Extortion, firm's size and the sectoral allocation of capital*

Luigi Balletta† Mario Lavezzi‡

PRELIMINARY AND INCOMPLETE. DO NOT CIRCULATE.

April 29, 2014

Abstract

Extortion of firms is a typical activity of organized crime such as Mafia. We develop a simple principal-agent model to find the Mafia-optimal extortion as a function of firm's observable characteristics, specifically firm's size. We test the predictions of the model on a unique dataset on extortion in Sicily, the Italian region where Mafia is most active. Our empirical findings show that i) extortion moderately increases with firm's size ii) extortion is *regressive*, the average extortion rate ranging from approximately 40% of operating profits for small firms to 2% for large firms iii) extortion turns average cost function decreasing, therefore influencing market competition.

Keywords: Organized Crime, Economic Structure, Sicilian Mafia, Asymmetric Information, Principal-Agent Theory.

JEL Classification Numbers: C72, D86, K42.

1 Introduction

In this paper we provide a theoretical and empirical analysis of extortion, a typical activity of organized crime (see, e.g., Gambetta, 1993, and Konrad and Skaperdas, 1998). In particular, we study extortion in the framework of a principal-agent model, where the criminal organization cannot perfectly observe firms' productivity. The theoretical predictions of the model are tested on a unique database on extortion in Sicily, the Italian region

*We thank the *Fondazione Chinnici* for providing us with the database on extortion and the Chamber of Commerce of Palermo for allowing us access on firms' data.

†Università di Palermo, Italy.

‡Università di Palermo, Italy.

where the criminal organization commonly known as Cosa Nostra is active. This database contains actual data on the amounts of money paid as *pizzo* to the Mafia,[1] and data from balance sheets of firms victims of extortion.

The main findings of the paper are: i) the amount of *pizzo* moderately increases with firms' size. This is in line with the predictions of the theoretical model. ii) Mafia "taxation" is regressive: in particular we find that the average "tax" rate applied to small firms amounts to approximately 40% of operating profits, and decreases to approximately 2% for large firms; iii) inclusion of Mafia "taxation" in the average cost function makes average costs become decreasing.

These findings suggest specific microeconomic channels through which organized crime negatively affects economic growth: i) by stifling small firms which, in an environment where access to credit can be difficult because of crime itself[2] have little resources to expand; ii) by erecting barriers to entry.[3]

The paper is organized as follows. In section 2 and 3 we describe and solve the basic theoretical model of a monopolistic Mafia optimally choosing extortion as a function of some observable variable. In section 4 we describe our newly built database containing information on the amounts paid and characteristics of the firm. In section 5 we perform our empirical analysis on the dataset. Appendix A analyzes the issue of sample validation, given its non-randomness.

2 The Model

In this section we describe the basic theoretical model of a monopolistic Mafia choosing optimal extortion as a function of observable characteristics of firms. There is a continuum of firms indexed by a productivity parameter θ, distributed according to $G(\cdot)$ with strict positive density $g(\cdot)$ on the interval $[\theta_L, \theta_H]$. Moreover G satisfies monotone hazard rate property. As in standard hidden information models, θ is private information of the firm but its distribution is common knowledge. Firm's profits, gross of any amount she is forced

[1] *Pizzo* is the Sicilian definition of the money paid to the Mafia.

[2] Analyzing Italian regions, Bonaccorsi di Patti (2009) finds that the presence of organized crime in a region increases interest rates payed by firms on loans by approximately 30 basis points.

[3] Previous literature (e. g. Gambetta, 1993) has emphasized that organized crime creates local monopolies.

to pay to Mafia, are written as $\theta f(K) - rK$, where K is a choice variable, which can be understood to be capital, as a measure of size of the firm, and r is a parameter, which can be understood as cost of capital. The function f is assumed to be increasing and strictly concave. This assumption guarantees single-crossing.

The timing we have in mind is the following. Mafia is a Stackelberg leader and its existence is commonly acknowledged by firms. Its offer for "protection" is summarized by a function $x(K)$ representing the amount of money asked to a firm which has chosen K. Mafia also makes clear the punishment in the case of non-compliance. We assume its monetary equivalent is a positive real number z, which is exogenous. The idea is that z represents "firepower". Higher values indicate a strong organization which is able to inflict much damage to a non-compliant firm.[4] Firms know both $x(K)$ and z when choosing whether to open for business. Once in the market, they choose the amount of observable K optimally. Moreover, they have the choice to refuse to pay the extortion amount, in which case they don't pay x but they suffer the cost z. For the purpose of this basic model, we assume that Mafia has enough reputation concerns that it finds always optimal not to renegotiate the extortion in case of compliance and to inflict the punishment to a non-compliant firm. The sequence of moves is therefore the following:

1. Nature extracts type θ that is observed only by the firm.
2. Mafia proposes an extortion function $x(K)$, given the level of firepower z.
3. Firm decides whether to open for business having observed both $x(K)$ and z. Outside option is normalized to 0, independent of firm's type.
4. If the firm is in business, she chooses K to maximize profits.
5. Firm decides whether to be compliant, in which case she pays $x(K)$, or to refuse to pay, in which case she suffers the cost z.

In the following sections our objective will be to characterize theoretically the optimal choice of the extortion function $x(K)$ and to compare it to the empirical counterpart in our dataset.

[4] In an extension of the model, z might be optimally chosen at the beginning of the game as a costly investment in punishment capacity.

3 Optimal extortion function

As a preliminary step, we note that the game induced by our timing is one of incomplete information where the uninformed party moves first and commits to a function $x(K)$ of the optimal choice of the follower. The assumption on commitment allows the problem to be mapped into a principal-agent framework in which the Mafia is the principal and the Firm is the agent. To solve for the optimal extortion function we can appeal to a revelation principle to first find the optimal direct mechanism $\{K(\theta), x(\theta)\}$, imposing standard incentive compatibility conditions, and then recovering the indirect mechanism $x(K)$ that implements the optimal choice.

We begin the analysis from the compliance choice of the firm. It is clear that the firm will choose to be compliant with extortion if $x(\theta) \leq z$. In stage 4, anticipating the compliance decision, firm will choose optimally a type declaration solving $\max_{\hat{\theta}}(\theta f\left(K\left(\hat{\theta}\right) - rK\left(\hat{\theta}\right)\right) - \min\{x\left(\hat{\theta}\right), z\})$. Denoting with $\pi(\theta)$ the value of *net profits* following a true declaration of type, the following lemma is standard:

Lemma 1 *A direct mechanism $\{K(\theta), x(\theta)\}$ is incentive compatible if and only if*
(i) $\pi'(\theta) = f(K(\theta))$
(ii) $K'(\theta) \geq 0$.

Again anticipating future decisions, in stage 3 firm will decide to enter the market if $\pi(\theta) \geq 0$, since outside option is independent of firm's type.

The above discussion allows the optimization problem of the Mafia to be written in the following compact form:

$$\max_{K(\theta),x(\theta)} \int_{\theta_L}^{\theta_H} x(\theta)\, dG \tag{SMP}$$
$$s.t.\ \pi'(\theta) = f(K(\theta))$$
$$\pi(\theta) \geq 0, x(\theta) \leq z, K'(\theta) \geq 0.$$

3.1 A powerful Mafia

In this section we first find the optimal solution to the Mafia optimization problem assuming that z is large enough so the constraint on x is never binding. The solution is found applying standard techniques in contract theory. The following proposition follows:

Proposition 2 *Suppose $z > x^{PM}(\theta_H)$. The optimal solution to (SMP) satisfies:*

(i) $\left(\theta - \frac{1-G(\theta)}{g(\theta)}\right) f'\left(K^{PM}(\theta)\right) = r$,

(ii) $x^{PM}(\theta) = \theta f\left(K^{PM}(\theta)\right) - rK^{PM}(\theta) - \int_{\theta_L}^{\theta} f\left(K^{PM}(s)\right) ds$.

(iii) Denoting with $\theta^{PM}(K)$ the inverse of $K^M(\theta)$ the optimal solution can be implemented via an extortion function $x^{PM}(K) = x^{PM}\left(\theta^{PM}(K)\right)$.

Figure 1 describes the functions $K^M(\theta)$ and $x^{PM}(K)$ when $f(K) = 2\sqrt{K}$, G is uniform on $[1, 2]$ and $r = 1$. In black we report the undistorted (first best) level of K.

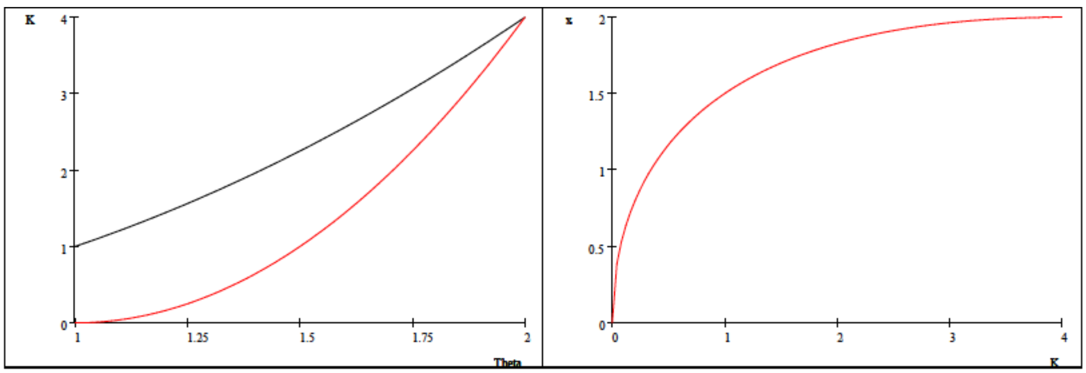


Figure 1: Capital as function of θ and pizzo as function of capital: powerful Mafia

3.2 A weak Mafia

In this subsection we describe graphically the optimal solution when Mafia has not enough power to implement the unconstrained solution of Proposition 2. We have the following:

Proposition 3 *Suppose $z < x^{PM}(\theta_H)$, where x^{PM} is the solution described in Proposition 2. The optimal solution to (SMP) is such that there exists a threshold $\widetilde{\theta}$ with*

(i) $\left(\theta - \frac{G(\widetilde{\theta})-G(\theta)}{g(\theta)}\right) f'\left(K^{WM}(\theta)\right) = r$ *for* $\theta \leq \widetilde{\theta}$.

(ii) $K^{WM}(\theta) = \widetilde{K}$ *for* $\theta > \widetilde{\theta}$, *where* $\widetilde{K}$ *satisfies* $\widetilde{\theta} f'\left(\widetilde{K}\right) = r$.

(iii) $z - \widetilde{\theta} f\left(\widetilde{K}\right) + r\widetilde{K} + \pi(\theta_l) + \int_{\theta_l}^{\widetilde{\theta}} f(K(\theta)) d\theta = 0$

We display in Figure 2 the optimal K^{WM} and x^{WM} in green, for the same parameters as in the previous pictures and for $z = 3/2$. We compare the solution to the solution for a powerful Mafia in red.

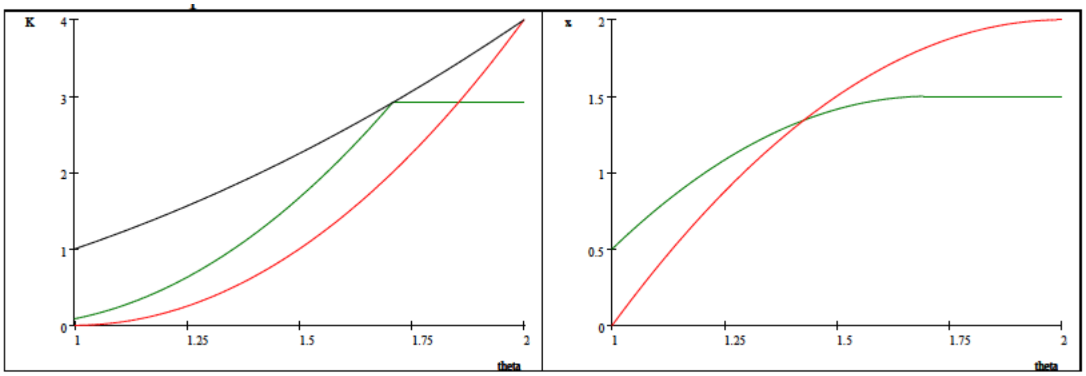


Figure 2: Capital and pizzo as function of θ: weak Mafia

It can be noted that: the distortion in capital investment changes. After a critical level of θ, the capital stock becomes constant. This is due to the fact that, after a critical level of θ, the amount extorted becomes constant and equal to the constraint z. Note that the implication for the *pizzo* function is that it becomes *less dispersed* and constant after a certain productivity level.

4 The Database

The database on extortion analyzed in this paper was contstructed by researchers of the *Fondazione Chinnici* of Palermo in 2007.[5] The database contains information on extortionary activities by the Sicilian mafia, *Cosa Nostra*, in the nine provinces of Sicily in the period 1987-2007. The main source of evidence is court documents, but also interviews with magistrates were conducted.[6] For a recorded extortion, the database contains information on: 1) the identity of extorted firm; 2) its sector; 3) its administrative location

[5]The construction of the database was part of a project on "The Costs of Illegality" (*I costi dell'illegalità*), whose results are published in La Spina (2008). The aim of the project was to estimate the costs for the Sicilian economy due to extortionist activities, i. e. the amount of resources subtracted to the legal economy by the mafia. Part of the results are also presented in Asmundo and Lisciandra (2008).

[6]Approximately, 200 documents were examined and 45 interviews were conducted.

(Town, Province, Address); 4) its *Mandamento*[7]; 5) the amount of *pizzo*; 6) the period of the payments;[8] 7) the type of payment: monthly, annual, one-off; 7) the presence of additional impositions (e. g. forced supply, forced hiring of workers, etc.) 8) some reference on the source of data (excerpts from documents, the name of police operation/investigation that originated the documentation). We built our database by matching this pieces of information with other data on these firms.

Overall, the database on extortion contains information on approximately 2300 episodes of extortion, but the amount of information for each recorded episode differs, as well as the quality of the information for every recorded item. In particular, the amounts paid to organized crime are reported only for a subset of cases, that is: 489 observations for monthly payments, 314 and 137 for, respectively one-off and annual payments. With respect to monthly payments, we were able to *exactly* identify the extorted firm in the database of the Chamber of Commerce[9] for 335 observations: 145 (corresponding to 134 firms, 11 observations refer to firms appearing more than once) for limited liability companies (*Società di capitali*), and 190 (corresponding to 189 firms, 1 observation refers to a firm appearing twice) for parterships (*Società di persone*).[10]

For both groups we collected other firms' data (i.e. legal form, initial year of activity, initial capital, number of employees, number of local units, whether the firm is still active), while for the limited liability firms we also collected balance sheets' data, as only the latter are required by the Italian law to file a copy of the balance sheets at the local Chamber of Commerce.[11]

The first obvious problem with this dataset is its non-randomness. We defer to Appendix A a description of the method we followed to assess the capacity of the dataset to represent the phenomenon of organized crime and the population of firms. Let us just remark that, from the way the database was assembled, the self-selection bias should be

[7] A *Mandamento* is: "a district incorporating an average of three mafia families" (Paoli, 2003, p. 45)

[8] The period of payment is most of the time indicated as an upper limit. That is, the typical information is whether *pizzo* was paid *until* year t. In few cases, the period corresponds to an interval, i.e. 1995-1998, in even fewer cases to individual years. Therefore, for all firms we are able to identify one year in which the *pizzo* was paid with certainty, i.e. the upper limit of the specified interval, or the individual year. We will denote this year as *pizzoyear*.

[9] Every legal economic activity must be officially recorded at the local Chamber of Commerce.

[10] Therefore, Approximately 68% of observations (335/489) were matched to firms' data from the Chamber of Commerce.

[11] We transformed all monetary values in costant 1995 prices.

somewhat attenuated. In particular: different sources were utilized, court documents and interviews. In addition, the way the information on an episode of extortion were recorded varies. An episode can be recorded because the extorted firm reported the crime or because, during an investigation, evidence was found of the case, or because a declaration was provided by an investigated person about the extorted firms.

In the following section we provide an empirical analysis of the dataset. In this paper we only focus on monthly payments. This type of payment is more consistent with the theoretical model and, in addition, the nature of the other forms of payments probably refer to another logic of extortion. In particular, "one-off" payments are typical for the construction sector and refer to payments demanded when the construction yard is open. "Annual payments" are not demanded regularly but in specified occasions, such as Christmas, Easter and during the month of August, the typical month for summer holidays.[12] The reason is that in such periods, for many firms (e.g. bars and restaurants), it is difficult to hide the high level of their business activity. Preliminary analysis showed that these amounts cannot be easily compared with monthly payments by simply computing the total and divinding by twelve. The analysis and comparison of these different methods of payments is deferred to further research.

5 Empirical Analysis

In this section we provide an empirical analysis of extortion. Our aim is to characterize this activity with the available data and to analyze the relationship between the amount of money extorted by organized crime and firms' characteristics, in particular their size.

5.1 Descriptive statistics

In Figure 3 we plot a nonparametric estimation of the density of the whole set of observations on monthly extortion, along with the density of the 335 observations referred to identified firms.

[12] *Ferragosto*, the religious recurrence of August 15th is the peak of the season and is a typical day for the payment, along with Christmas and Easter.

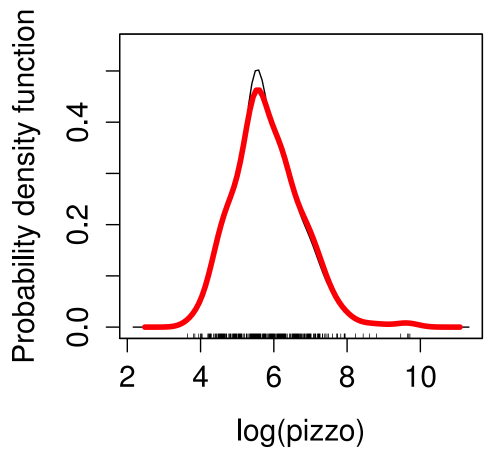


Figure 3: Distribution of (log) pizzo values: all observations vs values for “identified” firms

From Figure 3 we note that the shape of the estimated density is almost normal[13], approximating a lognormal distribution. This reflects the fact that, along with a large number of observations whose value is concentrated, there are few observations much larger than the former set. Moreover, we note that the distribution of the smaller set of 335 observations doest not show remarkable qualitative differences.[14]

Figure 4 presents the comparison between the density for the whole set of observations and the density for limited liability companies and the density for the values of *pizzo* paid by partnerships.

[13] A test for normality is rejected.

[14] The average value of *pizzo* in the two sets is, respectively, 618 and 599 (current) euros. The corresponding value in constant euros for the latter figure, which will be considered in the following, is 557 euros.

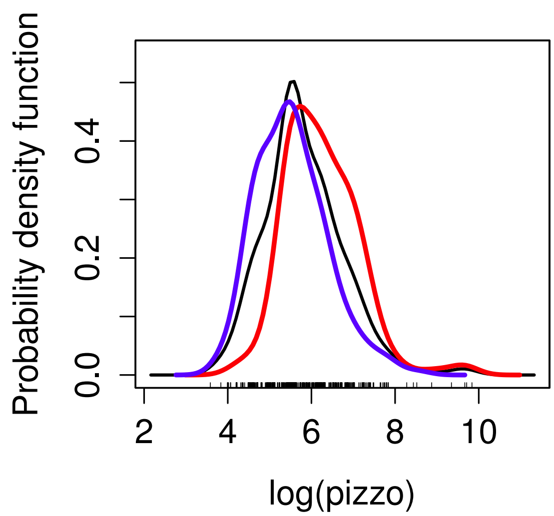


Figure 4: Distribution of (log) pizzo values: all observations vs limited liability firms (red) and partnerships (blue)

In Figure 4 it can be observed that the distributions are quite different. In particular, the mean (standard deviation) of the two subsets of observations are: 882 (1872) and 324 (526). Overall, therefore, the average *pizzo* paid by limited liability firms is much higher and dispersed than the *pizzo* paid by partnerships. In the following section we focus on the former group, and investigate the relationship between the amounts paid and some firms' characteristics.

5.2 Econometric analysis

In this section we provide an econometric analysis of the relationship between the *pizzo* paid by the firms in our sample and firms' characteristics, in particular their size. We will mostly focus on limited liability companies as data on balance sheets are available only for them.

5.2.1 Cross-section analysis

First of all we present a cross-section analysis of 119 firms for which we have balance sheets data. We begin by considering the amounts paid in the *pizzoyear*, with respect to the averages of all the values of balance sheets data we collected.[15] Figure 5 presents the distribution of the observations per province, sector and year.[16]

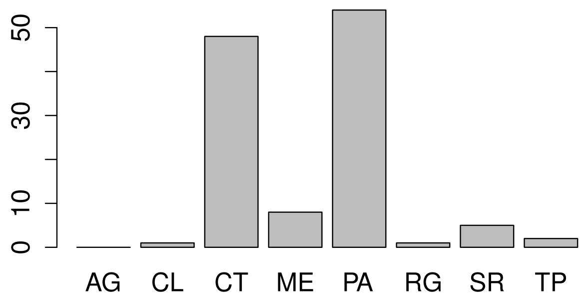
It can be observed that most of the observations come from the provinces of Palermo (PA) and Catania (CT).[17] In addition, the most represented sectors are: "food products", "Construction", "motor vehicules repair", "Wholesale trade", "Retail trade", "Hotels and restaurants" and "Land transport". This distribution suggests that, in line with Lavezzi (2008), extortion affects in particular traditional sectors and small/medium firms.[18] Finally, we can see that the observations spread unevenly in time, with most observations concentrated between the second half of the nineties and the early 2000.

[15]That is we abstract from considering whether these data refer to the period before or after the *pizzoyear*. As noted in footnote (8), most of the *pizzoyear* refer to the ending period of the payment. Therefore, the correct imputation of balance sheet data to the payment should consider only balance sheets for years not greater than *pizzoyear*. The available balance sheets, however, do not fully cover, on average, the periods preceding the years of payment. This is due to the fact that a relevant number of observations refer to the early nineties, while the collection of data in electronic form by the Italian chambers of commerce started in the early nineties as well, but reached a satisfactorily coverage only at the end of the decade (personnal communication from Chamber of Commerce staff in Palermo). The consequence is that when we restrict the analysis to the "before payment" years, we lose around 30% of observations and the estimates we obtain are little precise. If we compute, however, the correlation of the averages computed before and after the year of payment, for the relevant measures of size we utilize in the following (physical capital, total fixed assets and revenues), we find a value of approximately 0.8. Therefore, in the remaining of the cross-section analysis we will consider averages computed on any available year, under the assumption that missing values for the averages on years before payment are are well proxied by their values computed averaging after the year of payment.

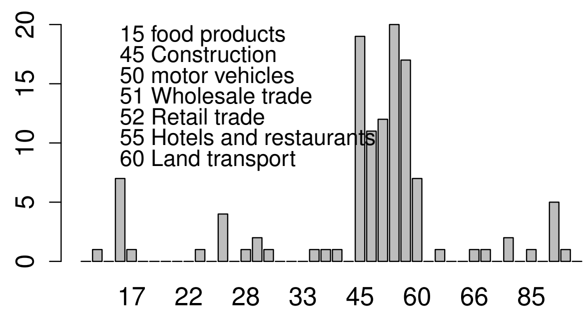
[16]The corresponding distributions for the larger samples of 335 and 489 observations do not show remarkable differences.

[17]The Sicilian administrative provinces are: Trapani (TP), Palermo (PA), Messina (ME), Agrigento (AG), Caltanissetta (CL), Enna (EN), Catania (CT), Ragusa (RG), Siracusa (SR)

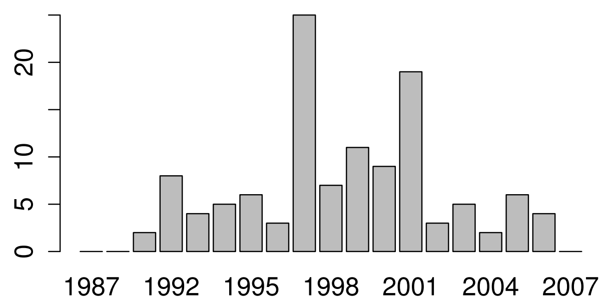
[18]Hi-tech/hi-skill sectors such as those referred to computer equipment of financial services are not represented.



(a) Distribution per province (119 firms)



(b) Distribution per sector (119 firms)



(c) Distribution per year (119 firms)

Figure 5: Distribution of observations per province, sector and year

Table 1 contains average values of *pizzo* for the most represented sectors.

Sector	Average *pizzo*
45 Construction	1579
60 Land transport	880
52 Retail trade	681
50 Motor vehicules repair	451
55 Hotels and restaurants	395
15 Food products	304
51 Wholesale trade	292

Table 1: Average *pizzo* in most represented sectors

It can be observed that these values are significantly different, highlightning a remarkable differentiation in the extortion payments across sectors.[19]

Next we analyze the relation between the amount of *pizzo* and different measures of the size of the firm: physical capital (as the main variable considered in the theoretical model), total fixed assets, revenues and number of workers.[20] Table 2 contains the values of the correlations among these variables.[21], while Figure 6 shows the relationship between the amount of *pizzo* and these measures.

	Phys. Cap.	Tot. fix. assets	revenues	# Workers
Phys. Cap.	1			
Tot. fix. assets	1	1		
Revenues	0.58	0.58	1	
# Workers	0.86	0.86	0.74	1

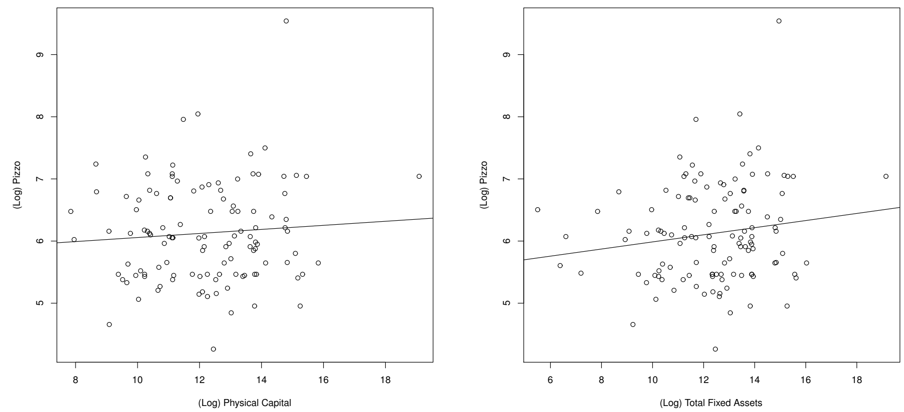
Table 2: Correlations among measures of firm's size

[19]This type of evidence is present also in Asmundo and Lisciandra (2008) who, however, do not analyze firm level data as in this paper.

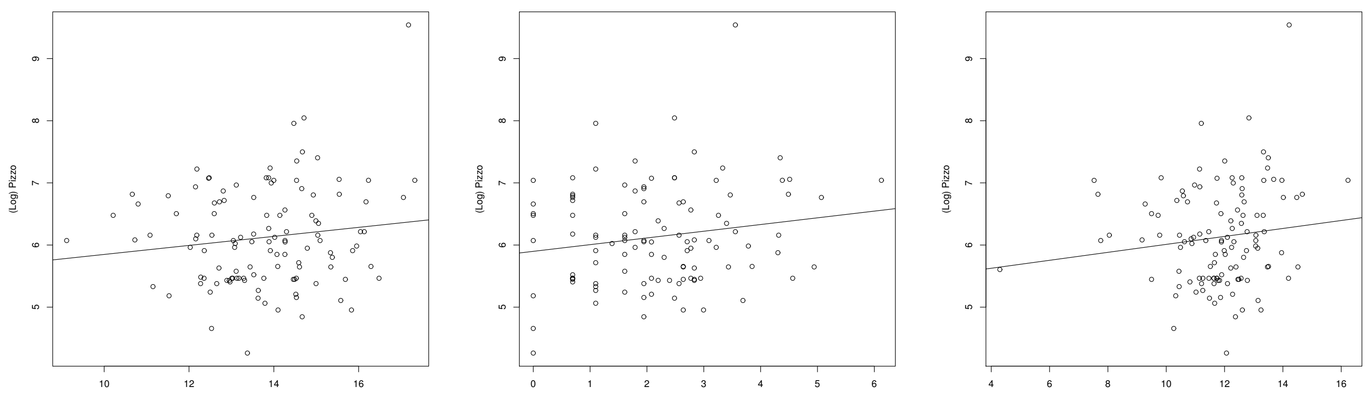
[20]The number of workers in our database does not come from balance sheets as firms are not obliged to report it in such document. This number is available typically for one year in the Chamber of Commerce registration form (*visura camerale*). Here we consider the available observations abstracting from the year to which they refer. A linear fit of the average personnel costs computed for all available years and the number of workers for the firms in our sample returns a highly significant linear relation, suggesting that these individual observations are good proxies of the number of workers actually employed in the period of observation.

[21]All reported values are significantly different from zero.

From Table 2, we note that the correlation between physical capital and total fixed assets is perfect, probably because the firms in our sample are not significantly developed financially and do not possess high levels of immaterial capital stock.[22]



(a) Pizzo and Physical Capital (b) Pizzo and Total Fixed Assets



(c) Pizzo and Revenues (d) Pizzo and Number of Workers (e) Pizzo and Personnel Costs

Figure 6: The relation between the amount of *pizzo* and various measures of firms' size

From Figure 6 It can be observed that the relation between *pizzo* and size is positive,

[22] Financial and Immaterial assets are the two other items regularly indicated in balance sheets of Italian firms in the capital stock.

although the effect on the amount of (log) pizzo as (log) size increases is quite low.

In Tables 3, 4, 5 and 6 we present the results of OLS regressions estimating the elasticities of the amount of *pizzo* on measures of the size of the firm We consider time, province and sector dummies.[23]

	(1)	(2)	(3)	(4)	(5)	(6)
physical capital	0.02 (0.04)	0.06 (0.04)	0.03 (0.04)	0.05 (0.04)	0.04 (0.04)	0.04 (0.04)
province dummies	YES	-	-	YES	-	YES
sector dummies	-	YES	-	YES	YES	YES
time dummies	-	-	YES	-	YES	YES
$\bar{R}^2$	0.07	0.17	0.12	0.16	0.30	0.30
Obs	115	115	115	115	115	115

Table 3: Dependent variable: log(pizzo). Main regressor: log(physical capital). Significance: * 10%, **5%, *** 1%

	(1)	(2)	(3)	(4)	(5)	(6)
total fixed assets	0.05 (0.03)	0.09*** (0.03)	0.04 (0.05)	0.09** (0.04)	0.08** (0.03)	0.08** (0.03)
province dummies	YES	-	-	YES	-	YES
sector dummies	-	YES	-	YES	YES	YES
time dummies	-	-	YES	-	YES	YES
$\bar{R}^2$	0.08	0.20	0.14	0.19	0.34	0.34
Obs	119	119	119	119	119	119

Table 4: Dependent variable: log(pizzo). Main regressor: log(total fixed assets). Significance: * 10%, **5%, *** 1%

[23] Some observations on the independent variable having value equal to zero were dropped in the estimation.

	(1)	(2)	(3)	(4)	(5)	(6)
revenues	0.06 (0.05)	0.15 *** (0.05)	0.03 (0.05)	0.14 *** (0.05)	0.09 * (0.05)	0.09 * (0.05)
province dummies	YES	-	-	YES	-	YES
sector dummies	-	YES	-	YES	YES	YES
time dummies	-	-	YES	-	YES	YES
$\bar{R}^2$	0.07	0.22	0.12	0.21	0.32	0.31
Obs	117	117	117	117	117	117

Table 5: Dependent variable: log(pizzo). Main regressor: log(revenues). Significance: * 10%, **5%, *** 1%

	(1)	(2)	(3)	(4)	(5)	(6)
# Workers	0.08 (0.06)	0.13 ** (0.06	0.07 (0.06)	0.14 ** (0.06)	0.05 (0.07)	0.05 (0.07)
province dummies	YES	-	-	YES	-	YES
sector dummies	-	YES	-	YES	YES	YES
time dummies	-	-	YES	-	YES	YES
$\bar{R}^2$	0.07	0.18	0.11	0.16	0.27	0.25
Obs	109	109	109	109	109	109

Table 6: Dependent variable: log(pizzo). Main regressor: log(number of workers). Significance: * 10%, **5%, *** 1%

	(1)	(2)	(3)	(4)	(5)	(6)
personnel cost	0.05 (0.05)	0.11 ** (0.05)	0.04 (0.05)	0.10 ** (0.05)	0.05 (0.05)	0.05 (0.05)
province dummies	YES	-	-	YES	-	YES
sector dummies	-	YES	-	YES	YES	YES
time dummies	-	-	YES	-	YES	YES
$\bar{R}^2$	0.08	0.17	0.13	0.17	0.30	0.30
Obs	116	116	116	116	116	116

Table 7: Dependent variable: log(pizzo). Main regressor: log(personnel cost). Significance: * 10%, **5%, *** 1%

The results in Tables 3, 4, 5, 6 and 7 show that the elasticity of *pizzo* with respect to the size of the firm is always positive but its absolute value is low. The most significant results appear when total fixed assets and revenues are considered. The significance is in particular increased when sectoral dummies are considered, as it is clear from the comparison of models (1), (2) and (3).[24]

5.2.2 On the incidence of extortion on profits

So far we identified the slope of the relationship between the amounts of *pizzo* and the size of the firm. Our findings show that the relationship is somewhat concave in levels, with the marginal *pizzo* paid for an euro invested in the size of the firm decreasing moderately as size increases. A further question, which is still unsettled in the literature, is what is the actual fraction of gross profits paid as extortion, and how this incidence is correlated with size. In other words, we want to measure the degree of progressivity of Mafia "taxation" .

To this purpose, we compute the average *pizzo rate*, given by the ratio of *pizzo* to the operating profits of the firms,[25] and evaluate it as a function of the size of the firm.[26] Figures 7-12 present the results.

[24]In the regressions presented, no evidence of heteroschedasticity is detected. Removing outliers (in particular the highest pizzo level) reduces the coefficients and the significance of the estimated effect of size. For example, the estimated coefficients of the effect of size on *pizzo* with total fixed assets and revenues decrease, respectively, from approx. 0.10 and 0.15 (signif. at 1%) to approx. 0.7 and 0.10 (significant at 5%).

[25]This value refers to the difference between the revenues of the firm and the production costs. It abstracts, therefore, from other costs, such as interests and from the consideration of taxes. As such, it provides the "purest" measure of the profits from the typical activity of the firm. We computed the rate for firms having positive average operating profits.

[26]Given the high correlation between physical capital and total fixed assets, we consider only the latter in the discussion that follows.

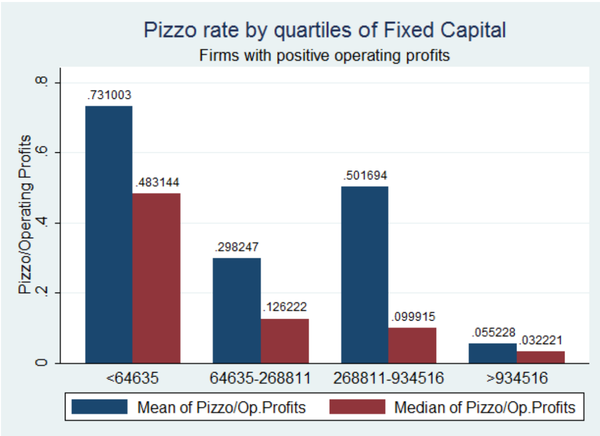


Figure 7: Pizzo rate and Total Fixed Assets: mean and median values per quartile

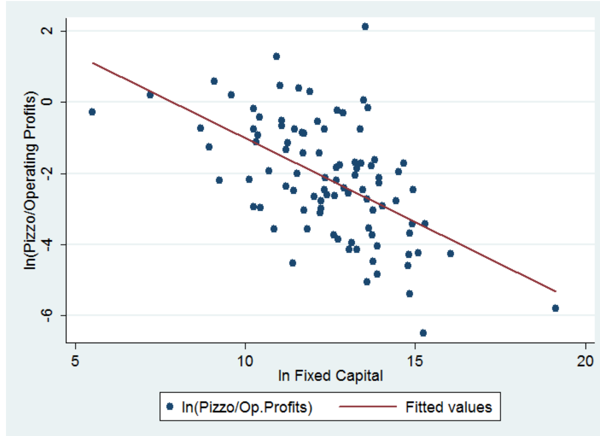


Figure 8: Pizzo rate and Total Fixed Assets

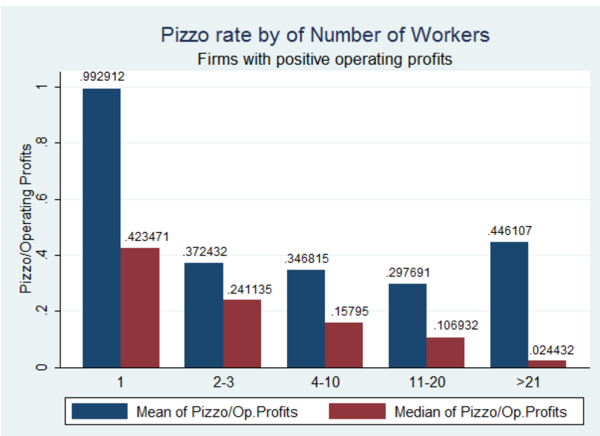


Figure 9: Pizzo rate and number of workers: mean and median values per quartile

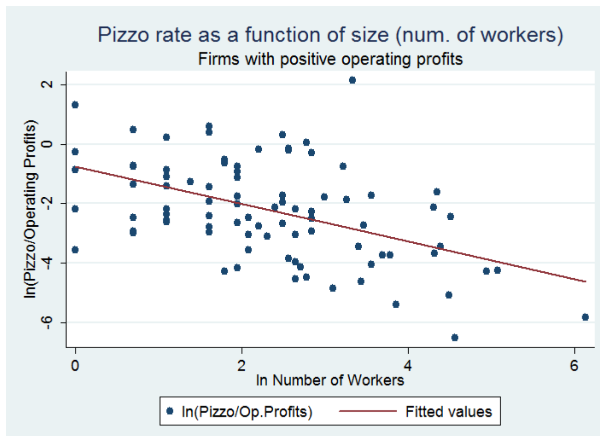
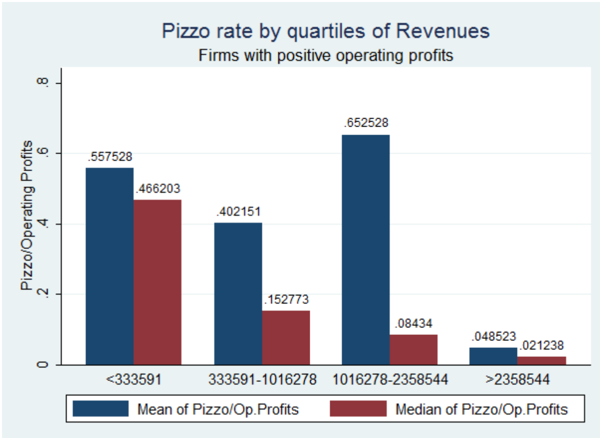


Figure 10: Pizzo rate and number of workers



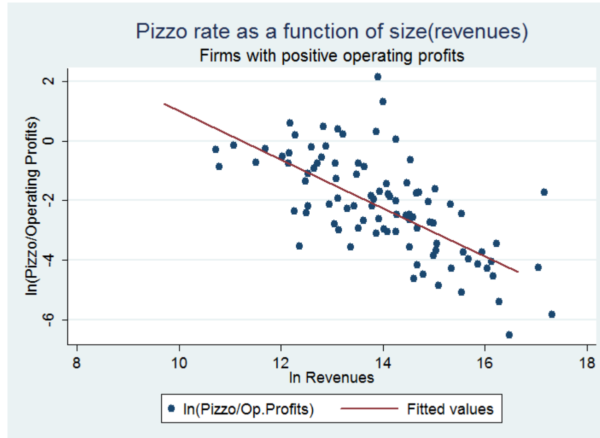


Figure 11: Pizzo rate and revenues: mean and median values per quartile

Figure 12: Pizzo rate and number of workers

In the left panel of each figure, we divide firms in quartiles of size, measured as Total Fixed Assets, Number of Workers and Revenues, and for each class we compute the mean and the median of *pizzo rate*.[27] In the right panel we report the log-log scatter plot of *pizzo rate* against the measure of size. From Figures 7-12 we note that the incidence of *pizzo* strongly decreases with the size of the firm, starting from very high values for the smallest firms.[28] In Table 8 we perform a simple OLS regresson of the *pizzo rate* on the same measures of size, with the same set of control dummies we used in the previous section. The results confirm that the relationship is significantly negative with all measures of firm's size and it is not driven by province, sector or time differences.

[27]We highlight the class size of 1 worker as it is commonly reported in official statistics on firm's size. See also Lavezzi (2008, p. 209).

[28]The non-monotonicity of the mean values is due to a single observation only, having a very large value of *pizzo*.

Variable	(1)	(2)	(3)
Total fixed assets	-0.53^{***} (0.11)		
Revenues		-0.81^{***} (0.12)	
# Workers			-0.82^{***} (0.18)
province dummies	YES	YES	YES
sector dummies	YES	YES	YES
time dummies	YES	YES	YES
R^2	0.70	0.77	0.72
Obs	94	94	88

Table 8: Dependent variable: log(pizzo/operating profits). Independent variables in logs. Significance: * 10%, **5%, *** 1%.

Thie exercise allows us to asses quantitatively the relevance of extortion for firms' profitability. There is substantial heterogeneity across class sizes and the incidence of pizzo decreases as size increases. The difference between the values of the *pizzo rate* for the smallest and largest firms in our sample appears striking: the smallest firms in our sample are taxed for a median value of more than 40%, while the largest are subjected to an extortion rate around 2%. , as largest firms are "taxed" for only approximately 2% of their operating profits. Overall, we find clear evidence that *Mafia taxation is regressive.*

5.2.3 On *pizzo* and average costs

Another aspect of the presence of organized crime on a territory is the possible deterrence of firms from starting their business. In other words, the presence of organized crime may represent a barrier to entry into a market. Criminal organization offer protection in various forms (Varese, 2006, p. 412), protection from competitors being on of them. This service can be offered through direct means, such as intimidation or the control of markets for labor and intermediate input. A more subtle, and economically interesting, barrier to entry is the imposition of a structure of extortion, as in the present paper, that modifies the minimum efficient scale of the market and hence the structure of competition. We already showed that the amount of *pizzo* does not seem to strongly react to changes in size. Here we perform a preliminary exercise to assess the impact of extortion on the cost structure of firms. If extortion is to have an effect on competition it should be the case that the shape of average cost function is affected by theinclusion of *pizzo*. In Table 9 we report the results of a regression of average costs on the level of production, excluding and including *pizzo* in the costs.[29]

	Dependent Variable	
Regressor	log av costs (without pizzo)	log av costs (with pizzo)
log(revenues)	-0.01 (0.16)	-0.04^{**} (0.02)
province dummies	YES	YES
sector dummies	YES	YES
time dummies	YES	YES
R^2	0.36	0.41
Obs	117	117

Table 9: (Heteroscedasticity-consistent) standard errors in parenthesis. Significance: * 10%, **5%, *** 1%.

It can be observed that in our sample average costs are approximately constant in output if *pizzo* is not included in costs. With the inclusion of extortion in costs, the average cost function becomes *decreasing* and the coefficient is statistically significant at the 5% level. It appears, therefore, that *pizzo* acts as a sort of fixed costs and may therefore influence the competitive structure of the economy.

5.2.4 Analysis of Sectoral Effects on Extortion

TO BE WRITTEN

6 Concluding remarks

TO BE WRITTEN

[29]The level of production is proxied by the level of revenues.

A Sample Validation

In this Appendix we describe the procedure we followed to assess the validity of our sample. In Section A.1 we evaluate its capacity to capture the phenomenon of interest, i.e. the presence and activity of the Sicilian Mafia on the territory. The focus is on the data on *pizzo*. In Section A.2 we separately analyze the sample of firms, to evaluate if it provides a non distorted picture of a population of firms. In this case the focus is on the data collected from balance sheets.

A.1 On the Representivity of the Phenomenon of Organized Crime

The procedure followed to build the sample makes it clearly non-random. As remarked, the presence of information of an extortive activity in the database depends on: i) being present in a court document, i. e. having been reported or discovered by investigation; ii) the corresponding document having being examined by the researchers and/or the piece of evidence being mentioned by a judge in an interview. We highlighted that the bias due to self-selection of entrepreneurs denouncing extortionists should be attenuated by the fact that some records originate from declarations of state witnesses or from independent investigations by prosecutors. It is not possible, however, to quantify the role of these different sources in the data analyzed in this paper.

Given these sources of bias, in this section we assess to what extent the sample is a good representation of the phenomenon of interest, i.e. the activity of Cosa Nostra in Sicily. Specifically, we examine the distribution of the observations in the sample across the Sicilian provinces and then compute correlations of such distibution with respect to indicators of the presence of Mafia on the territory, coming from external sources.

Table 10 reports the distribution of the observations: i) in the original database (2264 observations); ii) in the subsample of observations with reported *pizzo* value for monthly payments (489 observations); iii) in the subsample of observations with reported *pizzo* value for identified firms (335 observations); iv) in the subsample of observations with reported *pizzo* value for identified limited liability firms (145 observations) and, finally, v) in the subsample of observations with reported *pizzo* value for identified firms that we used for the cross-section analysis (119 observations).

	TP	PA	ME	AG	CL	EN	CT	RG	SR	TOT
Obs.	241	847	163	81	157	9	439	182	145	2264
%	10.6	37.4	7.2	3.6	6.9	0.4	19.4	8.0	6.4	100
Obs.	4	142	41	1	34	0	229	1	37	489
%	0.8	29.0	8.4	0.2	7.0	0.0	46.8	0.2	7.6	100
Obs.	3	98	28	1	11	0	169	1	24	335
%	0.9	29.3	8.4	0.3	3.3	0.0	50.4	0.3	7.2	100
Obs.	2	60	9	0	5	0	62	1	6	145
%	1.4	41.4	6.2	0.0	3.4	0.0	42.8	0.7	4.1	100
Obs.	2	54	8	0	1	0	48	1	5	119
%	1.7	45.4	6.7	0.0	0.8	0.0	40.3	0.8	4.2	100

Table 10: Distribution of observations across provinces: full sample and various subsamples. The Sicilian provinces are: Trapani (TP), Palermo (PA), Messina (ME), Agrigento (AG), Caltanissetta (CL), Enna (EN), Catania (CT), Ragusa (RG), Siracusa (SR)

It can be noted that observations are unevenly distributed across the provinces, with the provinces of Palermo and Catania being the most represented, both in the full sample and in the subsamples.[30] With respect to the largest subsample, in the smallest subsample of 119 firms, there is a clear underepresentation of the provinces of Trapani (TP), Caltanissetta (CL) and Ragusa (RG).

To evaluate whether our sample faces a severe distortion in the representation of the intensity of Mafia activities across the Sicilian provinces, we compare the distributions of observations to the distribution of various measures of the presence of the Mafia at provincial level. These measures are constructed by considering data on mafia-related crimes (including extortion), other indicators of the presence of the mafia (e. g. confiscated properties), or other socio-economic indicators.[31] Since this type of measurement necessarily involves some degree of arbitrariness, we report different indicators taken from Calderoni (2011).[32]

Table 11 contains the values of different indicators of Mafia presence in the nine Sicilian administrative provinces.[33] The indices *Mirate* and *Mirank* were constructed by Calderoni

[30]Palermo and Catania are the largest Sicilian cities.

[31]See Calderoni (2011) for a thorough discussion.

[32]Van Dijk (2007) provides a similar analysis at cross-country level.

[33]The reported indices are: *IPM* (*Indice di penetrazione mafiosa*): constructed by *Eurispes* (*Istituto di Studi Politici Economici e Sociali*) in 2010. It is based on Mafia-related crimes (extortion, mafia association, drug smuggling, etc.) and socio-economic indicators (unemployment, trust in institutions, etc.), and aims at measuring the capacity of the Mafia to penetrate a terrritory (see Calderoni , 2011, fn. 11). *POPM*: constructed by *Censis* (*Centro Studi Investimenti Sociali*) for the years 2004-2006. It measures

	IPM	POPM	ICC	Mirate	Mirank
PA	35.50	90.90	45.90	50.37	83.22
CL	33.10	95.20	44.20	42.20	84.50
CT	52.40	79.70	33.70	32.12	82.50
TP	29.40	91.00	31.60	29.42	77.86
AG	28.90	95.90	23.20	23.52	71.75
EN	29.20	73.80	30.40	17.21	57.74
SR	38.60	88.70	16.60	12.74	50.71
RG	28.40	57.50	25.60	17.83	61.82
ME	31.90	57.10	21.10	15.44	60.82

Table 11: Indicators of Mafia presence at provincial level

(2011) to overcome various shortcomings of previously constructed indices such as *IPM*, *POPM* and *ICC*. Therefore, they should be considered the most accurate.

The order of the provinces in Table 11 is based on the average ranking of provinces in the indices presented.[34]. This order suggests that the "quantity" of Mafia in the provinces of Palermo (PA), Caltanissetta (CL) and Catania (CT) is relatively high, while it is lower in the provinces of Siracusa (SR), Ragusa (RG) and Messina (ME).

Table 12 contains the values of the correlation between the distribution of observations across the provinces and the indices of mafia intensity for the full sample of 2264 observations. In the second row of the table we considered the distibution of observations on extortion normalized by the average population size over the period of interest.[35]

the population of each province living in municipalties with recorded Mafia activities, as a percentage of total provincial population. ICC (*Indice di contesto criminale*), was proposed by Calderoni and Caneppele (2009) in a study of infiltration by organized crime in public procurement, and is based on an average of various mafia-related murders, including indicators of infiltration in public procurement, for a period approximately covering the decade 1995-2005. *Mirank* and *Mirate* are new indices proposed by Calderoni (2011) based on four dimensions of Mafia activities: mafia-type associations, mafia murders, city councils dissolved for mafia infiltration, assets confiscated to organized crime. *Mirate* is based on averages of the four indicators, while *Mirank* is based on the average *ranking* of each province along the four dimensions. The period covered goes from the early eighties to the interval 2007-2009. Index *POPM* is related to other two *Censis* indices, the percentage of municipalties with recorded mafia activities, and their surface. Since these indices are quite rough, we report only what appears to be the most sensible. Results are not affected by this choice. Other indices exist but are defined at regional level. For further details and references see Calderoni (2011).

[34]In particular, PA and CL have the same average ranking, as well as SR and RG

[35]In particular, for each province we averaged the values of provincial total population for the period 1992-2004, which contains approximately 88% of the observations in the full sample. Data on population were downloaded from the Italian Institute of Statistics (ISTAT) website (`http://demo.istat.it/`).

IPM	POPM	ICC	Mirate	Mirank
0.43 (0.25)	0.21 (0.60)	0.61* (0.08)	0.75** (0.02)	0.58 (0.10)
0.31 (0.42)	0.54 (0.13)	0.67* (0.05)	0.81** (0.01)	0.71** (0.03)

Table 12: Correlation across the distribution of observations and indicators of Mafia activity. *: significant at 10%, **: significant at 5%. P-values in parenthesis. Full Sample

It can be noticed that the correlation is fairly high and significant for *ICC*, *Mirate* and *Mirank*.

Table 13 contains the results of this analysis for the subsample of 119 observations used in the cross-section analysis.

IPM	POPM	ICC	Mirate	Mirank
0.71** (0.03)	0.12 (0.76)	0.51 (0.16)	0.62* (0.08)	0.53 (0.14)
0.80** (0.01)	0.11 (0.78)	0.42 (0.26)	0.54 (0.13)	0.47 (0.20)

Table 13: Correlation across the distribution of observations and indicators of Mafia activity. *: significant at 10%, **: significant at 5%. P-values in parenthesis. Sample of 119 limited liability companies

From Table 13 we note that, with respect to the full sample which contained a much larger number of observations, the correlations are generally lower and less significant. with the exception of IPM. However, one correlation with $Mirate$ remains significant, while the p-values when the null is not rejected for $Mirate$ and $Mirank$ are not extremely high.[36]

We conclude that our sample does not offer a serious misrepresentation of the presence of organized crime across the Sicilian provinces. The biggest shortcoming seems to be a lack of observations for the province of Caltanissetta (CL), where high "levels" of Mafia are recorded.

[36]The index which performs worse is POPM. As noted, this index and the two related ones (percentage of municipalties with recorded mafia activities and their surface) are quite inaccurate with respect to the others. The correlations with the latter two (unreported) are never significant.

A.2 Estimation of Production Functions

As a complementary step in the evaluation of the representativity of the sample used in this paper, we separately analyze the sample of firms. In particular, we estimate production functions using data from their balance sheets to assess if the estimated coefficients have economically meaningful and significant values.

Specifically, we utilize the panel dimension of the dataset to estimate the following production function:

$$Y = Af(K, N, M) = AK^{\alpha}N^{\beta}M^{\gamma},$$

where K is capital, N is labor and M refers to intermediate inputs. Taking logs, we obtain:

$$y = a + \alpha k + \beta n + \gamma m.$$

The dataset is an unbalanced panel of 119 firms. Output is measured with gross nominal revenues, labor as the cost of labor. We measure capital with either the book value of total fixed capital (which includes tangible, intangible and financial capital) or tangible fixed capital alone. The cost of intermediate inputs is computed as the difference:

$$M = \text{Revenues-Cost of Labor-Amortization-Operating Profits.}$$

The results are reported in Table A.2. Columns (1) and (2), labeled means, are regressions of population averages over time. That is:

$$\mathbb{E}_t y_{ijt} = a + \alpha \mathbb{E}_t k_{ijt} + \beta \mathbb{E}_t n_{ijt} + \gamma \mathbb{E}_t m_{ijt} + d_j + \varepsilon_{ij},$$

where d_j are dummy variables controlling for province and sector. Columns (3) and (4) report the results for fixed effects regressions, that is:

$$y_{ijt} = a + d_{ijt} + \alpha k_{ijt} + \beta n_{ijt} + \gamma m_{ijt} + \varepsilon_{ijt}.$$

Columns (5) and (6), finally, report the results of estimations based on the approach of Levinsohn and Petrin (2003).[37]

[37] Their estimation algorithm is summarized in Petrin et al. (2004).

Dep.Var. Gross Revenues	(1) Means	(2) Means	(3) FE	(4) FE	(5) LevPet	(6) LevPet
Tangible Fixed Capital		0.0116		0.0336**		0.0888*
		(0.0146)		(0.0142)		(0.0467)
Labor	0.224***	0.232***	0.274***	0.266***	0.221***	0.235***
	(0.0245)	(0.0246)	(0.0201)	(0.0199)	(0.0370)	(0.0373)
Materials	0.752***	0.760***	0.617***	0.647***	0.743***	0.496**
	(0.0238)	(0.0244)	(0.0197)	(0.0201)	(0.225)	(0.219)
Total Fixed Capital	0.0272*		0.0543***		0.0316	
	(0.0153)		(0.0144)		(0.0594)	
Constant	0.822**	0.765*	1.640***	1.605***		
	(0.385)	(0.397)	(0.241)	(0.246)		
Observations	950	940	950	940	950	940
R-squared	0.986	0.985	0.777	0.776		
Number of firms	114	112	114	112		

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

(1) and (2) include dummies for Province and sector.

(5) and (6) Province and sectors are included as free variables.

Table 14: Production Function Estimates

The results in Table A.2, in particular those in Columns (5) and (6), show that the coefficient on materials is generally higher than the coefficient on labor and capital, the latter being the smallest. These results seem to be consistent with the values generally obtainable in this type of estimations (see Van Beveren, 2012, p. 117). We conclude, therefore, that the sample of firms used in this paper does not feature severe distorsions in the representation of a population of firms.

References

Asmundo, A., and M. Lisciandra (2008), "The Cost of Protection Racket in Sicily", *Global Crime* 9, 221-240.

Bonaccorsi di Patti, E. (2009), "Weak institutions and credit availability: the impact of crime on bank loans", *Occasional Paper Nr 52*, Bank of Italy.

Bowman, A. W. and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis*, Oxford, Clarendon Press.

Calderoni, F. (2011) "Where is the Mafia in Italy? Measuring the Presence of the Mafia Across Italian Provinces", *Global Crime* 12, 41-69.

Calderoni, F. and S. Caneppele, eds., (2009), "La geografia criminale degli appalti: le infiltrazioni della criminalit organizzata negli appalti pubblici nel Sud Italia". Milano: Franco Angeli.

Gambetta, D. (1993), *The Sicilian Mafia: the Business of Private Protection*, Harvard University Press.

La Spina, A. (ed.) (2008), *I costi dell'illegalità. Mafia ed estorsioni in Sicilia*, Il Mulino.

Lavezzi, A. M. (2008), "Economic Structure and Vulnerability to Organized Crime: Evidence from Sicily", *Global Crime* 9, 198-220.

Levinsohn, J., and A. Petrin (2003), "Estimating production functions using inputs to control for unobservables", *The Review of Economic Studies*, 70(2), 317-341.

Paoli, L. (2003), *Mafia Brotherhoods. Organized Crime, Italian Style*, Oxford University Press.

Petrin, A., Poi, B. P., and J. Levinsohn, (2004), "Production function estimation in Stata using inputs to control for unobservables", *Stata Journal*, 4, 113-123.

Konrad, K. A. and S. Skaperdas (1998), "Extortion", *Economica* 65, 461-77.

Van Dijk, J. (2007), "Mafia Markers: Assessing Organized Crime and its Impact upon Societies", *Trends in Organized Crime* 10, 39-56.

Van Beveren, I. (2012), "Total factor productivity estimation: A practical review", *Journal of Economic Surveys*, 26(1), 98-128.

Scaglione, A. (2008), "Le interviste agli imprenditori", in La Spina, A. (ed.), *I costi dell'illegalità. Mafia ed estorsioni in Sicilia*, Il Mulino.

Varese, F. (2006a), "How Mafias Migrate: The Case of the'Ndrangheta in Northern Italy", *Law and Society Review* 40, 411-444