Testing Financing Constraints on Firm Investment using Variable Capital*

Andrea Caggese
Pompeu Fabra University

First version: 12 February 2003

Abstract

We develop a structural model of a firm that uses a technology with both fixed and variable capital and is subject to both borrowing constraints and irreversibility of fixed capital. We show that the premium of marginal productivity over the user cost of variable capital is a theoretically consistent indicator of the intensity of financing constraints. We use this result to develop a new test of the presence of financing constraints on firm investment: if a firm is subject to borrowing constraints, then the indicator should be monotonously decreasing in the financial wealth of the firm, conditional on the productivity shock and on the stock of fixed capital. We test this hypothesis on a sample of small and medium Italian manufacturing firms. The indicator is estimated using a panel of balance sheet data and the financing constraints hypothesis is not rejected for all the firms in the sample except the larger ones. Importantly, the validity of this test is strongly supported by an independent source of qualitative information: firms with a high value of the indicator are three times more likely to state problems in financing investment than firms with a low value, even conditional on their size.

*The paper is based on the 5th chapter of my PhD thesis at the London School of Economics. I am most grateful to Nobu Kiyotaki for his encouragement and valuable feedback on my research. I would like also to thank Orazio Attanasio, Steven Bond, Heski Bar-Isaac, Vicente Cunat, Francois Ortalo-Magne' and Steve Pischke for their valuable comments and suggestions on earlier versions of this paper, as well as the participants to the LSE and FMG Ph.D. seminars. All errors are, of course, my own responsibility. Research support from the Financial Markets Group and from Mediocredito Centrale are gratefully acknowledged. Please address all correspondence to: andrea.caggese@econ.upf.es or Pompeu Fabra University, Department of Economics, Room 1E58, Calle Ramon Trias Fargas 25-27, 08005, Barcelona, Spain

1 Introduction

In order to explain the aggregate behaviour of investment and production, it is important to understand the factors that determine the investment decisions of firms. Some authors argue that the availability of internal finance may be important in explaining these decisions. It may affect the ability of firms to invest when external finance is not available. Financiers may be unwilling to fund profitable investment opportunities because once the funds have been handed to the firms, contractual and/or informational problems may prevent the financiers from appropriating their share of the revenues from the investment's output (Besanko and Thakor (1986), Milde and Riley (1988), Hart and Moore (1998), Albuquerque and Hopenhayn (2000)). It is therefore important to verify empirically whether or not financing constraints affect the investment decisions of firms. A literature started by Fazzari, Hubbard and Petersen (1988) focuses on the consideration that, if firms are unable to raise external financing, they only invest when internally generated funds become available. Several studies[1] show that this seems to be the case: investment is significantly correlated with proxies for changes in net worth or internal funds, and such correlation is most important for firms likely to face capital-market imperfections.

The motivation of this paper is that this result has been seriously questioned as evidence of financing constraints on firm investment. Kaplan and Zingales (1997 and 2000) find that the cash flow-investment correlation is stronger for firms which are financially very wealthy and, according to their selection criteria, surely not financially constrained[2]. More generally, Kaplan and Zingales claim that there is no theoretical support for the fact that the cash flow-investment sensitivity is monotonously increasing in the intensity of financing constraints. This claim has been proved recently by Gomes (2001) in a general equilibrium framework, while both Gomes (2001) and Ericson and Whited (2001) show that measurement errors are the most likely cause of the positive correlation between investment and cash flow.

In this paper we adopt a different approach. We develop a structural model of firm

[1] See Hubbard (1998) for a review of this literature.
[2] Similar evidence is showed by Cleary (1999), who studies a larger sample of 1317 US firms.

investment with both financing constraints and irreversibility of fixed capital, and we use it to derive a theoretically consistent procedure to test for the presence of financing constraints on firm investment. We consider a firm which has the opportunity to invest in a risky technology which generates output using two complementary factors of production, fixed and variable capital. Both factors take one period to become productive. Fixed capital cannot be disinvested unless the firm is liquidated, while variable capital is reversible. Because of an enforceability problem, the firm can obtain external financing only if it secures it with collateral. The only collateral accepted by the lenders is the physical capital used in the production. This implies that the firm needs some downpayment to finance investment, and that its borrowing capacity depends on its financial wealth. We determine the conditions under which, because of the uncertainty about productivity, the firm has a positive probability of facing financing constraints in equilibrium.

Since variable capital investment is reversible, then the "premium" of expected marginal productivity over user cost of variable capital reflects the tightness of current and future expected financing constraints. More specifically the financing constraints hypothesis implies that this premium is monotonously decreasing in the financial wealth of the firms, conditional on their fixed capital stock and their productivity shock. This test has two important properties: our structural model directly relates the value of the premium to the intensity financing constraints[3]. Therefore it is robust to the Kaplan and Zingales (1997) critique. Moreover the test maintains its power of discriminating the financing constraints hypothesis from the perfect markets hypothesis in the presence of two potential misspecification problems: the presence of adjustment costs in investment and the misspecification of the stochastic process for the productivity shock. We argue that this is because in the presence of these two problems the test is biased towards "rejecting the financing constraints hypothesis when it is true" rather than "accepting it when it is false".

We test the financing constraints hypothesis by estimating empirical measures of the

[3]Indeed such indicator can be consistent with other models of investment with financing imperfections. for example a similar procedure has been suggested also by Albuquerque and Hopenhayn (2000): "*... the theory predicts that short run financing constraints can only be identified by estimating the process for excess marginal return to production*".

productivity shocks and of the expected marginal productivity of variable capital for a sample of 561 small and medium Italian manufacturing firms from Mediocredito Centrale. We use the information from 11 years of balance sheet data, from 1982 to 1992, to estimate the indicator of the intensity of financing constraints for each firm-year observation. A unique feature of this sample is the availability, for the same firms, of a rich survey with qualitative information about their financial decisions and especially about the financing problems they faced in funding investment in the 1989-1991 period. This information is used to perform an independent robustness check which strongly supports the validity of our estimated indicator of the intensity of financing constraints: firms with an high value of the indicator are three times more likely to state financing problems in funding investment than firms with a low value. We then use the estimated indicator to verify the prediction of the model, and we show that the financing constraints hypothesis is not rejected for all the firms in the sample but the larger ones.

This paper is organised as follows: section 2 describes the model; section 3 defines the financing constraints test and the estimation strategy, showing the results of the production function estimation; section 4 verifies the validity of our indicator of the intensity of financing constraints using the qualitative data from Mediocredito Centrale; section 5 tests the financing constraints hypothesis; section 6 summarises the conclusions.

2 The model

The aim of this section is to provide a theoretical framework that supports the financing constraints test performed in the remaining sections of the paper. We will analyse a model of firm investment based on chapter 4 in Caggese (2002). We consider a risk neutral manager of a firm which has the objective to maximise the discounted sum of future expected dividends. The discount factor is equal to $1/R$, where $R = 1 + r$, and r is the lending/borrowing risk free interest rate. In an environment with limited access to external funds these assumptions imply that the firm never distributes dividends and retains earnings in the form of financial assets until there is even the smallest chance of facing future financing constraints. Therefore in order to allow for the presence of financially constrained firms in equilibrium, we include the following assumption[4]:

Assumption 1: A fixed share of output $\eta > 0$ is nontradable, and cannot be reinvested in the activity, but only distributed as dividend:

$$d_t = d_t^* + \eta y_t \tag{1}$$

$$d_t^* \geq 0 \tag{2}$$

We define $(1+\eta)y_t$ as total output. This is composed by ηy_t, the nontradable output, and y_t, which is "financial" output that can be distributed as financial dividend d_t^*or reinvested in the firm. d_t is total dividends. ηy_t can be interpreted as private benefits accruing to the shareholders of the firm. Its presence implies that the firm behaves like an empire builder a la Jensen (1986). If the firm has free cash flow available it invests in projects which are inefficient in financial terms, in order to increase output and the share of nontradable output. This overinvestment is counterbalanced by the presence of financing constraints. When financial wealth is low, future expected financing constraints induce the firm to downsize investment in order to increase financial profits. This increases cash flow, reduces future expected financing constraints and induces the firm to expand activity again. Thus the firm never accumulates so much financial wealth to become unconstrained forever. Finally, we assume that the expected lifetime of the firm is finite. We consider a parameter γ, arbitrarily close to 1, such that each period with probability $1-\gamma$ the firm's technology becomes useless. In this case the firm is liquidated[5].

Regarding the technology, the firm operates with two inputs, k_t *and* l_t, which are respectively fixed and variable capital, installed at or before time $t-1$, which will generate output at time t. Variable capital represents variable inputs such as materials and work in progress, while fixed capital represents fixed inputs such as plant and equipment. The production function is the following:

$$y_t = \theta_t k_t^\alpha l_t^\beta \text{ with } \alpha + \beta < 1 \tag{3}$$

[4]Caggese (2003) shows that a similar result is obtained in the context of a firm owned and managed by a risk averse entrepreneur who discounts future at a rate higher than the market interest rate.

[5]This assumption in conjunction of assumption 1 is necessary to allow for the presence of financing constraints in equilibrium. See Caggese (2002) for details

All prices are assumed constant and normalised[6] to 1. θ_t is a productivity shock that follows a first order stationary autoregressive stochastic process:

$$\begin{aligned} \theta_t &= \overline{\theta} + \rho\theta_{t-1} + \zeta_t; \text{ with } 0 \leq \rho < 1 \\ \zeta_t &\sim iid\left(0, \sigma_\zeta^2\right) \end{aligned} \tag{4}$$

The difference between the inputs is that variable capital is nondurable, while fixed capital is durable:

$$1 = \delta_l > \delta_k \tag{5}$$

δ_l and δ_k are the depreciation factors of variable and fixed capital respectively. Variable capital is reversible, while fixed capital is irreversible, and can only be disinvested if the whole firm is sold. Therefore conditional on continuation the firm is subject to the following constraints:

$$k_{t+1} \geq (1 - \delta_k)\, k_t \tag{6}$$

Irreversibility of fixed capital is justified by the fact that plant and equipment usually do not have a secondary market because they cannot be easily converted to other productions. Yet we allow fixed capital to be used as collateral by assuming that such conversion is easier if the whole of the assets is sold. The assumption that fixed capital is irreversible conditional on continuation is consistent with the empirical evidence on a very large sample of US manufacturing plants analysed by Caballero, Engel and Haltiwanger (1995).

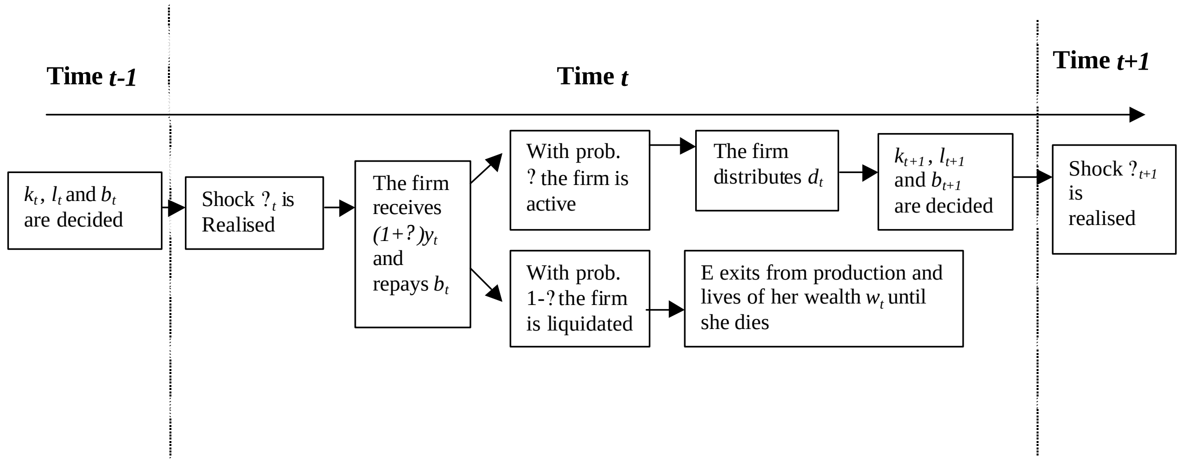
Financial imperfections are introduced by assuming that new shares issues and risky debt are not available. At time t the firm can borrow from (and lend to) the banks one period debt, with face value b_{t+1}, at the market riskless interest rate r. A positive (negative) b_{t+1} indicates that the firm is a net borrower (lender). Banks only lend secured debt, and the only collateral they accept is the next period residual value of physical capital. Therefore at time t the amount of borrowing is limited by the following constraint:

$$b_{t+1} \leq \tau_k k_{t+1} \tag{7}$$

$$\tau_k \leq (1 - \delta_k) \tag{8}$$

[6] This simplifying assumption will obviously be relaxed in the next section.

Figure 1: The timing of the model



τ_k is the share of fixed capital value that can be used as collateral[7]. From equation (5) it follows[8] that $\tau_l = 0$. The rationale for constraint (7) is that the firm can hide the revenues from the production. Being unable to observe such revenues the lenders can only claim, as repayment of the debt, the value of the firm's physical assets (Hart and Moore, 1998). Therefore the firm can only lend or borrow one period secured debt at the market interest rate r offered by the banks.

The presence of irreversibility of fixed capital and of financing constraints implies that in some situations the firm may be forced to be liquidated to repay the debt. If revenues allow the firm to repay the debt and to continue, it may still be optimal to liquidate it if expected short term return is so low as to offset long term gains from continuing activity. Although the interactions between financing constraints and entry-exit dynamics of firms is an interesting topic to explore in future research[9], it goes beyond the scope of this paper. Hence we restrict the parameter space to the values such that forced or voluntary exit never happens in equilibrium[10]. The timing of the model is illustrated in figure 1. At the beginning of period t the firm inherits from time $t-1$ the stocks of fixed and variable capital k_t and l_t. Then θ_t is realised, $(1+\eta)\,y_t$ is produced and b_t repaid. Residual wealth

[7] $\tau_k < 1-\delta_k$ implies that the firm can 'steal' a $1-\tau_k$ fraction of the residual value of capital $(1-\delta_k)\,k_t$.

[8] This assumption is not essential. The results would hold if we instead assumed that variable capital can be collateralisable, as long as some downpayment is needed to finance investment.

[9] Caggese (2003) analyses how the interactions between irreversibility and financing constraints are useful in explaining several stylised facts about investment dynamics

[10] See Caggese (2002), chapter 4 and appendix 1 for details. This restriction does not change the implications of the model for the test of the financing constraints hypothesis, it is only included to make the model's analysis easier.

w_t, net of non tradable output ηy_t is the following:

$$w_t = y_t + (1-\delta_k)k_t - b_t \tag{9}$$

After producing the firm is liquidated with probability γ. Conditional of continuation the firm borrows new one period debt and allocates the net wealth plus the new borrowing between dividends, investment in fixed capital and investment in variable capital. The problem is interesting for the values of η high enough so that the firm never accumulates enough financial resources to eliminate the probability to face future financing constraints. It is possible to prove the following:

Proposition 1 *if $\eta \geq \eta_{\min}$ then an active firm has always some probability of facing future financing constraints and never distributes financial dividends, i.e. $d_t = \eta y_t$ and $d_t^* = 0$ for $t = 0, 1, ..., \infty$.*

Proof: see Caggese[11] 2002.

Proposition 1 implies that conditional on continuation financial dividends are always zero. Hence the firm is subject to the following budget constraint:

$$l_{t+1} + k_{t+1} = w_t + b_{t+1}/R \tag{10}$$

Let's denote the value at time t of an active firm, after θ_t is realised, by $W_t(w_t, \theta_t, k_t)$:

$$W_t(w_t, \theta_t, k_t) = \eta y_t + \underset{\begin{pmatrix} k_{t+1+j} = k(w_t, \theta_t, k_t) \\ l_{t+1+j} = l(w_t, \theta_t, k_t) \\ b_{t+1+j} = b(w_t, \theta_t, k_t) \end{pmatrix}_{j=0,...\infty}}{MAX} E_t\left\{\sum_{t=0}^{\infty}\left(\frac{\gamma}{R}\right)^j\left\{\frac{1}{R}\left[\eta y_{t+1+j} + (1-\gamma)w_{t+1+j}\right]\right\}\right\} \tag{11}$$

The firm maximization problem is now defined by (11) subject to (6), (7) and (10). These constraints define a compact and convex feasibility set for l_{t+1}, k_{t+1} and b_{t+1}, and the law of motion of w_{t+1} conditional on w_t, k_t and θ_t is continuous. Therefore, given the assumptions on θ_t and the concavity of the production function, a solution

[11] Caggese (2002) shows that, for a large range of reasonable parameter space, $\eta_{\min}$ is aroud 15%-20%.

to the problem exists and is unique[12]. In order to describe the optimality conditions of the model, let μ_t, λ_t and ϕ_t be the Lagrangian multipliers associated respectively to constraints (6), (7) and (10). Taking the first order conditions of (11) with respect to b_{t+1}, l_{t+1} and k_{t+1} it is possible to show that the solution is given by the optimal sequence of $\{b_{t+1}, k_{t+1}, l_{t+1}, \lambda_t, \mu_t, \phi_t \mid k_t, w_t, \theta_t; \Theta\}_{t=1}^{\infty}$ that satisfies (6), (12), (13), (14) and (15) plus the standard Kuhn-Tucker complementary slackness conditions on λ_t and μ_t:

$$Dk_{t+1} + l_{t+1} \leq w_t \tag{12}$$

$$\phi_t = 1 + RE_t\left[\sum_{j=0}^{\infty} \gamma^j \lambda_{t+j}\right] \tag{13}$$

$$(1+\eta)\, E_t\left(MPL_{t+1}\right) = UL + R^2\lambda_t + E_t\left(\text{-}\ _{l,t+1}\right) \tag{14}$$

$$(1+\eta)\, E_t\left(MPK_{t+1}\right) = UK + DR^2\lambda_t + E_t\left(\text{-}\ _{k,t+1}\right) + (1-\delta)\,\gamma E_t\left(\mu_{t+1}\right) - R\mu_t \tag{15}$$

$D = 1 - \tau_k/R$ is the downpayment required to purchase one additional unit of fixed capital. Equation (12) combines together the budget constraint (10) and the collateral constraint (7) and implies that the downpayment necessary to buy k_{t+1} and l_{t+1} must be lower than the firm's net worth. Equation (13) is obtained by solving recursively forward the first order condition for b_{t+1}. Equations (14) and (15) are obtained using the first order condition for b_{t+1} to substitute ϕ_t in the first order conditions for k_{t+1} and l_{t+1}. $UL \equiv R$ and $UK \equiv R - (1-\delta_k)$ are the user cost of variable and fixed capital respectively. $E_t\left(MPL_{t+1}\right) \equiv \beta E_t\left(\theta_{t+1}\right) k_{t+1}^{\alpha} l_{t+1}^{\beta-1}$ is the marginal productivity of variable capital and $E_t\left(MPK_{t+1}\right) \equiv \alpha E_t\left(\theta_{t+1}\right) k_{t+1}^{\alpha-1} l_{t+1}^{\beta}$ is the marginal productivity of fixed capital. $E_t\left(\text{-}\ _{k,t+1}\right)$ and $E_t\left(\text{-}\ _{l,t+1}\right)$ are the premiums in the return of fixed and variable capital respectively. They are required by the firm to compensate for the cost of future expected financing constraints, and are defined as follows:

$$E_t\left(\text{-}\ _{z,t+1}\right) \equiv \gamma E_t\left[\left(\phi_{t+1} - 1\right)\left(UZ - MPZ_{t+1}\right)\right] \text{ for } z \in \{k, l\} \tag{16}$$

λ_t is positive when the collateral constraint (7) is binding, and is equal to zero otherwise. It represents the shadow cost of not being able to increase investment because of the

[12] See Stokey and Lucas (1989), Chapter 9.2.

lack of additional funds. $(1-\delta)\,\gamma E_t\left(\mu_{t+1}\right)$ is the cost of future expected irreversibility problems. μ_t is positive when the irreversibility constraint (6) is binding, and is equal to zero otherwise.

In this model the interactions between irreversibility and financing constraints generate a rich set of implications for the behaviour of firm investment in the business cycle. These are discussed in Caggese (2002) and Caggese (2003). In this paper we focus on the intensity of financing constraints. The method used is illustrated in the following section.

3 A new test of financing constraints on firm investment

Equation (14), shows that the marginal productivity is equal to the marginal cost of variable capital. The marginal cost can be divided into three components: UL is the user cost of purchasing one additional unit of variable capital. Without financing imperfections this would be the only relevant cost. λ_t is the shadow cost of a binding financing constraint. It is positive when equation (12) is binding with equality, meaning that all available resources are invested in l_{t+1} and k_{t+1} , but there are still some profitable investment opportunities.

E_t (- $_{l,t+1}$) is the cost of future expected financing constraints. Equation (16) shows that it is the product between the value of money in terms of its ability to reduce future expected financing problems[13], $\phi_{t+1}-1$, multiplied by the loss in monetary profit caused by the overinvestment problem, $UK - MPK_{t+1}$. E_t (- $_{t+1}$) is equal to zero if there are no expected financing constraints and $\phi_t = 1$ for $t = 0, 1, ..., \infty$. Otherwise it is positive and measures the opportunity value of reducing investment in the risky technology to increase financial earnings and to reduce future financing constraints. We now rewrite equation (14) adding the subscript i for the *i-th* firm:

$$E_t\left(\Psi_{i,t+1}\right) = (1+\eta)\,E_t\left(MPL_{i,t+1}\right) - UL_{i,t} \tag{17}$$

[13] This is because, from equation (13), ti follows that $\phi_{t+1} - 1 = RE_t\left[\sum_{j=0}^{\infty}\gamma^j\lambda_{t+1+j}\right]$. If there are no future expected financing constraints then $\lambda_{t+1} = 0$ for any t and $\phi_{t+1} = 1$. Otherwise $\phi_{t+1} - 1$ is monotonously increasing in the cost of future expected financing constraints.

Where $E_t(\Psi_{i,t+1})$ is defined as follows:

$$E_t(\Psi_{i,t+1}) \equiv E_t(- _{l,i,t+1}) + R^2\lambda_{i,t} \tag{18}$$

$E_t(\Psi_{i,t+1})$ is the premium in the expected marginal productivity of variable capital induced either by the cost of current financing constraint $\lambda_{i,t}$, or by the cost of future expected financing constraints $E_t(- _{i,t+1})$. It is possible to state the following proposition:

Proposition 2 *For any $t \geq 0$ and for a given θ_t and k_t:*

$$(i) \lim_{w_t \to w_t^{\min}} 0 < E_t(\Psi_{i,t+1}) < \infty$$

$$(ii) \lim_{w_t \to w^{MAX}} E_t(\Psi_{i,t+1}) = 0$$

$$(iii) \frac{\partial E_t(\Psi_{t+1} \mid \theta_t, k_t)}{\partial w_t} \mid_{w_t \leq w^{MAX}} < 0$$

Proof: see Caggese (2002)

Proposition 2 states that $E_t(\Psi_{i,t+1})$ is a monotonously decreasing function of financial wealth, conditional on $\theta_{i,t}$ and $k_{i,t}$. Moreover it implies that the higher is $E_t(\Psi_{i,t+1})$, the higher are current and future expected financing constraints. Therefore $E_t(\Psi_{i,t+1})$ represents an indicator of the intensity of financing constraints which is theoretically consistent and robust to the Kaplan and Zingales (1997) critique. This property depends crucially on the fact that variable capital is reversible, and hence it can be reduced proportionally to the intensity of current and future expected financing constraints. For the same reason fixed capital does not satisfy the same property, because of the presence of the irreversibility constraint. Thus we formulate the financing constraints hypothesis in the following way:

H0) if firms are subject, now or in the future, to financing constraints, then we expect a monotonously decreasing relationship between $E_t(\Psi_{i,t+1})$ and $w_{i,t}$, conditional on $\theta_{i,t}$ and $k_{i,t}$.

H1) If financing constraints are irrelevant then we expect no systematic relationship between $E_t(\Psi_{i,t+1})$ and $w_{i,t}$, conditional on $\theta_{i,t}$ and $k_{i,t}$.

An important property of this test is that it is robust to two potential misspecification problems that could affect our estimation of the empirical counterparts of $E_t(\Psi_{i,t+1})$ and

$\theta_{i,t}$: the misspecification of the stochastic process for the productivity shock and the presence of adjustment costs in variable capital. Since we allow only for a persistent source of uncertainty, we implicitly assume that any positive productivity shock which increases $\theta_{i,t}$ also increases $E_t(\theta_{t+1})$ and investment. But if a shock in $\theta_{i,t}$ is transitory, and does not affect $E_t(\theta_{t+1})$, then it should not affect the investment in variable capital l_{t+1} of unconstrained firms. Conversely if there are adjustment costs in variable capital, it is possible that positive persistent productivity shocks increase $\theta_{i,t}$ and $E_t(\theta_{t+1})$, but do not also immediately increase investment in variable capital l_{t+1}.

In both cases these shocks would increase the excess marginal productivity of variable capital, because $E_t(\Psi_{i,t+1})$ increases in θ_t, for a given level of investment. But such high value of $E_t(\Psi_{i,t+1})$ would not be related to financing constraints. However it would also be accompanied by positive cash flow and by an increase in net worth $w_{i,t}$ at time t. Therefore if this problem is very severe we would observe that the value of $E_t(\Psi_{i,t+1})$ is increasing rather than decreasing in $w_{i,t}$. This means that such misspecification problems are likely to bias our test toward rejecting the financing constraints hypothesis when it is true, rather than the opposite.

In the next section after estimating the empirical counterpart of $E_t(\Psi_{i,t+1})$, called ${}_t\widehat{\Psi}_{i,t+1}$, we first confirm its validity as a financing constraints indicator using the direct qualitative information about financing problems available for our sample of firms. Then we test the financing constraints hypothesis *H0*: we discretise the steady space of $\theta_{i,t}$ and $k_{i,t}$ and we perform, conditional on these variables, a nonparametric regression of ${}_t\widehat{\Psi}_{i,t+1}$on $w_{i,t}$. If financing constraints are irrelevant we expect no systematic negative relation, at firm level, between $w_{i,t}$ and $E_t(\Psi_{i,t+1} \mid \theta_{i,t}, k_{i,t})$ and *H0* should be rejected by the data.

Although this nonparametric approach suffers from a dimensionality problem, we believe that this strategy is feasible because we can discretise $\theta_{i,t}$ and $k_{i,t}$ in a small number of intervals. This is because $w_{i,t}$ is the only variable that directly affects $E_t(\Psi_{i,t+1})$ by determining the probability of present and future expected financing constraints. If $w_{i,t}$ is low, then $E_t(\Psi_{i,t+1})$ is expected to be positive, regardless of $\theta_{i,t}$ being low or high, unless the persistency of $\theta_{i,t}$ is very high. Since the estimated persistency of $\theta_{i,t}$ is quite low (see figure 5 in section 4), we can condition with respect to $\theta_{i,t}$ by discretising for a very small number of intervals. At the same time $k_{i,t}$ affects $E_t(\Psi_{i,t+1})$ only when $E_t\left(\mu_{i,t+1}\right) > 0$, and/or $\mu_{i,t} > 0$. In this case the ratio $\frac{k_{i,t}}{l_{i,t}}$ increases and this amplifies the sensitivity[14] of $E_t(\Psi_{i,t+1})$ with respect to $w_{i,t}$. This effect disappears when $k_{i,t}$ is low, so that $\mu_{i,t} = 0$ and $E_t\left(\mu_{i,t+1}\right)$ is small. Hence we can eliminate the distortion effect caused by $E_t\left(\mu_{i,t+1}\right) > 0$ and/or $\mu_{i,t} > 0$ by focusing on firm year observations with relatively smaller fixed capital/variable capital ratios.

4 Estimation strategy

In order to test the financing constraints hypothesis we need empirical estimates of the expected marginal productivity of variable capital $E_t(MPL_{i,t+1})$ and of the user cost of variable capital $UL_{i,t}$. The production function considered in this section is the following:

$$Y_{i,t}^T = cA_i\theta_{i,t}K_{i,t}^{\alpha}L_{i,t}^{\beta}N_{i,t}^{\gamma} \tag{19}$$

with $\alpha > 0$; $\beta > 0$; $\gamma > 0$ and $\alpha + \beta + \gamma < 1$. With respect to the theoretical section, we maintain the decreasing return to scale Cobb-Douglas production function, but we include labour $(N_{i,t})$. The inclusion of an additional factor of production does not modify the theoretical results showed in the previous section[15]. c is the constant common to the whole sample. A_i includes all assets that are fixed in the time period used for the estimation[16]. The unobservable productivity shock $\theta_{i,t}$ is the product of three components:

$$\theta_{i,t} = \varepsilon_t\chi_{s,t}\theta_{i,t}^f \tag{20}$$

ε_t is an exogenous market wide shock, $\chi_{s,t}$ is an exogenous sector specific shock, and $\theta_{i,t}^f$ is a firm specific idiosyncratic shock. The subscript s refers to the s-*th* industrial sector. Following the specification of the theoretical model, $\ln\theta_{i,t}^f$ is a first order stochastic process:

[14] See Caggese (2002), chapter 4 for details.

[15] Labour could in principle be considered an additional variable factor of production, and hence used to test the financing constraint hypothesis in conjunction with variable capital. We prefer instead to focus only on variable capital because of the characteristics of the data we use for the estimation. During the sample period Italian firms were unable to freely reduce employment, and therefore labour was closer to an irreversible than to a reversible factor of production.

[16] One way to interpret this term is to define it as $A_i \equiv E_i^{1-\alpha-\beta-\gamma}$, where E_i is the quality of the management of the $i-th$ firm.

$\ln\theta^f_{i,t+1} = \rho\ln\theta^f_{i,t} + v_{i,t}$ with $v_{i,t} \sim iid(0,\sigma^2_v)$ and $\rho \geq 0$. Hence $\ln\theta_{i,t+1}$ can be either serially correlated ($\rho > 0$) or i.i.d. ($\rho = 0$).

We know that, under some regularity conditions, α, β and γ can be estimated as the factors shares of output: $\widehat{\beta} = \frac{UL_t l_{t+1}}{y_{t+1}}$, and $\widehat{\alpha} = \frac{UK_t k_{t+1}}{y_{t+1}}$. This requires that for each factor of production expected marginal productivity equals the user cost. This is not true in our model, because the user cost of capital does not include the cost of financing and irreversibility constraints. It is in fact possible to show that in our case $\widehat{\alpha}, \widehat{\beta}$ and $\widehat{\gamma}$ would not be consistent estimators of α, β and γ. For example, in the case of $\widehat{\beta}$:

$$P\lim\left(\widehat{\beta}\right) = \beta\frac{\left[1+\eta+\gamma E_t\left(\phi_{t+1}\right)\right]UL}{\left[1+\gamma E_t\left(\phi_{t+1}\right)\right]UL + R^2\lambda_t - cov(MPL_{t+1},\phi_{t+1})} \neq \beta \tag{21}$$

Therefore we choose to directly estimate the parameters of the production function, using an instrumental variable estimation technique.

5 Estimation results

We estimate equation (19) using the following data: $p^y_t Y^T_{i,t}$ is total revenues in monetary terms. $p^k_t K_{i,t}$ is the replacement value of plant, equipment and other intangible fixed assets. $p^l_t L_{i,t}$ is the nominal value of working capital. $p^n_t N_{i,t}$ is labour cost. Detailed information about these variables is provided in appendix 1. Given that land and building are not included elsewhere in the production function, A_i also proxies for the size of these assets[17]. In order to transform the variables in real terms, we divide each variable at time t by the ratio $\frac{p^z_1}{p^z_t}$, and we redefine $p^y_1 Y^T_{i,t} = y_{i,t}$ and $p^z_1 Z_{i,t} = z_{i,t}$, with $z \in \{k,l,n\}$ .Variables y, k, l and n are therefore valued at constant 1982 prices. Figure 2 reports summary statistics of $y_{i,t}$, $k_{i,t}$, $l_{i,t}$ and $n_{i,t}$.

[17] This formulation is correct only if the stock of land and building is constant during the time period used for the estimation (11 years). Although this is true for some firms in the sample, it is obviously not always the case. Nonetheless we prefer this formulation because balance sheet data do not provide a reliable valuation of the replacement value of land and building. In fact almost all the items in the balance sheets are valued at historic costs, and due to the occasional nature of the investment in land and building we cannot use the perpetual inventory method. We hope that any variation in such assets will be absorbed by a similar variation in $K_{i,t}$.

Figure 2: Summary statistics of the variables used to estimate the production function

Variable	Mean	St. Dev.	Min.	Max
$y_{i,t}$	33.105	68.002	1.095	1162.078
$l_{i,t}$	19.582	51.121	0.093	1200.405
$n_{i,t}$	11.303	19.475	0.343	235.296
$k_{i,t}$	8.179	18.454	0.067	259.543

Values are in billions of Italian liras, 1982 prices. 1 billion liras was equal to 0.71 million US$ at the 1982 exchange rate

By taking logs, we have the following linearised version[18] of equation (19):

$$\ln y_{i,t} = c + \ln A_i + \ln \varepsilon_t + \ln \chi_{s,t} + \alpha \ln k_{i,t} + \beta \ln l_{i,t} + \gamma \ln n_{i,t} + \ln \theta^f_{i,t} \tag{22}$$

In order to allow some heterogeneity in the technology employed by firms in different sectors, equation (22) is separately estimated for seven groups of firms. Each group is composed of firms with as homogeneous as possible production activities. Figure 3 shows

Figure 3: Composition of the selected groups

Groups composition

Two Digits (ISTAT*) Sectors	Num	Two Digits (ISTAT*) Sectors	Num
Group 1: Industrial machinery	78	Group 5: Metallic products	80
Group 2: Electronic Machinery, Precision instruments	49	Group 6: Food, Sugar and Tobaccos, Paper and Printing	66
Group 3: Textiles, Shoes and Clothes, Wood furniture	117	Group 7: Non-metallic minerals, Other manufacturing	108
Group 4: Chemicals, Rubber and Plastics	63		

* Italian National Statistic Institute

their composition. Equation (22) is estimated using the Generalised Method of Moments (GMM)[19]. A discussion of the specification tests adopted to select the estimation method are present in Appendix 2. Figure 4 illustrates the estimation results. The first column is relative to the whole sample, while the next seven columns show estimates of $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\gamma}$ for the seven groups separately. The Wald test shows that the restriction $\widehat{\alpha} + \widehat{\beta} + \widehat{\gamma} = 1$ is rejected in favour of $\widehat{\alpha} + \widehat{\beta} + \widehat{\gamma} < 1$ for all groups except group 7. Since this assumption is necessary to derive the financing constraint hypothesis to be tested, we exclude the observations in group 7 from the empirical estimation of ${}_t\widehat{\Psi}^w_{i,t+1}$.

The estimated output elasticity of variable capital $\widehat{\beta}$ ranges between 0.29 and 0.56, and in three groups is higher than the output elasticity of labour $\widehat{\gamma}$. These high estimates

[18] Since we assume prices to be deterministic, any shock in relative prices $\frac{p^z_t}{p^y_t}$ is going to be absorbed by the dummies ε_t and $\chi_{s,t}$ and by the firm specific shock $\theta^f_{i,t}$.

[19] Such method is used for a similar problem by Hall and Mairesse (1996).

Figure 4: Production function estimation results

Estimated coefficients of the production function

	All Firms	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
? ??	0.111	0.105	0.062	0.114	0.081	0.038	0.040	0.198
	(0.02)	(0.02)	(0.015)	(0.03)	(0.02)	(0.022)	(0.01)	(0.02)
? ??	0.389	0.377	0.289	0.424	0.454	0.393	0.562	0.406
	(0.02)	(0.01)	(0.013)	(0.03)	(0.01)	(0.017)	(0.01)	(0.024)
???	0.441	0.494	0.468	0.348	0.193	0.491	0.350	0.401
	(0.03)	(0.02)	(0.023)	(0.04)	(0.01)	(0.034)	(0.01)	(0.05)
Test of overid. restr.	65.50	38.90	25.78	39.87	39.71	38.20	45.18	33.64
Degrees of freedom	37	37	27[3]	37	37	36[2]	36[2]	36[2]
p-value	0.00	0.38	0.53	0.34	0.35	0.37	0.14	0.58
Chi square[4]		29.7	41.7	814.6	217.2	9.61	11.35	0.01
p-value		0.00	0.00	0.00	0.00	0.00	0.00	0.91
n. firms	561	78	49	117	63	80	66	108
n. of observations	4488	624	392	936	504	640	528	864

1) Standard deviation in parenthesis; 2) One coefficient relative to a two-digit sector dummy variable is estimated here;
3) Only 6 years of lagged instruments used for the estimation of this group, due to the reduced number of observations;
4) Wald test of the following restriction: ???????

of β are quite common in firm-level estimates of the production function (see for example Hall and Mairesse, 1996). Output elasticity of fixed capital $\widehat{\alpha}$ ranges between[20] 0.04 and 0.11. The overidentifying restrictions are rejected for the estimation of the whole sample, but not for single groups estimations.

Before proceeding to estimate $E_t\left(MPL_{i,t+1}\right)$, the expected marginal productivity of variable capital, it is important to mention a problem which could have affected the estimates obtained above. Blundell and Bond (1998) show that weak instruments can cause large finite sample biases when performing GMM estimation on data transformed using first differencing. The same problem can affect the GMM estimates on the data transformed using the forward orthogonal transformation method, which is equivalent to first differencing when all moment conditions are used (Arellano and Bover, 1995). Indeed Blundell and Bond (2000) show that these biases are the likely reason why the fixed capital coefficient in the production function, estimated using GMM on the first differenced equations, is not significant (and even negative in some specifications) both in Hall and Mairesse (1996) and in their paper. Blundell and Bond (2000) propose a more efficient "System GMM" estimation method that includes lagged first difference of the

[20] This range of values is reasonable and consistent with the factor shares of output, given the amount of fixed capital as opposed to variable capital used in the production (see figure 2), and the difference in the user costs of fixed and variable capital generated by the difference in depreciation factors (see appendix 1).

series as instruments for the level equations. Even though this problem could have affected the precision of our estimates, it is important to note that our estimated coefficients of the production function do not exhibit a downward bias the same size as the one of the estimates of Hall and Mairesse (1996). In fact we estimate a positive and significant fixed capital coefficient, with a magnitude consistent with the factor shares, as argued above[21].

In order to obtain an empirical estimate of $E_t\left(MPL_{i,t+1}\right)$, we first use the estimates $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\gamma}$ to compute the total factor productivity for all the years from '82 to '92: $\widehat{TFT}_{i,t} = \ln y - \widehat{\alpha}\ln k_{i,t} + \widehat{\beta}\ln l_{i,t} - \widehat{\gamma}\ln n_{i,t}$. Where $\widehat{TFT}_{i,t} \equiv \ln c + \ln A_i + \ln \varepsilon_t + \ln \chi_{s,t} + \ln \theta^f_{i,t}$. Then we perform a panel data regression with fixed effects, year and sector dummy variables, to estimate $\ln \widehat{c}$, $\ln \widehat{A}_i$, $\ln \widehat{\chi}_{s,t}$ and $\ln \widehat{\varepsilon}_t$. The estimated residual from this regression is $\widehat{\theta}^f_{i,t}$, and we use it to estimate the autocorrelation coefficient $\widehat{\widehat{\rho}}$[22] separately for the seven groups of firms, as shown in figure 5.The $\widehat{\widehat{\rho}}$ estimates are positive and significant, but relatively

Figure 5: Autocorrelation coefficient estimation results

Estimated first order autocorrelation coefficient of the firm specific productivity shock

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
?^^	0.338	0.357	0.342	0.339	0.338	0.332	0.344
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Test of overid. restr.	1.599	1.67	1.60	1.59	1.59	1.61	1.60
Degrees of freedom	3	3	3	3	3	3	3
p-value	0.66	0.64	0.66	0.65	0.65	0.65	0.66
n. of observations	390	305	585	315	400	330	480
? ~	0.22	0.36	0.37	0.30	0.25	0.35	0.32

*Alternative estimator of ? based on the estimated transition probabilities of ?i,t discretised in a two states transition matrix.

low, ranging from 0.33 to 0.36. These values are broadly consistent with an alternative estimator $\widetilde{\rho}$ simply based on the transition probabilities of $\widehat{\theta}^f_{i,t}$ discretised in a two states transition matrix. Using (19) we define the expected marginal productivity of variable capital as follows:

$$E_t\left(MPL_{i,t+1}\right) = E_t\left(\frac{\partial Y_{i,t+1}}{\partial L_{i,t+1}}\right) = \beta E_t\left(cA_i\theta^f_{i,t+1}\varepsilon_{t+1}\chi_{s,t+1}K^{\alpha}_{i,t+1}L^{\beta-1}_{i,t+1}N^{\gamma}_{i,t+1}\right) \qquad (23)$$

As we mentioned before, the term A_i absorbs between-firm differences in $\theta^f_{i,t}$. Given that $Y_{i,t}$ is decreasing return to scale in $K_{i,t}$, $L_{i,t}$ and $N_{i,t}$, A_i mainly represents permanent

[21] See footnote n.20

[22] We compute $\widehat{TFT}'_{i,t} = \widehat{TFT}_{i,t} - \ln\widehat{\chi}_{s,t} - \widehat{\varepsilon}_t$, we apply to it the same forward orthogonal transformation described before to eliminate A_i, and we regress the transformed $TFT^{*'}_{i,t}$ on $TFT^{*'}_{i,t-1}$, using $t-2$ to $t-5$ lags as instruments.

dimensional differences between firms. Hence, in order to compare marginal productivity of variable capital across firms, we eliminate this effect and consider instead the following variable:

$$E_t \left(MPL_{i,t+1}\right)^W = \beta E_t \left(\theta^f_{i,t+1}\right) E_t \left(c\varepsilon_{t+1}\chi_{s,t+1} \frac{k^\alpha_{i,t+1} n^\gamma_{i,t+1}}{l^{1-\beta}_{i,t+1}} \right) \tag{24}$$

Equation 24 also assumes[23] that $cov\left(\theta^f_{i,t+1}, \frac{k^\alpha_{i,t+1} n^\gamma_{i,t+1}}{l^{1-\beta}_{i,t+1}}\right) = 0$. Hence our estimator of $E_t \left(MPL_{t+1}\right)^W$, called ${}_t\widehat{MPL}^W_{t+1}$, is the following:

$$\widehat{MPL}^w_{i,t+1} = \widehat{c}\widehat{\beta} \left({}_t\widehat{\theta}_{i,t+1} \frac{k^{\widehat{\alpha}}_{i,t+1} n^{\widehat{\gamma}}_{i,t+1}}{l^{1-\widehat{\beta}}_{i,t+1}} \right) \tag{25}$$

where ${}_t\widehat{\theta}_{i,t+1} \equiv e^{\widehat{\widehat{\rho}} \ln \widehat{\theta}^f_{i,t}} \widehat{\varepsilon}_t \widehat{\chi}_{s,t}$

In order to derive (25) from (24) we implicitly assume that investment is planned one period in advance. Therefore $k_{i,t+1}$, $l_{i,t+1}$ and $n_{i,t+1}$ are predetermined at time t. Regarding the shock, $e^{\widehat{\widehat{\rho}} \ln \widehat{\theta}^f_{i,t}}$ is the estimate[24] of $E_t\left(\theta^f_{i,t+1}\right)$, while $E_t(\varepsilon_{t+1})$ and $E_t(\chi_{s,t+1})$ are simply approximated by the estimated fixed effects $\widehat{\varepsilon}_t$ and $\widehat{\chi}_{s,t}$. The user cost of capital $UL_{i,t}$, in monetary units relative to 1982 prices, is the following:

$$UL_{i,t} = \frac{p^l_t}{p^l_1}\left(1 + r_t\right) - \frac{p^l_{t+1}}{p^l_1}\left(1 - \delta_l\right) \tag{26}$$

we apply to it the same normalization on prices applied to $l_{i,t}$, so that we define $ul_{i,t} = \frac{p^l_1}{p^l_t} UL_{i,t}$. Furthermore $\delta_l = 1$ by construction, because both in the model and in the empirical application we include in $l_{i,t+1}$ only variable capital consumed during time $t+1$ in order to produce time $t+1$ output (see appendix 1 for details). Hence the user cost simplifies to:

$$ul_{i,t} = 1 + r_t \tag{27}$$

where r_t is the real interest rate at time t, measured as the nominal riskless short term interest rate (average nominal interest rate, during period t of the three months treasury bills) minus inflation rate (change in the consumer price index between the fourth quarter

[23] This assumption is not likely to cause a relevant bias in the estimates of $E_t\left(MPL_{i,t+1}\right)^W$, because the estimated correlation between the empirical counterparts $\widehat{\widehat{\rho}}\widehat{\theta}^f_{i,t}$ and $\frac{k^{\widehat{\alpha}}_{i,t+1} n^{\widehat{\gamma}}_{i,t+1}}{l^{1-\widehat{\beta}}_{i,t+1}}$ is quite low. It ranges from a minimum of 0.05%, for firms in group 6, to a maximum of 4.83%, for firms in group 3.

[24] Since $\widehat{\theta}^f_{i,t}$ is relatively close to 0, then $e^{\widehat{\widehat{\rho}} \ln \widehat{\theta}^f_{i,t}}$ is a good approximation for $E_t\left(\theta^f_{i,t+1}\right)$

of period $t-1$ and the fourth quarter of period t). In order to use (27) as the equation that defines the user cost of variable capital, we adopt a series of simplifying assumptions[25]. Moreover we maintain the assumption made in the theoretical section that firms only borrow riskless debt. This implies that the user cost of capital is independent on the risk faced by the firms' projects. In reality some firms also borrow risky debt, and their user cost of capital should include a component above the riskless interest rate representing the price of risk. We believe that, even though the inclusion of such risk premium may be relevant, its omission would not substantially affect our findings. This is because the magnitude of the cost of financing constraints predicted by the model, and confirmed by the estimations in the next section, is much larger than the likely size of a risk premium component. Therefore the bias in the estimation of $E_t\left(\Psi_{i,t+1}\right)$ caused by the omission of the risk premium component is likely to be small.

Given (27), the estimated financing constraints indicator $E_t\left(\Psi_{i,t+1}\right)^W$ is the following[26]:

$$
{}_t\widehat{\Psi}^w_{i,t+1} = \left({}_t\widehat{MPL}^w_{i,t+1} - ul_{i,t+1}\right) \tag{28}
$$

Comparing equation (28) to (17) it follows that:

$$
E_t\left({}_t\widehat{\Psi}^w_{i,t+1}\right) = \frac{E_t\left(\Psi_{i,t+1}\right)^W}{1+\eta} < E_t\left(\Psi_{i,t+1}\right)^W \tag{29}
$$

Hence ${}_t\widehat{\Psi}^w_{i,t+1}$ is biased downwards by the unobservable parameter η. In the theoretical model we assume such parameter to be constant across firms. We interpret it as private (non financial) benefits accruing to the owners of the firm. Even if the value of η is

[25]This formulation is considerably simplified by the fact that we do not formally treat taxes. If we allow for taxation differentials, then $UL_{i,t}$ would be multiplied by one minus a term that represents the expected tax benefit of one additional unit of investment at time t. Such tax benefit is mainly given by the "debt tax shield", because tax credits are usually associated with fixed capital investment. An explicit treatment of this issue is beyond the scope of this paper, also because we do not have accurate information on the incidence of tax exhaustion in order to measure the effective tax parameters facing individual firms. Even though we agree that tax differentials play a relevant role in determining Italian firms' capital structure, we follow Bond and Meghir (1994) in assuming that fluctuations in the user cost of capital due to tax distortions are mainly absorbed by firm and year specific effects, captured by A_i and ε_t. All of the results presented in the following part of the paper are based on deviations from firm averages that are independent on A_i, while the exclusion of ε_t does not affect the results in any relevant way.

[26]The estimator of $E_t\left(\Psi_{i,t+1}\right)^W$ does not include fixed effects A_i as well. Since by construction $\sum_{i=1}^{N}\ln\widehat{A}_i = 0$, and since the exponential is a convex function, it follows that $\sum_{i=1}^{N}\widehat{A}_i > \exp\left(\sum_{i=1}^{N}\ln\widehat{A}_i\right) = 1$, and ${}_t\widehat{\Psi}^w_{i,t+1}$ is expected to slightly underestimate $E_t\left(\Psi_{i,t+1}\right)$. This bias is expected to be small, and in any case it is constant at firm level.

Figure 6: Summary statistics for the estimated excess expected marginal productivity of variable capital (${}_t\widehat{\Psi}^w_{i,t+1}$)

Mean	St. dev.	Min.	Max.
0.026	0.727	-1.05	3.013

non constant but rather a function of unobservable characteristics of the firm's owners, and such characteristics are randomly distributed across the sample, then this distortion should not affect the results[27]. Figures 6 and 7 show basic statistics[28] and the kernel estimation of the distribution function of ${}_t\widehat{\Psi}^w_{i,t+1}$. Figure 7 shows that ${}_t\widehat{\Psi}^w_{i,t+1}$has an asymmetric distribution with a thicker tail corresponding to higher than average values. The financing constraint hypothesis implies that these are firm-year observations where, because of financing constraints, the firm could not increase variable capital to exploit profitable investment opportunities. We verify this hypothesis in the next two sections.

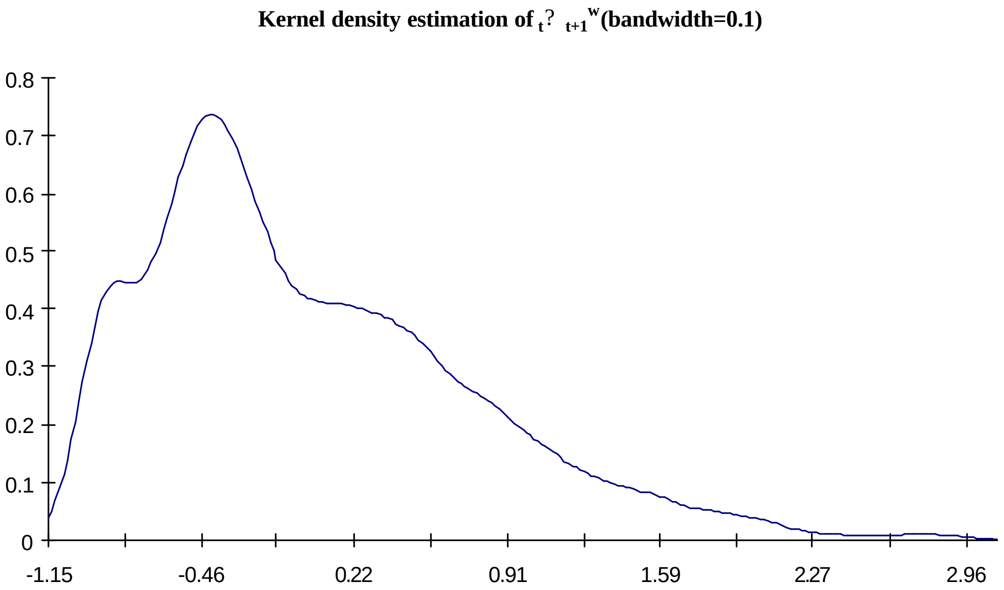
5.1 Empirical evidence of financing constraints on investment

Before testing the financing constraints hypothesis, we verify the validity of ${}_t\widehat{\Psi}^w_{i,t+1}$ as an estimate of the intensity of financing constraints. We use the direct information about financing problems available in the Mediocredito Centrale survey. In the survey firms were asked whether they had any of the following problems regarding the financing of new investment projects in the 1989-91 period: Q1) lack of collateral; Q2) lack of medium-long term financing; Q3) too high cost of banking debt. Such problems are directly related to the value of the variable ${}_t\widehat{\Psi}^w_{i,t+1}$. The bigger is ${}_t\widehat{\Psi}^w_{i,t+1}$, the higher is the shadow value of additional funding for the *i-th* firm and the higher the probability that it answers positively to one of the questions regarding financing constraints. Among the 561 firms considered,

[27]Another possibility is to assume that η is a decreasing function of the size of the firm. For example in a recent paper Hamilton (2000) shows that high levels of private (non financial) benefits for self employed entrepreneurs are needed to explain the estimated earning differential between self employment and paid employment, and self employed entrepreneur are more likely to maange smaller businesses. If this is the case, then we expect this bias not to significantly distort our estimation results, because we control for firm's size in all the following analysis,

[28]This is filtered from outliers. We first exclude observations that deviate from the mean by more than 8 times the standard deviation, then recompute the mean and exclude all observations that deviate more than 4 times from the standard deviation. Out of the initial 4821 observations, we eliminate 51 observations for ${}_t\widehat{\Psi}^w_{i,t+1}$.

Figure 7: Density estimation of the excess expected marginal productivity of variable capital



21.2% of them indicate one of the three problems in accessing bank credit during the 1989-1991 period. We construct 4 dichotomous variables, $ration_i^j$ with $j = \{1, 2, 3, 4\}$, that have value 0 if the $i-th$ firm does not state any financing problem, 1 if it answers positively to questions Q1, Q2 and Q3 (respectively $j = 1, 2$ and 3) or states any of the three problems ($j = 4$). We verify the reliability of ${}_t\widehat{\Psi}_{i,t+1}^w$ as a valid indicator of the intensity of financing constraints by estimating the following equation:

$$\mathbf{ration}_i^j = \alpha_0 + \alpha_1 \overline{\Psi}_i^w + \alpha_2 \dim_i \tag{30}$$

$\overline{\Psi}_i^w$ is the average value of ${}_t\widehat{\Psi}_{i,t+1}^w$ in the period covered by the Mediocredito Centrale survey:$\overline{\Psi}_i^w = \sum_{t=1989}^{1992} \left({}_t\widehat{\Psi}_{i,t+1}^w\right)$. The time interval used to compute $\overline{\Psi}_i^w$ includes 1989, 1990 and 1991, the period which the questions refer to, and 1992, the year in which the questionnaire has been compiled. $\dim_i$ is the size of the $i-th$ firm in number of employees, and is included to control for size effects. Figure 8 shows estimation results. The first column is relative to the whole sample and to $ration_i^4$ as dependent variable. The coefficient relative to $\overline{\Psi}_i^w$, α_1, is positive and significant. The second and third columns repeat the same regression for larger (more than 300 employees, 19% of the sample) and smaller (less than 300 employees, 81% of the sample) firms. The cutting point between small and

Figure 8: Relation between stated financing problems and the financing constraints indicator

Probit regression: $ration^j_i = a_0 + a_1 ?\ ^w_i + a_2 dim_i$

	Whole sample	Larger firms[1]	Smaller firms[2]			
Dependent variable	*Ration*[4] (all problems)	*Ration*[4] (all problems)	*Ration*[4] (all problems)	*Ration*[1] (Low collateral)	*Ration*[2] (Lack of bank credit)	*Ration*[3] (High cost of debt)
? 0	-0.64*** (0.09)	-0.69 (0.42)	-0.78*** (0.16)	-2.05*** (0.31)	-1.48*** (0.20)	-1.04*** (0.17)
? 1	0.24** (0.11)	-0.16 (0.36)	0.30*** (0.12)	0.36* (0.20)	0.28** (0.14)	0.29** (0.13)
? 2	-0.0006* (0.0003)	-0.0007 (0.0007)	0.0005 (0.001)	0.001 (0.002)	0.002* (0.001)	0.0005 (0.001)
Obs with ration=0	341	70	271	341	310	296
Obs with ration=1	92	11	81	11	42	56
(% of total)	(21.2%)	(13.6%)	(23%)	(3.1%)	(11.9)	(15.9%)
Total obs	433	81	352	352	352	352

Standard error in parenthesis; 1: More than 300 employees; 2: Less than 300 employees; * Significant at 90% confidence level; ** significant at 95% confidence level; *** significant at 99% confidence level; $ration_i$ = 1 if the entrepreneur stated financing constraints, and 0 otherwise; ?? $^w{}_i$ = average value of the premium in the expected productivity of variable capital; dim_i = dimension in number of employees;

large firms is suggested by figure 9, which shows the tri-dimensional smoothing of $ration^4$ with respect to both $\overline{\Psi}_i^w$ and $\dim_i$. Figure 9 shows that the positive correlation between the probability of stating financing problems and $\overline{\Psi}_i^w$ is strong for all the firms except the larger ones, so that on average "small-medium" firms with a high value of $\overline{\Psi}_i^w$ are three times more likely to declare financing constraints than firms with a low value of $\overline{\Psi}_i^w$. Such relation tends to disappear for firms bigger than 250-300 employees.

In order to interpret this result, we note that in our estimation the assumption that $\theta_{i,t}^f$ is stationary, plus the condition that $\alpha+\beta+\gamma<1$, imply that we assume different steady states sizes for different firms, according to their fixed effects A_i. Each firm evolves around its steady state depending on the realisations[29] of the idiosyncratic shock $\theta_{i,t}^f$. Therefore the result illustrated in figures 8 and 9 is consistent with the assumption that the higher the average size of a firm, the less likely it is to face the informational or contractual problem which cause the financing constraint (7) to be binding. This finding is not surprising, because large Italian firms usually have strong links with financial intermediaries, and the assumption that they have access only to fully collateralised credit is not realistic for them.

The strong correlation between $ration_i^j$ and $\overline{\Psi}_i^w$ for small and medium firms below 300

[29] This stationarity assumption is reasonable in this context, given that the time series is 11 years only.

Figure 9: Probability of stating financing problems as a function of size and financing constraints indicator

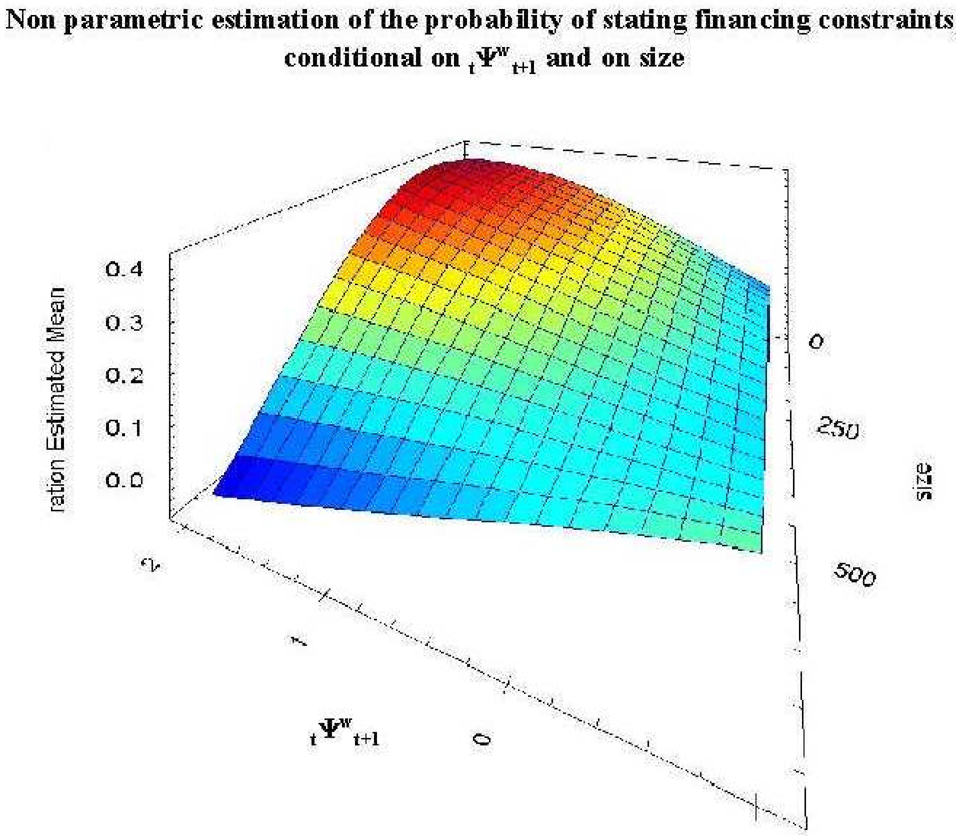


Figure 10: Sectorial distribution of financing constraints

Industrial sectors and financing problems

Sector	N firms	% with financing problems	% without financing problems
Whole sample	897	23	77
Mechanic materials and machineries	139	21.6	78.4
Metallic products	130	26.9	73.1
Textiles	87	21.8	78.2
Shoes and clothes	65	10.8	89.2
Electric and electronic materials	60	28.3	71.7
Paper, printing and publishing	57	33.3	66.7
Non metallic minerals	56	21.4	78.6
Other manufacturing	53	28.3	71.7
Wood and wooden furniture	50	22	78
Rubber and plastic	50	22	78

employees is confirmed in the last four columns of figure 8: the $\overline{\Psi}_i^w$ coefficient is positive and strongly significant, especially for the specification ($j = 4$) that pools together the three different questions. This finding demonstrates that ${}_t\widehat{\Psi}_{i,t+1}^w$ is a valid indicator of the intensity of financing constraints, and supports the validity of our theoretical model and our empirical approach. Moreover this finding is robust because of at least three reasons: first, the qualitative and quantitative information come from different sources. This reduces the probability that those firms that declared financing constraints also manipulated their balance sheets data to show that their investment was inefficiently low; second, we condition for firms size, thus ruling out the possibility that $\overline{\Psi}_i^w$ is on average higher for small firms, which are also more likely to state financing constraints; third, the result is not driven by sectorial differences: table 10 shows that financing constraints are equally distributed in the different industrial sectors[30].Given that ${}_t\widehat{\Psi}_{i,t+1}^w$ is a noisy measure of the intensity of financing constraints, because of the estimation problems mentioned in the previous subsection, this consistency result with our qualitative information is very important.

[30]The sample used to calculate this table is composed by 897 firms, while only 561 firms have been used for the estimations, because they provide a richer set of balance sheet information.

6 A formal test of the financing constraints hypothesis

After verifying that ${}_t\widehat{\Psi}^w_{i,t+1}$ is positively correlated to directly revealed financing problems, we test the financing constraints hypothesis by estimating the relation between ${}_t\widehat{\Psi}^w_{i,t+1}$ and net financial wealth $w^w_{i,t}$, conditional on $E_t\left(\theta^f_{i,t+1}\right)$ and $k_{i,t}$:

$$ {}_t\widehat{\Psi}^w_{i,t+1} = g\left(w^w_{i,t} \mid \widehat{\theta}^f_{i,t+1}, k^w_{i,t}\right) \tag{31} $$

$g(.)$ is a nonlinear function. If the financing constraints hypothesis is not rejected by the data, we expect to find a negative slope of the conditional mean of ${}_t\widehat{\Psi}^w_{i,t+1}$ with respect to $w^w_{i,t}$. Such a slope should be convex: steeper when $w^w_{i,t}$ is very low, then gradually flatter as $w^w_{i,t}$ increases. Figure 11 illustrates the relationship between $E_t\left(\Psi^w_{i,t+1}\right)$ and $w^w_{i,t}$ for a given $E_t\left(\theta_{i,t+1}\right)$ and $k_{i,t}$, predicted by the model[31]. The highest values of $E_t\left(\Psi^w_{i,t+1}\right)$, in correspondence with the lowest values of $w^w_{i,t}$, are due to a very high cost of a binding financing constraint ($\lambda_{i,t}$). In this region the slope is very steep because the strict concavity of the production function implies a very high marginal productivity of variable capital. As wealth increases, $E_t\left(\Psi^w_{i,t+1}\right)$ decreases, at a gradually slower pace.

The function $g\left(w^w_{i,t} \mid \widehat{\theta}^f_{i,t+1}, k^w_{i,t}\right)$ in equation (31) is estimated using a nonparametric estimation method[32]. The variables used are the following:

Financial wealth: we consider two alternative variables: i) $w^1_{i,t}$ = net financial wealth at the beginning of year t (liquidity plus short term financial assets[33] minus the loans that have to be repaid before the end of time t), plus the new cash flow generated during time t.ii) $w^2_{i,t}$ = net financial wealth at the beginning of year t.

$w^1_{i,t}$ would be the best estimator of net financial wealth available for investment at time t, if time t investment would be productive only from time $t+1$ on. In reality this is

[31] We find an explicit numerical solution of the theretical model for a set of parameters. Then we generate simulated data for both $E_t\left(\Psi^w_{i,t+1}\right)$ and $w^w_{i,t}$, applying to the simulated data the same within-transformation described above. The parameters are calibrated in order to match the average level of output and capital in the sample of italian firms. The variance of θ_t is directly estimated from $\widehat{\theta}^f_t$. For details about the calibation of the other parameters see Caggese 2002.

[32] The estimation of (31) is performed using a local polinomial regression method that fits a locally weighted least squares regression using raw data near each target observed point. We used the software package Glassbox.

[33] The implicit assumption we make, in order to focus only on financial variables, is that the entrepreneurs use the collateral value of the firms' assets at the end of time $t-1$ to borrow secured debt up to the limit, and then maintain the additional resources in the form of financial assets.

not always the case, as time t financial wealth is increased by revenues partly from past investment and partly from time t investment in variable capital. The more the revenues from time t investment increase time t financial wealth, the more our nonlinear estimation of (31) is biased. Therefore $w^2_{i,t}$ is a less precise but more robust estimator, because it excludes by construction the time t cash flow. We eliminate the size effect from both $w^1_{i,t}$ and $w^2_{i,t}$ by scaling these variables by the average size of firm i during the sample period. Finally, since the estimators of $E_t\left(\theta^f_{i,t+1}\right)$ and $E_t\left(MPL_{t+1}\right)$ do not include the firm specific productivity A_i, we apply the same procedure to $w^1_{i,t}$ and $w^2_{i,t}$, and consider the values $w^{1w}_{i,t} = w^1_{i,t} - \sum_{t=1}^{T} w^1_{i,t}$ and $w^{2w}_{i,t} = w^2_{i,t} - \sum_{t=1}^{T} w^2_{i,t}$ that are the deviations form firm specific means.

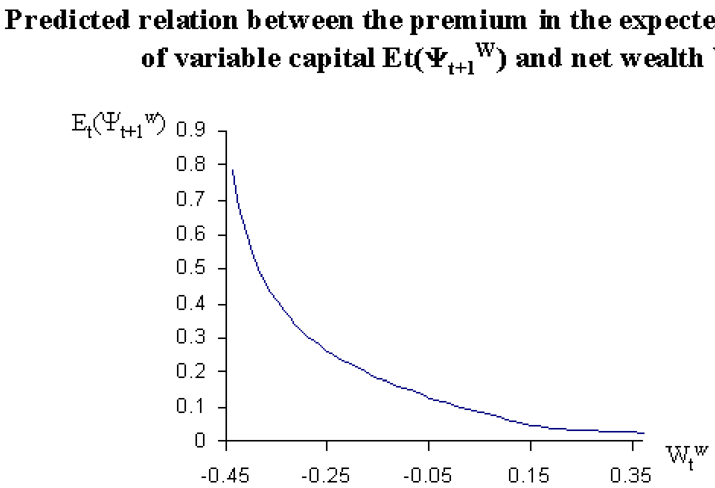
Expected productivity: we consider the empirical counterpart of $E_t\left(\theta^f_{i,t+1}\right)$, that is ${}_t\widehat{\theta}^f_{i,t+1}$.

Fixed capital: from $k_{i,t}$ we compute $k^w_{i,t}$, following the same within transformation applied to $w^w_{i,t}$.

In order to estimate equation (31) we condition by $E_t\left(\theta^f_{i,t+1}\right)$ by discretising the state space of its estimator ${}_t\widehat{\theta}^f_{i,t+1}$ in 3 equally spaced intervals and by estimating equation (31) for each interval. Moreover we condition by $k_{i,t}$ by excluding from our analysis the observations in the fourth quartile of $k^w_{i,t}$. Figures 12, 13 and 14 show the estimation of equation (31) for small firms (less than 300 employees) with respectively low, medium and high productivity shock, using both $w^{1w}_{i,t}$ and $w^{2w}_{i,t}$. The shaded lines represent the boundaries of the 90% confidence interval. The downward sloping relationship predicted by the financing constraints hypothesis is confirmed for the low and medium productivity shock observations[34] for both $w^{1w}_{i,t}$ and $w^{2w}_{i,t}$ (Figures 12 and 13). For $w^{1w}_{i,t}$ the model predictions are not confirmed for high values of financial wealth, for which we observe an upward sloping relationship instead of a flat one. This could be due to one of the two misspecification problems mentioned in section 2. Suppose a firm receives a positive shock between time $t-1$ and time t, which increases revenues and financial wealth at

[34] The minimum of ${}_t\widehat{\Psi}^w_{i,t+1}$ is lower than zero, and this could be due to one of the reasons mentioned before: i)the effect of the unobservable η; ii) the bias induced by the elimination of A$_i$ (see footnote n.26); iii) the overestimation of the user cost of capital, given that we do not explicitly measure tax differentials (see footnote n.??).

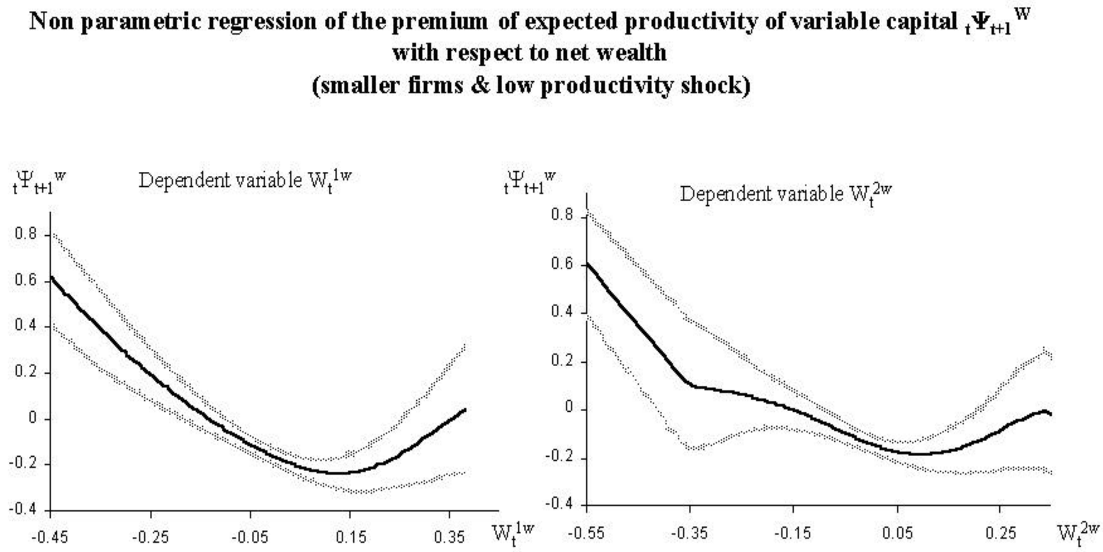
Figure 11: Expected relation between the intensity of financing constraints and financial wealth



the beginning of time t. It could be that this positive shock does not increase investment during time t either because the shock is entirely transitory, or because of the presence of convex adjustment costs. For this firm we would observe an high ${}_t\widehat{\Psi}^w_{i,t+1}$, because after a positive shock capital has not increased, and we would observe also high financial wealth. This positive correlation between the two variables would lead us to reject the financing constraints hypothesis. This bias should be stronger when using $w^{1w}_{i,t}$ than when using $w^{2w}_{i,t}$, because the latter does not include time t cash flow. This can explain why a positive relationship between ${}_t\widehat{\Psi}^w_{i,t+1}$ and $w^w_{i,t}$ for high net wealth values, is absent when using $w^{2w}_{i,t}$. It is important to note that if this problem is severe then we increase the chances to reject rather than to accept the financing constraints hypothesis. This property increases the reliability of this test, and hence the downward sloping relationship between ${}_t\widehat{\Psi}^w_{i,t+1}$ and $w^w_{i,t}$ estimated in figures 12 and 13 is a strong evidence in favour of the financing constraints hypothesis.

In the previous section we argued that the omission of the risk premium in the calculation of the user cost of capital could affect these findings. This omission has the effect of overestimating ${}_t\widehat{\Psi}^w_{i,t+1}$ for observations with higher risk premium in the cost of capital. Therefore an alternative explanation of the downward sloping relationship between ${}_t\widehat{\Psi}^w_{i,t+1}$ and $w^w_{i,t}$, which does not have to do with financing constraints, is that high ${}_t\widehat{\Psi}^w_{i,t+1}$ observations could be relative to firms with riskier projects, which are on average smaller and hence less wealthy. A problem with this explanation is that $w^w_{i,t}$ is not an absolute, but a relative measure of financial wealth. But even ignoring this point, we think that this alternative explanation cannot explain our results, because of the magnitude of our estimated ${}_t\widehat{\Psi}^w_{i,t+1}$. In fact figure (11) shows that the model predicts, for firms with low net financial wealth, that a binding financing constraint may cause a premium in the expected marginal productivity of variable capital up to 80%. Such prediction is important because it is based on technological parameters calibrated from the same panel data of Italian firms used for the empirical estimation. Figures (12) and (13) not only confirm the downward sloping relationship between ${}_t\widehat{\Psi}^w_{i,t+1}$ and $w^w_{i,t}$, but also they show a similar magnitude of ${}_t\widehat{\Psi}^w_{i,t+1}$, around 60%-80%, for firms with low financial wealth. Even if the omitted risk factor contributes to this result, it probably does not fully explain it, given its size.

Figure 12: Estimated relation between the intensity of financing constraints and financial wealth - small firms - low productivity



Observations with high productivity shock in figure 14 instead exhibit a slightly upward sloping relationship between $w^w_{i,t}$ and ${}_t\widehat{\Psi}^w_{i,t+1}$. Also in this case such relationship is

Figure 13: Estimated relation between the intensity of financing constraints and financial wealth - small firms - medium productivity

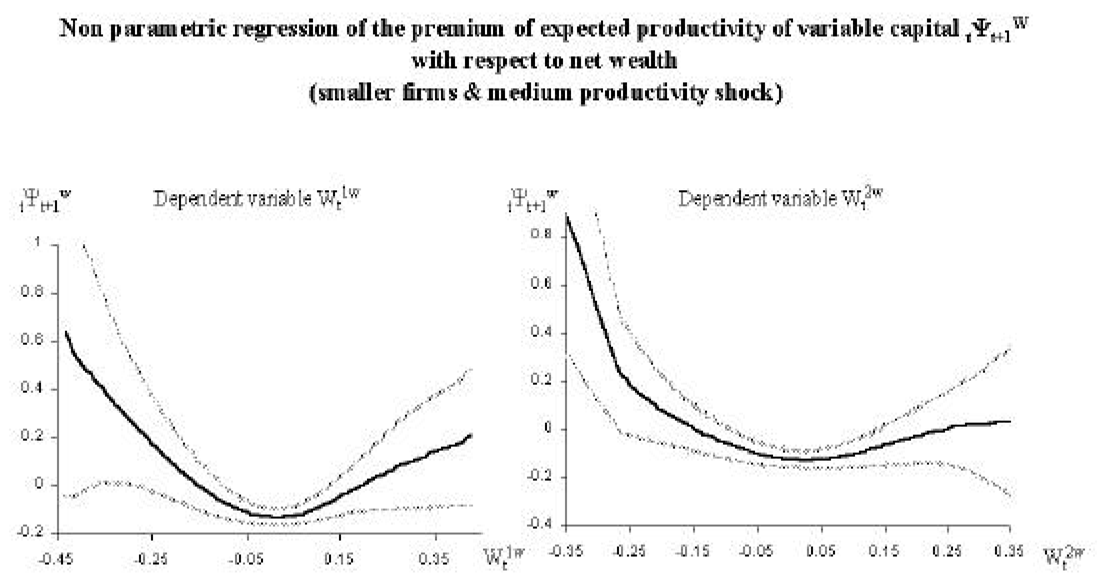
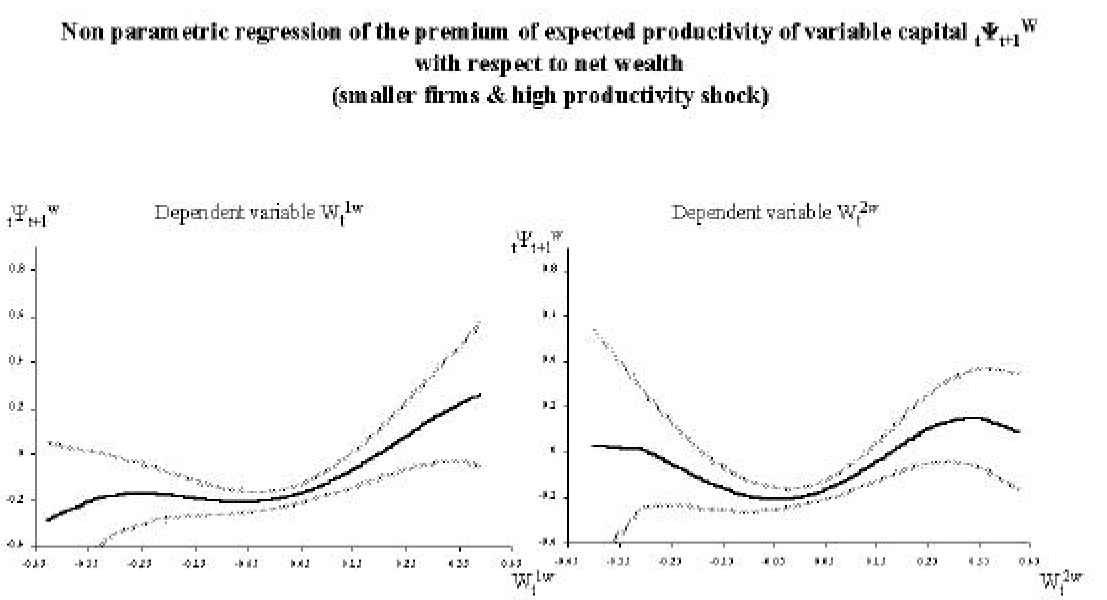


Figure 14: Estimated relation between the intensity of financing constraints and financial wealth - small firms - high productivity



more pronounced for $w_{i,t}^{1w}$. This is probably caused by the fact that, for these firm-year observations with higher productivity shocks, the two misspecification problems mentioned in section 2 are more severe, and as a consequence we reject the financing constraints hypothesis. This interpretation is supported by the qualitative information in figure 15: in the subset of observations of small firms with high productivity shock the probability of stating financing constraints is not correlated with the value of $\overline{\Psi}_i^w$.

Figure 15: Relation between stated financing problems and the financing constraints indicator - small firms

Probit regression and productivity shock levels: $ration^j_i = a_0 + a_1 ?\ ^w_i + a_2 dim_i$

	Smaller firms[2] Dependent variable: Ration[4] (all problems)		
	Low productivity shock	Medium productivity shock	High productivity shock
? 0	-0.77***	-0.58***	-1.09***
	(0.27)	(0.24)	(0.32)
? 1	0.28	0.43**	0.09
	(0.19)	(0.21)	(0.27)
? 2	0.0006	0.0001	0.0009
	(0.0018)	(0.0017)	(0.0022)
Obs with ration=0	96	97	78
Obs with ration=1	30	36	15
(% of total)	(23.8%)	(27.1%)	(16.1%)
Total obs	126	133	93

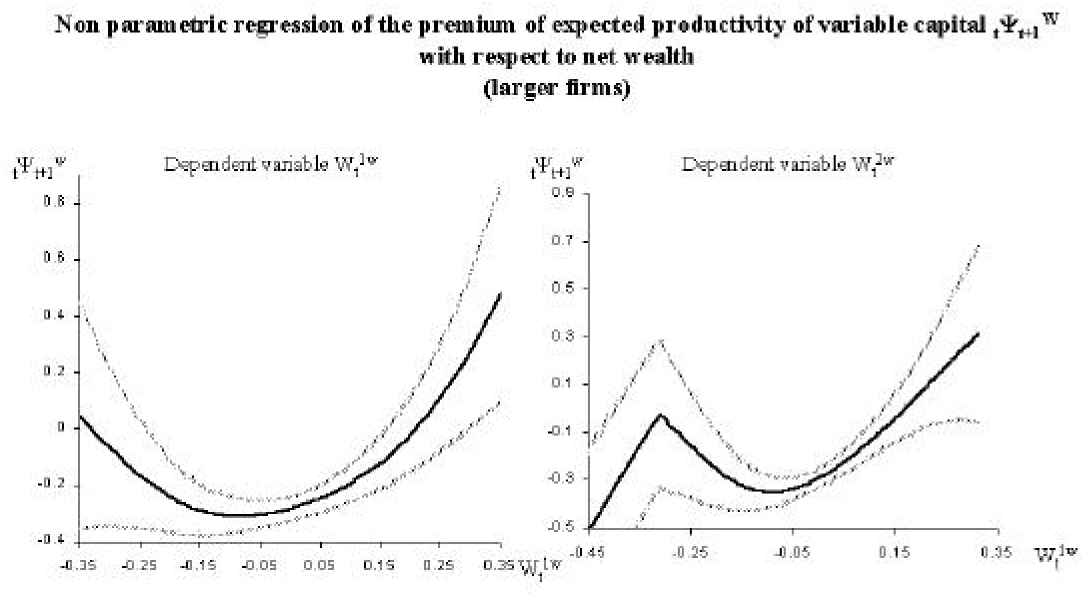
Standard error in parenthesis; 1: More than 300 employees; 2: Less than 300 employees; * Significant at 90% confidence level; ** significant at 95% confidence level; *** significant at 99% confidence level; $ration_i$ = 1 if the entrepreneur stated financing constraints, and 0 otherwise; ?? $^w{}_i$ = average value of the premium in the expected productivity of variable capital; dim_i = dimension in number of employees;

The regression results in figure 8 also suggest that the financing constraints hypothesis should be rejected for larger firms, for which $\overline{\Psi}_i^w$ is not related to financing constraints. Figure 16 confirms this. The estimation of equation (31) rejects the predictions of the model, showing instead an upward sloping relationship between ${}_t\widehat{\Psi}_{i,t+1}^w$ and $w_{i,t}^w$. As before this upward sloping relation is steeper for $w_{i,t}^{1w}$ than for $w_{i,t}^{2w}$ and it could again be caused by the two misspecification problems mentioned before.

7 Conclusions

In this paper we illustrated a structural model of firm investment with financing and irreversibility constraints. The model indicates that the premium of expected marginal productivity over the user cost of variable capital is the best available indicator of financing

Figure 16: Estimated relation between the intensity of financing constraints and financial wealth - large firms



constraints on firm investment. We estimated this indicator for a sample of small and medium Italian manufacturing firms, and we tested and did not reject the presence of financing constraints for all the firms in the sample except the larger ones. The robustness of this result has been confirmed by an independent source of qualitative information available for the same sample: conditional on their size, firms with an high value of the financing constraints indicator are three times more likely to state problems in financing new investment projects than firms with a low value.

The financing constraints test proposed in this paper has two major advantages with respect to previous empirical studies. First, it is theoretically consistent, and robust to the criticism raised by Kaplan and Zingales (1997 and 2000) to the previous studies which detected the presence of financing constraints focusing on the correlation between cash flow and fixed investment. Secondly, even if we ignore this criticism, another important advantage of our approach is that, in the presence of estimation errors, the test is likely to be biased towards rejecting the financing constraints hypothesis when it is true, while the cash flow-investment type of test is likely to be biased towards not rejecting it when it is false. This is because the studies that focus on the correlation between cash flow and fixed investment carry out the following test: if financing constraints affect firms fixed investment decisions, then we predict that the expected marginal productivity of fixed capital (summarised by Tobin's Q) is not a sufficient statistic for determining investment decisions, which are also positively affected by cash flow. The problem is that in practice, with or without financing constraints, investment in fixed capital, Tobin's Q and cash flow are all highly positively correlated, due to the presence of adjustment costs. Therefore if Tobin's Q is not properly estimated, this is likely to increase the significance of cash flow in the investment equation, and hence to bias the test towards not rejecting the financing constraints hypothesis when this is false.

On the contrary our test is based on estimating the correlation between marginal productivity of variable capital and net financial wealth. The financing constraints hypothesis predicts a negative relationship between the two variables. But since variable capital is less subject to adjustment costs than fixed capital then in the absence of financing constraints the two variables should not be correlated. Moreover even if adjustment costs are present, or if we estimate with error the productivity shock, it is more likely that we detect a positive relationship between marginal productivity of variable capital and net financial wealth, rather than the opposite, and hence we would tend to reject the financing constraints hypothesis when this is true.

References

[1] Albuquerque, R. and H.A. Hopenhayn (2000), "Optimal Dynamic Lending Contracts with Imperfect Enforceability", mimeo, University of Rochester.

[2] Arellano, M. and O. Bover (1995), "Another Look at the Instrumental Variable Estimation of Error-Components Models", *Journal of Econometrics*, 68, 29-51.

[3] Besanko, D. and A.V. Thakor (1986), "Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets", *International Economic Review*, 28, 671-89.

[4] Blundell, R. and S. Bond (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics,* 87, 115-43.

[5] Blundell, R. and S. Bond (2000), "GMM Estimation with Persistent Panel Data: An Application to Production Functions", *Econometric Reviews,* 19, 321-40.

[6] Caballero, R.J., Engel, E. and J. Haltiwanger (1995), "Plant-Level Adjustment and Aggregate Investment Dynamics", *Brookings Papers on Economic Activity,* 1-54.

[7] Caggese, A., 2002, "Financing Constraints and Firm Dynamics", PhD Thesis, London School of Economics and Political Science.

[8] Caggese, A., 2003, "Financing Constraints, Irreversibility, and Investment Dynamics", Financial Markets Group Working Paper, London School of Economics.

[9] Cleary, S. (1999), "The Relationship between Firm Investment and Financial Status" *Journal of Finance*, 54, 673-692.

[10] Erickson, T. and T. Whited, 2000, Measurement Error and the Relationship between Investment and q, Journal of Political Economy, 108, 1027-57

[11] Fazzari S.M., Hubbard G.R. and B.C. Petersen (1988), "Financing Constraints and Corporate Investment", Brooking Papers on Economic Activity, 1, 141-195.

[12] Gomes, J., 2001, Financing Investment, American Economic Review, 91, 1263-1285.

[13] Hall, B.H. and J. Mairesse (1996), "Estimating the Productivity of Research and Development: An Exploration of GMM Methods Using Data on French & United States Manufacturing Firms", NBER WP #5501.

[14] Hamilton, B.H., (2000), "Does Entrepreneurship Pay? An Empirical Analysis of the Returns of Self-Employment", Journal of Political Economy, 108, 604-631.

[15] Hart,O and J.H. Moore (1998), "Default and Renegotiation: a Dynamic Model of Debt", *Quarterly Journal of Economics*, 113, 1-41.

[16] Hubbard, G.R. (1998), "Capital-Market Imperfections and Investment", *Journal of Economic Literature*, 36, 193-225.

[17] Jensen, M. (1986), "Agency Costs of Free Cash Flow, Corporate Finance and Takeovers", *American Economic Review,* 76, 323-29.

[18] Kaplan, S.N. and L. Zingales (1997), "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?", *Quarterly Journal of Economics,* 112, 169-215.

[19] Milde, H. and J.G. Riley (1988), "Signalling in Credit Markets", *Quarterly Journal of Economics*, 103, 101-129.

Appendix 1

We describe here the variables used in the estimation of the production function:

$p_t^y Y_{i,t}$: total revenues realised during year t, at current prices.

$p_t^k K_{i,t}$: we compute $p_t^k K_{i,t}$ as the sum of the replacement value of two different kind of fixed capital: i) plants and equipment; ii) intangible fixed capital (Software, Advertising, Research and Development). In the theoretical model we assume that it takes one period for fixed and variable capital to become productive. In reality the time lag necessary to install the new capital is less than one year, and most likely around or less than six months. Therefore we include in $p_t^k K_{i,t}$ all capital purchased before the end of time t. Balance sheet data about fixed assets do not reflect their replacement value, for at least two reasons: first, the depreciation rate applied for accounting purposes is very variable and does not coincide with the physical depreciation rate; second, all values are "historical", and do not take into account the appreciation of the assets in nominal terms. Hence, to compute the replacement value of capital we prefer to adopt the following perpetual inventory method:

$$p_{t+1}^{kj} K_{i,t+1}^j = p_t^{kj} K_{i,t}^j (1+\pi_t^j)(1-\delta^j) + p_{t+1}^{kj} I_{i,t+1}^j$$

J={1,2} , where 1=plant and equipment and 2= intangible fixed capital . $\pi^1 = \%$ change in the producer prices index for agricultural and industrial machinery (source: OECD, from Datastream); $\pi^2 = \%$ change in the producer prices index (source: OECD, from Datastream). δ^j are estimated separately for the 20 manufacturing sectors using aggregate annual data about the replacement value and the total depreciation of the capital (source: Italian National Institute of Statistic). Given that within each sector depreciation rates vary only marginally between years, we conveniently used the yearly average: δ^1 ranges from 9.3% to 10.7%, and δ^2 from 8.4% to 10.6%.

$p_t^l L_{i,t}$: this variable is computed in the following way: beginning of the period t working capital inventories (materials, work in progress and finished products), plus new purchases of materials in period t, minus end of period t working capital inventories. Also in this case the time lag necessary to transform variable inputs in revenues is much less than one year. Therefore we assume that all the variable inputs that are in stock at the beginning of year t will contribute to generate year t revenues. By subtracting the end of period t working capital inventories we also assume that a fraction of the new purchases of materials during period t contributes to period t revenues, while the remaining part represents investment in the variable capital that will become productive in period $t+1$.

$p_t^n N_{i,t}$: this variable includes the total cost of the labour and the services used in year t.

In order to transform the variables in real terms, we used the following price indexes (source: ISTAT):

Output $Y_{i,t}^T$: consumer prices index relative to all products excluding services.

Fixed capital $K_{i,t}$: producer price index of durable inputs.

Labour $N_{i,t}$: wage earnings index of the manufacturing sector.

Variable capital $L_{i,t}$: wholesale price index for intermediate goods.

Appendix 2

This appendix describes the testing procedure adopted to determine the appropriate estimation method of equation (22). The time dimension of the data, 11 annual observations, is too short to allow the consistent estimation of $\ln A_i$ and of the moments of the distribution of $\ln \chi_{s,t}$ and $\ln \varepsilon_t$. Given that the number of firms in the sample is large we can estimate $\ln \varepsilon_t$ and $\ln \chi_{s,t}$ as fixed effects. We can also transform the data to eliminate the unobservable $\ln A_i$. The firm idiosyncratic shock $\ln \theta_{i,t}^f$ can neither be estimated as a fixed effect, nor eliminated through a transformation of the data. In the theoretical model we assume that $\ln \theta_{i,t}^f$ is not observed by E before she decides $k_{i,t}$ and $l_{i,t}$ at time t. If this is true, and if $\rho = 0$, then $cov(\ln \theta_{i,t}^f, \ln z_{i,t}) = 0$ for $z \in \{k, l, n\}$. Unfortunately this is not necessarily true in reality. Even assuming that $\rho = 0$, we can still expect $\ln \theta_{i,t}^f$ to be at least partially correlated with $\ln z_{i,t}$. This is because the duration of a cycle of production is most likely lower than one year, that is the frequency of our data. In order to correct this problem we use an instrumental variables estimation technique. Lagged $\ln z_{i,t-j}$ with $z \in \{k, l, n\}$ and with $j \geq 1$ are natural candidates as instruments, but their validity depends on the degree of serial correlation in $\ln \theta_{i,t}^f$. In practice some of the persistency in productivity shocks is likely to be captured by the economy wide and industry specific shocks $\ln \varepsilon_t$ and $\ln \chi_{s,t}$. Moreover the permanent differences in $\ln \theta_{i,t}^f$ between firms are captured by the fixed effect A_i. Therefore the residual persistency of $\ln \theta_{i,t}^f$ should be quite low, and this means that lagged right hand side variables can be valid instruments. We test the exogeneity of $\{z_{i,1}, ..., z_{i,T}\}$, for $z \in (k, l, n)$ by estimating the linearised system (22) with a GMM estimator. This allows, when the number of instruments is greater than the number of parameters to estimate, to test the validity of the instruments with the Sargan test of overidentifying restrictions. We choose as instruments two lags of the right hand side variables[35].

In figure 17 we compare the tests for the overidentifying restrictions obtained using lags -1 and -2 with the one obtained using lags 0 and -1. If independent variables $\{z_{i,1}, ..., z_{i,T}\}$ are contemporaneously correlated with $\theta_{i,t}^f$, but the persistency of $\ln \theta_{i,t}^f$ is not very high, then we expect only the 0 and -1 instruments to be rejected. Figure 17 shows that both sets of instruments are not rejected in four out of seven groups, while in the remaining three the lags 0&1 specification is close to rejection, with a P-value around 0.10-0.18. Given that the j-test is usually biased towards accepting the model when it should be rejected, we interpret this result as evidence of some endogeneity problem, and we decide to adopt the lags 1&2 specification, and to use it on all the groups, for homogeneity.

[35] We prefer not to increase the number of lags because additional lagged instruments did not improve the efficiency of the estimates. Therefore, given that the number of firms per group is relatively small, we prefer not to reduce excessively the number of degrees of freedom.

Figure 17: Test of the validity of the instruments

Sargan test of overidentifying restrictions

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
				Instruments: $z_{i,t-s}$ with s=0,1 and z={k,l,n}			
j test	35.1	44.5	35.1	43.5	47	45.9	35.7
d.f.	37	37	37	37	36*	36*	36*
Pr.(j)<?^{2}(d.f.)	0.56	0.18	0.56	0.22	0.1	0.12	0.48
				Instruments: $z_{i,t-s}$ with s=1,2 and z={k,l,n}			
j test	38.9	37.8	39.87	39.71	41.26	42.2	33.64
d.f.	37	37	37	37	36*	36*	36*
Pr.(j)<?^{2}(d.f.)	0.38	0.43	0.34	0.35	0.25	0.22	0.58
N. firms	78	49	117	63	80	66	108
N. observations	624	392	936	504	640	528	864

*One coefficient relative to a two-digit sector dummy variable is estimated

The discussion above and figure 17 justify the choice of the Generalised Method of Moments (GMM) as the estimation method[36]. We first eliminate the firm specific effect A_i. The within-firm transformation would be the obvious choice to do it, but unfortunately it is not consistent when we use lagged right hand side variables as instruments to correct for the correlation between $\ln \theta^f_{i,t}$ and $\ln z_{i,t}$, for $z \in \{k, l, n\}$. We therefore adopt the forward orthogonal transformation proposed by Arellano and Bover (1995). This is equivalent to a "forward within transformation" that remains consistent when lagged instruments are used[37]. We then stack the observations in 8 cross sectional equations, for the years 1985 to 1992. This means that we exclude year 1982, in order to diminish possible distortions caused by the perpetual inventory method, and we have the data from 1983 and 1984 available as instruments. We estimate the system (32), where the symbol * denotes the transformed variables, imposing the equality of parameters across equations:

$$\begin{array}{l} \ln y^*_{i,92} = c_{92} + d_{ts} + \alpha \ln k^*_{i,92} + \beta \ln l^*_{i,92} + \gamma \ln n^*_{i,92} + \ln \theta_{i,92} \\ \ln y^*_{i,91} = c_{91} + d_{ts} + \alpha \ln k^*_{i,91} + \beta \ln l^*_{i,91} + \gamma \ln n^*_{i,91} + \ln \theta_{i,91} \\ . \\ \ln y^*_{i,85} = c_{85} + d_{ts} + \alpha \ln k^*_{i,85} + \beta \ln l_{i,85} + \gamma \ln n_{i,85} + \ln \theta_{i,85} \end{array} \tag{32}$$

d and c are two digit I.S.T.A.T. sector and year specific dummy variables respectively, and capture the effect of $\ln \chi_{s,t}$ and $\ln \varepsilon_t$. Figure 4 reports estimation results using $\ln k_{i,t-j}, \ln l_{i,t-j}$ and $\ln n_{i,t-j}$ with $j \in \{1, 2\}$, as instrumental variables[38].

[36]Such method is used for a similar problem by Mairesse and Hall (1996).

[37]The transformed variable is the following: $z^*_{i,t} = \left(\frac{T-t+1}{T-t+2}\right)^{1/2} \left[z_{i,t-1} - \frac{1}{T-t+1}\left(z_{i,t} + z_{i,t+1} + ... + z_{i,T}\right)\right]$.

[38]Whenever possible one group is composed by one specific two digit I.S.T.A.T. sector. This is the case for groups 1 and 5. Hence the coefficient d_{ts} is omitted, in that it would be perfectly collinear with the constant c. The other groups are composed by firms in more than one 2-digits sector, because each sector has a too low number of firms. Here we include the coefficient d_{ts} only if it shows a significant deviation from the constant.