}$ , and on the number of establishments above or below the 20 employees size threshold.

The other feature of the size distribution of plants in a given province-industry that we want to characterize is the degree of dispersion around the mean. For this reason, we consider a set of regressions where the dependent variable is the standard deviation of weighted plant size in industry  $i$  and province  $j$ . It is defined as:

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

This variable puts more weight in the computation of dispersion on the size bins that carry out the bulk of economic activity in a given province-industry. The counterpart of (2) which will be employed in the robustness section is:

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

Finally, notice that 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.<sup>16</sup>

In order to highlight the difference between the weighting schemes employed, Figure 1 provides a graphical representation of the empirical size distribution of plants. The two graphs show the size distribution for the Manufacture of Tools (Nace rev. 1 code 2862) in two provinces of Tuscany, Siena and Pistoia.

[Insert Figure 1 about here]

Size bins are plotted on the horizontal axis, with the value on top of the bars indicating average plant size in that bin; that is,  $N_{b,i,j}^{emp}/N_{b,i,j}^{estab}$ . 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). By putting more weight on large plants, the employee-based weighting scheme increases both the average and standard deviation of plant size. For example, 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 idea behind our paper is to assess whether, for provinces belonging to the same region (say, Tuscany) and the same 4-digit industry (say, Manufacture of Tools) we can establish if part of the change in the plant size distributions like those of Figure 1 can be attributed to the variation in the stock of civic capital at the provincial level.

The maps displayed in Figure 2 and 3 show the geographic variation of  $APS$  and  $SDPS$  across Italian provinces.

[Insert Figures 2, 3 about here]

Not surprisingly,  $APS$  is the highest in the metropolitan areas of Milan, Rome and Turin. Moreover, the map reveals that average plant size is generally larger in the Northern part of the country. This pattern is even more evident in the case of the dispersion in plant size, as shown in Figure 3. The three

<sup>16</sup>The 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.

metropolitan areas just mentioned stand out, and the higher frequency of light-gray-colored provinces in the North reveals a substantial North-South difference in the dispersion of plant size. The only three light-gray-colored provinces in the South pertain to the metropolitan areas of Naples, Bari and Palermo.

### 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, and electoral participation in referenda over the period between 1946 and 1987.<sup>17</sup> 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.<sup>18</sup> 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.<sup>19</sup> 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 higher in the Central and Northern part of the country.

[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.<sup>20</sup>

[Insert Table 1 about here]

In order to remove the noise, we look at the common component of the three proxies. The last row in Table 1 shows the correlation of the three proxies with their first principal component.<sup>21</sup> The fact that the

---

<sup>17</sup>In Appendix 8.1 we provide a detailed description of the variables and their sources.

<sup>18</sup>For example, donating blood imposes a substantial physical limitation in the short-run, voting requires information gathering and personal evaluation of alternatives.

<sup>19</sup>See 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).

<sup>20</sup>For 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 population 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.

<sup>21</sup>In Appendix 8.2 we briefly review how to derive the first principal component.

correlation with each of the individual trust variables is strong, with only slight differences in magnitude, suggests 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.<sup>22</sup>

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 regions in the Center-North, like Emilia-Romagna, and the lowest in the Southern mainland and Sicily. However, 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. 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.<sup>23</sup> 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).<sup>24</sup> In section 6 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.<sup>25</sup> 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 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).

<sup>22</sup>Another example where the first principal component is used to summarize common cultural traits at the regional level is Tabellini (2010).

<sup>23</sup>For a detailed description consult the Appendix 8.1 and Putnam (1993).

<sup>24</sup>The 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 area that was under the Holy Roman Empire at the beginning of the second Millennium (basically, the Center-North of Italy). For this reason, in the analysis with this set of instruments we are forced to drop observations from the South Italy and from the Islands (Sicily and Sardinia).

<sup>25</sup>A detailed description of the data and the corresponding sources is provided in Appendix 8.1

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.<sup>26</sup> The second measure is provincial gross domestic product.

The degree of urbanization could be correlated with the plant size distribution: even after controlling for the intensity of final demand, plant size could still be the outcome of the overall degree of urbanization.<sup>27</sup> Correlation between the plant size distribution and urbanization could also go through the different geographic characteristics of the provinces, even within the same region. This translates 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.<sup>28</sup>

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. Working on Mexican regions, Laeven and Woodruff (2007) show that firm size is increasing with the quality of the legal system. We proxy the quality of the legal system computing the average number of days it took to complete first-degree trials in labor-related affairs which ended in 2001, in each of the 165 Italian labor courts.

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.<sup>29</sup> 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, we investigate whether the impact of civic capital on the plant size distribution is heterogeneous across industries. To this end, we characterize the production process of each industry according to the intensity in human interactions. To quantify this dimension we have picked up the capital-labor ratio at the four-digit industry level. We believe that this is a good proxy to measure the sensitivity of a given industry to human relations, or, symmetrically, the extent to which it depends on standardized routines and process automation.

In Table 2 we provide the full set of descriptive statistics for our data.

[Insert Table 2 about here]

---

<sup>26</sup>See Appendix 8.1 for further details.

<sup>27</sup>A classical reference linking plant growth and urbanization is Jacobs (1969).

<sup>28</sup>It 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.

<sup>29</sup>However 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.

## 4 The effect of civic capital: OLS results

### 4.1 Average plant size

We begin the regression analysis with the ordinary least squares estimation of the relationship between average plant size and civic capital. The equation that we estimate is the following:

$$\ln APS_{i,j} = \alpha_0 + \alpha_1 \ln CC_j + \ln X_j' \bar{\alpha}_2 + \gamma_r + \gamma_i + \epsilon_{i,j} \quad (5)$$

where  $\ln APS_{i,j}$  is the logarithm of employment-weighted average plant size in province  $j$  and industry  $i$ ,  $\ln CC_j$  is the log of the measure of civic capital in province  $j$ ,  $\ln X_j$  is the log of a vector of provincial controls,  $\gamma_r$  denotes the region's dummy,  $\gamma_i$  is a 4-digit unobserved industry effect, and  $\epsilon_{i,j}$  is the error term.<sup>30</sup> We apply to equation (5) 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 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.<sup>31</sup>

Results confirm our hypothesis that the level of civic capital is positively correlated with average plant size. In columns from (1) to (4) of Table 3 we regress 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. In addition to the region dummy variables, the only control that is added at the provincial level in these estimates is the strength of final demand. Also final demand turns out to affect positively average 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).

In columns from (5) to (8) of Table 3 we include additional determinants of plant size at the provincial level. In order to claim that civic capital increases plant size by increasing trust and cooperation at the plant level, we have to exclude that results are driven by certain characteristics of the environment in which a plant operates that are correlated with both civic capital and the plant size distribution. These confounding factors are dealt with by including appropriate provincial controls, described in section 3.3.

Turning to the results, they show little sensitivity to the inclusion of controls. This is partially reassuring, though of course it does not completely solve the issue of unobserved heterogeneity. With the exception of

---

<sup>30</sup>When we measure civic capital by the principal component our equation is the following:

$$\ln APS_{i,j} = \alpha_0 + \alpha_1 PC_j + \ln X_j' \bar{\alpha}_2 + \gamma_r + \gamma_i + \epsilon_{i,j}.$$

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

<sup>31</sup>The 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  $\gamma_i$ . 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).

referenda turnover, both the number of blood donations and that of volunteers are still statistically significant at the 5% level. The same is true for the first principal component, which is statistically significant at the 1% level. It should be kept in mind that the coefficient  $\alpha_1$  is actually an elasticity, because both the dependent variable and the regressors are logarithms. Then, in the case of blood donations, an estimated elasticity equal to 0.07 means that a 10% increase in the number of blood donations at the provincial level is associated to a 0.7% increase in weighted average size of plants at the level of 4-digit industries. The value is very similar 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,  $\gamma_r$ , there is not enough variability left in provincial turnout to allow a coefficient estimate which is statistically different from zero.<sup>32</sup> 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. The estimated coefficient equal to 0.05 implies that average plant size would be 25% larger if Naples had the stock of civic capital of Milan.<sup>33</sup> This is a sizeable impact given that the observed difference is such that plants are, on average, 47% larger in Milan.

Among the provincial controls, a coefficient being different from zero in a statistically significant way is the number of bank branches. The fact that we get a negative sign can be explained by a certain prevalence of younger and, for this reason, smaller plants where financial development is higher.<sup>34</sup> In our estimates the role of civic capital seems to outweigh that of formal institutions, since the quality of contract enforcement, expressed by the length of trials, is no longer statistically different from zero.

## 4.2 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 equation that we estimate is now the following:

$$\ln SDPS_{i,j} = \alpha_0 + \alpha_1 \ln CC_j + \ln X_j' \bar{\alpha}_2 + \gamma_r + \gamma_i + \epsilon_{i,j} \quad (6)$$

where  $\ln SDPS_{i,j}$  is the logarithm of the standard deviation of plant size in industry  $i$  and province  $j$ , computed according to equation (3). As before, we perform a fixed effect analysis where the panel dimension is in terms of 4-digit industries.

The results provided in Table 4, in the columns from (1) to (4), are obtained with just one provincial control (final demand). In columns from (5) to (8) we add the full set of controls. The following results

---

<sup>32</sup>When the proxy of civic capital is the first principal component of the above mentioned three variables, it is less straightforward to provide an interpretation in terms of elasticity of the coefficient of civic capital. As we outline in Appendix 8.2, 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 particularly 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.

<sup>33</sup>The 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 ( $\ln APS_{i,Milan} - \ln APS_{i,Naples}$ ) is then equal to 0.229, which corresponds to a difference of 25% in terms of average size measured in levels.

<sup>34</sup>However, this control variable could be endogenous in equation (5). Guiso et al. (2004) stress the importance of civic capital for the development of financial markets in the Italian context.

stand out. As before, all proxies of civic capital are statistically significant, with the exception of electoral turnout in referenda. Moreover, there is little sensitivity of the estimates with respect to the inclusion of the full set of provincial controls. The elasticity of the standard deviation with respect to civic capital is 0.08 for blood donations and for volunteers.

The results suggest that the impact of civic capital, although smaller in magnitude, is qualitatively similar to that of market demand: it increases both the first and the second moment of the distribution of plant size. As to the other controls, financial development, measured by the number of bank branches, is associated to a less dispersed distribution in terms of size (probably due to a larger number of small plants in the province-industry, something which makes the distribution more even).

### 4.3 A graphical discussion of the results

In order to discuss our results, we go back to Figure 1, where the distribution for NACE 2862 in Siena and Pistoia is plotted. 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 highlights that civic capital increases the number of plants at the top end of the distribution: comparing the height of the dark gray bar in size bin number 6 reveals that relatively large plants are more frequent in Siena than in Pistoia. Moreover, Siena hosts establishments which have no size counterpart in Pistoia (there is no plant in size bin number 7 in Pistoia). We do not find such a pattern in the left tail: in size bins number 2, 3, and 4 (where the total number of employees is small, ranging from 1 to 5), the distribution is actually fatter in Pistoia than in Siena.<sup>35</sup>

The mechanism we have in mind to explain why civic capital increases both average plant size and the standard deviation of plant size is coherent with this stylized example. Our argument hinges upon a larger mass of big plants being active in the local areas with high civic capital. This happens because of the different channels outlined in the conceptual framework, and gives rise to a magnification of the local plant size distribution which ultimately increases also the standard deviation of size. In the following section we provide further evidence on this.

## 5 Extensions and robustness of the analysis

### 5.1 Heterogeneity across industries, and small vs. large plants

In Table 5, we show that the role of civic capital for the first and second moment of the plant size distribution is not homogeneous across industries. In section 2 we have stressed the importance of civic capital for team production and cooperative behavior in one-shot interactions. Consequently, we expect civic capital to be more important in industries where the production process is more intensive in human interactions. To test this, we have added an interaction term to our baseline specification where the first principal component of

---

<sup>35</sup>Qualitatively, we get the same results looking at the light gray bars, where the weights are given by the number of plants that each size bin contains.



civic capital is interacted with the capital-labor ratio measured at the 4-digit industry level.<sup>36</sup> Accordingly, the equation for APS becomes

$$\ln APS_{i,j} = \alpha_0 + \alpha_1 PC_j + \alpha_2 PC_j \times \ln KL_i + \ln X_j' \bar{\alpha}_3 + \gamma_r + \gamma_i + \epsilon_{i,j} \quad (7)$$

where  $PC_j$  is the value of the principal component in province  $j$ , and  $KL_i$  is the capital-labor ratio of industry  $i$ . The corresponding results are shown in columns (1) and (4) of Table 5. The negative estimate for the coefficient  $\alpha_2$  of the interaction between the capital-labor ratio and the measurement of civic capital is in line with our reasoning: the magnitude of the impact of civic capital on the plant size distribution decreases as the industry's capital-labor ratio goes up. The explanation is that problems of coordination and shirking behavior, which civic capital alleviates, are more important for output production in sectors with a low degree of automation and a high intensity of human interactions.

In order to get a more precise picture of the heterogeneous effect of civic capital across industries we evaluate the marginal effects for different values of the capital-labor ratio. The industry with the median value of capital-labor intensity is Manufacture of ceramic tiles and flags ( $\ln KL_i = 4.034$ ). The estimate of the marginal effect of civic capital on APS in this specific industry is positive and equal to 0.052.<sup>37</sup> Performing an F-test on the linear restriction  $(\alpha_1 + \alpha_2 \times 4.034) = 0$  reveals that this effect is statistically different from zero at the 1% level. Repeating this exercise, we find that the marginal effect of civic capital is no longer different from zero at the 5% level only in the case of the top 15% industries in terms of the capital-labor ratio. Hence, this exercise shows that civic capital seems to have no impact for the size of plants only in industries which are very intensive in the use of capital.

Another 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. Results are again displayed in Table 5, in columns (2), (3), (5) and (6). For both manufacturing and services, civic capital increases average plant size and the standard deviation. The effect seems to be slightly more pronounced for manufacturing. It is interesting to observe that the coefficients of the control variables such as historical urbanization, university graduates and bank branches display substantial differences in both sign and significance<sup>38</sup> among the two sectors. However, we can conclude that the positive effect of civic capital on both APS and SDPS is not driven by a particular sector, be it manufacturing or services.

We then perform another exercise, testing, in two separate regressions, the impact of civic capital on the log of total employment in a province-industry, and on the log of the total number of establishments. The corresponding results, displayed in columns (7) and (8) of Table 5, provide evidence that there is a positive effect on total employment, while the effect on the number of establishments is not statistically different from zero. Taking the last regression as a starting point, we then ask whether there is some difference

<sup>36</sup>See Appendix 8.1 for a description of the way the capital-labor ratio is obtained.

<sup>37</sup>To obtain the estimate of the impact of civic capital in a specific industry one simply needs the value for  $\ln KL_i$ , and the estimates for  $\alpha_1$  and  $\alpha_2$ . The value is then equal  $(\alpha_1 + \alpha_2 \times \ln KL_i)$ .

<sup>38</sup>Urbanization in 1861 and university graduates have a statistically significant positive coefficient in the regression for services. This makes sense from an economic point of view, as long as larger plants in the service sector are located in more urbanized areas, and service provision is comparatively more intensive in skilled workers, so that larger service providers can be expected to locate where many college graduates live. The number of bank branches is negatively related to APS and SDPS only in manufacturing.

between small and large plants. Our conceptual framework suggests 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) should be more important for the functioning of large plants. On the contrary, in small establishments, cooperation can be sustained by other mechanisms such as direct monitoring or kinship bonds. Moreover, in small plants interactions between workers become more frequent, so that relational contracts are an efficient enforcement device. Therefore, we expect that especially the number of large plants should be raised by civic capital. In order to address this question, we count for each province-industry the number of plants with less than twenty employees (small plants), and those with at least twenty employees (large plants).<sup>39</sup> The results are presented in columns (9) and (10) of Table 5, and point to a positive effect of civic capital which is statistically different from zero in the case of large plants only. The point estimate is larger when the dependent variable is the number of big plants.<sup>40</sup> This is in line with our conceptual framework.

## 5.2 Further robustness and sensitivity checks

Other robustness checks that we have performed are provided in Table 6. The first five columns present the result for APS, while the last five present the result for SDPS. In columns (1) and (6) the dependent variable is the average plant size of equation (2), and the standard deviation of equation (4), respectively. Although somewhat reduced in magnitude, the effect of civic capital is still positive and significant. We can conclude that our results are not driven by the choice of the weighting scheme in the construction of the dependent variables.

In columns (2) and (7) 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. Again, the estimated coefficients are just slightly smaller in magnitude than in the baseline regressions.

In columns (3) and (8) we add yet another provincial control, namely GDP in 2001. 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 (4) and (9) 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 actual change is tiny.

Finally, in columns (5) and (10) we run regressions after excluding the provinces of Rome, Milan and Turin. As we show in Figures 2 and 3, these are the provinces with the highest average size and the highest

---

<sup>39</sup>The choice of this threshold is arbitrary. We have tried with other size thresholds such as 50 employees and the results do not change.

<sup>40</sup>Since the number of plants above or below a certain size threshold in a province-industry is a count variable, this analysis is suitable for a Poisson regression. In this case, the routine in Stata allows only for one-way clustering in the standard errors (in terms of 4-digit industries). This notwithstanding, the results are qualitatively close to the one presented in the table, and point to an impact of civic capital which is larger in magnitude in the case of large plants, with the coefficient still being statistically different from zero at the 1% level.

standard deviation of size. Therefore, it could be that our results are driven by them. Our estimates prove that their exclusion from the sample bears no particular change to our results.

## 6 The effect of civic capital: evidence from 2SLS estimation

### 6.1 Motivation and identifying assumptions

Apart from 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 us from interpreting results in Tables 3 and 4 as causal. 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, similarly to Buonanno et al. (2009) and de Blasio and Nuzzo (2010). 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 (1993)'s theory, Guiso et al. (2008a) shows that the free-city state experience has 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 equations (5) and (6) for *APS* and *SDPS*. 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.

## 6.2 Results

### 6.2.1 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 a 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.

Using lagged proxy variables of civic capital we again 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.<sup>41</sup> 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.<sup>42</sup> We also run a robustness check in which we instrument civic capital only by the membership rate in mutual aid societies in 1873.

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:

$$PC_j = \omega_0 + \omega_1 \ln CC_r^{Aid} + \omega_2 \ln CC_r^{Turn} + \ln X_j' \bar{\omega}_3 + \gamma_{mr} + \epsilon_j, \quad (8)$$

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 elections in the 1920s. These two terms are meant to capture the persistent component of civic capital. Since

---

<sup>41</sup>It 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.

<sup>42</sup>However, 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.

these measures are only available at the regional level we cannot include regional dummy variables. Instead, we replace them with macro-region dummies,  $\gamma_{mr}$ .<sup>43</sup> The vector  $\ln X'_j$  contains all provincial covariates, including provincial GDP.

The bottom part of Table 7 (Panel B) contains the results from the first stage. Column (2) refers to APS, while (6) refers to SDPS. 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  $R^2$ . 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 (8). 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.<sup>44</sup>

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-region instead of region dummies. The corresponding results are shown in columns (1) and (5). 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. For both APS and SDPS the 2SLS estimates for civic capital are larger than their OLS counterparts. Given our concerns of reversed causality and omitted variables, this is not necessarily what we expected. Actually, if results in Tables 3 and 4 were driven by reversed causality, then the 2SLS coefficients should have been lower than the OLS coefficients from columns (1) and (5). Two possible reasons might explain the increase of 2SLS estimates with respect to the OLS ones. First, despite the computation of the principal component, civic capital could still be measured with error. The 2SLS helps removing the attenuation bias. Second, the increase in the estimates may be a sign of violation of the exclusion restriction. Given that we have more than one instrument we can perform a test of overidentification to assess this issue. The p-value of the test statistic is reported at the bottom of Panel A. It is sufficiently high in order 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.<sup>45</sup>

In columns (3) and (7) we consider an alternative specification where the membership rate in mutual aid societies is the only instrument for our two-stages approach. The first stage relationship remains strong. As for the second stage result, there is hardly any change with respect to columns (2) and (6).

### 6.2.2 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

<sup>43</sup>The macro-regions are five: North-West, North-East, Center, South, Islands.

<sup>44</sup>Actually, we get a negative relationship between historical urbanization and the current level of civic capital in the first stage. This mirrors the negative correlations that we have shown above between historical urbanization and past civic capital.

<sup>45</sup>The 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.

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.

We exploit information on free-city states in 1300 provided by Guiso et al. (2008a). More specifically, for each province, we count the number of cities that were a communal republic in the year 1300, which gives a minimum of zero and a maximum of three free-city states per province.<sup>46</sup> Figure 9 shows the number of free-city states per province across Italy. The information is only available for the Center-North, so we exclude the South and the Islands from this exercise.

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 is the following:

$$PC_j = \omega_0 + \omega_1 City_j^1 + \omega_2 City_j^2 + \omega_3 City_j^3 + \ln X_j' \bar{\omega}_4 + \gamma_{mr} + \epsilon_j \quad (9)$$

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. The macro-region dummies and provincial covariates are also included. Columns (4) and (8) in Panel B show that also these instruments are good predictors for the current level of civic capital. As expected, ceteris paribus, provinces with one free-city state in 1300 display a higher level of civic capital than provinces with no free-city states. The impact of having two free-city states in 1300 is even larger. Performing an F-test reveals that the instruments are jointly significant (see the bottom part of Panel B). Compared to the columns (2) and (6) the first-stage  $R^2$  is somewhat reduced.

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. Quantitatively, the estimates of civic capital are very close to the OLS results.<sup>47</sup> The p-values of the overidentification tests do not reveal any violation of the exclusion restriction, neither for average size nor for the standard deviation of size.

## 7 Conclusion

Employing census data on 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 trust increases the average and the dispersion of the size distribution in 4-digit industries across provinces. The effect of civic capital holds for both manufacturing and services, and is more important for industries relying more heavily on labor as an input. Our econometric specifications address several potential threats. First, a set of

<sup>46</sup>Obviously, 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.

<sup>47</sup>When we drop the South and the Islands from the sample and re-run the regression specifications of columns (1) and (5) of Table 7, we get a point estimate of 0.081 for *APS* and 0.105 for *SDPS*. Both estimates are statistically different from zero at the 1% level. The comprehensive results are available upon request.

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 the sources of endogeneity of trust. On the one hand, extracting the first principal component of three well-established measures of trust minimizes the problem of measurement error. On the other hand, 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 serve as instruments for current trust. We get the conclusion that it is unlikely that our results are driven by reversed causality or omitted variables bias.

Recently, Algan and Cahuc (2010) and Tabellini (2010) have established a causal effect of trust on aggregate economic outcomes, such as per-capita GDP and growth rates. While these studies provide important new insights, they do not suggest specific channels of how trust enhances the growth potential of economies. Our study hints that one mechanism of how trust translates into growth might be that it allows plants to grow large. Indeed, Rajan and Zingales (1998) show that two thirds of industry growth can be attributed to the growth in existing establishments rather than to the formation of new ones.

We think that more evidence is needed on the relation between trust and the growth of organizations. Based on a cross-section, our analysis does not allow to derive dynamic patterns. Moreover, observing shifts over time could uncover the impact of trust with respect to firm entry and exit. Another interesting extension consists in identifying how trust precisely affects firm organization. We have presented several, mutually not exclusive, channels without being able to identify the relative impact of each of them. In order to open this black box, survey data which shed light on the decision making process and internal organization of firms, such as those used in Bloom et al. (2009), are very promising.

## 8 Appendix

### 8.1 Detailed description of the data

#### 8.1.1 Dependent variables

The construction of the dependent variables is described in detail in the main text. The information is taken from the 8th Istat Census of Industry and Services, corresponding to the year 2001.

#### 8.1.2 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 Superior Institute of 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) on data from the Superior Institute of Health.

*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).

### 8.1.3 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 the final demand coming from 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 and are taken with reference to the year 2001.

*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](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. The data refer to the year 2001 and are from Istat.

*Bank branches*: It is the number of bank branches per 1000 inhabitants, disaggregated by province. The data refer to the year 2001 with the exception of the provinces of Vibo Valentia and Verbano-Cusio-Ossola. For those two provinces, the values corresponding to the year 1998 are taken. 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 year 2001 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 of the province the value of the court to which the city belongs. This information is then averaged for all the cities belonging to the same province to get a provincial variable.

*GDP*: It is the provincial nominal gross domestic product in 2001, expressed in thousands of euros. The source is Istat.

*Capital-labor ratio*: It is calculated at the 4-digit industry level from the balance sheet data of 86,000 firms operating in Italy in 2001. For each firm, we compute the capital-labor ratio as the sum of tangible and intangible assets (measured in thousands of euros) over the total number of employees. We then take the average value for all the firms belonging to the same industry. The data source is Bureau van Dijk’s AIDA data set.

## 8.2 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  $\alpha_1$  consists of  $p$  constants,  $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}$ . Consider the linear function  $\alpha_1' \mathbf{x}$ :

$$\alpha_1' \mathbf{x} = \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1p}x_p = \sum_{j=1}^p \alpha_{1j}x_j \quad (10)$$

Finding the first principal component amounts to determine the elements of  $\alpha_1$  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  $\alpha_1$  is constrained to have unit length, which implies that  $\alpha_1' \alpha_1 = 1$ . The corresponding Lagrange maximization function takes the following form:

$$\alpha_1' \mathbf{S} \alpha_1 - \lambda(\alpha_1' \alpha_1 - 1). \quad (11)$$

Maximizing (11) with respect to  $\alpha_1$  gives

$$(\mathbf{S} - \lambda I_p) \alpha_1 = 0, \quad (12)$$

in which the Lagrange multiplier  $\lambda$  is the eigenvalue of  $\mathbf{S}$  and the corresponding eigenvector is  $\alpha_1$ .  $I_p$  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

- 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.
- Banfield, E., 1958: *The Moral Basis of a Backward Society*. The Free Press, New York.
- 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, 2009: The organization of firms across countries. NBER Working Papers 15129, National Bureau of Economic Research.
- 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.
- 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.
- 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, 2010: Civic Capital as the Missing Link. NBER Working Papers 15845, National Bureau of Economic Research.
- 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.
- 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.
- 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.
- Lu, Y. and Z. Tao, 2009: Contract enforcement and family control of business: Evidence from China. *Journal of Comparative Economics*, **37**(4), 597–609.
- 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.
- Percoco, M., 2011: Entrepreneurship, Social Capital and Institutions: Evidence from Italy. Working paper, Bocconi University, Milan.
- 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.
- Rothstein, B., 2000: Trust, social dilemmas and collective memories. *Journal of Theoretical Politics*, **12**(4), 477–501.
- 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.
- Syversen, 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.

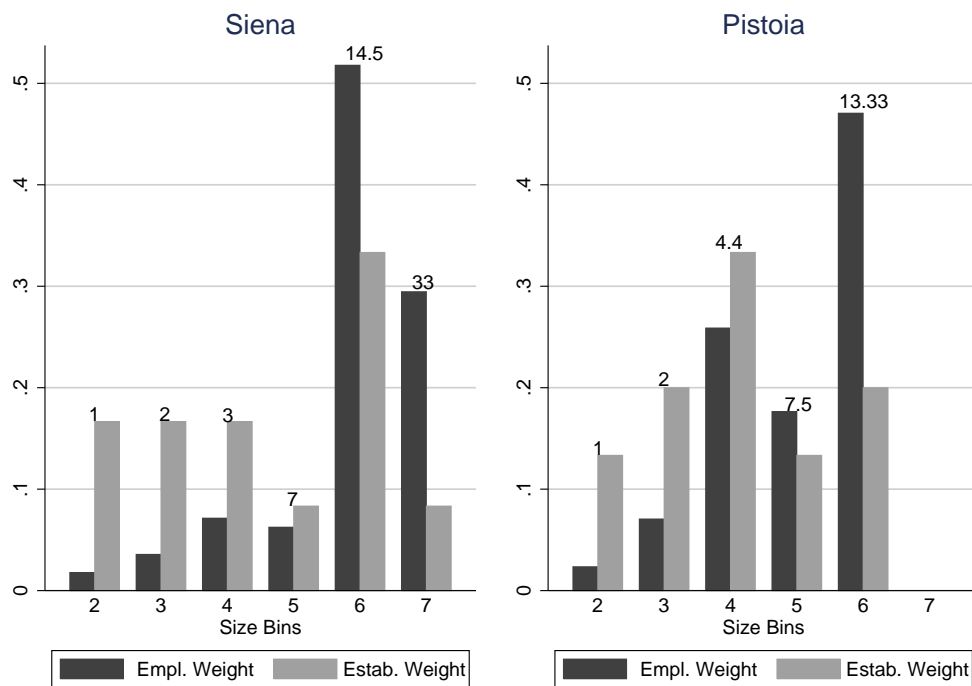


Figure 1: Empirical distribution of plant size in NACE industry 2862 (Manufacture of tools) in two provinces of Tuscany (Siena and Pistoia). On the top of each couple of bars there is the average size of that bin, given by the total amount of employment divided by the total amount of establishments.

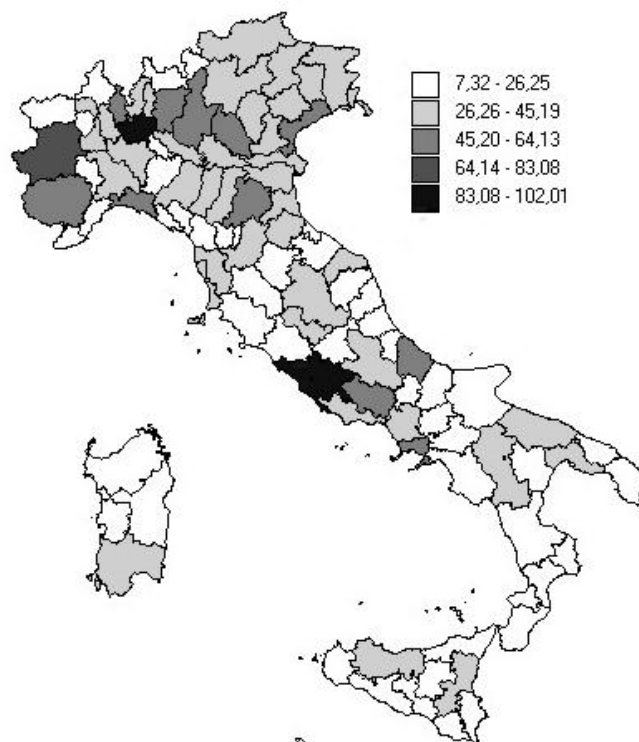


Figure 2: Average plant size, weighted by the number of employees in each size bin. Values are provincial averages over 4-digit industries.

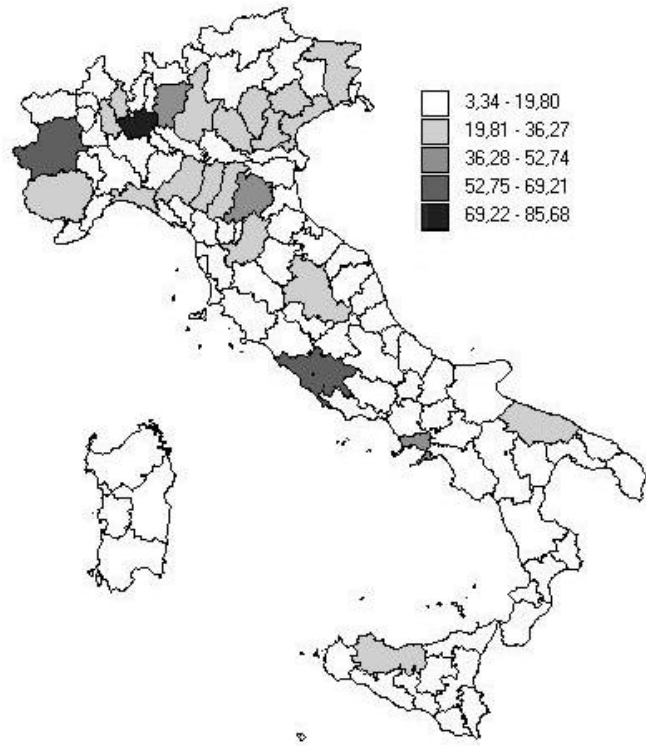


Figure 3: Standard deviation of size, weighted by the number of employees in each size bin. Values are provincial averages over 4-digit industries.

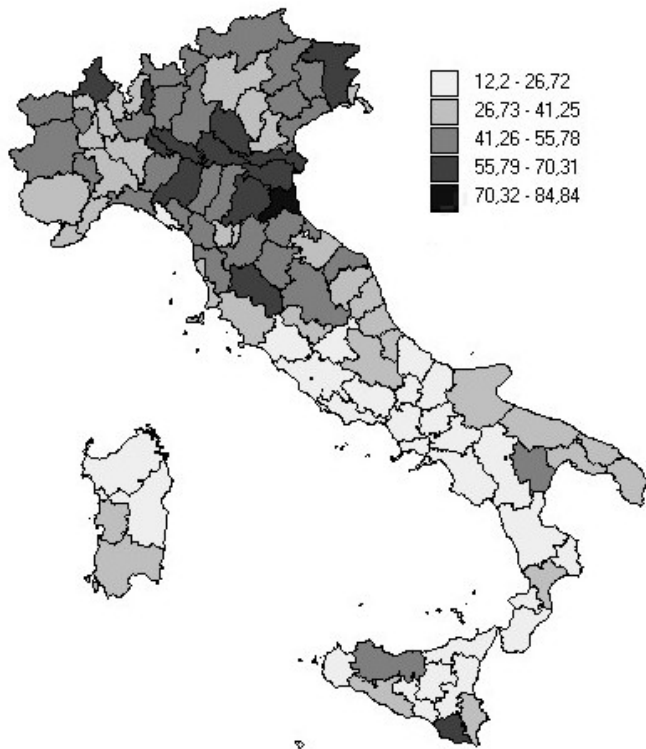


Figure 4: Blood donations per 1000 inhabitants.

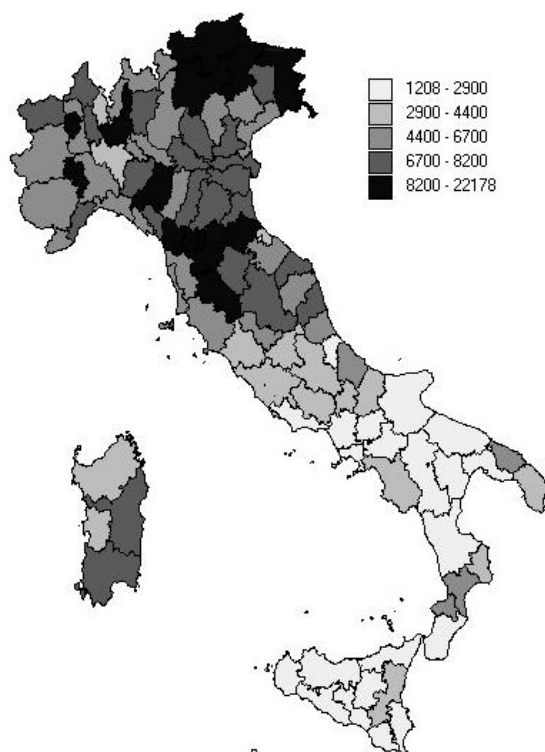


Figure 5: Number of volunteers in non-profit organizations per 100,000 inhabitants.

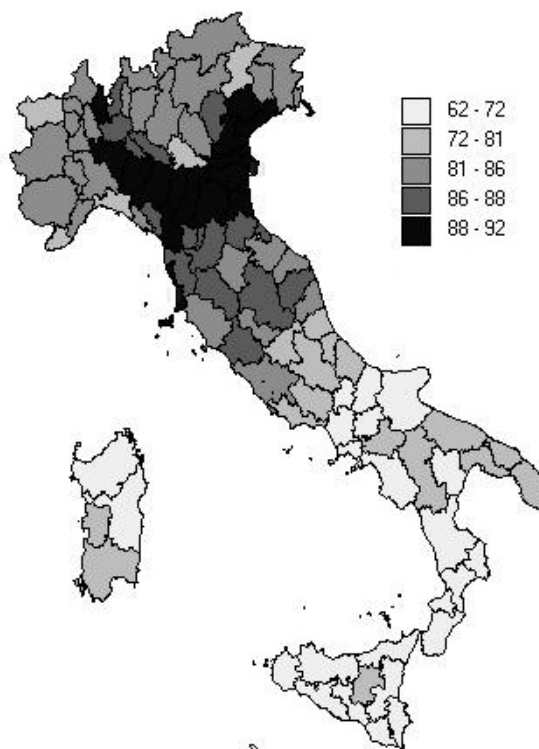


Figure 6: Electoral turnout in referenda, averaged over 7 referenda that took place between 1946 and 1987.

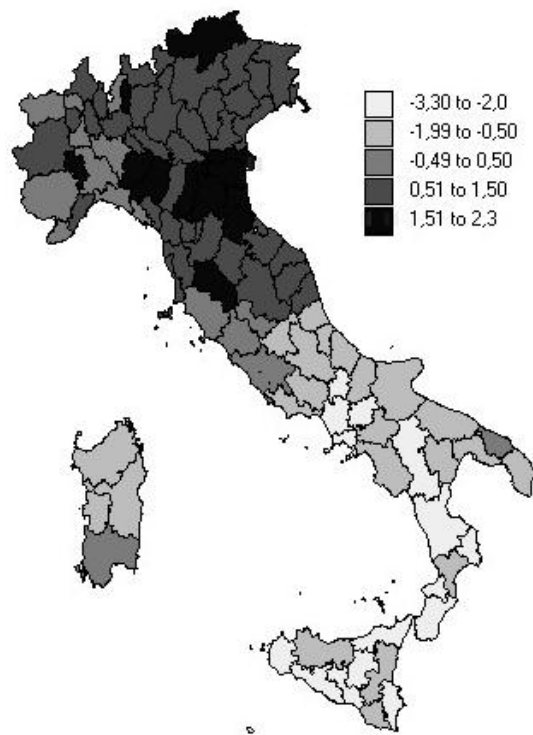


Figure 7: Map of civic capital measured by the first principal component of blood donations, volunteering, and electoral turnout.

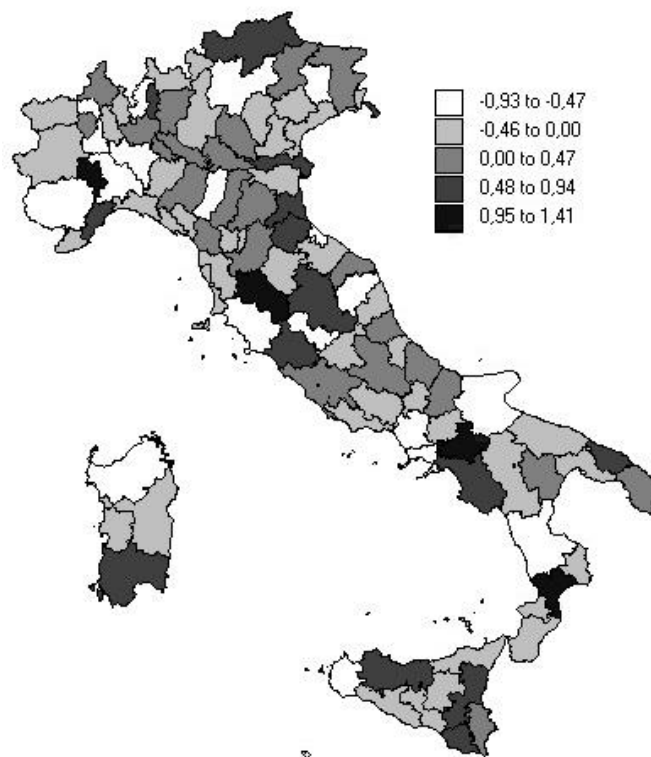


Figure 8: Map of civic capital measured by the variation of the first principal component within regions. The figure plots the residuals of a regression of the provincial principal component on regional dummies.

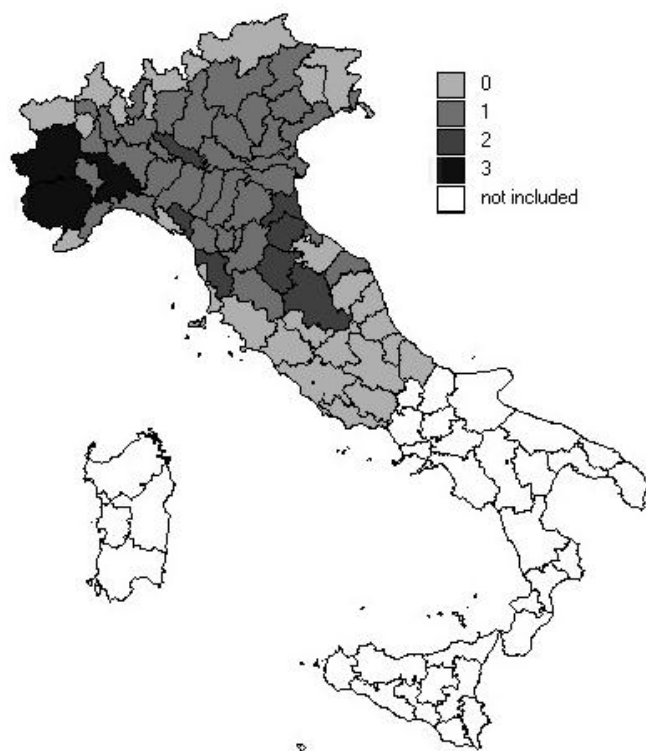


Figure 9: Number of free-city states by province in 1300 (Source: Guiso et al., 2008a). South of Italy and Islands are not included.



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 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</i>								
Average Plant Size (log)	32,071	2.212	1.398	0	8.708	1.111	2.011	3.091
Standard Deviation Plant Size (log)	31,410	1.826	1.479	-1.892	7.632	0.754	1.696	2.815
Average Plant Size - Estab. Weight (log)	32,071	1.383	1.015	0	8.303	0.596	1.160	1.946
Standard Deviation Plant Size - Estab. Weight (log)	31,410	1.420	1.368	-2.227	7.956	0.406	1.304	2.292
Total employment (log)	32,071	4.806	1.664	1.099	10.731	3.611	4.779	5.969
Total establishments (log)	32,071	3.422	1.613	1.099	9.782	2.079	3.219	4.522
N. of small plants (less than 20 empl., log)	32,008	3.343	1.669	0	9.778	2.079	3.135	4.477
N. of large plants (at least 20 empl., log)	15,319	1.031	1.049	0	6.004	0	0.693	1.609
<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 (log)	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</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
Urbanization in 1861 (log)	103	2.366	0.896	0	4.276	1.919	2.245	2.91
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
K-L ratio (log)	481	4.169	0.819	1.177	7.045	3.666	4.034	4.563

*Note:* The table provides descriptive statistics for the variables used in the regressions. The dependent variables are: *Average Plant Size* is the log of average size at the province-industry level, 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 at the province-industry level, with each size bin weighted by the corresponding share of employees; *Average Plant Size - Estab. Weight* is the log of average size, with each size bin weighted by the corresponding share of plants; *Standard Deviation Plant Size - Estab. Weight* is the log of the standard deviation of plant size at the province-industry level, with each size bin weighted by the corresponding share of plants; *Total employment* is the log of the total number of employees; *Total establishments* is the log of the total number of establishments; *N. of small plants* is the log of the number of plants with less than 20 employees; *N. of large plants* is the log of the number of plants with at least 20 employees. 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 are: *Final demand* is the log of weighted population, with weights obtained from Input-Output Use tables; *Urbanization in 1861* is the log of the share of inhabitants living in cities with at least 10,000 people in 1861; *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; *K-L ratio* is the log of the average capital-labor ratio at the industry level. Unless otherwise stated, the data refer to the year 2001.

Table 3: Determinants of average plant size: baseline regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Blood	Volunteers	Referenda	Principal comp.	Blood	Volunteers	Referenda	Principal comp.
	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se
Blood donations (log)	0.076** (0.032)				0.071** (0.033)			
Volunteers (log)		0.061** (0.027)				0.066** (0.027)		
Referenda turnout (log)			-0.067 (0.279)				-0.199 (0.283)	
Principal component (log)				0.052*** (0.016)				0.050*** (0.017)
Final demand (log)	0.394*** (0.022)	0.396*** (0.021)	0.395*** (0.022)	0.395*** (0.021)	0.356*** (0.023)	0.357*** (0.022)	0.354*** (0.022)	0.359*** (0.022)
Urbanization in 1861 (log)					0.024 (0.015)	0.029** (0.015)	0.025 (0.015)	0.027* (0.015)
University graduates (log)					0.168* (0.091)	0.169* (0.090)	0.191** (0.088)	0.145 (0.091)
Bank branches (log)					-0.035** (0.016)	-0.034** (0.016)	-0.028* (0.016)	-0.038** (0.016)
Length of labor trials (log)					0.001 (0.025)	0.002 (0.025)	0.006 (0.026)	-0.003 (0.025)
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$R^2$ (within)	0.121	0.121	0.120	0.121	0.122	0.122	0.121	0.122
Obs.	32,060	32,060	32,060	32,060	32,060	32,060	32,060	32,060

*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. Dependent variable is the log of  $APS_{ij}$ , the weighted average plant size. 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; *Urbanization in 1861* is the log of the share of inhabitants living in cities with at least 10,000 people in 1861; *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. Unless otherwise stated, the data refer to the year 2001. \*\*\*, \*\*, \* denote significance at the 1%, 5%, 10% level, respectively.

Table 4: Determinants of the standard deviation of plant size: baseline regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Blood	Volunteers	Referenda	Principal comp.	Blood	Volunteers	Referenda	Principal comp.
	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se
Blood donations (log)	0.0833** (0.037)				0.0777** (0.039)			
Volunteers (log)		0.078*** (0.029)				0.084*** (0.029)		
Referenda turnout (log)			0.146 (0.326)				0.011 (0.329)	
Principal component (log)				0.067*** (0.018)				0.066*** (0.019)
Final demand (log)	0.609*** (0.026)	0.612*** (0.024)	0.610*** (0.025)	0.610*** (0.024)	0.567*** (0.026)	0.568*** (0.025)	0.567*** (0.026)	0.571*** (0.025)
Urbanization in 1861 (log)					0.027	0.033**	0.028	0.031*
University graduates (log)					(0.017)	(0.017)	(0.017)	(0.017)
Bank branches (log)					0.185*	0.184*	0.196**	0.152
					(0.100)	(0.097)	(0.099)	(0.099)
					-0.036**	-0.036**	-0.029*	-0.041**
					(0.016)	(0.017)	(0.017)	(0.017)
Length of labor trials (log)					0.000	0.002	0.005	-0.005
					(0.030)	(0.029)	(0.031)	(0.029)
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$R^2$ (within)	0.177	0.177	0.177	0.178	0.178	0.178	0.178	0.178
Obs.	31,398	31,398	31,398	31,398	31,398	31,398	31,398	31,398

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. Dependent variable is the log of  $SDAPS_{ij}$ , the weighted standard deviation. 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; *Urbanization in 1861* is the log of the share of inhabitants living in cities with at least 10,000 people in 1861; *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. Unless otherwise stated, the data refer to the year 2001. \*\*\*, \*\*, \* denote significance at the 1%, 5%, 10% level, respectively.

Table 5: Determinants of the size distribution: extensions

	Average size			Standard deviation of size			Total empl. and establ., number of small vs. large plants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	K-L ratio Coef./se	Manuf. Coef./se	Services Coef./se	K-L ratio Coef./se	Manuf. Coef./se	Services Coef./se	Total Empl. Coef./se	Total Establ. Coef./se	N. of small plants Coef./se	N. of large plants Coef./se
Principal component	0.125*** (0.041)	0.067** (0.029)	0.048** (0.020)	0.161*** (0.046)	0.087*** (0.031)	0.058** (0.023)	0.049*** (0.018)	0.020 (0.013)	0.015 (0.013)	0.061*** (0.023)
Princ. comp. $\times$ K-L ratio (log)	-0.018** (0.009)			-0.023** (0.010)						
Final demand (log)	0.359*** (0.022)	0.365*** (0.040)	0.372*** (0.025)	0.570*** (0.026)	0.551*** (0.046)	0.606*** (0.030)	1.019*** (0.024)	0.917*** (0.019)	0.921*** (0.019)	0.628*** (0.030)
Urbanization in 1861 (log)	0.027* (0.015)	0.007 (0.031)	0.050*** (0.015)	0.030* (0.016)	0.010 (0.034)	0.058*** (0.017)	0.013 (0.018)	-0.000 (0.014)	-0.004 (0.015)	0.017 (0.016)
University graduates (log)	0.134 (0.094)	-0.153 (0.176)	0.425*** (0.096)	0.143 (0.102)	-0.196 (0.189)	0.497*** (0.107)	-0.022 (0.092)	-0.041 (0.071)	-0.041 (0.075)	0.019 (0.112)
Bank branches (log)	-0.038** (0.015)	-0.074*** (0.028)	0.007 (0.033)	-0.042*** (0.015)	-0.093** (0.037)	0.013 (0.038)	-0.031 (0.032)	-0.002 (0.032)	0.003 (0.034)	-0.065*** (0.023)
Length of labor trials (log)	-0.000 (0.025)	-0.019 (0.043)	-0.020 (0.023)	-0.003 (0.029)	-0.027 (0.050)	-0.011 (0.029)	-0.013 (0.028)	-0.014 (0.022)	-0.014 (0.022)	-0.003 (0.029)
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$R^2$ (within)	0.121	0.141	0.170	0.177	0.177	0.240	0.415	0.496	0.486	0.310
Obs.	30,523	12,352	16,779	29,894	12,133	16,411	32,060	32,060	31,998	15,305

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. The variable employed to measure civic capital is *Principal component*, which 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 1990. We also include the following provincial controls: *Final demand* is the log of weighted population, with weights obtained from Input-Output Use tables; *Urbanization in 1861* is the log of the share of inhabitants living in cities with at least 10,000 people in 1861; *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; *K-L ratio* is the log of the average capital-labor ratio at the industry level. The dependent variable in column (7) is the log of total employment, in column (8) it is the log of the total number of establishments, in column (9) it is the log of the number of plants with less than 20 employees, in column (10) it is the log of the number of plants with at least 20 employees. Unless otherwise stated, the data refer to the year 2001. \*\*\*, \*\*, \* denote significance at the 1%, 5%, 10% level, respectively.

Table 6: Determinants of the plant size distribution: further robustness and sensitivity analysis

	Average size					Standard deviation of size				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Principal component	$\ln \overline{APS}$ Coef./se	Two var. Coef./se	GDP Coef./se	Full sample Coef./se	TO, MI, RM Coef./se	$\ln \overline{SDPS}$ Coef./se	Two var. Coef./se	GDP Coef./se	Full sample Coef./se	TO, MI, RM Coef./se
	0.029** (0.011)	0.045*** (0.014)	0.043*** (0.015)	0.050** (0.020)	0.047*** (0.017)	0.053*** (0.017)	0.053*** (0.016)	0.056*** (0.016)	0.068*** (0.020)	0.063*** (0.019)
Principal component (two var, log)										
		0.045*** (0.014)								
Final demand (log)	0.102*** (0.012)	0.357*** (0.022)	0.095 (0.065)	0.410*** (0.022)	0.352*** (0.024)	0.327*** (0.020)	0.568*** (0.025)	0.173** (0.082)	0.592*** (0.026)	0.569** (0.028)
GDP (log)			0.251*** (0.061)					0.379*** (0.076)		
Urbanization in 1861 (log)	0.013 (0.010)	0.027* (0.015)	0.019 (0.014)	0.034** (0.015)	0.025 (0.016)	0.021 (0.015)	0.030* (0.017)	0.018 (0.015)	0.034* (0.017)	0.027 (0.017)
University graduates (log)	0.019 (0.061)	0.160* (0.088)	0.096 (0.087)	0.104 (0.103)	0.110 (0.094)	0.102 (0.085)	0.173* (0.097)	0.076 (0.091)	0.110 (0.105)	0.118 (0.099)
Bank branches (log)	-0.030*** (0.011)	-0.040** (0.016)	-0.041* (0.022)	-0.059** (0.023)	-0.036** (0.016)	-0.039** (0.017)	-0.042** (0.017)	-0.046* (0.026)	-0.047** (0.019)	-0.038** (0.016)
Length of trials (log)	0.001 (0.017)	-0.001 (0.024)	-0.006 (0.023)	0.010 (0.027)	-0.004 (0.027)	-0.002 (0.025)	-0.003 (0.028)	-0.009 (0.027)	-0.001 (0.031)	-0.006 (0.031)
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$R^2$ -r2 (within)	0.073	0.122	0.123	0.119	0.096	0.113	0.178	0.180	0.180	0.139
Obs	32,060	32,060	32,060	39,122	30,810	31,398	31,398	31,398	33,473	30,153

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 (in columns (5) and (10) we use 5 macro-regional dummies). Dependent variable is the log of APS (column (1)), log of APS (column (2)-(5)), the log of  $\overline{SDAPS}$  (column (6)), log of SDPS (columns (7)-(10)). In each regression, standard errors are clustered at the province and industry level. The variable employed to measure civic capital is *Principal component*, which is the log of 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 1990. We also include the following provincial controls: *Final demand* is the log of weighted population, with weights obtained from Input-Output Use tables; *Urbanization in 1861* is the log of the share of inhabitants living in cities with at least 10,000 people in 1861; *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. Unless otherwise stated, the data refer to the year 2001. \*\*\*, \*\*, \* denote significance at the 1%, 5%, 10% level, respectively.

Table 7: Instrumental variables analysis

	Panel A: Second stage				Standard deviation of size			
	Average size							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Principal component	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se	Coef./se
GDP (log)	0.076*** (0.017)	0.112*** (0.036)	0.116** (0.051)	0.081* (0.049)	0.094*** (0.020)	0.138*** (0.040)	0.127** (0.055)	0.116** (0.055)
Final demand (log)	0.338*** (0.074)	0.310*** (0.070)	0.307*** (0.075)	0.261*** (0.085)	0.462*** (0.085)	0.429*** (0.081)	0.437*** (0.087)	0.365*** (0.107)
Urbanization in 1861 (log)	0.009 (0.080)	0.038 (0.077)	0.041 (0.083)	0.123 (0.092)	0.091 (0.090)	0.126 (0.087)	0.117 (0.094)	0.233** (0.114)
University graduates (log)	-0.013 (0.015)	-0.007 (0.016)	-0.007 (0.016)	-0.002 (0.017)	-0.018 (0.017)	-0.011 (0.017)	-0.013 (0.017)	-0.003 (0.017)
Bank branches (log)	0.050 (0.079)	0.050 (0.080)	0.050 (0.080)	-0.011 (0.093)	0.045 (0.087)	0.043 (0.088)	0.044 (0.088)	-0.032 (0.098)
Length of labor trials (log)	-0.021 (0.018)	-0.041* (0.023)	-0.043 (0.028)	-0.023 (0.029)	-0.035 (0.022)	-0.058** (0.028)	-0.052 (0.032)	-0.038 (0.034)
Macro-regional dummies	-0.014 (0.023)	-0.016 (0.024)	-0.016 (0.024)	-0.021 (0.026)	-0.011 (0.027)	-0.012 (0.027)	-0.012 (0.027)	-0.013 (0.031)
Hansen test of overidentif. (p-value)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R <sup>2</sup> (within)	0.118	0.118	0.118	0.104	0.177	0.176	0.176	0.168
Obs.	32,060	32,060	32,060	21,635	31,398	31,398	31,398	21,299
Panel B: First stage (Dependent var. is Principal component)								
GDP (log)	0.546 (0.364)	0.509 (0.424)	0.866** (0.363)	0.535 (0.364)	0.498 (0.424)	0.858** (0.363)	0.498 (0.424)	0.858** (0.363)
Final demand (log)	-0.621 (0.388)	-0.559 (0.464)	-1.026*** (0.386)	-0.616 (0.388)	-0.554 (0.464)	-1.020** (0.386)	-0.554 (0.464)	-1.020** (0.386)
Urbanization in 1861 (log)	-0.179** (0.077)	-0.220** (0.088)	-0.041 (0.071)	-0.178** (0.077)	-0.219** (0.088)	-0.039 (0.071)	-0.219** (0.088)	-0.039 (0.071)
University graduates (log)	0.445 (0.366)	0.411 (0.394)	0.174 (0.388)	0.459 (0.364)	0.428 (0.391)	0.176 (0.387)	0.428 (0.391)	0.176 (0.387)
Bank branches (log)	0.327* (0.196)	0.351* (0.188)	0.434*** (0.159)	0.323 (0.195)	0.346* (0.186)	0.432*** (0.158)	0.346* (0.186)	0.432*** (0.158)
Length of labor trials (log)	0.220 (0.155)	0.160 (0.147)	-0.094 (0.165)	0.218 (0.155)	0.159 (0.147)	-0.096 (0.165)	0.159 (0.147)	-0.096 (0.165)
Aid societies in 1873 (log)	0.540** (0.211)	0.857*** (0.196)	0.544** (0.211)	0.544** (0.211)	0.861*** (0.196)	0.544** (0.211)	0.861*** (0.196)	0.544** (0.211)
Turnout in 1920s (log)	3.690*** (1.179)		3.673*** (1.184)	3.673*** (1.184)		3.673*** (1.184)		3.673*** (1.184)
1 free-city state in 1300		Yes	0.318* (0.173)	Yes	0.318* (0.173)	Yes	0.318* (0.173)	Yes
2 free-city states in 1300		0.000	0.666*** (0.169)	0.002	0.666*** (0.169)	0.002	0.666*** (0.169)	0.002
3 free-city states in 1300		0.847	-0.035 (0.297)	0.513	-0.035 (0.297)	0.513	-0.035 (0.297)	0.513
Macro-regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of joint signif. (p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
R <sup>2</sup>	0.847	0.830	0.830	0.846	0.846	0.846	0.829	0.512
Obs.	32,060	32,060	21,635	31,398	31,398	31,398	31,398	21,299

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, available at the regional level, 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. Unless otherwise stated, the data refer to the year 2001. \*\*\*, \*\*, \* denote significance at the 1%, 5%, 10% level, respectively.



Alma Mater Studiorum - Università di Bologna  
DEPARTMENT OF ECONOMICS

Strada Maggiore 45  
40125 Bologna - Italy  
Tel. +39 051 2092604  
Fax +39 051 2092664  
<http://www.dse.unibo.it>